

## ON THE RELATIONSHIPS BETWEEN STATISTICS AND OTHER SUBJECTS IN THE K-12 CURRICULUM

Zalman Usiskin

The University of Chicago, Chicago, IL USA

[z-usiskin@uchicago.edu](mailto:z-usiskin@uchicago.edu)

*Sixty years ago, statistics barely touched the school experience of a typical student. In the study of social science, students might encounter data. In a science laboratory experience, students might collect data. In a mathematics classroom, students would be expected to know how to calculate the mean of a data set. In contrast, today it is becoming prevalent to expect increasing numbers of students to learn several measures of central tendency and spread, to encounter theoretical and actual distributions, and to discuss topics such as randomness, statistical tests, and statistical significance that in the past were introduced at the college level. As one of the mathematical sciences, the study of statistics in grades K-12 naturally has been considered as a part of the school mathematics curriculum. This has great advantages, as mathematics is the second most important academic school subject, behind reading and language arts. But as statistics has become more important, its connections with everyday literacy, science, health, and the social sciences suggest teaching statistics across the curriculum in addition to a reconsideration of its relationships with mathematics.*

The views in this paper stem from three premises. The first is a truism to most who will read this paper. (1) Statistics are found and applied today in almost every domain of human activity. However, the second premise is probably not believed by all readers of this paper. (2) The fundamental ideas of statistics, including variability, randomness, the laws of probability, etc., are accessible to virtually all students. Even if you think not all students can learn the basic ideas of statistics, I believe school policy should operate as if they can and give everyone the opportunity. (3) The third premise comes from the GAISE document of the American Statistical Association. “Every high school graduate should be able to use sound statistical reasoning in order to cope intelligently with the requirements of citizenship, employment and family, and to be prepared for a healthy, happy and productive life” (Franklin, 2007).

In most places in the world, despite wonderful strides over the past generation, it is not currently the case that all high school graduates are *familiar* with the tenets of sound statistical reasoning, let alone *able to use* them. In many places only a small minority of students are exposed to these ideas. Statistics is not a traditional subject at the K-12 level. The subjects that are in the curriculum compete for time.

The basic question that I attempt to treat here is how and where should we teach statistics when there are so many other subjects competing for this time. My approach to the answer is somewhat historical, looking at the past, observing the present, and looking to the future to sustain and grow our subject. It is, for the most part, a United States perspective.

### TEACH STATISTICS WITHIN MATHEMATICS

A half-century ago and still in many school situations, the answer to this question was simple. Statistics was viewed as a branch of mathematics like algebra or geometry, so teach statistics within mathematics. This view is reflected in the TIMSS international studies and its predecessors (e.g., see TIMSS, undated; Medrich & Griffith, 1992), and in the U.S. Common Core State Standards Initiative (2010). In this view, we teach statistics within the time frame that is allocated to mathematics.

Why would we want to do this?

- (1) Statistical formulas involve mathematics, and so why would anyone not think of statistics as a part of mathematics?
- (2) Statistics is in many if not most schools a relatively new area of study in the curriculum. On the other hand, mathematics is a fixture of the K-12 curriculum. The school day is limited in time and there is competition for the time that is available. If we fit statistics within mathematics,

although we have to compete with other areas of mathematics, we do not have to carve out a separate time in the day for statistics.

- (3) Mathematics is tested everywhere. So it is taught everywhere. Topics that are not tested are easy to omit and it is difficult to build on previous years' work if you cannot be certain that this work was actually taught. So, if we place statistics within mathematics, we can be more confident that statistics will be tested and taught.

The view of statistics as a branch of mathematics is the view of statistics that I was educated to have. When I went to school in the 1950s, very little statistics was taught in mathematics courses in grades K-12 in the United States. By the end of grade 8, a student could expect to have seen bar graphs, circle graphs, and line graphs, but the notion of collecting your own data for such graphs was absent. The only statistical measures taught were the mean and median. In grades 9 to 12, non-college intending students would likely see these topics reviewed, but college-intending students would see no statistics in their mathematics courses and no statistics outside those courses, either.

Through my undergraduate career as a mathematics major and a statistics minor, all of the statistics courses I took were inside a mathematics department. I never saw a data set. In graduate school, though I used the statistics commonly found in the educational field, such as hypothesis testing with t-tests, analysis of variance, and the like, these were taught as statistics for specialists and not viewed as the kind of stuff you would learn *before* college.

I have believed strongly that all students need to encounter a significant amount of statistics in their schooling, starting in the earliest grades, and the most likely way that students would encounter this content would be to embed the content within the mathematics curriculum and not to teach statistics in separate units that could easily be omitted. Reflecting this view, in my curriculum work for over 40 years I have tried to match traditional topics in the mathematics curriculum with related ideas from statistics, and going in the opposite direction, looked at basic ideas from statistics for the mathematics involved in them.

Table 1 shows a number of traditional mathematics and statistics topics that go together.

Table 1: Traditional K-12 mathematics topics and related statistical ideas

<b>Traditional mathematics topic</b>	<b>Statistics topic to motivate or apply</b>
Distance	Deviation
Absolute value	Mean absolute deviation
Graphing in two dimensions	Bivariate data
Subscripts	Identifying data
Function	Distribution
Function notation $f(x)$	Probability notation $P(E)$
Area under a curve	Normal distribution
Line through two points	Line of best fit through $n$ points
Summation	Mean
Squares of binomials	Chi-square and standard deviation
Square root	Standard deviation
Translation (geometric transformation)	Translation of data
Scale change transformation (stretch)	Scaling data
Base of natural logarithms ( $e$ )	Normal distribution

There is nothing particularly inventive about finding these connections, but teaching them together often requires careful thought and raises non-trivial questions. For instance, consider an easy connection. Euclidean geometry tells us that there is exactly one line through 2 points and, if those points are identified on a coordinate plane, it is a standard algebra topic to find an equation for that line. It is appropriate mathematical thinking to ask: what about finding an equation of a line through 3 points? Typically those three points will not lie on the same line, so what about finding

an equation of a line that is closest to the 3 points? It is a beautiful and natural way to get into the line of best fit, and it generalizes to  $n$  points as we all know.

However, the algorithm for finding the line of best fit is not at all like the algorithm for finding the equation of the line through two points. In grades 8-10, the line of best fit cannot be obtained with the same sort of attention to understanding the algorithm as the line through 2 points. Furthermore, the geometry is not the same. Why do we examine the square of the vertical distance from a point to a line rather than just the perpendicular distance from a point to a line? Why is the line of best fit for  $n$  points  $(x,y)$  usually not the same as the line of best fit for  $n$  points  $(y,x)$ , that is with the points with coordinates switched? The geometry has different criteria for distance than the statistics, where a level of prediction is involved, and predicting  $y$  from  $x$  is quite different from predicting  $x$  from  $y$ .

Sometimes there are surface connections between the mathematics and the statistics that are not easy to explain, such as why  $e$ , the base of natural logarithms, is the base of the exponential expression in the formula for the normal distribution.

On the other hand, there are times when the statistics can precede the corresponding traditional mathematics topic. The idea of mean absolute deviation can precede the introduction of absolute value and distance between two points on a number line. That is, the statistics idea can motivate the mathematical idea. At other times, more than motivation is involved. Subscripts to identify student scores on a classroom test as  $s_1, s_2$ , and so on, are easier for students to grasp than the traditional  $x_1, x_2, y_1, y_2$  that are – at least in the United States - traditionally introduced when formula for the slope (gradient) of a line is to be studied. Why is it easier? I think it is because when there are scores  $s_1, s_2, s_3, \dots s_{25}$ , the subscripts make it obvious that these identify 25 scores, but when there are only the two subscripts  $x_1$  and  $x_2$ , the naïve student confuses the subscript with the exponent and looks to do some sort of calculation with the 1 and the 2.

At still other times, the traditional application of an idea and a use of the idea in statistics are quite complementary, helping each other out. Scaling data and scaling graphs of functions and relations involve the same mathematics. The formula for standard deviation helps students to realize that square roots are found in other places than right triangles.

The examples in Table 1 are from the latter half of the K-12 experience, but there are numerous connections between statistics and the mathematics of earlier grades, some of which are in Table 2.

Table 2: Traditional K-6 mathematics topics and related statistical ideas.

<b>Traditional mathematics topic</b>	<b>Statistics topic to motivate or apply</b>
Number line	Bar graph
Counting	Frequency
Subtraction as Comparison	Error
Rate	Mean
Place value	Stem-and-leaf display
Ordering numbers	Ranking
Time (duration)	Collecting time data that varies
Length (distance)	Collecting data such as the distances from homes to school
Fractions between 0 and 1	Circle graphs

All in all, the many places in which the mathematical ideas in statistics can be associated with topics in the traditional mathematics curriculum are quite fruitful for the teaching not just of statistics, but also for the teaching of traditional mathematics. And it has been a major impetus for introducing some statistics into the curriculum.

But how well does statistics fare in an environment where it competes with other mathematical areas, such as arithmetic, algebra, and geometry?

Consider the most recently published TIMSS, from 2011. At grade 4, there are three item content categories: number, geometric shapes and measures, and data display. Data display items constitute 15% of the total score points available. At grade 8, there are four content categories:

number, algebra, geometry, and data and chance, and the data and chance items constitute 19% of the total score points available (TIMSS, undated, p. 426).

For another comparison, consider the Common Core State Standards in the United States. At grades 9-12, the number of standards in statistics and probability is about 20% of all standards; at grades 6-8 about 23%, and at grades 1-5 about 5%. Except for grades 1-5, this is an extraordinary inroad into a mathematics curriculum in the U.S. that in the past has tended to be very traditional and difficult to change. It is hard to imagine the mathematics curriculum giving much more attention to statistics – except at grades 1-5, where the Common Core has simply missed the boat.

**TEACH STATISTICS AS APPLIED MATHEMATICS**

A second way to view statistics in the K-12 curriculum is as applied mathematics. Teaching statistics as applied mathematics recognizes that, like other areas that we identify as applied mathematics, such as operations research, computer science, or even physics, the issues and problems statistics treats originate from outside mathematics and are not taught simply because we think it is nice for students to learn about ideas such as center, spread, variability, distribution, and the like.

The process of mathematical modeling used in applied mathematics has often been depicted in a diagram similar to the one shown in Figure 1.

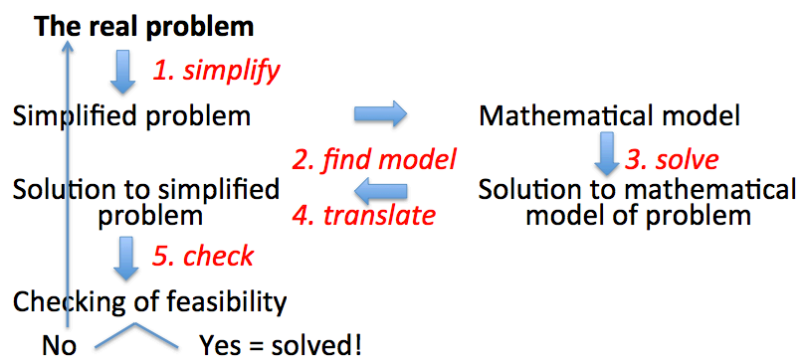


Figure 1: The modeling process

In the modeling process, the basic idea is that one goes from the “real world” into the “world of mathematics”, works with the model in the mathematical world, and then goes back out to check with reality.

How does teaching statistics as applied mathematics differ from teaching statistics within mathematics? In teaching a statistical idea within mathematics, the motivation for the idea might come from a mathematical idea that we wish to reinforce. For instance, we might introduce the idea of mean absolute deviation just after or just before we teach absolute value, and examine the idea for its theoretical properties without much concern for the application. But if we are to teach mean absolute deviation as applied mathematics, then the problem needs to arise from outside mathematics. So we look for one or more real situations in which the concept of spread is meaningful.

One of my favorite modeling problems for grades 6-9 is to examine the question of when a person should leave from home to the nearest airport in order to be on time for a flight. Let us suppose that the departure time for a plane flight is ten o’clock, a person is driving to the airport and will leave a car there. Let  $L$  be the time to leave. Call the departure time  $D$ . So  $D = 10$ . We can identify some variables. Let  $G$  be the length of time before departure you want to be at the gate. Let  $T$  be the time to travel from home to airport, and let  $P$  be the time to park the car. Let  $B$  be the time for getting a boarding pass and let  $S$  be the time it takes to get through security. It seems like this problem is a pure subtraction problem, working back from 10. We have a simple mathematical model:

$$L = D - G - T - P - B - S = 10 - G - T - P - B - S = 10 - (G + T + P + B + S).$$

Statistics in this model comes from the fact that  $T$ ,  $P$ ,  $B$ , and  $S$  all vary! The travel time  $T$  from home depends on the route and the traffic. If someone else drives you to the airport, then  $P = 0$ . If you print your boarding pass from home and do not check bags, then  $B = 0$ . The security time  $S$  may range from a few minutes to 30 minutes or more.

The problem begs for some actual data. How does  $T$  vary? With Google Maps or MapQuest or other travel time software, today it is possible to collect data about travel times on various routes over a period of days without having to do the travel. It is also possible and would be interesting to collect data on the length of time  $S$  it takes passengers to get through security at the airport nearest one's home. For both  $T$  and  $S$ , we would be particularly interested in measuring the spread of the distribution of times. Measures of spread become appropriate because of the need to factor in a person's tolerance for risk and provide one way to approach mean absolute deviation and other measures of spread.

As nice as so many examples like this one can be in a mathematics classroom, there are problems with approaching statistics either as a subfield of mathematics or applied mathematics. The mathematics curriculum is crowded; it is not a curriculum with loads of room begging for topics to fill it. Many mathematics teachers are uncomfortable with the lack of certainty arising from statistical situations, and with statistics in general. Even statistical units within mathematics courses tend to become devoured by other mathematics.

However, perhaps the most difficult problem of incorporating statistics within mathematics is that statistics has its own body of knowledge separate from mathematics. It is difficult to incorporate into traditional algebra, geometry, or analysis (1) questions of sample selection, (2) more than the simplest kinds of exploratory data analysis, (3) the interplay between the degree of confidence in a statistic and the length of the confidence interval, and (4) many other rather basic ideas of statistics – even mathematical ideas – that take some time to develop. As a result, within a mathematics course, one can develop a taste for statistics, discuss basic notation and some language, and show how one needs mathematical knowledge to treat statistics, but for intense study, statistics needs its own courses even at the pre-college level.

#### TEACH STATISTICS AS AN INDEPENDENT SUBJECT

At the college level, the idea of treating statistics separate from mathematics or applied mathematics gained some traction after the second world war. I do not know the first university in the world to form a statistics department, but in the U.S. the first statistics department was formed at North Carolina State University in 1941 by Gertrude Cox, and soon thereafter departments were initiated at Columbia University in 1946, at Chicago in 1949, and at Harvard in 1957. Perhaps this has been more a U.S. phenomenon because it took until 1988 for Oxford to have a department of statistics. At the same time, many other colleges and universities broadened the name of the mathematics department to “mathematics and statistics”, “mathematical sciences”, or even “mathematics, computer science, and statistics”.

Shortly after the department of statistics at the University of Chicago was created, perhaps to emphasize the difference between statistics and mathematics, two of its faculty, Allen Wallis and Harry Roberts, described statistics in a way that does not mention mathematics. They wrote, “Statistics is a body of methods for making wise decisions in the face of uncertainty.” This view of statistics was quite a departure for its time. They continued,

This modern conception of the subject is a far cry from that usually held by laymen. Indeed, even the pioneers in statistical research have adopted it only within the past decade or so.” Today, instead of “uncertainty”, which invokes immediately the mathematical idea of probability, the word “variability” is more commonly used – it serves better in distinguishing statistics from mathematics. (Wallis & Roberts, 1956, p. 11)

The view of statistics as its own body of knowledge, related to but separate from mathematics (Cobb & Moore, 1997), dominates many introductory courses at the college level today. It is reflected in the simple description of statistics found in one popular college textbook, “Statistics is the science of data” (Yates, Moore, & Starnes, 2003, p. 4).

The GAISE report emphasizes this break from mathematics.

A major objective of statistics education is to help students develop statistical thinking. Statistical thinking, in large part, must deal with this omnipresence of variability; statistical problem solving and decision making depend on understanding, explaining and quantifying the variability in the data. It is this focus on *variability in data* that sets statistics apart from mathematics. (Franklin et al., p. 5)

This position of statistics relative to mathematics is somewhat like that of physics relative to mathematics. It is impossible to do much physics without knowing mathematics, and it could be said that physics is the application of mathematics to the study of matter and energy and their interactions with various forces. Yet there is so much to physics that is not mathematics that almost everywhere physics is studied as a subject in its own right, not as a part of a mathematics course.

This belief in statistics as its own field of study in high school, separate from mathematics, began to take hold in the United States within the last generation. In 1990, only 1% of high school graduates with at least one semester of statistics; by 2009, 11% had. This compares with 35% who have credits in precalculus mathematics and 16% with at least one credit in calculus (NCES, 2013).

These increases go along with increases in participation in the Advanced Placement (AP) program, by which students take the equivalent of selected college courses in high school and are given national tests on each course. The AP program began in 1955 and calculus was there from the beginning. The AP program in statistics began in 1997 and since the year 2000 the number of students taking the exam has increased by about 10% a year (Table 3). By the 2012-13 school year, the most recent year for which data has been published (College Board, 2014), almost 170,000 students, or about 5% of all high school graduates in the U.S., were taking the advanced placement statistics exam, and perhaps an equal number are taking the course without taking the exam.

Table 3. Number of advanced placement test takers in the U.S. in calculus and statistics (in 000s)

Year	Statistics	AB Calculus	BC Calculus
1999	25	124	30
2001	41	147	38
2003	58	167	46
2005	77	186	54
2007	98	212	64
2009	117	231	73
2011	143	255	85
2013	170	283	104

Sources: The College Board, Program Summary Report, various years

When the AP statistics exam was first offered, I am certain there were those who felt that enrollments in the AP calculus exam would fall. But, as the data in Table 3 show, the increases in advanced placement statistics have not come at the expense of taking advanced placement calculus. Since 2005, each year more students in the U.S. have been enrolled in calculus courses in high school than in college.

But these are college-level courses even if taught in high schools. Full courses devoted to statistics at the high school level for all students seem to be less common. Furthermore, by including statistics as part of the Common Core standards in the United States, the movement for a separate course may have been slowed down.

Nevertheless, the history of school mathematics around the world is that the content we teach now to students in primary and secondary schools used to be taught to much older students. New content often appears first at the college or advanced secondary level and then filters down over the years, decades, and centuries. So it has been with arithmetic, algebra, and geometry, and so it is with statistics. It is quite likely that within a couple of generations most of the content of the course in introductory statistics at the college level will filter down to become standard school content.

## TEACH STATISTICS ACROSS THE CURRICULUM

At the same time that the movement towards statistics courses in the secondary school was beginning, there was a concerted movement to reach students who in the past dropped out of their study of mathematics quite early in schooling, say after grade 9. In the U.S. the slogan was embodied in the title of the report *Everybody Counts* (National Research Council, 1989); worldwide there were the mantras “mathematics for all”, “algebra for all”. In this arena there arrived “statistics for all”.

In the same year as *Everybody Counts*, the view that statistics is important in the early grades was represented in a UNESCO publication (Morris, 1989), based on presentations at ICOTS-2. A strong theme of this publication is statistics for all, with articles on the exploration of data (e.g., see Biehler, 1989; Dunkels, 1989), on quantitative literacy (e.g., see Burrill, 1989), and many other interesting contributions.

It is significant that these themes were laid out before the first public access to the internet in 1990, before public e-mail, before Facebook and other social networks, before mobile phones and apps, that is, before the accessibility of data to everyone and before big data. In the work of these and many other people we see the move to introduce the collection and discussion of data early in primary school and the desire for grades 6-8 to be a particularly useful time for introduction of more formal statistical concepts. As the years ago on, we see more and more confirmation of the belief that, because data is present everywhere, the study of statistics has implications for the study of virtually all other disciplines.

This is nothing new for statistics at higher levels. But as the years have gone on, consideration has been given to younger and younger students (e.g., see Royal Statistical Society (2005)), to teach statistics across the curriculum.

Even at the earliest grades, a strong case can be made for treating some basic statistical concepts inside disciplines other than mathematics. First, it is important for all students to learn reasons for statistics early, and these reasons are imbedded in the understanding of other subject areas. Second, statistics has different associations with the major subject areas of early education. Third, particularly in the early grades, in most places many of these subjects are taught by the same teacher, making cross-subject teaching easier to implement.

Here we are speaking of students of ages 6 to 12, young students. The school curriculum for students of ages 6 to 12 consists mainly of:

- Language arts (reading, writing, speaking)
- Mathematics
- Social studies (history, government, geography, economics)
- Science (physical science, earth science, biological science)
- Health and physical education.

When we think of the school curriculum at these early grade levels, we think of what *all* people need to know for survival and for citizenship as well as grounding for more advanced work.

Other subjects may include the study of foreign languages, fine arts, applied arts such as woodworking and cooking and, in some places, moral and religious education. I cannot make the case for incorporating statistics within all of these areas, but I would like to mention a few ways to support the general theme of statistics across the curriculum in the main subject areas.

Language arts covers the understanding and the use of one’s native language through reading, writing, and speaking. Throughout the world, language arts understandably dominates the school curriculum in the first few grades. In the United States, the time devoted to language arts is two to three times the time devoted to mathematics, the subject receiving the second most attention. It is in language arts that students are expected to attain basic literacy. We want citizens to be able to read and listen with understanding to a fictional story or nonfiction news account, to be able to digest critically the information they hear or see. It is not coincidence that, in the GAISE report, the first reference to statistical literacy concerns what is reported in a newspaper and in other media (Franklin et al., 2007, p. 1).

The amount of information in a single issue of a daily newspaper can be staggering. How many numbers are there in a newspaper or on a typical page of the newspaper? This is an

interesting statistical question. What do we mean by newspaper? That is, what is the population? How should we select a sample? I decided not to do a statistical study and just examine one day's newspaper and count all the numbers myself.

I picked what I considered to be a typical day of a large city newspaper that I receive daily: Thursday, February 6<sup>th</sup> of this year of the *Chicago Tribune*. I then tried to determine the number of numbers on each of the 64 pages of the main six sections of the newspaper – News, Business, Sports, Arts and Entertainment, Dining, and AutoMart. I did not include 70 pages of advertising inserts.

Now that I had selected the newspaper and the pages, the next question was “What is a number?” Is a word like “first”, or “three”, or “millions” a number? Is a telephone number 10 numbers, 4 numbers, 1 number, or not a number at all? Is the score of a football game to be counted as two numbers or one? And so on. Table 4 shows some of the decisions I made.

Table 4: Counting appearances of “numbers” in a newspaper

Appearance	General type	How many numbers?
three hundred	word name	1
1-800-555-1212	telephone number	1
20-30%	interval	2
first	ordinal	1
several	estimate	1
millions	estimate	1
July 3, 2014	date	1
B-29	product code name	1
none	absence of something	not a number
one	a single object	1
one	an abstract person	not a number
double	multiplier	1
half	fraction	1
7-3	score	2

The number of numbers on a page ranged from 13 on the first page of the dining section, a page dominated by a large picture of a dinner plate with food, to 2278 on a page containing sports results. The total number of numbers was 13,518, so the mean number of numbers on a page was just over 211, with a standard deviation of about 426. The median was 72.5.

Thus there are loads of numbers in this newspaper issue, but is there statistics? Numerical literacy does not guarantee statistical literacy. Fortunately for those who teach statistics, the newspaper is rich with examples of discussion of decisions made based on analyses of data. Here are some examples identified by section-page together with statistical questions the examples might engender for young students. The examples are mostly understandable at early grades. The questions can range from those grades to the college level.

- *1-4* A report on the Chicago area mass transit system includes two maps showing census block areas underserved by mass transit. What is a census block area? Why are the data organized in this way?
- *1-8* A summary of changes in police staffing reports a slew of data: “...the department had 7,078 rank-and-file officers and supervisors assigned to work in its 22 patrol districts as of December 8<sup>th</sup>, a decline of 10 percent, or 779 beat officers, since fall 2011.” Also included is a bar graph of staffing over the past five years. Why is the time line so important?
- *1-16* Reported are the results of a survey of Hispanic Americans regarding their views on upward mobility in the economy. Unusual is that the actual survey questions and choices given as responses are shown with the percents of the sample who chose each response.



- 1-20 The editorial is about the question of government-sponsored health insurance and the incentive to work, responding to a report from the Congressional Budget Office (CBO) in Washington. The subhead of the editorial is: “What’s behind the CBO numbers?” A teacher does not even need to frame the question! What answers are given in the editorial?
- 2-1 The lead article of the Business Section 2 is about the decision by CVS, a major pharmacy chain, to stop selling tobacco products this year. There are graphs of the number of cigarettes sold and the total retail value of cigarettes for each year since 1998 and a circle graph of the distribution of sales of cigarettes in various types of retail stores in 2012. What trends do these show and how did these trends fit into the CVS decision?
- 2-5 Half of a business page is devoted to a summary of the stock market action from the preceding day, with a line graph, three bar graphs, and lots of data about most active stocks, those most held, those with the largest capitalization, and more. How are these data collected? Why are these particular attributes of the data chosen to be presented?
- 3-6 Of course, the sports pages contain a lot of data. One unusual chart is a listing of the top 10 football signing classes, that is, a ranking of the universities that are thought to have recruited the best high school graduates this year according to the website *247sports.com*. The *247sports* website is said to have combined several rankings to get the composite ranking shown. How are these ranking systems created? What is the significance of Arrow’s paradox for combining ratings into a composite ranking?
- 3-many In addition to the results of games, the sports section contains a huge amount of statistics about teams and players. Why are particular statistics displayed? Which statistics are most important?
- 4-1 In the Arts and Entertainment section, the lead article concerns the artwork of a deceased painter whose family wishes to sell his paintings. How does statistics play into the development of asking and selling prices for paintings?
- 4-10 The weather page is filled with forecasts, temperatures, amounts of precipitation, some bar graphs about record snowfalls, etc. (Chicago had a relatively severe winter this year.) Since weather varies from spot to spot even within a small geographic area, the numbers can provoke a discussion about their accuracy and variability.

Thus in this single day’s newspaper there is much data available for examination appropriate not only for the early grades of schooling but also later. This demonstrates that both numeracy and statistical literacy are necessary for general literacy and demands that we promote the teaching of some statistical ideas inside the teaching of language arts, in the areas of critical reading and critical listening. We want to attune students from the earliest grades to ask: Where did this number or these numbers originate? Is the number an exact value or an estimate? If an estimate, how might the estimate have been determined? What does the number tell us about the situation? How could the value of the number or the reliability of the estimate be verified? Etc.

The subtleties of language found in surveys provide wonderful opportunities for the teacher of the native language. In that increased sophistication there needs to be a discussion of how the wording of a survey item can significantly influence a response. When I tell my colleagues who teach literature that people in the mathematical sciences are at least as attuned to the meanings of words and the subtleties of interpretation as are poets and novelists, they typically are surprised. This is a role for the teacher of language, grammar and writing in the upper grades of schooling.

Why should we who are passionate about statistics education want to encourage putting this content into language arts? It is to show that the collection of data is a normal activity, done by people in all walks of life – in government, in business, in entertainment, in sports, in everyday activities. It is to encourage students to collect data in situations that are pertinent to their lives. It is to counter the notion that statistics are just for those who are good with numbers. It is also to counter the notion that statistics are routinely collected and used for sinister purposes – the all-too-popular notion that people routinely use statistics to lie, the notion perpetuated so often that there is a Wikipedia article titled “lies, damned lies, and statistics”! When I googled this phrase, I got 592,000 hits.

Now let us move to the social sciences. As ubiquitous as the uses of statistics are in everyday events, so are statistics everywhere in history, geography, economics, and politics, that is, in the social sciences. Even when I was a student over a half-century ago, it was in social studies, not mathematics, that we learned to read bar graphs, circle graphs, and line graphs along with the reading of maps.

What are some broad statistical principles that can be best brought out in social studies? One is the variability within groups and the danger of characterizing an entire group by its central tendency or by a distorted sample of its membership. Long-term and short-term trends over time are often eye-openers for young children, from the human population on our planet to the number of cell phones in use. (A helpful source of data in some countries can come from the CensusAtSchool International project.) The analyses can become more sophisticated than is reasonable for young students; as students go up the grades, their look at social data should become more sophisticated.

Whereas the wording and grammar of opinion surveys is appropriate for language arts study, the intent of an opinion survey is more in the area of social studies. Why are there polls before elections? How do pollsters determine who they poll? How are TV ratings gathered? And so on.

The teaching of health and physical education calls for a discussion of statistics about longevity, about what might cause and what might cure disease, and about what data supports a healthy life style. It also begs for a discussion of anecdotal evidence, of the occurrence of events with very low probability for individuals in a large population. On this topic a book has been published just this year by David Hand. Both in health and in the social studies, one can discuss the dangers of extrapolating results from one population to other populations.

I have left science for last because the collection of data in science courses is a long-standing practice. From simulations of Mendel's experiments with peas to replicating Galileo's experiments with inclined planes in physics, I think that it is the case in many schools so little statistics is taught within the mathematics courses that more statistics is taught in science than in mathematics.

In the United States, unlike the situation with the Common Core, statistics is found in all grades K-5 in the Next Generation Science Standards, released last year. Here is a section of the *kindergarten* NGSS standards:

In the kindergarten performance expectations, students are expected to demonstrate grade-appropriate proficiency in asking questions, developing and using models, planning and carrying out investigations, analyzing and interpreting data, designing solutions, engaging in argument from evidence, and obtaining, evaluating, and communicating information. Students are expected to use these practices to demonstrate understanding of the core ideas. (NGSS Lead States, 2013)

A similar pair of sentences appears in each of grades 1 through 5, and in each grade these expectations are detailed.

#### SPANNING THE CURRICULUM

Since the 1980s, there has been a movement both at the school level and at the college level called "writing across the curriculum". Last month the 12<sup>th</sup> International Writing across the Curriculum Conference was held, with speakers from over a dozen countries [at least Canada, Sweden, Singapore, Germany, U.S. Virgin Islands, France, UAE, Turkey, China, Qatar, Lebanon, England, Denmark, and the U.S.]. That conference was aimed at the college level, but the school movement has also been quite strong, with the goal of using writing as a learning tool in the various disciplines as well as recognizing that, in the disciplines, there are characteristic and sometimes unique writing conventions. For instance, writing a mathematical argument is quite different from writing a biography.

Statistics can be viewed as similar to writing in spanning the curriculum. Writing has its home in language arts, and we can say that statistics has a home in mathematics. And just as the style of writing varies dependent on the discipline in which it is being used, so does statistics vary.

This is not a new idea. Over twenty years ago, a supplement to the mathematics standards of the National Council of Teachers of Mathematics in the U.S. was titled *Data Analysis and Statistics Across the Curriculum* (Burrill et al., 1992). The intent of that supplement was to show that there were examples of data analysis and statistics in a large variety of subjects taught in school.

I think a next step is needed – to go beyond examples within the time allotted for the teaching of statistics into the time allotted for the teaching of language arts, social studies, science, and health, so as to teach students that it is not just for learning statistics that one looks at these examples, but that for learning these other subjects statistics is necessary.

#### SUMMARY

Statistics has at least four kinds of places in the K-12 curriculum relative to the other subjects: (1) its traditional place as part of mathematics; (2) a more accurate place as applied mathematics, (3) as the science of data studied for its own sake, and, (4) across the other main subjects of the curriculum, inside the teaching of those subjects, to help in understanding them. We need to strive in all of these ways to increase the probability that all students, whether or not they intend to go on to higher education, in order to reflect its full importance in society and so that all students will have the understandings to deal wisely with the data that surrounds every aspect of their lives.

#### REFERENCES

- Biehler, R. (1989). Educational perspectives on exploratory data analysis. In R. Morris (Ed.), *Studies in mathematics education, Vol. 7: The teaching of statistics* (pp. 185-201). Paris: UNESCO.
- Burrill, G. (1989). Quantitative literacy in the United States. In R. Morris (Ed.), *Studies in mathematics education, Vol. 7: The teaching of statistics* (pp. 175-184). Paris: UNESCO.
- Burrill, G., Burrill, J. C., Coffield, P., Davis, G., de Lange, J., Resnick, D., & Siegel, M. (1992). *Data analysis and statistics across the curriculum*. Reston, VA: National Council of Teachers of Mathematics.
- CensusAtSchool. [www.censusatschool.org.uk](http://www.censusatschool.org.uk)
- Cobb, G. W., & Moore, D.S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, 104(9), 801-823.
- College Board (2014 (and earlier)). *Advanced Placement Program Summary Report 2013* (and earlier). New York: The College Board.
- Common Core State Standards Initiative (2010). *Common Core State Standards for Mathematics*. [www.corestandards.org](http://www.corestandards.org)
- Dunkels, A. (1989). Exploratory data analysis in the primary classroom – graphing and concept formation combined. In R. Morris (Ed.), *Studies in mathematics education, Vol. 7: The teaching of statistics* (pp. 19-26). Paris: UNESCO.
- Franklin, C., Kader, G, Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report*. Alexandria, VA: American Statistical Association.
- Hand, D. (2014). *The improbability principle: Why coincidences, miracles, and rare events happen every day*. New York: Scientific American/Farrar, Straus and Giroux.
- Lies, damned lies, and statistics (2014). [en.wikipedia.org/wiki/Lies\\_damned\\_lies\\_and\\_statistics](http://en.wikipedia.org/wiki/Lies_damned_lies_and_statistics)
- Medrich, E. A., & Griffith, J. E. (1992). *International mathematics and science assessments: What have we learned?* Washington: National Center for Education Statistics.
- Morris, R. (Ed.). (1989). *Studies in mathematics education, Vol. 7: The teaching of statistics*. Paris: UNESCO.
- National Center for Education Statistics (NCES) (2013). High School Transcript Study, 1990 and 2009. In *Digest of Education Statistics 2013*, table 225.40. Washington, DC: U.S. Department of Education
- National Research Council (1989). *Everybody Counts: A Report to the Nation on the Future of Mathematics Education*. Washington, DC: National Academy Press.

- NGSS Lead States (2013). *Next Generation Science Standards: For States, By States*. [www.nextgenscience.org](http://www.nextgenscience.org)
- The Royal Statistical Society (April 2005). *Teaching statistics across the 14-19 curriculum*. [www.rss.org.uk/uploadedfiles/documentlibrary/783.pdf](http://www.rss.org.uk/uploadedfiles/documentlibrary/783.pdf)
- The Royal Statistical Society (2014). *CensusAtSchool International Project*. London: Royal Statistical Society.
- Shaughnessy, M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*. New York: Macmillan.
- TIMSS (undated). *TIMSS: IEA's third international mathematics and science study. Released item set for the final year of secondary school: mathematics and science literacy, advanced mathematics, and physics*. Boston: Boston College, Lynch School of Education, TIMSS & PIRLS International Study Center.
- Wallis, W. A., & Roberts, H. V. (1956). *The nature of statistics*. New York: The Free Press.
- Yates, D. S., Moore, D. S., and Starnes, D. S. (2003). *The practice of statistics* (2<sup>nd</sup> ed). New York: W.H. Freeman.