```
                         Board A
1l  |
1h  |
2l  |
2h  |
3l  |
3h  |  55,70,80
4l  |  25,30,30,30,40,40,45,45,45
4h  |  50,55,55,60,60,70,75,75,75,75,75,85,90,95,95,95,95,95,95
5l  |  00,05,10,10,10,15,20,20,25,25,30,30,30,35,40
5h  |  50,65

6l  |
6h  |
7l  |
7h  |
8l  |
8h  |

                         Board B
1l  |
1h  |
2l  |
2h  |
3l  |  45
3h  |  50,75,80,90,90,95
4l  |  00,05,05,15,25,30,35
4h  |  50,50,50,55,55,60,70,70,70,85,85,85,90,95
5l  |  00,00,00,10,20,30,30,35,40,40
5h  |  50,60,60,70,70
6l  |  00,05,25
6h  |  65
7l  |
7h  |
8l  |
8h  |
                         Board C
1l  |
1h  |
2l  |
2h  |  70
3l  |  35
3h  |  50,60,70
4l  |  00,05,10,10,35,40
4h  |  50,50,70,70,75,75,80,80.85,90,95,95
5l  |  00,00,00,00,05,10,15,30,30
5h  |  50,50,50,65,70,70,80,95
6l  |  10,20,35,45
6h  |  50
7l  |  10
7h  |  95
8l  |
8h  |
```

Fig. 7
Stem-and-leaf plots of estimates, $n = 20$

TABLE 2

Summaries of Estimates

	Mean	Minimum	Lower Quartile (LQ)	Median	Upper Quartile (UQ)	Maximum	Range	Interquartile Range (IQR)
(n = 10)								
A	485	230	450	490	530	670	440	80
B	491	180	440	500	550	710	530	110
C	498	170	410	475	590	850	680	180
(n = 20)								
A	480	355	445	490	515	565	210	70
B	481	345	425	485	535	665	320	110
C	514	270	450	500	565	795	525	115

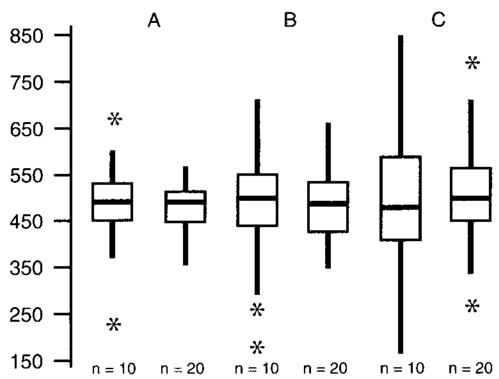


Fig. 8
Estimates from ninety-four samples, n = 10, and forty-seven samples, n = 20

of size n = 10, estimates from board A show the least variability and those from board C show the most variability. Similar differences are indicated for samples of size n = 20.

The simulated data suggest the following:

- Increasing the sample size decreases the variability of estimates.
- Sampling from board A produces the least variability of estimates, and sampling from board C gives the most.

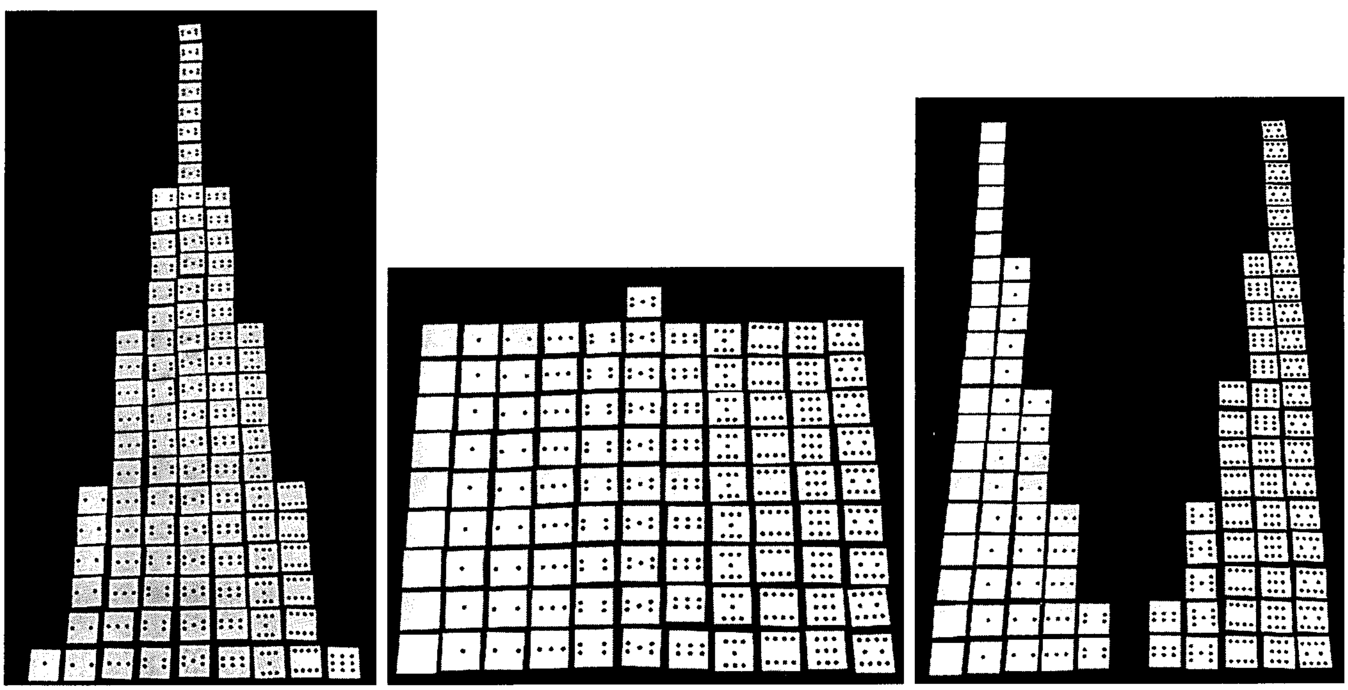
Can we explain the differences in results among the three boards? The answer lies in the differences among the distribution of dots on each board. The placement of the cards on the board is arbitrary and has no effect on the estimation because of random selection. To demonstrate these differences, the penguin cards from each board are arranged as shown in figures **9a, 9b,** and **9c.** When conducting the activity, do not label these three arrangements but rather allow the students to determine which is A, B, and C. Note that the distribution for board A is the most concentrated of the three around the middle. The variation around the mean is the smallest of the three distributions ($\sigma \approx 1.65$), which is why estimates based on samples from board A have the least variation. On the other extreme, the distribution for board C has the greatest variation about the mean ($\sigma \approx 4.07$), which is why estimates based on samples from board C have the most variability.

Increasing sample size decreases the variability of estimates

For some classes, it may be desirable to sample from only one board. In this situation, you would focus on the differences between sample sizes and not on the effect of the underlying population distribution. In other classes, you might consider only samples of size n = 10 and focus on the quality of estimation for one sample size from one particular population.

THE CENTRAL LIMIT THEOREM

It is of interest to note the approximate bell shape of each of the distributions of estimates represented



(a) (b) (c)

Fig. 9
Distribution of cards

by the stem-and-leaf plots in **figures 6** and **7** with each centered at $N = 500$. For instance, the ninety-four estimates from board B are shown in **figure 10.** This arrangement is a consequence of the central limit theorem: For "large" sample sizes, the distribution of sample means can be approximated by a normal curve with center at the population mean.

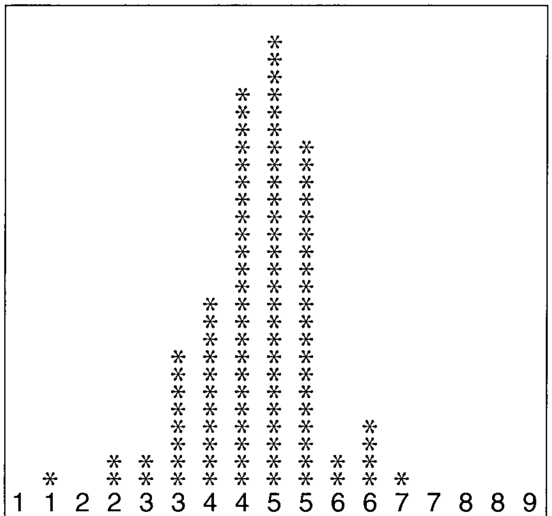


Fig. 10
Estimates of N from board B with $n = 10$

In our application, we are using a multiple of the sample means, but the principle still applies. The normal distribution and the central limit theorem are widely discussed in the literature (see, e.g., Moore [1995]). Students may not completely comprehend the concept of the central limit theorem initially. This activity could be extended to do additional simulations with the same sample sizes to see that replications of the experiment give similar results or to do additional simulations with different sample sizes to reinforce the concept.

CLASSROOM DISCUSSION

The activity may raise questions among students and generate discussion. This discussion may lead to further investigations. Potential questions for discussion follow:

- If we do the activity again, will we get the same or similar results?
- In what other situations would this estimation be appropriate?
- What would be some of the constraints or idiosyncrasies of these other situations?
- How many times would you expect to sample in such a simulation activity before you are fairly confident about your results?
- How does sampling without replacement affect results?

REFERENCES

Landwehr, J. M., and A. E. Watkins. *Exploring Data.* Rev. Student Ed. Palo Alto, Calif.: Dale Seymour Publications, 1994.

Moore, David S. *The Basic Practice of Statistics.* New York: W. H. Freeman & Co., 1995.

National Council of Teachers of Mathematics (NCTM). *Curriculum and Evaluation Standards for School Mathematics.* Reston, Va.: NCTM, 1989.

Scheaffer, R. L., W. Mendenhall, and L. Ott. *Elementary Survey Sampling.* Boston: PWS-Kent Publishing Co., 1990.

This work was supported by the National Science Foundation's Instructional Materials Development Program under grant no. ESI-9150117. The opinions expressed in this article do not necessarily represent the position, policy, or endorsement of the foundation.

Perry, Mike and Kader, Gary (February, 1998). Counting Penguins. *Mathematics Teacher, 91* (2), 110–116.