

I. Uses of this article

- A. Whitepaper
- B. Publish in PDMA Visions magazine
- C. Publish in RTM

II. Introduction

- A. **Scoring rules are a very powerful tool and many companies use them in to support New Product Development. Like many powerful tools, they are often applied in places where they do not really apply—the proverbial hammer looking for a nail. Used in inappropriate ways, they lead to increased cost, bureaucracy and ultimately the tool itself is rejected. These limitations appear most acutely when scoring rules are used to answer the questions of where and how much to invest. This article explores some of the limitation of the scoring rules approach and provides well-established principles for overcoming the limitations.**
- B. **There are many ways scoring rules are deployed in organizations. Most deployments share most of the following characteristics:**
 - 1. **Metric-based:** the company develops a variety of metrics, such as time-to-market, commercial potential, cost, fit to strategy, risk, etc. (See RTM article on 37 metrics).
 - 2. **Subjective Scoring:** projects are scored on each metric based on a 1-5 (or similar scale). For example, on technical success, the scale might be 1 = new to world technology, 3 = challenging but in our competency area, and 5 = straightforward engineering. There is a wide variety of definitions of the scales with varying degrees of sophistication. Some match your answers to questions to a statistical database of other projects to product a score. (reference RTM article)
 - 3. **Figure of merit for each project:** The scores on each of the metrics are combined, for example by adding or taking a weighted average, to produce a figure of merit for each project. This figure of merit is used to judge the attractiveness of the project compared to others in the portfolio.
 - 4. **Intricate charting:** The data is combined in a wide variety of ways to produce an array of portfolio displays, such as bubble charts. Often the data is used to “slice and dice” the portfolio—looking at different cuts. These charts are usually used to judge individual projects and set their priority, as well as to ensure “balance” in the portfolio.

III. This approach works very well in some circumstances.

- A. **For addressing large numbers of projects: Scoring rules are easy to understand. Administration can be straightforward (although not always easy)—you mainly ask project leaders to rate their projects. Therefore it can be applied fairly rapidly over a large number of projects.**
- B. **For rough screening: When you are just trying to separate the wheat from the chaff, scoring rules work well. It serves as a mental checklist, helping think through different aspects of their project. If an idea is weak on several dimensions, it**

probably needs to be discarded or refined, and the scoring helps people focus their thinking. For very early stage projects, scoring rules can be particularly helpful in driving conversations about what would otherwise be gut feel.

- C. For organizations at a low maturity level of portfolio management: If your company is just getting started, a simple scoring rules system can help get people talking about the right topics. As the maturity of the organization evolves, however, the limitations of scoring rules start to become apparent.**

IV. Scoring rules have several significant limitations:

- A. The CFO of a major NPD-intensive organization, Eric, once told me the following story.**

1. The NPD organization is presenting me with beautiful charts. They ranked the all the projects in the portfolio on every imaginable dimension. They showed that their portfolio was balanced. On the basis of this, the leader, Allen, asked for authorization to spend, and hinted that his organization could use more budget very productively.
2. I know that we have a lot of great projects, and think perhaps we are investing too much in other areas. NPD is our lifeblood and perhaps deserves higher budget. I had participated in the creation of the scoring rules to help guide NPD. Now things were getting real: we were talking about competing for resources. These scoring rules were designed to make this process easier. At this moment of truth, I realized it couldn't work, unless we could answer a few simple questions:
3. So I asked Allen my three simple questions: "What does this mean? What are the implications of funding a project with a score of 65%?" "Why should I believe this? What is the basis of the inputs?" and "Why did you draw the line where you drew it?"
4. Allen was unable to answer any of these questions:
 - a) **No meaning:** He could not translate the score into any real impact on the business. Since the scores combined so many dimensions, he could not tell me if a high score meant better financial performance, building an important core platform for future use, or anything else. The scores were meaningless.
 - b) **No basis:** He could not provide a basis for why I should believe the scores (even if I accepted them as a figure of merit). "We surveyed each of the project leaders and they ranked their projects. Then we discussed the differences in a group and settled on these scores." We had simply increased the sophistication of asking people for their opinion about the merit of the projects. But we had not really changed the basis of information for our decisions.
 - c) **No discrimination:** He could not tell me why he drew the line where he drew it. All the projects were tightly clustered, in the 60-70% range, with the cut off point at 65%. The difference between the project above the line and below the line was less than 1%. Allen agreed that the quality of our inputs did not support the level of discrimination we were making. We agreed that a project scoring 50% was worse than one scoring 70%. But the difference between a projects scoring 62% and one scoring 63% was absolutely meaningless. Projects scoring 50% had all been screened out long ago.

5. We need a better way. The scoring rule system got us started, but when it comes to making investments decisions about the portfolio, we need something that has real meaning, has a solid and auditable basis, and discriminates among projects.

B. Eric’s story illustrates the limitations of scoring rules when used too broadly. Each of these limitations is an inevitable consequence of the structure of scoring rule-based systems:

1. **No meaning:** When it comes to investment questions, you need to know what you get for your investment. By taking a weighted average of many dimensions, scoring rules provide a figure of merit that has no interpretation. Indeed, the mathematics of probability in business evaluation requires that some of the dimensions, like probability of success and commercial potential should be multiplied. Scoring rules typically add these dimensions. So not only does the figure of merit lack meaning, it is often wrong.
2. **No basis:** Asking project leaders for their assessments of their projects seems to provide data. But in a competitive situation, this process is so laden with bias and misinterpretation as to render it meaningless. Project leaders often feel that they are being asked to provide the rope that will be used to hang them. To see this, think about this scenario: you and twenty others are competing for funding for your projects. Half of you will not be funded. You are each asked to provide “unbiased assessments” of your projects on various dimensions: “commercial attractiveness”, “technical feasibility”, etc. What are you going to say? Your assessment is essentially unauditable, so the scoring rule approach just becomes another way to play the budgeting game.
3. **No discrimination:** Many of you may have discovered that the more dimensions you rate projects on, the less discriminating the scoring rules become. This is an inevitable consequence of a well-established statistical law called the law of large numbers. The law of large numbers basically says that if you add up scores for large number of dimensions for a project, all projects will converge towards the mean score. The more dimensions, the stronger the convergence. Practically speaking, this happens above around dimensions. To be sure, we must be careful about pushing this law too far: really bad ideas will be weeded out by a scoring rule system. But once projects are all reasonable, each will be better and worse on various dimensions, so these projects will converge most strongly. So exactly where we look to a scoring rule to provide the most insight—in helping to make the tough calls—is exactly where it breaks down!

V. A way forward:

- A. **The core idea behind scoring rules, of measuring and comparing projects, is essential for delivering high-impact projects and prioritizing among them. So how can we implement this crucial idea of measurement in the many places where the scoring rule approach breaks down? Here are a few principles to follow:**
- B. **Focus on a very few metrics and “true north”**
 1. Fewer, more meaningful metrics is the somewhat counter-intuitive answer to overcoming the limitations of scoring rules.

2. First, develop a single, overriding metric of value for your company, something we call “true north”. All other metrics are understood as waypoints towards this true north. True north in most situations should be a single expected net present value of cash flows type of measure. There are many poor ways of crafting this true north metric, and the details of a good one are beyond the scope of this article.
3. Second, have a few supporting metrics that directly tie to true north—say three or four. These do not substitute for true north, but clarify and illuminate how specific projects deliver value: some will be incremental sure things, others will be long shots. In virtually every situation I have seen, the following are completely sufficient: Commercial Potential, Probability of Technical Success, Resources and Timing.
4. Strategic fit sometimes comes up, but it is a debatable and dangerous metric. When people defend projects that are poor on commercial potential, have low probabilities of technical success, or that cost a lot on strategic grounds, they are usually hiding something. “Strategic Fit” is highly ambiguous and highly subject to manipulation, so use it with great care. That said, it can be helpful in rough screening of projects and in other circumstances.
5. When it comes to metrics, less is more.

C. Model, don't average

1. As we have seen, the weighted average approach implicit in score rules creates problems. The alternative to this is to model the business structure of opportunity. For the few metrics, arrange them in a logical fashion. For example, to get an overall true north metric use $\text{Value} = \text{Probability of Success} * \text{Commercial Potential} - \text{Cost}$. If you want a return on investment measure as your figure of merit, use $\text{Value} / \text{Cost}$. Note how differently these simple equations look from the scoring rule equation of $\text{Score} = (\text{Probability of Success} + \text{Commercial Potential} - \text{Cost}) / 3$. The equation that models the business opportunity has a meaningful interpretation; the scoring rule does not.
2. As part of this, you should also model the individual opportunities to get the few metrics of Commercial Potential, etc. While you may use metrics from your long list, like peak sales or market share, use them in a model appropriate to the opportunity. A new product opportunity needs a good discussion of its potential for market share; a process improvement does not. With the modeling approach, you only measure projects on the metrics that are relevant.
3. Many companies find that once they have developed a few templates for their opportunities, this modeling effort is very straightforward. In any given business area, there are only a handful of templates, for example “razors and blades business”, “cost reductions”, “incremental product improvement” and so on.
4. This approach leverages the process extremely efficiently in two important ways. First, the templates level the playing field because they embody a comparable approach to evaluation and because they often incorporate standardized forecasts for common uncertainties, such as the size of the market or prices of key commodities. Second, the templates scale efficiently to address more complex projects in an integrated way. Projects that are too novel or complex to use the templates are those

that deserve extra attention, so you put analysis and modeling effort where it is needed.

D. Treat uncertainty explicitly

1. Scoring rules often use risk or uncertainty as a dimension. This very much contributes to the limitations of scoring rules. It is much better to treat uncertainty explicitly as scenarios or ranges. This changes the fundamental question from “how risky is this project?” to “what is the range of possible outcomes?”
2. A quick illustration: Consider a very risky project, scoring a 10 on the risk scale. Suppose this risk is coming from market uncertainty—we think the project could produce between \$100M and a \$1B business. The risk score is high because there is an order of magnitude range on commercial success, and this penalizes the project. But that is the wrong answer—the project is a big, high-impact, high-upside project. The issue is not about pursuing the opportunity, but rather, about what we can do now to drive towards the upside and to reduce the range of uncertainty.
3. This approach improves the basis of information, because the scenarios are being developed based not on what will justify the project, but on an estimate of what is known. These forecasts are auditable and supportable, and while open to discussion, move the basis of information to the next level. The inputs are clearly grounded.
4. Once you have looked at the range of uncertainty for each of the inputs, run them through your model to find the impact of these uncertainties on your project. There are many well-established efficient decision-analysis based methods to do this, for example, monte carlo simulation, decision trees, and real options. The technical merits of each of these approaches is beyond the scope of this article.
5. Once you understand the implications of the uncertainties on the project, you should use it to improve the projects. Figure out ways to reduce big uncertainties early in the project plan and cheaply. Focus your thinking on ways you might be able to achieve the upside.

E. Make the process add value to everyone

1. Most scoring rule systems treat the project leaders as data sources supporting the portfolio process. In my experience, project leaders don't like being treated this way and simply play along because they have to. One of the keys to overcoming the limitations of scoring rules is to conduct your process so that it first adds value at the project level—so that project teams see the evaluation process as helpful. Then they want to participate, and see it is a way to improve their projects. Once project leaders are not feeling defensive, they take the analysis results more seriously. We have experience several situations where project leaders cancel their own projects because they realize it is unlikely to produce great benefits.
2. By conducting the project-level evaluations in the right way (using few metrics, modeling, and treating uncertainty explicitly), you will get the portfolio data needed as a natural side effect. And when you make decisions using this information, project leaders will be more supportive because they have had real help making their projects the best they could possibly be.

VI. Conclusion

- A. These principles have been successfully applied in many organizations, and have been extensively written up, for example in my book *The Smart Organization*. Done well, they produce clear benefits in a streamlined way. For many companies, projects (and portfolios) typically improve by 30% or more, creating millions in cost savings or improved revenue.**
- B. If you are looking to answer the questions of how much and where to invest, do what it takes to answer that question. Scoring rules taken too far will lead you astray and be far more difficult and costly to use. Use scoring rules for rough screening; use hammers for pounding nails. Set yourself and your organization up for success.**