TABLE 8.8:
Predictions of the Rudy G. Bee

	Micheala	*Juliana*	*Average*	*Outcome*
Maggie	6	10	8	6
Cole	3	7	5	5
Brody	5	1	3	1

convenient assumptions with good ones. By assuming independent signals, these scholars assume more diversity than may exist.[6]

THE DIVERSITY PREDICTION THEOREM

Now that we've worked through an example, we're ready to turn to the more general theorems that reveal the importance of diverse predictive models among the members of a crowd. Versions of these theorems can be found in computer science, statistics, and econometrics.[7] To describe these theorems, we'll need two measures. The first captures how much a collection of predictive models differs. The other captures how accurate the models are. Both are based on the same accuracy measure: *squared errors*. In statistics, errors are squared so that negative errors and positive errors do not cancel one another out. If errors were added, a person who was equally likely to overestimate or underestimate an amount on average would make no errors ($-5 + 5 = 0$). If we first square the errors, then the negative and positive errors do not cancel ($(-5)^2 + 5^2 = 25 + 25 = 50$). To build the logic of the theorem, we first construct an example. Suppose that Micheala and Juliana have developed models to predict where three students—Maggie, Cole, and Brody—will place in an upcoming spelling bee at Rudy Giuliani Elementary. Table 8.8 shows their individual predictions, their average prediction, and the actual outcome from the bee.

We first compute the squared errors of Micheala and Juliana's predictions. Michaela picks Maggie to take sixth place and she takes sixth, an error of zero. She picks Cole to take third and he takes fifth, an error of two. And she picks Brody to take fifth, but he takes first place, an error of four. Squaring these three errors gives zero, four, and sixteen. The sum of the her errors equals twenty.

Micheala's Individual Error: $(6-6)^2 + (3-5)^2 + (5-1)^2 = 0 + 4 + 16 = 20$

We next make the same calculation for Juliana. She misses Maggie's placement by four, she misses Cole's by two, and gets Brody's place exactly right. Squaring these errors gives sixteen, four, and zero, for a total squared error of twenty.

Juliana's Individual Error: $(10-6)^2 + (7-5)^2 + (1-1)^2 = 16 + 4 + 0 = 20$

The sum of each of their squared errors equals twenty, so their average sum of squared errors also equals twenty. We call this the *average individual error*. Here that's easy because their errors are the same.

Average Individual Error: *Average of the individual squared errors*

$$\frac{20+20}{2} = 20$$

We next compute the error of their *collective prediction:* the average of their individual predictions. They collectively predict that Maggie will take eighth place. She takes sixth, for an error of two. Their collective prediction for Cole, fifth place, is correct, and their prediction for Brody is off by two. Squaring these errors gives four, zero, and four, for a total of eight. We call this their *collective error.*

Collective Error: *Squared error of the collective prediction*

$$(8-6)^2 + (5-5)^2 + (3-1)^2 = 4 + 0 + 4 = 8$$

Notice that their collective prediction is more accurate than either of their individual predictions. The explanation for this can be found in the diversity of their predictions. When one of them predicts too high, the other predicts too low and their mistakes, while not canceling entirely, become less severe. To make this relationship between the diversity of their predictions and the accuracy of their collective prediction more formal, we calculate how much their predictions differ. We do this by calculating Juliana's squared distance from their collective prediction and Micheala's squared distance from their collective prediction. We then average these two numbers. Statisticians call this the *variance* of their predictions. We will call it the *prediction diversity*.

We first compute Micheala's squared distance from the collective prediction. The collective prediction for Maggie is eighth place. Micheala predicts sixth place for a difference of two. The collective prediction for Cole is fifth place, and she predicts third place for a difference of two. Finally, the collective prediction for Brody is third place, and she predicts fifth, a difference also equal to two. The squares of these differences are four, four, and four, which sum to twelve.

Micheala's Squared Distance from the Average:
$(6 - 8)^2 + (3 - 5)^2 + (5 - 3)^2 = 4 + 4 + 4 = 12$

As there are only two predictors in this example, Juliana's distance from the average in each case must be the same as Michaela's. That calculation can be made as follows:

Juliana's Squared Distance from the Average:
$(10 - 8)^2 + (7 - 5)^2 + (1 - 3)^2 = 4 + 4 + 4 = 12$

The *prediction diversity* equals the average of these two distances; in this case, it equals twelve.

Prediction Diversity: *Average squared distance from the individual predictions to the collective prediction.*

$$\frac{12 + 12}{2} = 12$$

Notice the relationship between the collective error (8), the average individual error (20), and the prediction diversity (12): *Collective error equals average error minus diversity.* This equality is not an artifact of our example. It is always true. And, even better, it holds for any number of predictors, not just two predictors as in our example. Thus, we call this the *Diversity Prediction Theorem*.

The Diversity Prediction Theorem: *Given a crowd of predictive models*

$$\textit{Collective Error} = \textit{Average Individual Error} \\ -\textit{Prediction Diversity}$$

We have to be careful not to over- or understate what this theorem means. It doesn't say that you don't want all accurate people. If individual people predict perfectly, they cannot be diverse. (If average individual error equals zero, then diversity must also equal zero.) Notice also that prediction diversity equals the *average* squared distance from the collective prediction, so adding someone who predicts differently need not increase overall prediction diversity. Prediction diversity increases only if the additional person's predictions differ by more, on average, than those of other people. This implies a limit to the amount of predictive diversity we can have. If a collection of people has an average individual error of one thousand, then their prediction diversity cannot exceed one thousand. Any more diversity and the collective error would become negative, an impossibility.

Fine, we've got some caveats. But they just reveal some of the theorem's subtleties. What's important is that we keep in mind the core insight: individual ability (the first term on the right-hand side) and collective diversity (the second term) contribute *equally* to collective predictive ability. *Being different is as important as being good.* Increasing prediction diversity by a unit results in the same reduction in collective error as does increasing average ability by a unit.

Contrasting the Diversity Trumps Ability Theorem with the Diversity Prediction Theorem reveals important differences. In making a prediction, a group of randomly selected predictors might or might not predict more accurately than a group of the best predictors. Randomly selected predictors will be more diverse, to be sure, but they will also be less accurate. The two effects work in opposite directions. So, we cannot expect that a random intelligent group will predict more accurately than the group of the best. Yet, that stronger claim holds in the problem-solving context. The reason why is that poor performers fail to drag down problem-solving teams. If we bring Larry, a social scientist, into our cheese-making business, his lack of relevant tools won't hurt our cheese making. We just ignore him. He may cause delay or frustration, but if he has only bad ideas—peppermint cheese—those ideas won't be adopted. However, if we're predicting how much cheese to make, we won't know that he doesn't know and his prediction gets averaged along with everyone else's. And he could make the crowd less wise.

An implication of the theorem is that a diverse crowd always predicts more accurately than the average of the individuals. This runs counter to our intuition. We can call this the Crowd Beats the Average Law.

The Crowd Beats the Average Law: *Given any collection of diverse predictive models, the collective prediction is more accurate than the average individual predictions*

Collective Prediction Error $<$ *Average Individual Error*

The Crowd Beats the Average Law follows from the Diversity Prediction Theorem. The Diversity Prediction Theorem says that collective error = average individual error − prediction diversity. Prediction diversity has to be positive if the predictions differ. Therefore, the collective error must be larger than the average individual error. There's no deep math going on. But the insight is powerful nonetheless.

We now have a logic for the wisdom of crowds. In an ideal world, these formal claims would replace pithy statements such as "two heads are better than one," but they may not be catchy enough. We can try though. We might replace the Diversity Prediction Theorem with "the wisdom of a crowd is equal parts ability and diversity" and the Crowd Beats the Average Law with "the crowd predicts better than the people in it." Not memorable, but accurate.

A Crowd of Draft Experts

To cement our understanding of the logic, let's consider some real data. Die-hard theorists prefer constructed examples because they are neater and cleaner. But sometimes even a theorist cannot help but peek out the window. So if we're going to look at data, we might as well look at something important: football draft selections. Table 8.9 shows predictions for the top dozen picks in the 2005 NFL draft from seven prognosticators. The players are listed in the order that they were selected. Each predictor provides a ranking of the draftees. We use the NFL draft because it has clean, integer-valued data, because it can be seen as a ramped-up version of our earlier example that involved Juliana and Michaela, and because these experts' predictions came from detailed analyses. They don't call them draft experts for nothing. These people, er men, devote long days and nights evaluating team needs, player skills, and a host of other factors.

If we look at their predictions, we see that they differ in their accuracy. The table reveals that some do far better than others. The last column, by the way, shows the crowd's prediction.[8] Here the crowd is just the collection of all seven predictors.

These data show the Crowd Beats the Average Law in full force. The average of the individual errors equals 137.3. The collective error, shown in the last column, equals about one-fourth of that, 34.4. In this example, the crowd predicts more accurately even than its most accurate member even though the Crowd Beats the Average Law makes no such claim.[9] The example also shows the power of diversity. These predictors are so diverse that they collectively predict well.

Even more amazing, note that this comparison between the crowd and its most accurate member is unfair. In selecting the best person after the fact, we stack the deck against the crowd. No one, other than perhaps Clark Judge himself, would have predicted Judge (despite his name) to be more accurate than the others. In the future, Judge may not be the best predictor. To take

TABLE 8.9:
Experts' predictions of 2005 NFL draft

	Predictor							
Player	*Wright*	*Alder*	*Fanball*	*SNews*	*Zimm*	*Prisco*	*Judge*	*Crowd*
Alex Smith	1	1	1	1	1	1	1	1.0
Ronnie Brown	2	2	4	2	2	5	2	2.7
Braylon Edwards	3	3	2	7	3	2	3	3.3
Cedric Benson	4	4	13	4	8	4	8	5.9
Carnell Williams	8	5	5	5	4	13	4	6.4
Adam Jones	16	9	6	8	6	6	9	8.1
Troy Williamson	13	14	12	12	13	7	7	9.7
Antrell Rolle	6	6	8	10	9	8	6	7.9
Carlos Rodgers	9	8	9	9	16	9	9	9.9
Mike Williams	7	7	7	6	7	12	12	8.0
Demarcus Ware	11	15	14	24	11	11	13	13.9
Shawn Merriman	12	11	3	11	12	10	11	10.1
Error^2	158	89	210	235	112	82	75	34.4

another example with higher stakes, successful investment funds differ from year to year. If at the beginning of the year, we could pick the fund that would do best at the year's end, investing would be fun and easy. But we cannot, so we diversify. By going with the crowd, we take on less risk. We should go with the expert only if we know that person to be far more accurate than the others and the others to make similar predictions.

Points and Ranges

Up until now, we have focused on the difference between the predictions and the outcomes. In many instances, we may want to know best- and worst-case scenarios. We want to know the range of possibilities. In building a stock portfolio, an investor may care about the range of possible prices. How high might the stock price go? How low might it go? In predicting a potential political uprising, a policy analyst may care less about having an accurate point prediction than about knowing worst- and

TABLE 8.10:
Range of predictions of 2005 NFL draft

Player	*Actual*	*Low*	*High*
Alex Smith	1	1	2
Ronnie Brown	2	2	5
Braylon Edwards	3	2	7
Cedric Benson	4	4	13
Carnell Williams	5	4	13
Adam Jones	6	9	16
Troy Williamson	7	7	14
Antrell Rolle	8	6	9
Carlos Rodgers	9	8	21
Mike Williams	10	6	12
Demarcus Ware	11	11	15
Shawn Merriman	12	3	13

best-case scenarios. We can look at the best and worst predictions and the actual outcomes (see table 8.10). In every case, the outcome falls within the range of predictions.

Amazing? No, not given the diversity of the predictions.

The Madness of Crowds

Up to now, we have not discussed communication among crowd members. If people can share predictions, then they might become less diverse. To paraphrase Socrates, it's much easier to go with the flow, and people often change their predictions to match those of others. And, rather than seeing wisdom emerge, we might see madness—we might see speculators buying tulips at crazy prices. We can use the Diversity Prediction Theorem to explain the madness of crowds. When we think of a crowd being mad, we think of a collection of people all taking an action that in retrospect doesn't make sense. The madness of crowds led people to drink the green Kool Aid. The madness of crowds leads people to burn cars and sometimes even houses after sporting events. The madness of crowds explains stock market bubbles and stock market crashes.[10]

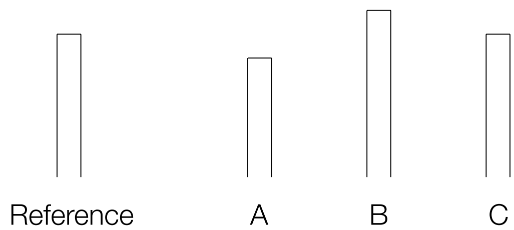

Figure 8.1 Asch's Lines

For a crowd to be mad, its members must systematically make the same bad decision. If people make these decisions in the heat of the moment—such as when burning a couch—we can chalk it up to the human tendency to join in, a topic we will return to in the epilogue. If though, people have time to construct what they believe to be reasonable predictive models, then we can often blame a lack of diversity. The Diversity Prediction Theorem implies that a crowd can make egregious errors only if the crowd members lack both accuracy and diversity.

Thus, the theorem shows the double-edged sword of deliberation. If people communicate with one another, if they share information and criticize one another's models, they can increase the accuracy of their models. However, they can also reduce their diversity. And it has been shown time and again that people often choose to abandon accurate predictive models in favor of inaccurate models. In a classic experiment, Solomon Asch asked people to compare the lengths of several lines. Each was given pictures with a reference line and three other lines marked A, B, and C.[11] Figure 8.1 provides an approximation of Asch's pictures.

Subjects were assembled together in a room and sequentially asked which lines were longer than the reference line, which lines were the same length as the reference line, and so on. The first subjects to answer were planted by Asch. They purposefully gave wrong answers. Asch found that others follow the majority—giving wrong answers—about one-third of the time. Given that people abandon their stated beliefs on the lengths of lines, we can hardly be surprised that they would abandon their beliefs in their predictions about the stock market, housing prices, or winning number combinations in the lottery.

More than just conformity leads to the madness of crowds. Often, in a group setting, people move too far in the direction of the majority opinion. So, if on average people think that prices are going to rise, then the group may work itself into a frenzy and begin to believe that because most people think prices are going up, prices are going to rise substantially.

Diversity's Free Lunch

To make the next step in our analysis of why and how crowds can be wise, we can build from an earlier insight that diverse interpretations lead to diverse predictive models. In Screening Success we saw how diverse interpretations lead to negatively correlated predictions using the Projection Property. This told us how crowds can sometimes be far wiser than we might expect.

To make this connection more explicit, we next analyze a class of examples in which people use diverse projection interpretations. In these examples, all of the interpretations rely on a common perspective. We then analyze an example in which the crowd members base their interpretations on different perspectives. We see that, in some cases, interpretations based on diverse perspectives can make a predictive task easier than it would be using predictive models that rely on interpretations based on either perspective alone. We have some magic after all. That magic results from diverse perspectives.

Different Parts of the Same Vision

If asked to make an important prediction such as who will win an election, whether an economy will grow, or the likelihood of armed conflict, we have to include many variables or attributes to make an accurate prediction. We've got to keep lots of variables in our heads. We might try a single-variable model, but they