

# The Outside View

A few years after my collaboration with Amos began, I convinced some officials in the Israeli Ministry of Education of the need for a curriculum to teach judgment and decision making in high schools. The team that I assembled to design the curriculum and write a textbook for it included several experienced teachers, some of my psychology students, and Seymour Fox, then dean of the Hebrew University's School of Education, who was an expert in curriculum development.

After meeting every Friday afternoon for about a year, we had constructed a detailed outline of the syllabus, had written a couple of chapters, and had run a few sample lessons in the classroom. We all felt that we had made good progress. One day, as we were discussing procedures for estimating uncertain quantities, the idea of conducting an exercise occurred to me. I asked everyone to write down an estimate of how long it would take us to submit a finished draft of the textbook to the Ministry of Education. I was following a procedure that we already planned to incorporate into our curriculum: the proper way to elicit information from a group is not by starting with a public discussion but by confidentially collecting each person's judgment. This procedure makes better use of the knowledge available to members of the group than the common practice of open discussion. I collected the estimates and jotted the results on the blackboard. They were narrowly centered around two years; the low end was one and a half, the high end two and a half years.

Then I had another idea. I turned to Seymour, our curriculum expert, and asked whether he could think of other teams similar to ours that had developed a curriculum from scratch. This was a time when several pedagogical innovations like "new math" had been introduced, and Seymour said he could think of quite a few. I then asked whether he knew the history of these teams in some detail, and it turned out that he was familiar with several. I asked him to think of these teams when they had made as much progress as we had. How long, from that point, did it take them to finish their textbook projects?

He fell silent. When he finally spoke, it seemed to me that he was blushing, embarrassed by his own answer: "You know, I never realized this before, but in fact not all the teams at a stage comparable to ours ever did complete their task. A substantial fraction of the teams ended up failing to finish the job."

This was worrisome; we had never considered the possibility that we might fail. My anxiety rising, I asked how large he estimated that fraction was. "About 40%," he answered. By now, a pall of gloom

was falling over the room. The next question was obvious: "Those who finished," I asked. "How long did it take them?" "I cannot think of any group that finished in less than seven years," he replied, "nor any that took more than ten."

I grasped at a straw: "When you compare our skills and resources to those of the other groups, how good are we? How would you rank us in comparison with these teams?" Seymour did not hesitate long this time. "We're below average," he said, "but not by much." This came as a complete surprise to all of us—including Seymour, whose prior estimate had been well within the optimistic consensus of the group. Until I prompted him, there was no connection in his mind between his knowledge of the history of other teams and his forecast of our future.

Our state of mind when we heard Seymour is not well described by stating what we "knew." Surely all of us "knew" that a minimum of seven years and a 40% chance of failure was a more plausible forecast of the fate of our project than the numbers we had written on our slips of paper a few minutes earlier. But we did not acknowledge what we knew. The new forecast still seemed unreal, because we could not imagine how it could take so long to finish a project that looked so manageable. No crystal ball was available to tell us the strange sequence of unlikely events that were in our future. All we could see was a reasonable plan that should produce a book in about two years, conflicting with statistics indicating that other teams had failed or had taken an absurdly long time to complete their mission. What we had heard was base-rate information, from which we should have inferred a causal story: if so many teams failed, and if those that succeeded took so long, writing a curriculum was surely much harder than we had thought. But such an inference would have conflicted with our direct experience of the good progress we had been making. The statistics that Seymour provided were treated as base rates normally are—noted and promptly set aside.

We should have quit that day. None of us was willing to invest six more years of work in a project with a 40% chance of failure. Although we must have sensed that persevering was not reasonable, the warning did not provide an immediately compelling reason to quit. After a few minutes of desultory debate, we gathered ourselves together and carried on as if nothing had happened. The book was eventually completed eight(!) years later. By that time I was no longer living in Israel and had long since ceased to be part of the team, which completed the task after many unpredictable vicissitudes. The initial enthusiasm for the idea in the Ministry of Education had waned by the time the text was delivered and it was never used.

This embarrassing episode remains one of the most instructive experiences of my professional life. I eventually learned three lessons from

it. The first was immediately apparent: I had stumbled onto a distinction between two profoundly different approaches to forecasting, which Amos and I later labeled the inside view and the outside view. The second lesson was that our initial forecasts of about two years for the completion of the project exhibited a planning fallacy. Our estimates were closer to a best-case scenario than to a realistic assessment. I was slower to accept the third lesson, which I call irrational perseverance: the folly we displayed that day in failing to abandon the project. Facing a choice, we gave up rationality rather than give up the enterprise.

## Drawn to the Inside View

On that long-ago Friday, our curriculum expert made two judgments about the same problem and arrived at very different answers. The *inside view* is the one that all of us, including Seymour, spontaneously adopted to assess the future of our project. We focused on our specific circumstances and searched for evidence in our own experiences. We had a sketchy plan: we knew how many chapters we were going to write, and we had an idea of how long it had taken us to write the two that we had already done. The more cautious among us probably added a few months to their estimate as a margin of error.

Extrapolating was a mistake. We were forecasting based on the information in front of us—WYSIATI—but the chapters we wrote first were probably easier than others, and our commitment to the project was probably then at its peak. But the main problem was that we failed to allow for what Donald Rumsfeld famously called the “unknown unknowns.” There was no way for us to foresee, that day, the succession of events that would cause the project to drag out for so long. The divorces, the illnesses, the crises of coordination with bureaucracies that delayed the work could not be anticipated. Such events not only cause the writing of chapters to slow down, they also produce long periods during which little or no progress is made at all. The same must have been true, of course, for the other teams that Seymour knew about. The members of those teams were also unable to imagine the events that would cause them to spend seven years to finish, or ultimately fail to finish, a project that they evidently had thought was very feasible. Like us, they did not know the odds they were facing. There are many ways for any plan to fail, and although most of them are too improbable to be anticipated, the likelihood that *something* will go wrong in a big project is high.

The second question I asked Seymour directed his attention away from us and toward a class of similar cases. Seymour estimated the base rate of success in that reference class: 40% failure and seven to ten years for

completion. His informal survey was surely not up to scientific standards of evidence, but it provided a reasonable basis for a baseline prediction: the prediction you make about a case if you know nothing except the category to which it belongs. As we saw earlier, the *baseline prediction* should be the anchor for further adjustments. If you are asked to guess the height of a woman about whom you know only that she lives in New York City, your baseline prediction is your best guess of the average height of women in the city. If you are now given case-specific information, for example that the woman's son is the starting center of his high school basketball team, you will adjust your estimate away from the mean in the appropriate direction. Seymour's comparison of our team to others suggested that the forecast of our outcome was slightly worse than the baseline prediction, which was already grim.

The spectacular accuracy of the outside-view forecast in our problem was surely a fluke and should not count as evidence for the validity of the *outside view*. The argument for the outside view should be made on general grounds: if the reference class is properly chosen, the outside view will give an indication of where the ballpark is, and it may suggest, as it did in our case, that the inside-view forecasts are not even close to it.

For a psychologist, the discrepancy between Seymour's two judgments is striking. He had in his head all the knowledge required to estimate the statistics of an appropriate reference class, but he reached his initial estimate without ever using that knowledge. Seymour's forecast from his inside view was not an adjustment from the baseline prediction, which had not come to his mind. It was based on the particular circumstances of our efforts. Like the participants in the Tom W experiment, Seymour knew the relevant base rate but did not think of applying it.

Unlike Seymour, the rest of us did not have access to the outside view and could not have produced a reasonable baseline prediction. It is noteworthy, however, that we did not feel we needed information about other teams to make our guesses. My request for the outside view surprised all of us, including me! This is a common pattern: people who have information about an individual case rarely feel the need to know the statistics of the class to which the case belongs.

When we were eventually exposed to the outside view, we collectively ignored it. We can recognize what happened to us; it is similar to the experiment that suggested the futility of teaching psychology. When they made predictions about individual cases about which they had a little information (a brief and bland interview), Nisbett and Borgida's students completely neglected the global results they had just learned. "Pallid" statistical information is routinely discarded when it is incompatible with

one's personal impressions of a case. In the competition with the inside view, the outside view doesn't stand a chance.

The preference for the inside view sometimes carries moral overtones. I once asked my cousin, a distinguished lawyer, a question about a reference class: "What is the probability of the defendant winning in cases like this one?" His sharp answer that "every case is unique" was accompanied by a look that made it clear he found my question inappropriate and superficial. A proud emphasis on the uniqueness of cases is also common in medicine, in spite of recent advances in evidence-based medicine that point the other way. Medical statistics and baseline predictions come up with increasing frequency in conversations between patients and physicians. However, the remaining ambivalence about the outside view in the medical profession is expressed in concerns about the impersonality of procedures that are guided by statistics and checklists.

## The Planning Fallacy

In light of both the outside-view forecast and the eventual outcome, the original estimates we made that Friday afternoon appear almost delusional. This should not come as a surprise: overly optimistic forecasts of the outcome of projects are found everywhere. Amos and I coined the term *planning fallacy* to describe plans and forecasts that

- are unrealistically close to best-case scenarios
- could be improved by consulting the statistics of similar cases

Examples of the planning fallacy abound in the experiences of individuals, governments, and businesses. The list of horror stories is endless.

- In July 1997, the proposed new Scottish Parliament building in Edinburgh was estimated to cost up to £40 million. By June 1999, the budget for the building was £109 million. In April 2000, legislators imposed a £195 million "cap on costs." By November 2001, they demanded an estimate of "final cost," which was set at £241 million. That estimated final cost rose twice in 2002, ending the year at

£294.6 million. It rose three times more in 2003, reaching £375.8 million by June. The building was finally completed in 2004 at an ultimate cost of roughly £431 million.

- A 2005 study examined rail projects undertaken worldwide between 1969 and 1998. In more than 90% of the cases, the number of passengers projected to use the system was overestimated. Even though these passenger shortfalls were widely publicized, forecasts did not improve over those thirty years; on average, planners overestimated how many people would use the new rail projects by 106%, and the average cost overrun was 45%. As more evidence accumulated, the experts did not become more reliant on it.
- In 2002, a survey of American homeowners who had remodeled their kitchens found that, on average, they had expected the job to cost \$18,658; in fact, they ended up paying an average of \$38,769.

The optimism of planners and decision makers is not the only cause of overruns. Contractors of kitchen renovations and of weapon systems readily admit (though not to their clients) that they routinely make most of their profit on additions to the original plan. The failures of forecasting in these cases reflect the customers' inability to imagine how much their wishes will escalate over time. They end up paying much more than they would if they had made a realistic plan and stuck to it.

Errors in the initial budget are not always innocent. The authors of unrealistic plans are often driven by the desire to get the plan approved—whether by their superiors or by a client—supported by the knowledge that projects are rarely abandoned unfinished merely because of overruns in costs or completion times. In such cases, the greatest responsibility for avoiding the planning fallacy lies with the decision makers who approve the plan. If they do not recognize the need for an outside view, they commit a planning fallacy.

## **Mitigating the Planning Fallacy**

The diagnosis of and the remedy for the planning fallacy have not changed since that Friday afternoon, but the implementation of the idea has come a long way. The renowned Danish planning expert Bent Flyvbjerg, now at Oxford University, offered a forceful summary:

The prevalent tendency to underweight or ignore distributional information is perhaps the major source of error in forecasting. Planners should therefore make every effort to frame the

forecasting problem so as to facilitate utilizing all the distributional information that is available.

This may be considered the single most important piece of advice regarding how to increase accuracy in forecasting through improved methods. Using such distributional information from other ventures similar to that being forecasted is called taking an “outside view” and is the cure to the planning fallacy.

The treatment for the planning fallacy has now acquired a technical name, *reference class forecasting*, and Flyvbjerg has applied it to transportation projects in several countries. The outside view is implemented by using a large database, which provides information on both plans and outcomes for hundreds of projects all over the world, and can be used to provide statistical information about the likely overruns of cost and time, and about the likely underperformance of projects of different types.

The forecasting method that Flyvbjerg applies is similar to the practices recommended for overcoming base-rate neglect:

1. Identify an appropriate reference class (kitchen renovations, large railway projects, etc.).
2. Obtain the statistics of the reference class (in terms of cost per mile of railway, or of the percentage by which expenditures exceeded budget). Use the statistics to generate a baseline prediction.
3. Use specific information about the case to adjust the baseline prediction, if there are particular reasons to expect the optimistic bias to be more or less pronounced in this project than in others of the same type.

Flyvbjerg’s analyses are intended to guide the authorities that commission public projects, by providing the statistics of overruns in similar projects. Decision makers need a realistic assessment of the costs and benefits of a proposal before making the final decision to approve it. They may also wish to estimate the budget reserve that they need in anticipation of overruns, although such precautions often become self-fulfilling prophecies. As one official told Flyvbjerg, “A budget reserve is to contractors as red meat is to lions, and they will devour it.”

Organizations face the challenge of controlling the tendency of executives competing for resources to present overly optimistic plans. A well-run organization will reward planners for precise execution and

penalize them for failing to anticipate difficulties, and for failing to allow for difficulties that they could not have anticipated—the unknown unknowns.

## **Decisions and Errors**

That Friday afternoon occurred more than thirty years ago. I often thought about it and mentioned it in lectures several times each year. Some of my friends got bored with the story, but I kept drawing new lessons from it. Almost fifteen years after I first reported on the planning fallacy with Amos, I returned to the topic with Dan Lovallo. Together we sketched a theory of decision making in which the optimistic bias is a significant source of risk taking. In the standard rational model of economics, people take risks because the odds are favorable—they accept some probability of a costly failure because the probability of success is sufficient. We proposed an alternative idea.

When forecasting the outcomes of risky projects, executives too easily fall victim to the planning fallacy. In its grip, they make decisions based on delusional optimism rather than on a rational weighting of gains, losses, and probabilities. They overestimate benefits and underestimate costs. They spin scenarios of success while overlooking the potential for mistakes and miscalculations. As a result, they pursue initiatives that are unlikely to come in on budget or on time or to deliver the expected returns—or even to be completed.

In this view, people often (but not always) take on risky projects because they are overly optimistic about the odds they face. I will return to this idea several times in this book—it probably contributes to an explanation of why people litigate, why they start wars, and why they open small businesses.

## **Failing a Test**

For many years, I thought that the main point of the curriculum story was what I had learned about my friend Seymour: that his best guess about the future of our project was not informed by what he knew about similar projects. I came off quite well in my telling of the story, in which I had the role of clever questioner and astute psychologist. I only recently realized that I had actually played the roles of chief dunce and inept leader.

The project was my initiative, and it was therefore my responsibility to ensure that it made sense and that major problems were properly discussed by the team, but I failed that test. My problem was no longer the planning fallacy. I was cured of that fallacy as soon as I heard Seymour's statistical summary. If pressed, I would have said that our earlier estimates



had been absurdly optimistic. If pressed further, I would have admitted that we had started the project on faulty premises and that we should at least consider seriously the option of declaring defeat and going home. But nobody pressed me and there was no discussion; we tacitly agreed to go on without an explicit forecast of how long the effort would last. This was easy to do because we had not made such a forecast to begin with. If we had had a reasonable baseline prediction when we started, we would not have gone into it, but we had already invested a great deal of effort—an instance of the sunk-cost fallacy, which we will look at more closely in the next part of the book. It would have been embarrassing for us—especially for me—to give up at that point, and there seemed to be no immediate reason to do so. It is easier to change directions in a crisis, but this was not a crisis, only some new facts about people we did not know. The outside view was much easier to ignore than bad news in our own effort. I can best describe our state as a form of lethargy—an unwillingness to think about what had happened. So we carried on. There was no further attempt at rational planning for the rest of the time I spent as a member of the team—a particularly troubling omission for a team dedicated to teaching rationality. I hope I am wiser today, and I have acquired a habit of looking for the outside view. But it will never be the natural thing to do.

## **Speaking of the Outside View**

“He’s taking an inside view. He should forget about his own case and look for what happened in other cases.”

“She is the victim of a planning fallacy. She’s assuming a best-case scenario, but there are too many different ways for the plan to fail, and she cannot foresee them all.”

“Suppose you did not know a thing about this particular legal case, only that it involves a malpractice claim by an individual against a surgeon. What would be your baseline prediction? How many of these cases succeed in court? How many settle? What are the amounts? Is the case we are discussing stronger or weaker than similar claims?”

“We are making an additional investment because we do not

want to admit failure. This is an instance of the sunk-cost fallacy.”