produced it does not have the right brand name. Perhaps the classic example of this is Gregor Mendel, who found his work on heredity ignored, at least in part, because he was an unknown monk and who, as a result, simply stopped publishing his results.

The point is not that reputation should be irrelevant. A proven record of achievement does—and should—confer credibility on a person's ideas. The point instead is that reputation should not become the basis of a scientific hierarchy. The genius of the scientific ethos, at least in theory, is its resolute commitment to meritocracy. As Merton wrote in a famous essay on scientific norms, "The acceptance or rejection of claims entering the lists of science is not to depend on the personal or social attributes of the protagonist; his race, nationality, religion, class, and personal qualities are irrelevant." Ideas are meant to triumph not because of who is (or who is not) advocating them but because of their inherent value, because they seem to explain the data better than any of the others. This is perhaps just an illusion. But it's a valuable one.

9.

COMMITTEES, JURIES, AND TEAMS: THE COLUMBIA DISASTER AND HOW SMALL GROUPS CAN BE MADE TO WORK

On the morning of January 21, 2003, the Mission Management Team (MMT) for NASA mission STS-107—the twenty-eighth flight of the space shuttle *Columbia*—held a teleconference, its second since the *Columbia*'s launch on January 16. An hour before the meeting, Don McCormack had been briefed by members of the Debris Assessment Team (DAT), a group of engineers from NASA, Boeing, and Lockheed Martin, who had spent much of the previous five days evaluating the possible consequences of a large-debris strike on the *Columbia*. During the shuttle's ascent into the atmosphere, a large piece of foam had broken off the left bipod area of the shuttle's external fuel tank and had smashed into the ship's left wing. None of the cameras that were tracking the shuttle's launch had provided a clear picture of the impact, so it was difficult to tell how much damage the foam might have caused. And although by January 21 a request had been made for on-orbit pictures of the *Columbia*, they had not been approved. So the DAT had done what it could with the information it had, first estimating the size of the foam and the speed at which it had struck the *Columbia*, and then using an algorithm called Crater to predict how deep a piece of debris that size and traveling at that speed would

penetrate into the thermal-protection tiles that covered the shuttle's wings.

The DAT had reached no conclusions, but they made it clear to McCormack that there was reason to be concerned. McCormack did not transmit that sense of concern to the MMT during its teleconference. The foam strike was not mentioned until two-thirds of the way through the meeting, and was brought up only after discussions of, among other things, a jammed camera, the scientific experiments on the shuttle, and a leaky water separator. Then Linda Ham, who was the MMT leader, asked McCormack for an update. He simply said that people were investigating the possible damage and what could potentially be done to fix it, and added that when the *Columbia* had been hit by a similar strike during mission STS-87, five years earlier, it had suffered "fairly significant damage." This is how Ham answered: "And I really don't think there is much we can do so it's not really a factor during the flight because there is not much we can do about it."

Ham, in other words, had already decided that the foam strike was inconsequential. More important, she decided for everyone else in the meeting that it was inconsequential, too. This was the first time the MMT had heard any details about the foam strike. It would have been logical for McCormack to outline the possible consequences and talk about what the evidence from past shuttles that had been struck with debris showed. But instead the meeting moved on.

Hindsight is, of course, twenty-twenty, and just as with the critiques of the U.S. intelligence community after September 11, it's perhaps too easy to fault the MMT at NASA for its failure to see what would happen to the *Columbia* when it reentered the Earth's atmosphere on February 1. Even those who have been exceptionally critical of NASA have suggested that focusing on this one team is a mistake because it obscures the deep institutional and cultural problems that plague the agency (which happen to be

many of the same problems that plagued the agency in 1986, when the *Challenger* exploded). But while NASA clearly is an object lesson in organizational dysfunction, that doesn't fully explain just why the MMT handled the *Columbia* crisis so badly. Sifting through the evidence collected by the Columbia Accident Investigation Board (CAIB), there is no way to evade the conclusion that the team had an opportunity to make different choices that could have dramatically improved the chances of the crew surviving. The team members were urged on many different occasions to collect the information they needed to make a reasonable estimate of the shuttle's safety. They were advised that the foam might, in fact, have inflicted enough damage to cause "burn-through"—heat burning through the protective tiles and into the shuttle's fuselage—when the shuttle reentered the Earth's atmosphere. The team's leaders themselves raised the possibility that the debris damage might have been severe. And yet the MMT as a whole never came close to making the right decision on what to do about the *Columbia*.

IN FACT, THE PERFORMANCE of the MMT is an object lesson in how not to run a small group, and a powerful demonstration of the way in which, instead of making people wiser, being in a group can actually make them dumber. This is important for two reasons. First, small groups are ubiquitous in American life, and their decisions are consequential. Juries decide whether or not people will go to prison. Boards of directors shape, at least in theory, corporate strategy. And more and more of our work lives are spent on teams or, at the very least, in meetings. Whether small groups can do a good job of solving complex problems is hardly an academic question.

Second, small groups are different in important ways from groups such as markets or betting pools or television audiences. Those groups are as much statistical realities as experiential ones.

Bettors do get feedback from each other in the form of the point spread, and investors get feedback from each other in the stock market, but the nature of the relationship between people in a small group is qualitatively different. Investors do not think of themselves as members of the market. People on the MMT thought of themselves as members of that team. And the collective wisdom that something like the Iowa Electronic Markets produces is, at least when it's working well, the result of many different independent judgments, rather than something that the group as a whole has consciously come up with. In a small group, by contrast, the group—even if it is an ad hoc group formed for the sake of a single project or experiment—has an identity of its own. And the influence of the people in the group on each other's judgment is inescapable.

What we'll see is that this has two consequences. On the one hand, it means small groups can make very bad decisions, because influence is more direct and immediate and small-group judgments tend to be more volatile and extreme. On the other hand, it also means that small groups have the opportunity to be more than just the sum of their parts. A successful face-to-face group is more than just collectively intelligent. It makes everyone work harder, think smarter, and reach better conclusions than they would have on their own. In his 1985 book about Olympic rowing, *The Amateurs*, David Halberstam writes: "When most oarsmen talked about their perfect moments in a boat, they referred not so much to winning a race but to the feel of the boat, all eight oars in the water together, the synchronization almost perfect. In moments like that, the boat seemed to lift right out of the water. Oarsmen called that the moment of *swing*." When a boat has swing, its motion seems almost effortless. Although there are eight oarsmen in the boat, it's as if there's only one person—with perfect timing and perfect strength—rowing. So you might say that a small group which works well has intellectual swing.

Swing, though, is hard to come by. In fact, few organizations have figured out how to make groups work consistently well. For all the lip service paid, particularly in corporate America, to the importance of teams and the need to make meetings more productive, it's still unusual for a small group to be more than just the sum of its parts. Much of the time, far from adding value to their members, groups seem to subtract it. Too often, it's easy to agree with Ralph Cordiner, the former chairman of General Electric, who once said, "If you can name for me one great discovery or decision that was made by a committee, I will find you the one man in that committee who had the lonely insight—while he was shaving or on his way to work, or maybe while the rest of the committee was chattering away—the lonely insight that solved the problem and was the basis for the decision." On this account, groups are nothing but obstacles, cluttering the way of people whose time would be better spent alone.

The performance of the MMT helps explain why. First, the team started not with an open mind but from the assumption that the question of whether a foam strike could seriously damage the shuttle had already been answered. This was, to be fair, partly a matter of bad luck, since one of the team's technical advisers was convinced from the beginning that foam simply could do no serious damage, and kept saying so to anyone who would listen. But there was plenty of evidence to suggest otherwise. Rather than begin with the evidence and work toward a conclusion, the team members worked in the opposite direction. More egregiously, their skepticism about the possibility that something might really be wrong made them dismiss the need to gather more information, especially in the form of pictures, leading to the DAT's requests for on-orbit images being rejected. Even when MMT members dealt with the possibility that there might be a real problem with *Columbia*, their conviction that nothing was wrong limited discussion and made them discount evidence to the contrary. In that sense,

Bettors do get feedback from each other in the form of the point spread, and investors get feedback from each other in the stock market, but the nature of the relationship between people in a small group is qualitatively different. Investors do not think of themselves as members of the market. People on the MMT thought of themselves as members of that team. And the collective wisdom that something like the Iowa Electronic Markets produces is, at least when it's working well, the result of many different independent judgments, rather than something that the group as a whole has consciously come up with. In a small group, by contrast, the group—even if it is an ad hoc group formed for the sake of a single project or experiment—has an identity of its own. And the influence of the people in the group on each other's judgment is inescapable.

What we'll see is that this has two consequences. On the one hand, it means small groups can make very bad decisions, because influence is more direct and immediate and small-group judgments tend to be more volatile and extreme. On the other hand, it also means that small groups have the opportunity to be more than just the sum of their parts. A successful face-to-face group is more than just collectively intelligent. It makes everyone work harder, think smarter, and reach better conclusions than they would have on their own. In his 1985 book about Olympic rowing, *The Amateurs*, David Halberstam writes: "When most oarsmen talked about their perfect moments in a boat, they referred not so much to winning a race but to the feel of the boat, all eight oars in the water together, the synchronization almost perfect. In moments like that, the boat seemed to lift right out of the water. Oarsmen called that the moment of *swing*." When a boat has swing, its motion seems almost effortless. Although there are eight oarsmen in the boat, it's as if there's only one person—with perfect timing and perfect strength—rowing. So you might say that a small group which works well has intellectual swing.

Swing, though, is hard to come by. In fact, few organizations have figured out how to make groups work consistently well. For all the lip service paid, particularly in corporate America, to the importance of teams and the need to make meetings more productive, it's still unusual for a small group to be more than just the sum of its parts. Much of the time, far from adding value to their members, groups seem to subtract it. Too often, it's easy to agree with Ralph Cordiner, the former chairman of General Electric, who once said, "If you can name for me one great discovery or decision that was made by a committee, I will find you the one man in that committee who had the lonely insight—while he was shaving or on his way to work, or maybe while the rest of the committee was chattering away—the lonely insight that solved the problem and was the basis for the decision." On this account, groups are nothing but obstacles, cluttering the way of people whose time would be better spent alone.

The performance of the MMT helps explain why. First, the team started not with an open mind but from the assumption that the question of whether a foam strike could seriously damage the shuttle had already been answered. This was, to be fair, partly a matter of bad luck, since one of the team's technical advisers was convinced from the beginning that foam simply could do no serious damage, and kept saying so to anyone who would listen. But there was plenty of evidence to suggest otherwise. Rather than begin with the evidence and work toward a conclusion, the team members worked in the opposite direction. More egregiously, their skepticism about the possibility that something might really be wrong made them dismiss the need to gather more information, especially in the form of pictures, leading to the DAT's requests for on-orbit images being rejected. Even when MMT members dealt with the possibility that there might be a real problem with *Columbia*, their conviction that nothing was wrong limited discussion and made them discount evidence to the contrary. In that sense,

the team succumbed to what psychologists call "confirmation bias," which causes decision makers to unconsciously seek those bits of information that confirm their underlying intuitions.

These problems were also exacerbated by the team's belief that it knew more than it did. For instance, when the shuttle managers turned down the request for pictures, one of the justifications they offered was that the resolution of the images would not be good enough to detect the small area where the foam struck. In fact, as the CAIB noted, none of the managers had the necessary security clearances to know how good the resolution of the photos would be, nor did any of them ask the Department of Defense—which would have taken the pictures—about picture quality. In other words, they were "making critical decisions about imagery capabilities based on little or no knowledge," and doing so with an air of complete assurance.

Social scientists who study juries often differentiate between two approaches juries take. Evidence-based juries usually don't even take a vote until after they've spent some time talking over the case, sifting through the evidence, and explicitly contemplating alternative explanations. Verdict-based juries, by contrast, see their mission as reaching a decision as quickly and decisively as possible. They take a vote before any discussion, and the debate after that tends to concentrate on getting those who don't agree to agree. The MMT's approach was practically, though not intentionally, verdict-based. You can see this especially clearly in the way Linda Ham asked questions. On January 22, for instance, the day after the meeting where the foam was first mentioned, Ham e-mailed two members of the team about whether the foam strike might, in fact, pose a threat to the shuttle's safety. "Can we say that for any ET [external tank] foam lost," she wrote, "no 'safety of flight' damage can occur to the Orbiter because of the density?" The answer that Ham wanted was built into the question. It was a way of deflecting genuine inquiry even while seeming to pursue it. As it happens,

one of the members of the team did not give Ham the answer she was looking for. Lambert Austin answered her question by writing, "NO," in capital letters, and then went on to explain that there was no way at that point to "PRECLUDE" the possibility that the foam might have seriously damaged the tiles. Yet Austin's cautionary note garnered little attention.

One reason for the team's lack of follow-through may have been its implicit assumption that if something was wrong, there was no possibility of fixing it. At that January 21 meeting, you'll remember, Ham said, "And I really don't think there is much we can do so it's not really a factor during the flight because there is not much we can do about it." Two days later, Calvin Schomburg, the technical expert who insisted throughout that the foam could not seriously damage the tiles, met with Rodney Rocha, a NASA engineer who had become the unofficial representative of the DAT. By this point, the DAT was increasingly concerned that the damage inflicted by the foam could potentially lead to burn-through on reentry, and Rocha and Schomburg argued over the question. At the end of the discussion, Schomburg said that if the tiles had been severely damaged, "Nothing could be done."

The idea that nothing could have been done if the damage to the tiles had been uncovered in time was wrong. In fact, as part of the CAIB investigation, NASA engineers came up with two different strategies that might have brought the *Columbia* crew back to earth safely (though the shuttle itself was doomed from the moment the foam struck). There was no reason for the MMT to know what those strategies were, of course. But here again, the team had made a decision before looking at the evidence. And that decision—which roughly amounted to saying, "If there is a problem, we won't be able to find a solution"—undoubtedly shaped the team's approach to figuring out whether there was a problem at all. In fact, the CAIB report includes personal notes from an unnamed NASA source that say that when Ham canceled the DAT's request

for pictures of the *Columbia*'s wing, "[she] said it was no longer being pursued since even if we saw something, we couldn't do anything about it." This was not exactly the ethos that brought *Apollo 13* safely back to Earth.

One of the real dangers that small groups face is emphasizing consensus over dissent. The extreme version of this, as we've already seen, is the kind of groupthink that Irving Janis described in his account of the planning of the Bay of Pigs, where the members of the group become so identified with the group that the possibility of dissent seems practically unthinkable. But in a more subtle way small groups can exacerbate our tendency to prefer the illusion of certainty to the reality of doubt. On January 24, the DAT engineers met again with Don McCormack, who had become their unofficial liaison to the MMT, to present the findings of their foam-strike study. The briefing room where the presentation took place was so crowded that engineers ended up out in the hallway, which said a lot about how worried people were. In any case, the DAT offered five different scenarios of what might have happened. The team's conclusion was that it was likely that the shuttle was safe. But they qualified their conclusion by saying that their analysis was profoundly limited by their tools and their lack of good information. Because the MMT had refused to authorize on-orbit images, the engineers did not know where exactly the foam had struck. And the Crater algorithm they were using had been designed to measure the impact of pieces of debris hundreds of times smaller than the one that hit *Columbia*, so there was no way to be sure that its results were accurate. The engineers stressed, in other words, how uncertain their analysis was. But NASA management focused instead on their conclusion.

An hour after the briefing, the MMT met, and McCormack summarized what the DAT had said. "They do show obviously there's potential for significant damage here, but thermal analysis does not indicate that there is potential for a burn-through," he said. "Obviously there is a lot of uncertainty in all this in terms of

the size of the debris and where it hit and the angle of incidence and it's difficult." This was a relatively obscure way of explaining that the engineers' analysis was built on a lot of untested assumptions, but it was at least an attempt at caution. Ham responded by again asking a question that answered itself: "No burn-through, means no catastrophic damage and the localized heating damage would mean a tile replacement?" McCormack said, "We do not see any kind of safety of flight issue here yet in anything that we've looked at." Ham came back with another nothing-is-wrong question: "No safety of flight and no issue for this mission nothing that we're going to do different, there may be a turnaround?" Then, after a short interchange between Ham and McCormack and Calvin Schomburg, one of the other team members on the conference call said that they hadn't been able to hear what McCormack had said. Ham summarized neatly: "He was just reiterating with Calvin that he doesn't believe that there is any burn-through so no safety of flight kind of issue, it's more of a turnaround issue similar to what we've had on other flights. That's it? Alright, any questions on that?" For all intents and purposes, when that meeting ended, the *Columbia*'s fate had been sealed.

What's most striking about that January 24 meeting is the utter absence of debate and minority opinions. As the CAIB noted, when McCormack summarized the DAT's findings, he included none of its supporting analysis nor any discussion of whether there was a division of opinion on the team about its conclusions. More strikingly, not one member of the MMT asked a question. Not one member expressed any interest in seeing the DAT study. One would have thought that when McCormack mentioned the uncertainties in the analysis, someone would have asked him to explain and perhaps even quantify those uncertainties. But no one did. In part, that may have been because Ham was so obviously anxious for the problem to be resolved, and so convinced that there was nothing to talk about. Her attempts to briskly summarize McCormack's conclusions—"No burn-through, means no catastrophic

damage"—effectively shut off discussion. And anyone who's ever been in a business meeting knows that "Alright, any questions on that?" really means "There are no questions on that, right?"

The MMT failed to make the right decision in part because of problems that are specific to the culture of NASA. Although we think of NASA as a fundamentally meritocratic, bottom-up culture, it is in fact deeply hierarchical. This meant that even though the DAT engineers had serious qualms from the beginning about the foam strike, their concerns—and, in particular, their insistence that they needed images of the Orbiter's wing before they could make a truly informed analysis—never received a serious hearing from the MMT. At the same time, the MMT violated nearly every rule of good group decision making. To begin with, the team's discussions were simultaneously too structured and not structured enough. They were too structured because most of the discussions—not just about the debris strike, but about everything—consisted of Ham asking a question and someone else answering it. They were not structured enough because no effort was made to ask other team members to comment on particular questions. This is almost always a mistake, because it means that decisions are made based on a very limited supply of analysis and information. One of the consistent findings from decades of small-group research is that group deliberations are more successful when they have a clear agenda and when leaders take an active role in making sure that everyone gets a chance to speak.

The team also, as I've mentioned, started with its conclusion. As a result, every new piece of information that came in was reinterpreted to fit that conclusion. This is a recurring problem with small groups that have a hard time incorporating new information. Social psychologist Garold Stasser, for instance, ran an experiment in which a group of eight people was asked to rate the performance of thirty-two psychology students. Each member of the group was given two relevant pieces of information about the students (say, their grades and their test scores), while two members of the group

were given two extra pieces of information (say, their performance in class, etc.), and one member of the group received another two. Although the group as a whole therefore had six pieces of useful information, their ratings were based almost entirely on the two pieces of information that they all shared. The new information was discounted as either unimportant or unreliable. Stasser has also shown that in unstructured, free-flowing discussions, the information that tends to be talked about the most is, paradoxically, the information that everyone already knows. More curiously, information can be presented and listened to and still make little difference, because its contents are misinterpreted. New messages are often modified so that they fit old messages, which is especially dangerous since unusual messages often add the most value. (If people are just saying what you expect them to say, they're hardly likely to change your thinking.) Or they are modified to suit a preexisting picture of the situation.

What was missing most from the MMT, of course, was diversity, by which I mean not sociological diversity but rather cognitive diversity. James Oberg, a former Mission Control operator and now NBC News correspondent, has made the counterintuitive point that the NASA teams that presided over the *Apollo* missions were actually more diverse than the MMT. This seems hard to believe, since every engineer at Mission Control in the late 1960s had the same crew cut and wore the same short-sleeved white shirt. But as Oberg points out, most of those men had worked outside of NASA in many different industries before coming to the agency. NASA employees today are far more likely to have come to the agency directly out of graduate school, which means they are also far less likely to have divergent opinions. That matters because, in small groups, diversity of opinion is the single best guarantee that the group will reap benefits from face-to-face discussion. Berkeley political scientist Charlan Nemeth has shown in a host of studies of mock juries that the presence of a minority viewpoint, all by itself, makes a group's decisions more nuanced and its decision-making

process more rigorous. This is true even when the minority viewpoint turns out to be ill conceived. The confrontation with a dissenting view, logically enough, forces the majority to interrogate its own positions more seriously. This doesn't mean that the ideal jury will follow the plot of *Twelve Angry Men,* where a single holdout convinces eleven men who are ready to convict that they're all wrong. But it does mean that having even a single different opinion can make a group wiser. One suspects that, had there been a single devil's advocate pushing the idea that the foam strike might have seriously damaged the wing, the MMT's conclusion would have been very different.

Without the devil's advocate, though, it's likely that the group's meetings actually made its judgment about the possible problem worse. That's because of a phenomenon called "group polarization." Usually, when we think of deliberation, we imagine that it's a kind of recipe for rationality and moderation, and assume that the more people talk about an issue, the less likely they will be to adopt extreme positions. But evidence from juries and three decades of experimental studies suggests that much of the time, the opposite is true.

Group polarization is still a phenomenon that is not well understood, and there are clearly cases where it has little or no effect. But since the 1960s, sociologists have documented how, under certain circumstances, deliberation does not moderate but rather radicalizes people's point of view. The first studies of the phenomenon tried to elicit people's attitudes toward risk, by asking them what they would do in specific situations. For instance, they were asked, "If a man with a severe heart illness is told that he must either change his way of life completely or have an operation that will either cure him or kill him, what should he do?" Or, "If an electrical engineer who has a safe job at a small salary is given the chance to take a new job that pays much better but is also less secure, should he move?" Individuals answered these questions privately at first,

then gathered into groups to reach collective decisions. At first, researchers thought that group discussions made people more likely to advocate risky positions, and they termed this the "risky shift." But as time went on, it became clear that the shift could be in either direction. If a group was made up of people who were generally risk averse, discussion would make the group even more cautious, while groups of risk takers found themselves advocating riskier positions. Other studies showed that people who had a pessimistic view of the future became even more pessimistic after deliberations. Similarly, civil juries that are inclined to give large awards to plaintiffs generally give even larger awards after talking it over.

More recently, University of Chicago law professor Cass Sunstein has devoted a great deal of attention to polarization, and in his book *Why Societies Need Dissent,* he shows both that the phenomenon is more ubiquitous than was once thought and that it can have major consequences. As a general rule, discussions tend to move both the group as a whole and the individuals within it toward more extreme positions than the ones they entered the discussion with.

Why does polarization occur? One reason is because of people's reliance on "social comparison." This means more than that people are constantly comparing themselves to everyone else (which, of course, they are). It means that people are constantly comparing themselves to everyone else with an eye toward maintaining their relative position within the group. In other words, if you start out in the middle of the group and you believe the group has moved, as it were, to the right, you're inclined to shift your position to the right as well, so that relative to everyone else you're standing still. Of course, by moving to the right you're moving the group to the right, making social comparison something of a self-fulfilling prophecy. What's assumed to be real eventually becomes real.

It's important to see, though, that polarization isn't just the result of people trying to stay in tune with the group. It also results, strangely, from people doing their best to figure out what the right answer is. As we saw in our discussion of social proof—remember the passersby who ended up staring into an empty sky—people who are uncertain about what they believe will look to other members of the group for help. That's the point of deliberating, after all. But if a majority of the group already supports one position, then most of the arguments that will be made will be in support of that position. So the uncertain people are likely to be swayed in that direction, in part simply because that's more of what they'll hear. Similarly, people who have more extreme positions are more likely to have strong, coherent arguments in favor of their positions and are also more likely to voice them.

This matters because all the evidence suggests that the order in which people speak has a profound effect on the course of a discussion. Earlier comments are more influential, and they tend to provide a framework within which the discussion occurs. As in an information cascade, once that framework is in place, it's difficult for a dissenter to break it down. This wouldn't be a problem if the people who spoke earliest were also more likely to know what they were talking about. But the truth is that, especially when it comes to problems where there is no obvious right answer, there's no guarantee that the most-informed speaker will also be the most influential. On juries, for instance, two-thirds of all foremen—who lead and structure deliberations—are men, and during deliberations men talk far more than women do, even though no one has ever suggested that men as a gender have better insight into questions of guilt and innocence. In groups where the members know each other, status tends to shape speaking patterns, with higher-status people talking more and more often than lower-status people. Again, this wouldn't matter as much if the authority of higher-status people was derived from their

greater knowledge. But oftentimes it doesn't. Even when higher-status people don't really know what they're talking about, they're more likely to speak. A series of experiments with military fliers who were asked to solve a logic problem, for instance, found that pilots were far more likely to speak convincingly in defense of their solution than navigators were, even when the pilots were wrong and the navigators were right. The navigators deferred to the pilots—even when they had never met the pilots before—because they assumed that their rank meant they were more likely to be right.

That kind of deference is important, because in small groups ideas often do not succeed simply on their own merits. Even when its virtues may seem self-evident, an idea needs a champion in order to be adopted by the group as a whole. That's another reason why a popular position tends to become more popular in the course of deliberations: it has more potential champions to begin with. In a market or even a democracy, champions are far less important because of the sheer number of potential decision makers. But in a small group, having a strong advocate for an idea, no matter how good it is, is essential. And when advocates are chosen, as it were, on the basis of status or talkativeness, rather than perceptiveness or keenness of insight, then the group's chance of making a smart decision shrinks.

Talkativeness may seem like a curious thing to worry about, but in fact talkativeness has a major impact on the kinds of decisions small groups reach. If you talk a lot in a group, people will tend to think of you as influential almost by default. Talkative people are not necessarily well liked by other members of the group, but they are listened to. And talkativeness feeds on itself. Studies of group dynamics almost always show that the more someone talks, the more he is talked to by others in the group. So people at the center of the group tend to become more important over the course of a discussion.

This might be okay if people only spoke when they had expertise in a particular matter. And in many cases, if someone's talking a lot, it's a good sign that they have something valuable to add. But the truth is that there is no clear correlation between talkativeness and expertise. In fact, as the military-flier studies suggest, people who imagine themselves as leaders will often overestimate their own knowledge and project an air of confidence and expertise that is unjustified. And since, as political scientists Brock Blomberg and Joseph Harrington suggest, extremists tend to be more rigid and more convinced of their own rightness than moderates, discussion tends to pull groups away from the middle. Of course, sometimes truth lies at the extreme. And if the people who spoke first and most often were consistently the people with the best information or the keenest analysis, then polarization might not be much of a problem. But it is.

THE OBVIOUS TEMPTATION IS to do away with or at least minimize the role that small groups play in shaping policy or making decisions. Better to entrust one reliable person—who at least we know will not become more extreme in his views—with responsibility than trust a group of ten or twelve people who at any moment, it seems, may suddenly decide to run off a cliff. It would be a mistake to succumb to that temptation. First of all, groups can be, as it were, depolarized. In a study that divided people into groups of six while making sure that each group composed two smaller groups of three who had strongly opposed views, it was found that discussion moved the groups from the extremes and toward each other. That same study found that as groups became less polarized, they also became more accurate when they were tested on matters of fact.

More important, as solid as the evidence demonstrating group polarization is, so too is the evidence demonstrating that nonpolarized groups consistently make better decisions and come

up with better answers than most of their members, and surprisingly often the group outperforms even its best member. What makes this surprising is that one would think that in a small group, one or two confused people could skew the group's collective verdict in the wrong direction. (The small group can't, in that sense, rely on errors canceling themselves out.) But there's little evidence of that happening.

One of the more impressive studies of small-group performance was done in 2000 by Princeton economists Alan S. Blinder and John Morgan. Blinder had been vice chairman of the Federal Reserve Board during the mid-1990s, and the experience had made him deeply skeptical of decision making by committee. (Interest-rate changes are set by the Federal Open Market Committee, which consists of twelve members, including the seven members of the Federal Reserve Board and five presidents of regional Federal Reserve banks.) So he and Morgan designed a study that was meant to find out if groups could make intelligent decisions and if they make decisions as a group quickly, since one of the familiar complaints about committees is that they are inefficient.

The study consisted of two experiments that were meant to mimic, crudely, the challenges faced by the Fed. In the first experiment, students were given urns that held equal numbers of blue balls and red balls. They started to draw the balls from the urns, having been told that sometime after the first ten draws, the proportions in the urn would shift, so that 70 percent of the balls would be red and 30 percent blue (or vice versa). The goal was to identify, as soon as possible, which color had become more prevalent. This was roughly analogous to the Fed's job of recognizing when economic conditions have changed and whether a shift in monetary policy is needed. To place a premium on making the right decision quickly, students were penalized for every draw they made after the changeover had happened. The students played the game by themselves first, then played together as a group with free dis-

cussion, played as individuals again, and finally once more as a group. (This was to control for the effect of learning.) The group's decisions were both faster and more accurate (the group got the direction right 89 percent of the time, versus 84 percent for individuals), and outperformed even the best individual.

The second experiment demanded more of the students. Essentially, they were asked to play the role of central bankers, and to set interest rates in response to changes in inflation and unemployment. What the experiment was really asking was whether they could detect when the economy had started to slow or was picking up steam, and whether they would move interest rates in the right direction in response. Once again, the group made better decisions than the individuals, who moved interest rates in the wrong direction far more often, and made them as quickly as the individuals. Most strikingly, there was no correlation between the performance of the smartest person in a group and the performance of that group. In other words, the groups were not just piggybacking on really smart individuals. They genuinely were smarter than the smartest people within them. A Bank of England study modeled on Blinder and Morgan's experiment reached identical conclusions: groups could make intelligent decisions quickly, and could do better than their smartest members.

Given what we've already seen, this is not shocking news. But there are two important things about these studies. The first is that group decisions are not inherently inefficient. This suggests that deliberation can be valuable when done well, even if after a certain point its marginal benefits are outweighed by the costs. The second point is probably obvious, although a surprising number of groups ignore it, and that is that there is no point in making small groups part of a leadership structure if you do not give the group a method of aggregating the opinions of its members. If small groups are included in the decision-making process, then they should be allowed to make decisions. If an organization sets up teams and then uses them for purely advisory purposes, it loses the true advantage

that a team has: namely, collective wisdom. One of the more frustrating aspects of the *Columbia* story is the fact that the MMT never voted on anything. The different members of the team would report on different aspects of the mission, but their real opinions were never aggregated. This was a mistake, and it would have been a mistake even had the *Columbia* made it home safely.