Limitations of the Ten Slide Cytology Proficiency Test Model

George K. Nagy, M.D.

Wadsworth Center
Albany, NY

The following comments are offered on the cytology proficiency testing as mandated by CLIA'88

The current model calls for a set of 10 test slides, and does not have the discriminatory power to differentiate between competent and incompetent test takers.

These tests may misclassify many test takers. A high proportion of competent test takers will fail the test, and a high proportion of incompetent test takers will pass.

If a competent individual fails a test, which we can call “type 1 error”, it is not a serious problem because the probability is high that that person will pass the second test event.

However, if incompetent individuals pass the test (type 2 error), it is worrisome because they will not be caught and they will continue to screen cytology slides until the next test cycle.

The statistics of “misclassification” can be explained with an example.

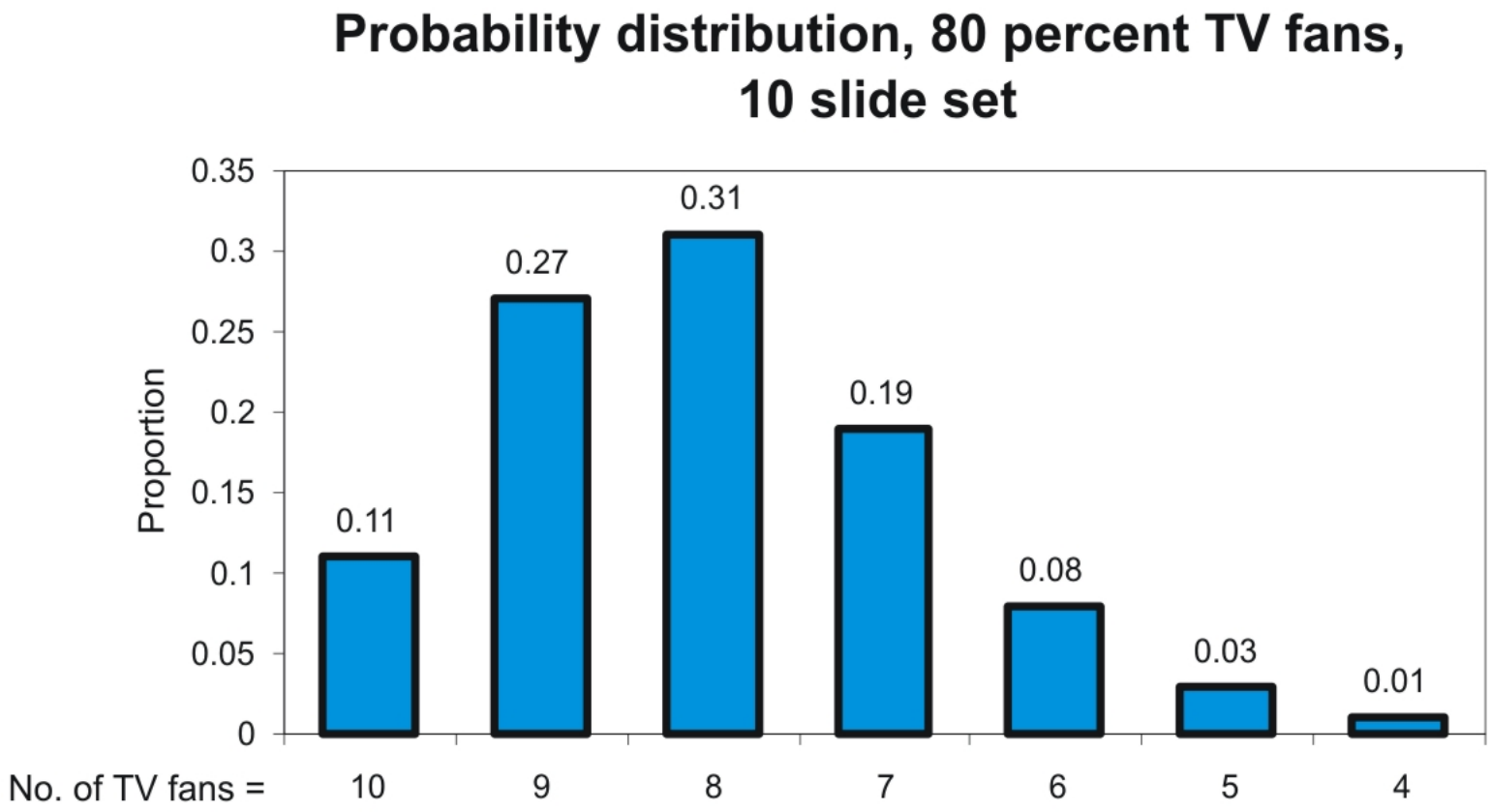
Let us suppose that in a large population a statistical survey finds that the proportion of people who like television is 80 percent.

The basic question on which the analogy with proficiency testing is based:

What can we expect if we ask 10 randomly selected individuals in this population about their attitude towards television?

We will get a so-called “probability distribution”. Most frequently there will be 8 persons in the sample of 10 who like television. But sometimes all 10 of them will be TV fans, and sometimes just 7 persons will like TV.

The probabilities can be calculated using the so-called binomial formula.


Probability distribution, 80 percent TV fans, 10 slide set
Proportion
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.11
0.27
0.31
0.19
0.08
0.03
0.01
No. of TV fans =
10
9
8
7
6
5
4

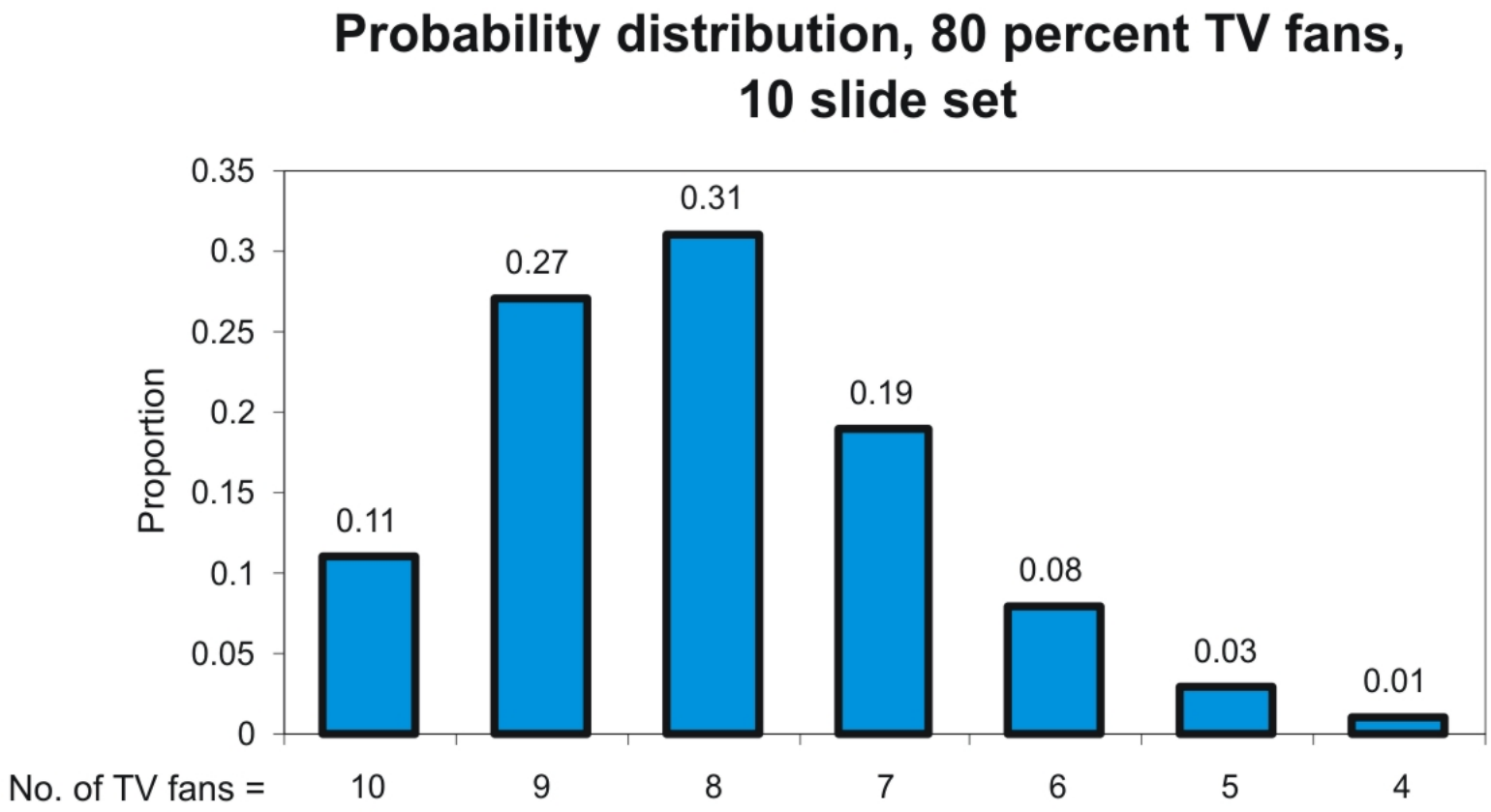
If we increase the size of the sample, we will get a different probability distribution. Larger samples give more accurate estimations.

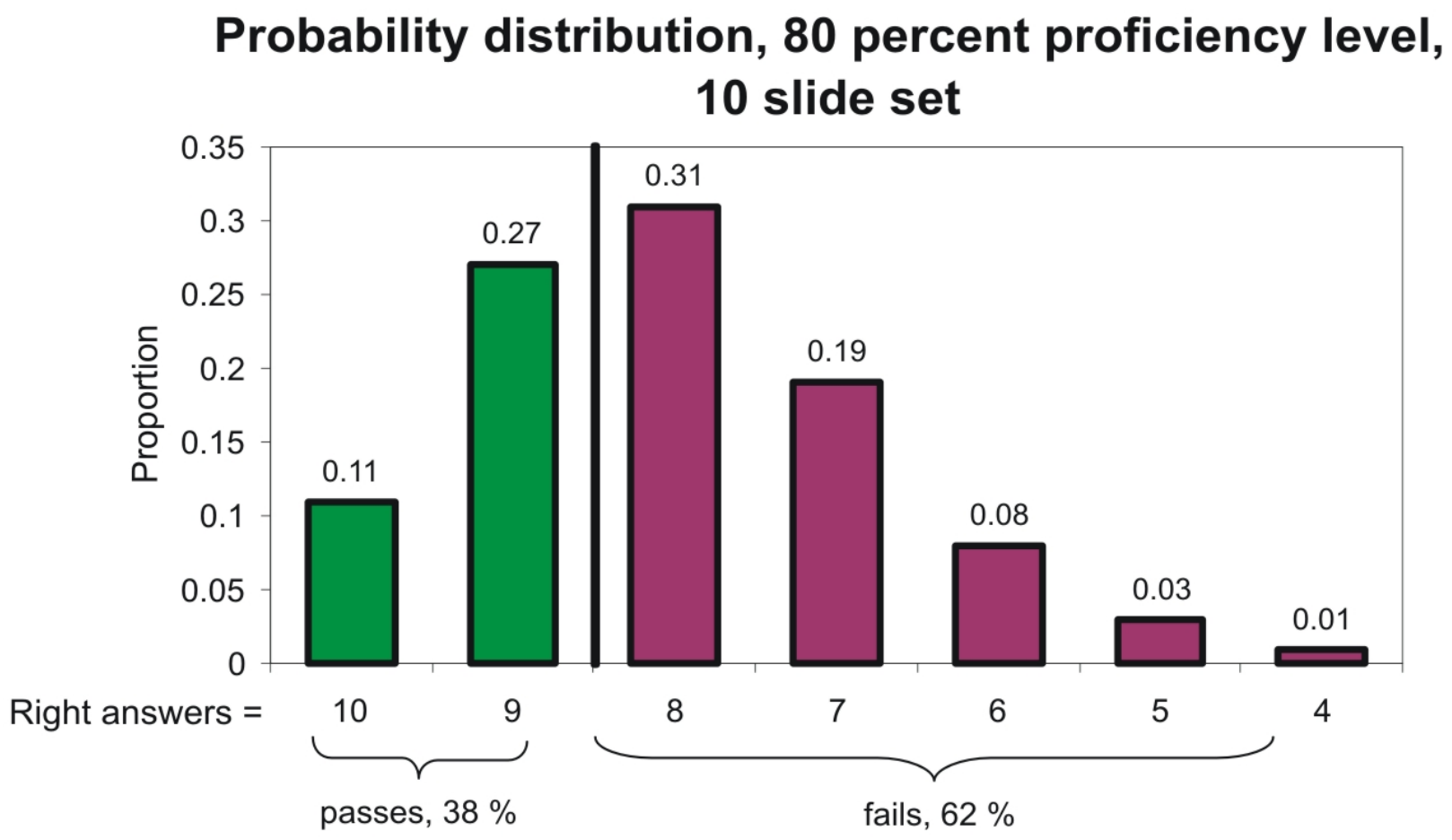
That is why political polls use large samples consisting of several thousand people.

If we use large samples in proficiency testing (test sets consisting of many test slides), we will substantially reduce misclassifications.

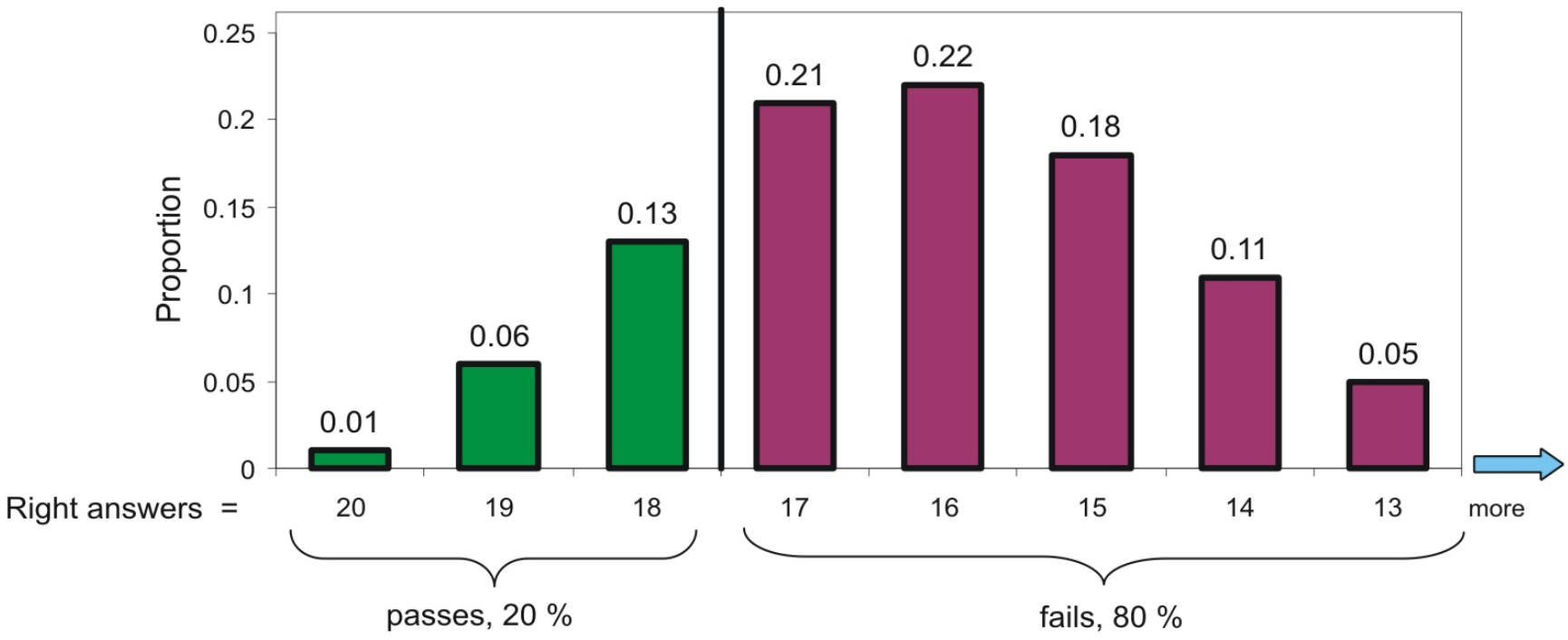
The binomial distribution does not depend on the specific circumstances. The only data necessary: **the probability and the sample size.**

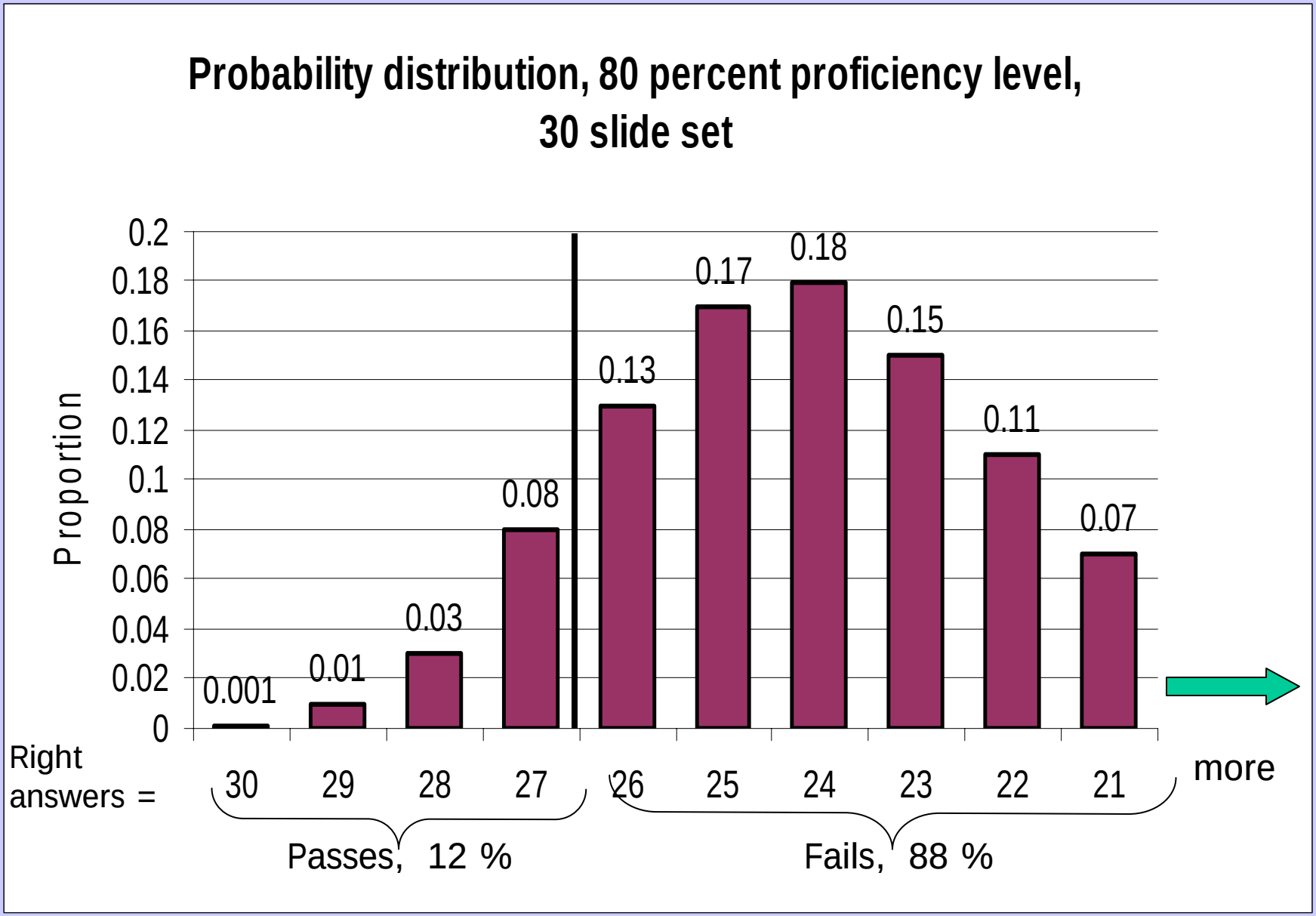
The percentages in the probability distributions are the same if we have 80 percent TV fans and samples of 10 persons, or 80 percent “competency level” and test sets consisting of 10 slides.


Probability distribution, 80 percent TV fans, 10 slide set
Proportion
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.11
0.27
0.31
0.19
0.08
0.03
0.01
No. of TV fans =
10
9
8
7
6
5
4

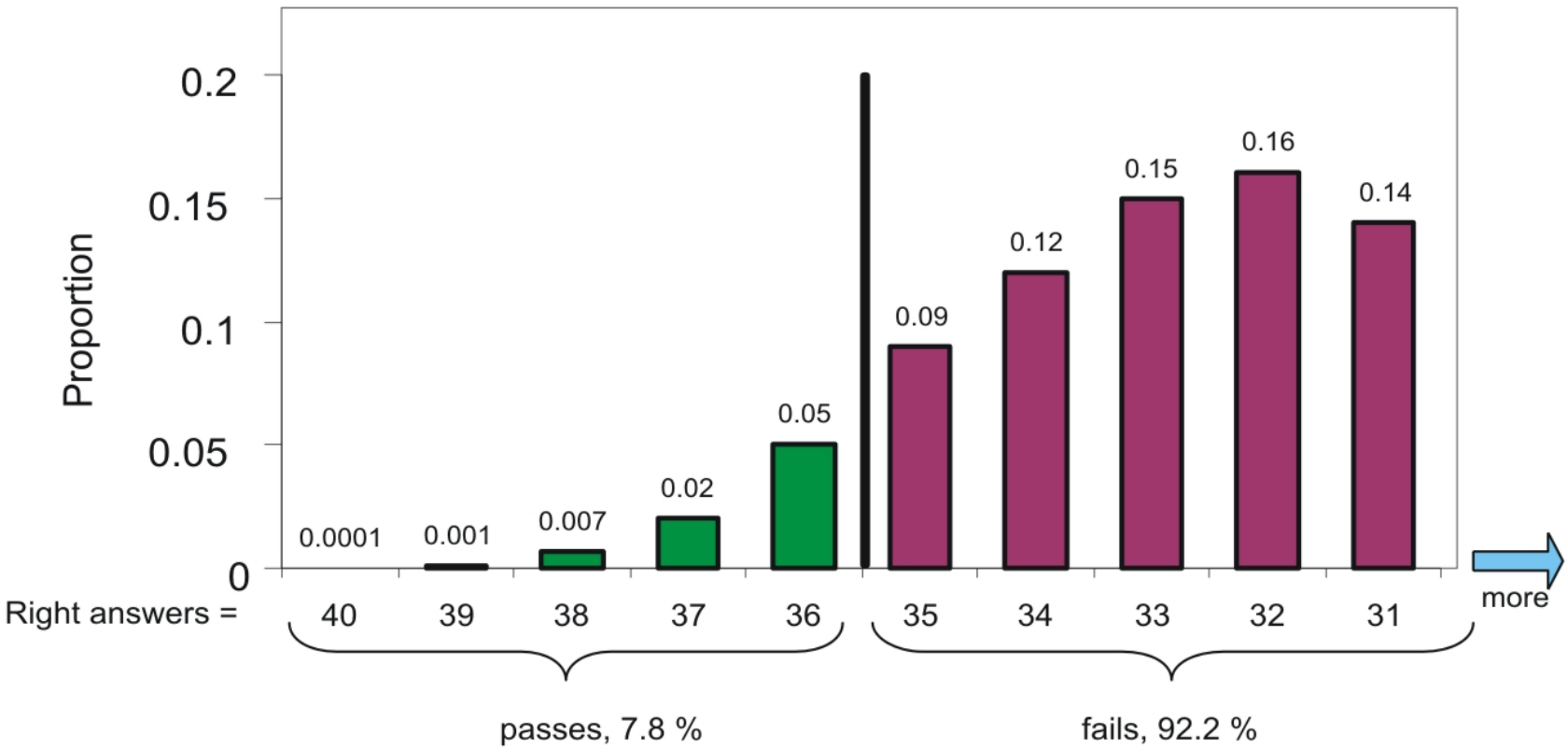

Probability distribution, 80 percent proficiency level, 10 slide set
Proportion
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.11
0.27
0.31
0.19
0.08
0.03
0.01
Right answers =
10
9
8
7
6
5
4
passes, 38 %
fails, 62 %

Probability distribution, 80 percent proficiency level, 20 slide set




Probability distribution, 80 percent proficiency level,
30 slide set
Proportion
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0.001
0.01
0.03
0.08
0.13
0.17
0.18
0.15
0.11
0.07
Right answers =
30
29
28
27
26
25
24
23
22
21
more
Passes, 12 %
Fails, 88 %

Probability distribution, 80 percent proficiency level, 40 slides



Probabilities of passing the test for varying competence levels

Competence levels

	Unacceptable			Acceptable		
No. of slides in test set	75%	80%	85%	90%	95%	97.50%
10 slides	0.24	0.38	0.54	0.74	0.91	0.98
20 slides	0.09	0.21	0.40	0.68	0.92	0.99
30 slides	0.04	0.12	0.32	0.65	0.94	0.99
40 slides	0.02	0.08	0.26	0.63	0.95	1.00

The main purpose of the increase in number of test slides would not be to reduce type 1 errors (as it can be seen in the table, for complicated statistical reasons, it is not easy).

But the reduction of type 2 errors is spectacular when we increase the number of slides in the test sets. That is the main purpose of cytology proficiency testing:
To catch those whose skills are in doubt, and remediate them.

One could say that this is just statistical speculation, but the reality of these speculations have been verified by the statistical results of the federally mandated proficiency tests in 2005.

According to CMS statistics, among 12,786 test takers 1,156 (9 percent) failed the initial testing event.

At the second testing events these individuals markedly improved. Only 10 percent of them failed. They went from 100 percent failure to 90 percent success.

Is it possible to improve so much in a few weeks?

Conclusions

A proficiency test consisting of a low number of test slides has a high misclassification rate

An initial test consisting of 40-60 slides administered at the onset of the cytotechnologists career would provide a more accurate assessment of competency

The ability to screen slides does not diminish over time, thus this initial assessment could be followed by proficiency tests at intervals of as much as five-ten years, as is currently the practice for recertification by many professional boards