9

Superteams

On the morning of January 10, 1961, as breakfast was being prepared across America, readers of the *New York Times* opened the newspaper on the kitchen table and read the front-page headline: U.S. HELPS TRAIN AN ANTI-CASTRO FORCE AT SECRET GUATEMALAN AIR-GROUND BASE. A little inland from Guatemala's Pacific coast "commando-like forces are being drilled in guerrilla warfare tactics by foreign personnel, mostly from the United States." The trainees were identified as Cubans. American aircraft using the base were identified. The American company that built the base was named. "Guatemalan authorities from President Miguel Ydigoras Fuentes down insist that the military effort is designed to meet an assault, expected almost any day, from Cuba," the *Times* reported, but "opponents of the Ydigoras Administration have insisted that the preparations are for an offensive against the regime of Premier Fidel Castro and that they are being planned and directed, and to a great extent paid for, by the United States. The United States embassy is maintaining complete silence on the subject."

In truth, the CIA was training Cuban exiles to land in Cuba and launch a guerrilla war against the new government of Fidel Castro. Secrecy was critical. Once the guerrillas landed, they had to look like an independent force of patriots coming to liberate the nation. To ensure this, no American soldiers would land with the guerrillas, and air support would be provided by old bombers without American markings. No one would know that the United States engineered the whole thing. At least that was the plan.

One might suppose that in Washington, DC, among the planners of this secret mission, the exposure of the scheme on the front page of the *New York Times* would cause concern and reconsideration. There was indeed concern—but no reconsideration. "Somehow the idea took hold around the cabinet table that this would not much matter so long as United States soldiers did not take part in the actual fighting," Arthur M. Schlesinger Jr. recalled. As an adviser to the new president, John F. Kennedy, Schlesinger was part of the inner circle that authorized the mission, and his recollections were filled with amazement at the blunders they made in planning what became known as the Bay of Pigs invasion.[1]

When the CIA-trained guerrillas landed, the Cuban army was waiting and the fourteen hundred men onshore were quickly surrounded by twenty thousand soldiers. Within three days they were all dead or taken prisoner.

The problem was not one of execution. It was the plan. It was harebrained. And that's not hindsight bias. The whole sorry saga has been dissected, and there is rare consensus among historians, left and right, that the plan was riddled with problems that the White House should have spotted but did not. A particularly blatant example was the contingency plan. The CIA assured the president's advisers that if the landing failed the guerrillas could escape to the Escambray Mountains, where they would join other anti-Castro forces. But that idea came from the first version of the plan, which would have landed the guerrillas on the shore at the base of the mountains. The planners had switched the landing site—but didn't consider what that switch meant for the

contingency plan. "I don't think we fully realized that the Escambray Mountains lay 80 miles from the Bay of Pigs, across a hopeless tangle of swamps and jungle," Schlesinger recalled.[2]

After the debacle, no one believed that the United States was not involved, and the consequences were immediate and severe. Traditional allies were embarrassed. Latin American nations were outraged. Anti-American protests erupted around the world. Liberals who had high hopes for the new Kennedy administration felt betrayed, while conservatives mocked the novice president's incompetence. Worst of all for the strategic interests of the United States, the Cuban government put itself more firmly inside the Soviet camp. Within eighteen months, an island off the coast of Florida was a base for five thousand Soviet soldiers and an array of Soviet intermediaterange nuclear missiles that could destroy Washington, DC, and New York City, and the two global superpowers were locked in a crisis that Kennedy estimated, in retrospect, had between a one-third and one-half chance of escalating into nuclear war.

The story of the Cuban missile crisis that followed from the Bay of Pigs fiasco is equally familiar, but the similarities end there. Over thirteen terrifying days in October 1962, the Kennedy administration considered a range of dangerous options to counter the Soviet threat—including outright invasion—before settling on a naval blockade. As Soviet ships approached the American red line, each side tried to figure out the other's intentions from its actions and back-channel communications. Finally an agreement was reached, war was averted, and the world exhaled.

If the Bay of Pigs was the Kennedy administration's nadir, the Cuban missile crisis was its zenith, a moment when Kennedy and his team creatively engineered a positive result under extreme pressure. Knowing this, we might assume Kennedy cleaned house after the Bay of Pigs and surrounded himself with far superior advisers in time for the missile crisis. But he didn't. The cast of characters in both dramas is mostly the same: the team that bungled the Bay of Pigs was the team that performed brilliantly during the Cuban missile crisis.

In his 1972 classic, *Victims of Groupthink*, the psychologist Irving Janis—one of my PhD advisers at Yale long ago—explored the decision making that went into both the Bay of Pigs invasion and the Cuban missile crisis. Today, everyone has heard of groupthink, although few have read the book that coined the term or know that Janis meant something more precise than the vague catchphrase groupthink has become today. In Janis's hypothesis, "members of any small cohesive group tend to maintain *esprit de corps* by unconsciously developing a number of shared illusions and related norms that interfere with critical thinking and reality testing."[3] Groups that get along too well don't question assumptions or confront uncomfortable facts. So everyone agrees, which is pleasant, and the fact that everyone agrees is tacitly taken to be proof the group is on the right track. We can't all be wrong, can we? So if a secret American plan to invade Cuba without apparent American involvement happens to be published on the front page of the *New York Times*, the plan can still go ahead—just make sure there are no American soldiers on the beach and deny American involvement. The world will believe it. And if that sounds implausible…well, not to worry, no one in the group has objected, which means everyone thinks it's perfectly reasonable, so it must be.

After the fiasco, Kennedy ordered an inquiry to figure out how his people could have botched it so badly. It identified cozy unanimity as the key problem and recommended changes to the decision-making process to ensure it could never develop again. Skepticism was the new watchword. Participants were to speak not only as specialists in their area of expertise but as generalists, with a license to question anything. Special counsel Theodore Sorensen and the president's brother Bobby were designated "intellectual watchdogs," whose job was to "pursue relentlessly every bone of contention in order to prevent errors arising from too superficial an analysis of the issues," Janis noted. "Accepting this role avidly, Robert Kennedy, at the expense of becoming unpopular with some of his associates, barked out sharp and sometimes rude questions. Often, he deliberately became the devil's advocate." Protocol and hierarchy would impede these freewheeling discussions, so they were set aside. New advisers were occasionally brought in to provide fresh perspectives. And John F. Kennedy would sometimes leave the room to let the group talk things through, knowing that there

was less true give-and-take when the president was present. That last consideration was crucial. Kennedy started the crisis thinking that, at a minimum, he had to authorize preemptive air attacks on the Soviet missile launchers, but he kept that to himself so it wouldn't be the focus of the discussion. As a result, "by the end of the first day of meetings the committee had seriously discussed ten alternatives," and the president's thinking started to change. It was never easy. There were constant disagreements. The stress was brutal. But it was a process that led to a negotiated peace, not nuclear war.[4]

How the Kennedy White House changed its decision-making culture for the better is a must-read for students of management and public policy because it captures the dual-edged nature of working in groups. Teams can cause terrible mistakes. They can also sharpen judgment and accomplish together what cannot be done alone. Managers tend to focus on the negative or the positive but they need to see both. As mentioned earlier, the term "wisdom of crowds" comes from James Surowiecki's 2004 bestseller of the same name, but Surowiecki's title was itself a play on the title of a classic 1841 book, *Extraordinary Popular Delusions and the Madness of Crowds*, which chronicled a litany of collective folly. Groups can be wise, or mad, or both. What makes the difference isn't just who is in the group, Kennedy's circle of advisers demonstrated. The group is its own animal.

TO TEAM OR NOT TO TEAM?

In the IARPA tournament, our goal was accuracy. Would putting forecasters on teams help? We saw strong arguments for both yes and no. On the negative side, the research literature—as well as my decades of experience on university committees—suggested that teams might foster cognitive loafing. Why labor to master a complex problem when others will do the heavy lifting? When this attitude is widespread it can sink a team. Worse, forecasters can become too friendly, letting groupthink set in. These two tendencies can reinforce each other. We all agree, so our work is done, right? And unanimity within a group is a powerful force. If that agreement is ill-founded, the group slips into self-righteous complacency.

But groups also let people share information and perspectives. That's good. It helps make dragonfly eye work, and aggregation is critical to accuracy. Of course aggregation can only do its magic when people form judgments independently, like the fairgoers guessing the weight of the ox. The independence of judgments ensures that errors are more or less random, so they cancel each other out. When people gather and discuss in a group, independence of thought and expression can be lost. Maybe one person is a loudmouth who dominates the discussion, or a bully, or a superficially impressive talker, or someone with credentials that cow others into line. In so many ways, a group can get people to abandon independent judgment and buy into errors. When that happens, the mistakes will pile up, not cancel out. This is the root of collective folly, whether it's Dutch investors in the seventeenth century, who became collectively convinced that a tulip bulb was worth more than a laborer's annual salary, or American home buyers in 2005, talking themselves into believing that real estate prices could only go up.

But loss of independence isn't inevitable in a group, as JFK's team showed during the Cuban missile crisis. If forecasters can keep questioning themselves and their teammates, and welcome vigorous debate, the group can become more than the sum of its parts.

So would groups lift superforecasters up or drag them down? Some of us suspected one outcome, others the opposite, but deep down, we knew we were all guessing. Ultimately, we chose to build teams into our research for two reasons. First, in the real world, people seldom make important forecasts without discussing them with others, so getting a better understanding of forecasting in the real world required a better understanding of forecasting in groups. The other reason? Curiosity. We didn't know the answer and we wanted to, so we took Archie Cochrane's advice and ran an experiment.

In year 1 (2011–12), before a single superforecaster had been tagged and classified, we randomly assigned several hundred forecasters to work alone and several hundred others to work together in teams. The team forecasters wouldn't meet face-to-face, of course, but we created online forums for discussion and team members could communicate by e-mail, Skype, or however else they wanted. They would still be scored as individuals, but individual scores would be pooled to create a team score. Forecasters would see how both they and their team were doing. Beyond that, forecasters could organize however they wished. The goal was accuracy. How they achieved it was up to them.

We also gave teams a primer on teamwork based on insights gleaned from research in group dynamics. On the one hand, we warned, groupthink is a danger. Be cooperative but not deferential. Consensus is not always good; disagreement not always bad. If you do happen to agree, don't take that agreement—in itself—as proof that you are right. Never stop doubting. Pointed questions are as essential to a team as vitamins are to a human body.

On the other hand, the opposite of groupthink—rancor and dysfunction—is also a danger. Team members must disagree without being disagreeable, we advised. Practice "constructive confrontation," to use the phrase of Andy Grove, the former CEO of Intel. Precision questioning is one way to do that. Drawing on the work of Dennis Matthies and Monica Worline, we showed them how to tactfully dissect the vague claims people often make. Suppose someone says, "Unfortunately, the popularity of soccer, the world's favorite pastime, is starting to decline." You suspect he is wrong. How do you question the claim? Don't even think of taking a personal shot like "You're silly." That only adds heat, not light. "I don't think so" only expresses disagreement without delving into why you disagree. "What do you mean?" lowers the emotional temperature with a question but it's much too vague. Zero in. You might say, "What do you mean by 'pastime'?" or "What evidence is there that soccer's popularity is declining? Over what time frame?" The answers to these precise questions won't settle the matter, but they will reveal the thinking behind the conclusion so it can be probed and tested.

Since Socrates, good teachers have practiced precision questioning, but still it's often not used when it's needed most. Imagine how events might have gone if the Kennedy team had engaged in precision questioning when planning the Bay of Pigs invasion:

"So what happens if they're attacked and the plan falls apart?"

"They retreat into the Escambray Mountains, where they can meet up with other anti-Castro forces and plan guerrilla operations."

"How far is it from the proposed landing site in the Bay of Pigs to the Escambray Mountains?"

"Eighty miles."

"And what's the terrain?"

"Mostly swamp and jungle."

"So the guerrillas have been attacked. The plan has fallen apart. They don't have helicopters or tanks. But they have to cross eighty miles of swamp and jungle before they can begin to look for shelter in the mountains? Is that correct?"

I suspect that this conversation would not have concluded "sounds good!"

Questioning like that didn't happen, so Kennedy's first major decision as president was a fiasco. The lesson was learned, resulting in the robust but respectful debates of the Cuban missile crisis—which exemplified the spirit we encouraged among our forecasters.

SUPERTEAMS

At the end of the year, the results were unequivocal: on average, teams were 23% more accurate than individuals.

When year 2 arrived, we all agreed teams should be an essential part of the research design. But we faced another choice. Having identified the top forecasters across experimental conditions, what should we do with these freshly anointed superforecasters? Should they be told of their status? Should they be put together in teams—and hope that superforecasters working with each other would produce superteams?

The risks were obvious. Tell someone they're exceptionally good at something and they may start taking their superiority for granted. Surround them with others who are similarly accomplished, tell them how special they are, and egos may swell even more. Rather than spur a superforecaster to take his game to the next level, it might make him so sure of himself that he is tempted to think his judgment must be right because it is his judgment. This is a familiar paradox: success can lead to acclaim that can undermine the habits of mind that produced the success. Such hubris often afflicts highly accomplished individuals. In business circles, it is called CEO disease.

Again we rolled the theoretical dice. We created teams of superforecasters, with a dozen people on each. We gave them more guidance about how high-performance teams function and created special forums to help them communicate online. The teams did not meet face-to-face, which had its own pluses and minuses. On the minus side, it's easier to disregard people we have never met. It could even foster conflict. Look at how quickly discussion on the Internet can degenerate into poisonous harangues. On the plus side, distance could make it easier to manage disputes and maintain a critical perspective.

Joining a team for the first time, superforecasters like Elaine Rich had more immediate concerns. "I was pretty intimidated by my team," she told me. Elaine lives in Washington, DC, and works as a pharmacist at the Walter Reed Medical Center. Some of the people on her team "announced huge, impressive credentials," she recalled. "And I had no credentials." At first she stayed quiet, making forecasts but rarely venturing opinions. It wasn't only that her teammates had credentials and confidence. She found it difficult to question the views of teammates who were, after all, strangers. People take things differently. What one person would consider a helpful inquiry another might take as an aggressive criticism. And some of the questions touched on issues many people feel passionately about, so talking about them felt like walking through a minefield. The Arafat-polonium discussion was the worst. "There was a lot of tension around that," Elaine said. "It was almost a taboo question."

"There was a lot of what I'll call dancing around," recalled Marty Rosenthal of his first year on a team. People would disagree with someone's assessment, and want to test it, but they were too afraid of giving offense to just come out and say what they were thinking. So they would "couch it in all these careful words," circling around, hoping the point would be made without their having to make it.

Experience helped. Seeing this "dancing around," people realized that excessive politeness was hindering the critical examination of views, so they made special efforts to assure others that criticism was welcome. "Everybody has said, 'I want push-back from you if you see something I don't,'" said Rosenthal. That made a difference. So did offering thanks for constructive criticism. Gradually, the dancing around diminished.

Research on teams often assumes they have leaders and norms and focuses on ensuring these don't hinder performance. The usual solutions are those the Kennedy administration implemented after the Bay of Pigs invasion—bring in outsiders, suspend hierarchy, and keep the leader's views under wraps. There's also the "premortem," in which the team is told to assume a course of action has failed and to explain why—which makes team members feel safe to express doubts they may have about the leader's plan. But the superteams did not start with leaders and norms, which created other challenges.

Marty Rosenthal is semiretired now, but for decades he was a management consultant who specialized in building teams. Doing that with no organizational structure is a challenge, he knew,

and doing that without meeting face-to-face is tougher still. Someone could step forward and start giving directions, but among strangers that can backfire. "I saw the gaps in how we were forming up as a team, wanted to address some of that, but also didn't want to be seen as, you know, taking over," he said. "And so a lot of what I did was what I think of as leading from behind. I just tried to lead by example." When Marty felt people weren't explaining their forecasts enough to get good discussions going, he explained his in greater detail and invited comments. He also organized a conference call to hash out workloads, with details handled by him—and most of the team signed up. "The feedback afterward was people loved it," Marty said. "I think people felt a little stronger commitment to the team coming out of that."

There were also two opportunities for superforecasters to meet teammates face-to-face, at the end of the second and third years, when GJP project manager Terry Murray hosted conferences at the Wharton School and the University of California, Berkeley. The official goal of the gatherings was to share knowledge: the researchers presented data and the superforecasters offered their views. The unofficial goal was to add a human dimension to the teams. Many superforecasters made the most of it. Marty lives less than a mile from the Berkeley campus so he invited his teammates—most came to the conference—to his house for a barbecue and beer. Modest as this and his other efforts were, Marty thinks they made a difference. "Definitely it's helped our ability to push back on each other and feel a commitment that we really need to step up and share information when we have it."[5]

That sense of belonging developed in Elaine Rich. She did well, boosting her confidence, and her sense of responsibility grew with it. "I felt that I had to be really careful that I was sharing, shouldering my part of the burden, rather than being a freeloader by reading what other people wrote," and not offering thoughts and research, "which is always a temptation."

Most teams have a nucleus of five or six members who do most of the work. Within that core, we might expect to see a division of labor that reduces the amount of effort any one person needs to invest in the task, at least if he or she approached forecasting as work, not play. But we saw the opposite on the best teams: workloads were divided, but as commitment grew, so did the amount of effort forecasters put into it. Being on the team was "tons more work," Elaine said. But she didn't mind. She found it far more stimulating than working by herself. "You could be supporting each other, or helping each other, or building on ideas," she said. "It was a rush."[6]

Committed superteams did some impressive digging. On a question about who would win the 2013 presidential election in Honduras, Paul Theron, a South African superforecaster—and an investment manager who hosts *Hot Stoxx* on CNBC Africa—located a political scientist who specializes in Honduran politics and was told, among other tidbits, that although polls showed a candidate named Castro with a slight lead, the polls were dodgy. Theron also found an analysis of Honduran politics on an obscure website and was so impressed by its thoroughness and the author's credentials that he e-mailed him and had an informative discussion. Paul changed his forecast, giving Castro's opponent Hernández the edge. Hernández won—and Paul's considerable effort paid off. And since Paul shared everything he learned with his teammates, they benefited too. "The team is so much more effective at gathering information than one person could ever be," Paul told me. "There is simply no way that any individual could cover as much ground as a good team does. Even if you had unlimited hours, it would be less fruitful, given different research styles. Each team member brings something different."[7]

The results speak for themselves. On average, when a forecaster did well enough in year 1 to become a superforecaster, and was put on a superforecaster team in year 2, that person became 50% more accurate. An analysis in year 3 got the same result. Given that these were collections of strangers tenuously connected in cyberspace, we found that result startling.

Even more surprising was how well superteams did against prediction markets.

Most economists would say markets are the most effective mechanism for collecting widely dispersed information and distilling it down to a single judgment. Markets do that with trading. If I

think a stock is a good value at a certain price, I may offer to buy yours. If you agree with my judgment, you won't sell. If you think I'm wrong, you will. Of course, in reality, trades happen for other reasons—you and I may have different financial needs steering us in different directions—but in general markets create incentives for people to relentlessly second-guess each other. The aggregation of all those judgments—and the information they are based on—is expressed in the price. If many people agree with me that a stock is worth more than it's selling for, they will try to buy it. Increasing demand pushes the price up. In that way, all the individual judgments of the buyers, and all the information guiding those judgments, becomes "priced in."

None of this means markets are perfect, or such efficient aggregators of information that no mortal should ever be so foolish as to aspire to beat them. That's the strong version of what economists call the efficient market hypothesis (EMH), and it's hard to square with what we have learned from psychology and experience. Markets make mistakes. Sometimes they lose their collective minds. But even if markets are far less efficient than ardent proponents of the EMH suppose, it is still very hard to consistently beat markets, which is why so few can plausibly claim to have done it.

Prediction markets are simply markets that trade in predictions, meaning traders buy and sell contracts on specified outcomes—such as "Hillary Clinton will be elected president of the United States in 2016." When the election of 2016 is held, that contract is settled. If Clinton loses, the contract pays out nothing. If she wins, it pays out $1. If the contract is currently selling for 40 cents and I think Clinton has a 60% or 70% chance of winning, I should buy. If lots of traders agree with me, demand for the contract will be strong and the price will rise—until it reaches a level where more traders think it's about right and buying slackens. If a new event suggests Clinton will not win, there will be a rush to sell and the price will decline. By aggregating all these judgments, the contract price should, in theory, closely track the true probability of Hillary Clinton winning.

Prediction markets like the famous Iowa Electronic Markets have an impressive track record. And they have a theory, backed by a battalion of Nobel laureates, going for them. So who would win in a battle between superteams and prediction markets? Most economists would say it's no contest. Prediction markets would mop the floor with the superteams.

We put that proposition to the test by randomly assigning regular forecasters to one of three experimental conditions. Some worked alone. Others worked in teams. And some were traders in prediction markets run by companies such as Inkling and Lumenogic. Of course, after year 1—when the value of teams was resoundingly demonstrated—nobody expected forecasters working alone to compete at the level of teams or prediction markets, so we combined all their forecasts and calculated the unweighted average to get the "wisdom of the crowd." And of course we had one more competitor: superteams.

The results were clear-cut each year. Teams of ordinary forecasters beat the wisdom of the crowd by about 10%. Prediction markets beat ordinary teams by about 20%. And superteams beat prediction markets by 15% to 30%.

I can already hear the protests from my colleagues in finance that the only reason the superteams beat the prediction markets was that our markets lacked liquidity: real money wasn't at stake and we didn't have a critical mass of traders. They may be right. It is a testable idea, and one worth testing. It's also important to recognize that while superteams beat prediction markets, prediction markets did a pretty good job of forecasting complex global events.

How did superteams do so well? By avoiding the extremes of groupthink and Internet flame wars. And by fostering minicultures that encouraged people to challenge each other respectfully, admit ignorance, and request help. In key ways, superteams resembled the best surgical teams identified by Harvard's Amy Edmondson, in which the nurse doesn't hesitate to tell the surgeon he left a sponge behind the pancreas because she knows it is "psychologically safe" to correct higher-ups.

Edmondson's best teams had a shared purpose. So did our superteams. One sign of that was linguistic: they said "our" more than "my."

A team like that should promote the sort of actively open-minded thinking that is so critical to accurate forecasting, as we saw in chapter 5. So just as we surveyed individuals to test their active open-mindedness (AOM), we surveyed teams to probe their attitudes toward the group and patterns of interaction within the group—that is, we tested the *team's* AOM. As expected, we found a correlation between a team's AOM and its accuracy. Little surprise there. But what makes a team more or less actively open-minded? You might think it's the individuals on the team. Put high-AOM people in a team and you'll get a high-AOM team; put lower-AOM people in a team and you'll get a lower-AOM team. Not so, as it turns out. Teams were not merely the sum of their parts. How the group thinks collectively is an emergent property of the group itself, a property of communication patterns among group members, not just the thought processes inside each member.[8] A group of open-minded people who don't care about one another will be less than the sum of its open-minded parts. A group of opinionated people who engage one another in pursuit of the truth will be more than the sum of its opinionated parts.

All this brings us to the final feature of winning teams: the fostering of a culture of sharing. My Wharton colleague Adam Grant categorizes people as "givers," "matchers," and "takers." Givers are those who contribute more to others than they receive in return; matchers give as much as they get; takers give less than they take. Cynics might say that *giver* is a polite word for *chump*. After all, anyone inclined to freeload will happily take what they give and return nothing, leaving the giver worse off than if he weren't so generous. But Grant's research shows that the pro-social example of the giver can improve the behavior of others, which helps everyone, including the giver—which explains why Grant has found that givers tend to come out on top.

Marty Rosenthal is a giver. He wasn't indiscriminately generous with his time and effort. He was generous in a deliberate effort to change the behavior of others for the benefit of all. Although Marty didn't know Grant's work, when I described it to him, he said, "You got it." There are lots more givers on the superteams. Doug Lorch distributed programming tools, which got others thinking about creating and sharing their own. Tim Minto contributed an analysis that showed how to make valuable automatic tweaks to forecasts with the passage of time. All are givers. None is a chump. In fact, Doug Lorch's individual score was the best in year 2, while Tim Minto topped the chart in year 3. And each man's team won the team competition.[9]

But let's not take this too far. A busy executive might think "I want some of those" and imagine the recipe is straightforward: shop for top performers, marinate them in collaborative teams, strain out the groupthink, sprinkle in some givers, and wait for the smart decisions and money to start flowing. Sadly, it isn't that simple. Replicating this in an existing organization with real employees would be a challenge. Singling out people for "super" status may be divisive and transferring people into cross-functional teams can be disruptive. And there's no guarantee of results. There were eccentric exceptions to the tendencies outlined above, such as the few teams who were not mutually supportive but who nonetheless did well. One of the best superforecasters even refused to leave comments for his teammates, saying he didn't want to risk groupthink.

This is the messy world of psychological research. Solid conclusions take time and this work, particularly on superteams, is in its infancy. There are many questions we have only begun to explore.

One involves the provocative phrase "diversity trumps ability," coined by my colleague (and former competitor in the IARPA tournament) Scott Page.[10] As we have seen, the aggregation of different perspectives is a potent way to improve judgment, but the key word is *different*. Combining uniform perspectives only produces more of the same, while slight variation will produce slight improvement. It is the diversity of the perspectives that makes the magic work. Superteams were fairly diverse—because superforecasters are fairly diverse—but we didn't design them with that in mind. We put ability first. If Page is right, we might have gotten even better results if we had made

diversity the key determinant of team membership and let ability take care of itself. Again, though, flag the false dichotomy. The choice is not ability or diversity; it is fine-tuning the mixes of ability and diversity and gauging which work best in which situations.

To appreciate this balancing act—and how promising it is—think back to President Obama asking each member of his team of advisers how likely it was that the unusually tall man in the mystery house in Pakistan was Osama bin Laden. The answers ranged from 30% to 95%, most well over 50%. Add them up and divide by the number of advisers and, from the sketchy reports available, it averages out to roughly 70%. That's the wisdom of the crowd. It's a hard-to-beat number that should have been given more respect than it got in that meeting. But could President Obama have done even better than that?

Our research suggests yes—depending on the diversity of his team. The more diverse his team, the greater the chance that some advisers will possess scraps of information that others don't. And since these scraps mostly point toward "it's bin Laden," if all the advisers were given all the scraps they don't have, they would individually raise their estimate. And that would boost the "wisdom of the crowd" figure—maybe to 80% or 85%.

That's the thinking behind the extremizing algorithm I mentioned in chapter 4. It works superbly, but its effectiveness depends on diversity.[11] A team with zero diversity—its members are clones and everyone knows everything that everyone else knows—should not be extremized at all. Of course no team matches that description. But some teams are good at sharing information and that reduces diversity somewhat. Superforecaster teams were like that, which is why extremizing didn't help them much. But regular forecasting teams weren't as good at sharing information. As a result, we got major gains when we extremized them. Indeed, extremizing gave regular forecaster teams a big enough boost to pass some superteams, and extremizing a large pool of regular forecasters produced, as we saw earlier, tournament-winning results.

These tools won't replace intelligence analysts or the officials who synthesize their conclusions. And they shouldn't. As far as I can see, there will always be a need for a chief executive to be surrounded by a smart team of advisers, as John F. Kennedy was during the Cuban missile crisis. But the tools are good enough that the remarkably inexpensive forecasts they generate should be on the desks of decision makers, including the president of the United States.