Case No. S200923

IN THE SUPREME COURT OF THE

STATE OF CALIFORNIA

SAM DURAN, MATT FITZSIMMONS, individually and on behalf of other members of the general public similarly situated,

Plaintiffs and Respondents,

vs.

U.S. BANK NATIONAL ASSOCIATION,

Defendant and Appellant.

Review of a Decision of the Court of Appeal, First Appellate District, Division One, Case Nos. A12557 and A12687, Reversing Judgment and Decertifying Class in Civil No. 2001-035537, Superior Court of the State of California, County of Alameda County, The Honorable Robert B. Freedman, Judge Presiding

---

BRIEF OF AMICUS CURIAE THE GALLUP ORGANIZATION IN SUPPORT OF POSITION OF U.S. BANK NATIONAL ASSOCIATION

ALLAN G. KING, *pro hac vice*
JULIE A. DUNNE, Bar No. 160544
MARGARET HART EDWARDS, Bar No. 65699
LITTLER MENDELSON, P.C.
501 W. Broadway, Suite 900
San Diego, CA 92101-3577
Telephone: 619.232.0441
Facsimile: 619.232.4302

Attorneys for AMICUS CURIAE
THE GALLUP ORGANIZATION

TABLE OF CONTENTS

TABLE OF CONTENTS
(CONTINUED)

PAGE

TABLE OF AUTHORITIES

Page(s)

CASES

OTHER AUTHORITIES

I. INTRODUCTION

The Gallup Organization has engaged in survey research for more than 75 years. Gallup has performed more than 100,000 surveys, in more than 150 countries. It has pioneered the use of surveys in a wide variety of circumstances and is acknowledged internationally as a leader in its field. It presently employs more than 2,000 people throughout the world, many of whom hold doctorates in statistics and survey methods. Gallup files this brief in order to share with the Court its learning regarding the statistical issues raised by this appeal, and because Gallup shares a strong interest with the public in having these methods applied appropriately, which requires understanding and acknowledging both their uses and limitations.

Science has a long history as an aid to judicial fact-finding. Its traditional role has been to aid fact-finders by making available to them the results of scientific experiments and the state of scientific knowledge. The instant case is different because it is one of the very few in which a court has invoked aspects of the scientific *method*, not scientific results, as the foundation for its trial plan and its factual conclusions.[1] This brief explains why "sampling," the principal scientific tool relevant to this case, is miscast in this litigation, both as applied by the trial court and

---

[1] *See, e.g. Cimino v. Raymark Indus., Inc.*, 151 F.3d 297 (5th Cir. 1998) (reversing on Seventh Amendment grounds district court's trial based upon sampled testimony).

conceptually. Accordingly, Gallup fully supports US Bank's brief urging affirmance of the decision by the Court of Appeal.

The fundamental premise of the scientific paradigm is that scrupulous adherence to objectivity and measurement is the pathway to scientific accuracy. Researchers are entrusted to assemble and utilize unbiased evidence and to draw conclusions that are supported by experimental results. Thus, subjects in scientific experiments are chosen from those who have no incentive to align their "testimony" with their self-interest. In addition, independent researchers should be able to replicate the same experiment and determine whether they can reproduce the reported results. The objectivity and transparency of each of these aspects of scientific research is what makes the findings meaningful and reliable.

For example, in the case of the double-blind experiment, neither the researcher nor the subject knows who is assigned to the treatment group or control group. For this reason, the double-blind experiment is the gold standard: it produces data that are unlikely to be tainted by biases of either the scientist or the subject. Additionally, transparency in the methods of gathering and analyzing that data, and the ability of other researchers to replicate the experiment and generate their own experimental results, ensures that any errors will be identified and corrected.

These protocols do not translate easily to litigation because

the premises underlying the legal system are quite different. Unlike researchers, lawyers are charged with zealously advocating on behalf of their clients—putting forward their client's best case, not the most even-handed case. Fundamental to the trial system is the belief that justice results from each side advancing a partisan view of the facts. Thus, while objectivity is a cornerstone of science, our legal system is premised on evidence marshaled by advocates who are duty-bound to put forward their client's strongest case.

These tensions are apparent in the contrast between neutral scientific sampling methods, which encourage objective researchers to collect data from disinterested, unbiased subjects, and an adversarial system in which lawyers strive to present testimony that is most favorable to their case. Indeed, class members themselves have a personal, financial interest in the outcome of the litigation which may lead them to recall the pertinent facts selectively. And while repetition of experiments across the research community ensures that findings converge on scientific facts, trials are one-time events and factual findings reached on just one occasion dispose of the issues in that particular case. Moreover, in this instance one judge—not several independent fact finders looking at discrete sets of facts—found the facts in each case that was tried. As a result, the scrupulous objectivity that renders experimental data reliable in scientific experiments is absent by design in litigation. What is more, the trial court's flawed sampling plan in

this case exacerbated the already uneasy alliance between science and the law.

Gallup of course recognizes that a courtroom is not a scientific laboratory and that some compromise with pristine notions of scientific purity may be both tolerated and required. However, a court should not be permitted, on the one hand, to invoke the principles of "science" to justify departing from traditional trial procedures, while on the other hand departing from the principles of science for the sake of a more expedient trial.

The Supreme Court recognized this tension in *Wal-Mart Stores, Inc. v. Dukes*,[2] and ruled that a trial court may not sacrifice a party's due process rights in favor of a sampling procedure that prevents defendants from contesting the claims of each individual claimant.

> The Court of Appeals believed that it was possible to replace such proceedings with Trial by Formula. A sample set of the class members would be selected, as to whom liability for sex discrimination and the backpay owing as a result would be determined in depositions supervised by a master. The percentage of claims determined to be valid would then be applied to the entire remaining class, and the number of (presumptively) valid claims thus derived would be multiplied by the average backpay award in the sample set to arrive at the entire class recovery—without further individualized proceedings. We disapprove that novel project.[3]

---

[2] 131 S. Ct. 2541, 2561 (2011).

[3] *Id.* Although the Supreme Court premised its objections on the Rules Enabling Act, which prohibits federal courts from adopting a procedural

More recently, the Supreme Court reiterated its distrust of the formulaic proof of class-wide damages, absent evidence that the proposed formula measured salient aspects of each of the disputed claims. Absent such evidence, damages must be proved with respect to each individual class member, which makes the case unsuitable for class treatment.[4]

Accordingly, Gallup agrees with U.S. Bank that, for the reasons stated in *Dukes*, the Trial by Formula applied by the trial court in this case precluded U.S. Bank from defending itself against each class member's claim. Moreover, should this Court ultimately find a role for sampling in a trial court's case-management toolkit, the trial court's methodology in this case was so flawed, and infused the case with such bias, that its judgment must be reversed and the decision of the Court of Appeal must be affirmed.

---

mechanism that either abridges or enlarges a party's substantive rights, that statute merely codified an existing principle of due process. The Rules Enabling Act appears to have no statutory counterpart in California, but its fundamental purpose—to ensure the separation of powers between the courts and the legislature—is a principle that applies with equal force to the states. Burbank, "Of Rules and Discretion: The Supreme Court, Federal Rules, and Common Law," 63 *Notre Dame L. Rev.* 693, 700 (1988) ("As a historical matter, there can be no doubt that the major purpose of those who wrote and defended the bill that became the Enabling Act was to allocate power to make federal law prospectively between the Supreme Court as rulemaker and Congress").

[4] *Comcast Corp. v. Behrend*, 2013 U.S. LEXIS 2544 (March 27, 2013).

II. THE TRIAL COURT'S SAMPLING METHODOLOGY WAS, IN FACT, NON-RANDOM

A. Initial observations regarding sampling

"Sampling" is a methodology grounded in the field of statistics, which itself is a branch of mathematical probability theory. In a nutshell, the theory underlying sampling is this: under certain circumstances, a researcher may draw scientifically meaningful inferences regarding various characteristics of a population from a randomly chosen subset (a smaller group drawn from members of that same population), with a predictable degree of precision.[5]

However, not all samples achieve the same accuracy. A sample is only as representative as the sampling plan that produces it. The well-worn maxim, "garbage in, garbage out," is apt. Two potential sampling errors are particularly relevant here: self-interest bias and selection bias.

B. Self-interest causes selection bias

One of the earliest uses of sample surveys in legal disputes involved "mall surveys" of shoppers to help decide questions arising under

---

[5] "Probability sampling ensures that within the limits of chance ... the sample will be representative of the sampling frame ... When these goals are met, the sample tends to be representative of the population. Data from the sample can be extrapolated to describe the characteristics of the population." D.H. Kaye & D.A. Freedman, "Reference Guide on Statistics," in Federal Judicial Center, *Reference Manual on Scientific Evidence* at 226-27 (3d ed. 2011).

the Lanham Act regarding "brand confusion."[6] Similar surveys were used to define the "relevant market" in antitrust litigation.[7] Because the shoppers who were surveyed were deemed representative of the larger universe of shoppers, and those who were interviewed had no knowledge of the survey's ultimate purpose, there was little concern that their answers were biased or unrepresentative of the views of the larger population.[8] In addition, and quite importantly, those surveyed were not parties to the litigation, so there was no concern that their answers reflected self-interest. Because these surveys, conducted for litigation purposes, incorporated principles that were well-established in non-litigation contexts, most courts

---

[6] S.S. Diamond, "Reference Guide on Survey Research," in Federal Judicial Center, *Reference Manual on Scientific Evidence*, at 366 (3d ed. 2011) ("A routine use of surveys in federal courts occurs in Lanham Act cases, when the plaintiff alleges trademark infringement or claims that false advertising has confused or deceived consumers"). "A survey of members of the Council of American Survey Research Organizations, the national trade association for commercial survey research firms in the United States, revealed that 95% of the in-person independent contacts in studies done in 1985 took place in malls or shopping centers." *Id.* at 382 n.102.

[7] *Id.* at 366 n.25 ("Surveys have long been used in antitrust litigation to help define relevant markets"). In *United States v. E.I. du Pont de Nemours & Co.*, 118 F. Supp. 41, 60 (D. Del. 1953), *aff'd*, 351 U.S. 377 (1956), a survey was used to develop the "market setting" for the sale of cellophane. In *Mukand, Ltd. v. United States*, 937 F. Supp. 910 (Ct. Int'l Trade 1996), a survey of purchasers of stainless steel wire rods was conducted to support a determination of competition and fungibility between domestic and Indian wire rod.").

[8] *Herman Miller v. Palazzetti Imports & Exports*, 270 F.3d 298, 312 (6th Cir. 2001) ("Because the determination of whether a mark has acquired secondary meaning is primarily an empirical inquiry, survey evidence is the most direct and persuasive evidence").

viewed survey evidence as a legitimate foundation for expert testimony.[9]

The move from surveys of disinterested shoppers to the testimony of interested class members is a far greater leap than the trial court appreciated. Third-party shoppers typically are voluntary participants in a survey. Their decision to participate is unrelated to any anticipated benefit because they may not be aware of the study's purpose and in any event they are unlikely to have a claim on any benefits the study provides. Although some may withdraw from the survey once it has begun, the likelihood is low that they are doing so because they fear that if they remained in the survey, their answers might jeopardize their self-interest. As a result, a shopper who exits the sample can be supplemented with an additional random shopper with no anticipated change in the representativeness of the findings.

The sampling of interested *parties* involves a very different calculus. Although the names of those initially included in the sample may be selected randomly, that merely is the first step in determining the group of cases that actually will be tried. Unlike the shoppers, who are neutrals, a class member who decides to participate or decides to withdraw from the sample may do so for reasons that are directly related to the strength of his or her claim. For example, if class counsel believes a class member's claim

---

[9] *See generally* Diamond, "Reference Guide to Survey Research," at 363-67.

is weak or non-existent, it would be advantageous to have that employee exit the sample and thereby avoid testifying to damaging facts. In that way, this individual avoids incurring the burdens of discovery and trial in pursuit of a weak claim, he or she retains the right to recover as a member of the class, and class counsel has strengthened the case of the entire class.[10]

C. Selection bias makes the sample unrepresentative

Selection bias exists when members of the sample are selected either according to a non-neutral criterion, or because members of the target population are selectively included or excluded from the sample, perhaps reflecting their own self-interest. A famous example in which relevant responses were excluded occurred with respect to a telephone survey conducted to forecast the winner of the 1936 Presidential election:

> After successfully predicting the winner of every U.S. presidential election since 1916, the [*Literary*] *Digest* used the replies from 2.4 million respondents to predict that Alf Landon would win the popular vote, 57% to 43%. In fact, Franklin Roosevelt won by a landslide vote of 62% to 38%. The *Digest* was so far off, in part, because it chose names from telephone books, rosters of clubs and associations, city directories, lists of registered voters, and mail order listings.

---

[10] Under the rules of the trial court in this case, that appears to be exactly what happened in the case of class member Borsay Bryant. Bryant was selected by the trial court to be a member of the Random Witness Group ("RWG"). Although he failed to appear at trial, he remains a passive member of the class. *See Duran v. United States Bank National Assn.*, A12555 & A126827, slip op. (Cal. App. 1st Dist. Feb. 6, 2012) at 29 n.39 and 31 (hereinafter "slip op."). Indeed, the trial court judgment awards Bryant approximately $50,000. Thus, at no cost to himself, Bryant was able to exit the RWG yet retain his personal claim, which was adjudicated based upon the presumptively stronger testimony of the remaining RWG.

> In 1936, when only one household in four had a telephone, the people whose names appeared on such lists tended to be more affluent. Lists that overrepresented the affluent had worked well in earlier elections, when rich and poor voted along similar lines, but the bias in the sampling frame proved fatal when the Great Depression made economics a salient consideration for voters.[11]

The pollsters designed the survey, unintentionally to be sure, so that the sample was selected by a criterion—telephone ownership—that varied among the population but unfortunately was correlated with voting preference (favoritism towards Landon). As a result, although the respondents may have been representative of the preferences of the telephone-owning public, telephone owners were not representative, at that time, of the voting public at large.

Similarly, those members of a television audience who respond to an invitation to text their vote for a favorite contestant also are self-selected. The fact that some are motivated to respond, whereas other audience members are not, may indicate that the first group differs from the more passive members of the audience in other ways as well. For example, a particular performer may stir up strong feelings among a minority of the audience, which motivates them to text their approval, as opposed to a lower-keyed competitor, who actually is preferred by a majority of the audience but not enough to overcome their passivity. Thus, the responses may be a good gauge of the intensity of viewer preferences, but a poor

---

[11] Kaye & Freedman, "Reference Guide on Statistics," at 225 n.32.

indicator of whom the majority of viewers actually prefer.

The critical point regarding these examples is that *the same behavior or motivation that causes an individual to be included in the sample is also reflected in that person's survey response.* If this effect is significant, the survey responses will be unrepresentative of the larger population and inferences drawn from that sample will be biased if they are extrapolated to the group as a whole. In litigation this principle means that if individuals are permitted to choose whether to be active or passive members of the class, those who choose to actively participate will be those with the most to gain by their participation.

In a true random sample, every member of the defined population, and every combination of such persons, must have an equal chance to be selected. If the manner of selection departs from those principles, either because the selection criteria are biased or because of self-selection, the results obtained with respect to the sample of cases may be unrepresentative and therefore uninformative regarding those who have been excluded from the sample. As a general rule, "Selection bias is acute when constituents write their representatives, listeners call into radio talk shows, interest groups collect information from their members, or attorneys choose cases for trial."[12] Moreover, selection bias cannot be overcome by increasing the size of the sample. As one eminent statistician observed,

---

[12] Kaye & Freedman, "Reference Guide on Statistics," at 225.

"When a selection procedure is biased, taking a large sample does not help. This just repeats the basic mistake on a larger scale."[13]

D. The trial court's methodology resulted in an unrepresentative sample of cases being tried.[14]

The trial court initially announced a trial plan that would have required the parties to litigate the rights and recovery of the entire class based on the claims and testimony of the named plaintiffs and a 20-person Random Witness Group ("RWG") only. However, as explained below, the trial court made several rulings that allowed class counsel to influence which cases were selected for trial and therefore skewed those "representative" trials in favor of the plaintiff class.

1. The trial court refused to hear the claims of weaker class representatives

The case initially was filed by a single plaintiff, Amina Rafiqzada. She withdrew as the proposed class representative and three alternative plaintiffs were substituted—Vanessa Haven, Abby Karavani, and Parham Shekarlab. These three then were replaced by Sam Duran and Matt Fitzsimmons, the two named-plaintiffs whose individual cases

---

[13] D. Freedman, R. Pisani & R. Purves, *Statistics* at 335 (4th ed. 2007). The pages of this book that are cited throughout this brief are included in the Attachment. The late Professor David Freedman, of the University of California at Berkeley, co-authored the "Reference Guide on Statistics" published in the Federal Judicial Center's *Reference Manual on Scientific Evidence, supra.*

[14] Because the parties amply present the facts and procedures in general, Gallup addresses here only the facts and procedures that confirm that the trial court's verdict was based upon a biased sample.

ultimately were decided by the court, although they were not chosen in a random drawing.

Although it is common to substitute class representatives, that ordinarily does not preclude the former named-plaintiffs from being called as witnesses and from permitting the defendant to expose the weaknesses in their cases. However, in this instance, the trial court did precisely that, in the mistaken belief that this was mandated by its misconceived sampling plan: because only named-plaintiffs and members of the RWG were to have their cases heard, and the former named-plaintiffs belonged to neither group, their claims were deemed irrelevant.[15] Thus, the court precluded testimony from the four former named-plaintiffs regarding their own job duties, and prohibited USB from referencing that testimony in its post-trial briefing.[16] Instead, the trial court permitted class counsel to substitute two newly-named plaintiffs whose cases were heard instead. In this manner, class counsel was permitted to stack the deck, as it were, and exclude four less favorable cases and add two stronger cases for the trial court's consideration.

2. The trial court permitted class counsel to influence which "random" cases were tried

Next, the trial court implemented its randomized trial plan by initially drawing the names of 20 class members plus five alternates, whom

---

15 Slip op at 54.

16 *Id.* at 26 & n.35.

it designated as the RWG. However, after drawing these names, the trial court granted the Plaintiffs' motion to dismiss their "legal" claims, leaving only their equitable claims in dispute, which would be tried to the court, not a jury. The trial court then ordered the parties to re-notify class members of that amendment and provided class members, *including the 20 whose cases were randomly selected for trial*, a second opportunity to opt-out of the case.

In the interim, class counsel could freely interview the RWG, determine the strengths and weaknesses of each of their cases, and cull the weakest of them by encouraging them to opt-out of the case. *One need not speculate whether this is true because RWG members Lewis and MacClelland testified that they were influenced to drop out of the case by class counsel.*[17] Ultimately, four of the 20 members of the RWG (*i.e.*, 20 percent) opted out, in contrast to five of the 240 non-testifying class members (*i.e.*, two percent) who also chose to opt out.[18]

USB urged the trial court to permit the four RWG members who opted out of the action to opt back in. In support of its motion, USB submitted the declaration of Dr. Phillip Gorman, who explained that, in light of the marked difference in the opt-out rates between the RWG and

---

[17] Mr. Lewis testified that plaintiffs' counsel called him several times to encourage him to opt out of the case; Mr. MacClelland testified that plaintiffs' counsel left him intimidating messages on his answering machine. *Id.* at 10 n.20.

[18] *Id.* at 10.

other class members (*i.e.*, 20 versus two percent), removing from the sample the four RWGs who opted out "would undermine the accuracy of any extrapolation process."[19]

In opposition, plaintiffs submitted the declaration of its statistical expert, Dr. Richard Drogin. Dr. Drogin asserted, despite testimony that RWG members were influenced to withdraw from the case, that it was "statistically acceptable" to substitute alternates because there was "no reason to infer that the sample is not representative, or that there is any bias in the sample."[20] In his view, "as long as the set of persons selected to testify at trial includes those in the original random selection made by the court, and is restricted to those in the class, the testifying group will be a random sample of the class."[21] The trial court therefore denied USB's motion to permit the RWGs to opt back into the suit, and deferred ruling on whether they would be allowed to testify at trial.[22]

Although the trial court was apparently swayed by Dr. Drogin's testimony, a simple example illustrates why his opinion is wrong. Imagine the good fortune of a black-jack player who is permitted to view each card he is dealt randomly, and then decide whether to keep it or discard it in favor of another randomly drawn card. Although each card

---

19 *Id.*
20 *Id.*
21 *Id.*
22 *Id.* at 11.

may be drawn randomly, obviously the hand the card-player finally is able to play has been selected card-by-card. This precisely is the advantage the trial court's methodology conferred on the class, and the Court of Appeal was correct to discredit Dr. Drogin's testimony on that issue.[23]

3. The trial court precluded four random witnesses from testifying about their job duties

Prior to trial, the court also ruled on the parties' motions *in limine*. The trial court ruled that non-RWG witnesses could testify only to impeach the claims of designated plaintiffs in the individual trials, and could not testify regarding their own job duties or those of other non-RWG class members.[24] The trial court also denied USB's motion to permit the testimony of the original RWG members who opted out of the case when given a second opportunity. As a result, Michael Lewis and Sean MacClelland were prevented from testifying about their own job duties, despite their declarations stating that they were influenced to opt out by class counsel.[25] Two other class members originally selected for the RWG also opted out of the case when given the second opt-out opportunity, but the Court of Appeal's opinion does not reflect their specific reasons for doing so.[26] Thus, the trial court drew from the five "alternates" to replenish the RWG. However, as explained, this rendered the trials of the RWG non-

---

23 Slip op. at 10-11.
24 *Id.* at 12.
25 *Id.* at 10 n.20.
26 *Id.*

random and therefore unrepresentative.

4. The trial court excluded a randomly-selected class member whom it perceived to be atypical

The trial court itself modified the initial random selections in one additional respect. It excluded one member of the RWG, Bryan Smith, from the group whose cases would be tried, on the grounds that his actual work activities differed from those of a "true" Business Banking Officer ("BBO").[27] In effect, the court considered Smith to be an "outlier," whose job duties were too atypical to remain in the sample. However, whether and to what extent class members performed their duties in a similar fashion was precisely the issue the court was to decide by means of the trial. By characterizing a randomly-selected class member as atypical, when in fact Smith may have been representative of a sizeable but unknown number of other BBO's in the class, the court imposed a greater degree of uniformity on the RWG than indicated by its own, ostensibly random, methodology.[28]

---

27 *Id.* at 37 n.46.

28 *See* D. Freedman, R. Pisani & R. Purves, *Statistics*, 4th ed. (2007) at 103 (quoting the National Bureau of Standards with regard to excluding "outliers"): "The problem [of outliers] is more often associated with conscious, or perhaps unconscious, attempts to make a particular process perform as one would like it to perform rather than accepting the actual performance...Rejection of data on the basis of arbitrary performance limits severely distorts the estimate of real process variability."

5. All told, the trial court departed from its random design in at least four ways

To summarize, the trial court made a series of decisions regarding the composition of the sample of cases to be tried that are inconsistent with the principles on which it purported to rely. Specifically:

(a) By permitting class counsel to juggle named-plaintiffs, the trial court permitted class counsel to determine which claims of non-randomly selected plaintiffs it would try;

(b) After the RWG was selected, and after these employees could be contacted by class counsel only, the trial court permitted four members of the RWG to opt out of the case (at least two at the urging of class counsel) and precluded their testimony regarding the amount of time they and other non-RWG members spent working outside the bank's facilities;

(c) One member of the 20-member RWG was permitted to exclude himself from testifying and still recover under the judgment;

(d) One member of the RWG was excluded because the trial court deemed him to be an "outlier," although there was no evidence that this class member was unrepresentative of a substantial segment of other class members; and

Thus, as a direct result of the actions of class counsel, the court, or the class members themselves, at least 10 cases (four named-plaintiffs plus six from the RWG) that otherwise would have been tried escaped the court's consideration

Based upon the remaining sample, 19 members of the RWG plus two named-plaintiffs, the court found that USB failed to carry its

burden of proving its claimed exemption with respect to the class. Despite the RWG being culled, it further concluded that the remaining members were typical and representative of the entire class, which "validates the ... use of the [RWG] process as part of the trial management plan of a wage and hour class action."[29] Extrapolating from this biased sampling, the trial court determined that all class members were non-exempt employees and eligible to receive backpay.

III. THE TRIAL COURT'S PROCEDURE WAS PREMISED ON THE UNPROVEN ASSUMPTION THAT THE TRIALS WOULD NOT PRODUCE A MIXED VERDICT

A. The trial court failed to specify the question its trial plan was designed to answer

Every sampling plan must begin with a clearly framed hypothesis to be tested, usually characterized as the null hypothesis.[30] This requires the researcher to specify, *before undertaking the experiment*, the findings that will be judged sufficient to answer the research question one way or the other. Based upon this specification, the researcher then is positioned to determine the sample size necessary to support the accept/reject decision with a predetermined degree of confidence.

---

29 *Id.* at 29.

30 D.R. Rubinfeld, "Reference Guide on Regression Analysis" in Federal Judicial Center, *Reference Manual on Scientific Evidence* 3d ed. (2011) at 311 ("Research begins with a clear formulation of a research question. The data to be collected and analyzed must relate directly to this question; otherwise, appropriate inferences cannot be drawn from the statistical analysis.").

In the context of litigation, these specifications essentially constitute the trial plan, or "study design" to be followed by the court.

> The interpretation of data often depends on understanding "study design" —the plan for a statistical study and its implementation. Different designs are suited to answering different questions. Also, flaws in the data can undermine any statistical analysis, and data quality is often determined by study design.[31]

It is critical for the trial court to specify the precise question to which its trial plan will provide an answer because that is the benchmark against which its proposed methodology must be assessed. That is, the questions posed to the fact-finder must accurately reflect the governing law, and any trial plan must be evaluated in terms of whether it is likely to generate representative evidence that will answer those legally-mandated questions. The trial court in this case failed to specify this question.[32]

B. The trial court failed to accommodate the likelihood of a mixed verdict

Although it may appear that the critical fact question in this case is obvious—whether the evidence establishes that employee class members spend more than half their time outside the bank?—in fact, this

---

31 Kaye and Freedman, Reference Guide on Statistics at 216.

32 In statistical terms, this means that the null hypothesis is unspecified. *See* Kaye & Freedman, Reference Guide on Statistics at 216. ("To test for (statistical) significance, a researcher develops a 'null hypothesis'—*e.g.*, the assertion that there is no relationship ... between [class membership and the claimed exemption]. The researcher then calculates the probability of obtaining the observed data (or more extreme data) if the null hypothesis is true (called the p-value)"). *See also Matrixx Initiatives, Inc. v. Siracusano*, 131 S. Ct. 1309, 1319 n.6 (2011).

question poses a false dichotomy with respect to class members' claims. It omits the possibility that some class members would be found to be exempt, if their cases were tried individually, while other class members would be found non-exempt—in essence, a mixed verdict.[33] Because the trial plan was not designed to accommodate this possibility, it can support a class-wide judgment only in the event that each mini-trial produces the same verdict. This constrains the trial court, which is both the architect of the trial plan and finder of fact, to reach the same finding in each and every sampled case, or else concede that the trial plan was misconceived.[34]

C. An employer is entitled to prove the reasonable likelihood of a mixed verdict

The employer must be provided an opportunity to prove that a mixed verdict is appropriate, unconstrained by a trial plan premised on the assumption that it is not. The question therefore is not whether most employees are exempt, or whether the "typical" employee is exempt, but whether *any* employees are likely to be exempt and whether the trial plan

---

[33] USB provided the court with at least 75 declarations stating that these declarants qualified for the claimed exemption. Thus, unless the trial court prejudged the veracity of these employees, the probability of a mixed verdict—some employees judged exempt and others non-exempt—was a likely possibility that should have been accounted for.

[34] Although it might be objected that this mixed result—that the employer would prevail in its defense against some class members but not others—failed to materialize, the fairness of the trial court's methods cannot be assessed after-the-fact, based upon findings entered by the same trial court judge who designed the trial plan. The trial court's objectivity is subject to the criticism that the court may have conformed the verdicts to the necessities of its trial plan, rather than vice versa.

affords the employer an opportunity to present that proof. Indeed, *Dukes* confirms that an employer has a due process right to assert its defenses to the claims of *each* class member.[35] Yet, the trial court's plan limited USB to proving its defense solely with respect to witnesses chosen by the court and class counsel, from which those class members most likely to be exempt were excluded by design.

For example, prior to class certification, USB obtained the declarations of 75 putative class members who testified that they spent more than 50 percent of their time outside the bank's offices. USB repeatedly proposed that it be permitted to prove its defense with respect to each class member, including of course these 75. This was rejected by the trial court, which instead determined that the exempt status of these 75 and the majority of other class members would be determined by the testimony of 20 randomly-chosen class members, presumably to economize on the time necessary to try the case.[36]

The random character of this sample then was changed significantly because of the selective withdrawal of at least six of these class members, the addition of two "third generation" named-plaintiffs, and the exclusion of four prior named-plaintiffs. In light of *Dukes*, the trial

---

[35] *Dukes*, 131 S. Ct. at 2560 (confirming the employer's "entitle[ment] to individualized determinations of each employee's eligibility for monetary relief," and cautioning against "trial-by-formula").

[36] Slip op. at 54.

court's decision to prelude USB from defending against each class member's individual claim was questionable; but it was an egregious violation of due process to deny USB the right to prove its defense against *any* class member of its choosing, while conferring on class counsel the right to select a significant fraction of the cases that would decide that issue.[37]

Moreover, the small sample selected by the trial court was prone to omit exempt employees even when they truly exist. Thus, if 10 percent of the class is exempt, and the court elects to try 19 cases selected at random, no exempt employees will be found among these cases around 14 percent of the time.[38] If five percent of the employees are truly exempt, then the court's methodology would fail to find any exempt class members

---

[37] "The nature of the adversarial system is that each side will emphasize what is favorable for it, and attack the other side's presentation." *Fischer v. Hill*, 2012 U.S. Dist. LEXIS 63448, *36 (E.D. Cal. January 9, 2012). *See, e.g., United States v. Cronic*, 466 U.S. 648, 655, 104 S. Ct. 2039 (1984) ("'Truth' is best discovered by powerful statements on both sides of the question."); and *Burdett v. Miller*, 957 F.2d 1375, 1380 (7th Cir. 1992) ("Ours is an adversarial system; the judge looks to the parties to frame the issues for trial and judgment. Our busy district judges do not have the time to play the 'proactive' role of a Continental European judge.").

[38] The relevant arithmetic is the following. By assumption, the probability of trying the case of a non-exempt employee is .9. The probability of selecting 19 such cases without any being found exempt is .9 raised to the 19th power. Similarly, when the assumed incidence of exempt employees is five percent, the probability that all 19 members of a random sample will be non-exempt is 38 percent. As explained in the following sections, these calculations are premised on the assumption that the verdicts in each case are "independent." Of course, USB submitted declarations indicating that a much larger fraction of the class was properly classified as exempt.

38 percent of the time. Because classes alleging misclassification can number well into the thousands,[39] the trial court's methodology, if approved by this Court, could regularly misjudge the exempt status of *hundreds* of class members.

IV. THE "MARGINS OF ERROR" OR "CONFIDENCE INTERVALS" ARE FAR GREATER THAN THE TRIAL COURT RECOGNIZED BECAUSE THE SAMPLED TRIALS WERE NOT "STATISTICALLY INDEPENDENT"

A. "Independence" is an essential characteristic of sampling but in this instance the trials were not independent

The standard assumption in extrapolating the results of a sample to the larger population, and determining the associated margin of error, is that each "response" or verdict is independent of all others.[40] In a coin-flipping experiment, this means that the chances of getting "heads" on any given coin flip must not depend on the outcome of any previous flips of the coin.[41] In a trial setting, this means that the chances of either side

---

[39] For example, in *Bell v. Farmers Insurance Exchange*, 115 Cal. App. 4th 715, 722 (2004), the class included 2,500 employees.

[40] "Two things are independent if the chances for the second given the first are the same, no matter how the first one turns out." Freedman et al, *Statistics* 4th ed., at 230.

[41] Thus, if coin flips are independent, and the probability on each flip that the coin will land on "heads" is 50 percent, then the probability of heads on two consecutive flips is, by the multiplication rule, .5 x .5, or .25. *See* Freedman *et al.*, *Statistics* 4th ed., at 229 ("The chance that two things will both happen equals the chance that the first will happen, multiplied by the chance that the second will happen given that the first has happened"). On the other hand, if the flips are perfectly dependent, heads on the first flip ensures that second flip also will come up heads. The probability of two

prevailing in the 20th case should be unaffected by decisions rendered in any of the previous 19.

The exigencies of trial make it extremely difficult to satisfy the independence condition. In a bench-tried case such as this, *the same trial court judge* is asked to decide the identical fact question in each sampled case, and in principle must keep his assessment of the evidence from cumulating. But, as Judge Posner observed:

> The *Mejdrech* decision, and *Bridgestone/Firestone* and *Rhone-Poulenc* more fully, discuss the danger that resolving an issue common to hundreds of different claimants in a single proceeding may make too much turn on the decision of a single, fallible judge or jury. The alternative is multiple proceedings before different triers of fact, from which a consensus might emerge; a larger sample provides a more robust basis for an inference. But that is an argument for separate trials on pecuniary relief, ...[42]

However, the trial court did not render its decision at the conclusion of *each* case, but instead reached a decision *only after all the cases were tried.* Thus, the court heard testimony regarding the 20th case before deciding the exemption at issue in the first. As a result, testimony in each case was permitted to influence the decision in every other—a trial plan strikingly at odds with the need for "independent" verdicts.

---

heads therefore is .5, *i.e.*, .5 x 1.0, the same as the probability of heads on the first flip. As explained subsequently, because the trials in this case were not independent, the accuracy with which the sampled trials approximate the verdicts that would be reached if the cases of all class members were tried separately is vastly overstated.

[42] *McReynolds v. Merrill Lynch, Pierce, Fenner & Smith*, 672 F.3d 482 (7th Cir.), *cert. denied*, 133 S. Ct. 338 (2012).

If verdicts are not independent in the sense described above, then the usual statistical criteria are inapplicable. For example, suppose the first case is so compelling that it determines the outcome in all subsequent cases. In that event, the number of cases in the sample is irrelevant and one trial is as good as 100 because, after the first, the results of all other trials are predetermined. In that event, the "law of averages"[43] has no relevance and the margin of error will be unrelated to the number of cases that are tried. Yet, the one statistical conclusion on which both parties seem to agree—that up to 13 percent of the class may be properly classified as exempt and yet 21 trials could produce no wins for USB five percent of the time[44]—rests on the unproved assumption that the verdicts in each trial are independent. However, if that assumption is false, then an untold number

---

[43] This is the tendency for the average of the sample to converge on the average for the entire group as the sample size increases. In terms of the coin-tossing experiment, as the number of tosses goes up, "the difference between the percentage of heads and 50% gets smaller." Freedman et al, *Statistics* 4th ed., at 276-77.

[44] Appellant's Opening Brief filed in the Court of Appeal, at 59. This calculation is based upon the "multiplication rule" that would apply if trials were independent. If each class member has a 13 percent chance of being exempt, then there is an 87 percent chance this employee is non-exempt. The probability that *any* class member will be found exempt in any of 21 trials is the flip side of the probability that the class winning all 21 cases. By the multiplication rule, the probability of the class wining all 21 cases is .87 raised to the 21st power, or five percent. Thus, in 95 percent of the trials the employer would win at least one case. *See also* note 41, *supra*, setting forth calculations that reflect the assumption that trials are "independent."

of non-testifying class members may be exempt.[45]

This Court appears to be the first in the nation to recognize the critical importance of "independence" in correctly assessing probabilities. Thus, in *People v. Collins*, this Court observed:

> But, as we have indicated, there was another glaring defect in the prosecution's technique, namely an inadequate proof of the statistical independence of the six factors. No proof was presented that the characteristics selected were mutually independent, even though the witness himself acknowledged that such condition was essential to the proper application of the "product rule" or "multiplication rule."[46]

Because the pertinent statistical evidence in this case is based upon the "multiplication rule" or one of its variations, and that rule, in turn, rests on the unproven assumption of "independence," the statistical properties imputed to the court's trial plan must be rejected.

Indeed, the trial court's recitation of the evidence in this case demonstrates the likelihood that the verdicts in each case in fact were conflated. The trial court's opinion consists, almost exclusively, of facts that ostensibly are common among the plaintiffs in the cases it tried, with

---

[45] When trials are not independent, the probability of a given number of wins or losses depends on the extent to which the outcome in each trial is correlated with the outcomes of all other trials, a fact that is essentially unknowable.

[46] 68 Cal. 2d 319, 438 P.2d 33, 66 Cal. Rptr. 497 (1968). "Thus, the *Collins* opinion correctly indicates that the product rule cannot be applied to identifying characteristics unless ... the mutual independence of each of the characteristics is established." R. N. Jonakait, "When Blood Is Their Argument: Probability in Criminal Cases, Genetic Markers, and, Once Again, Bayes' Theorem," 1983 *U. Ill. L. Rev.* 369, 375 (1983).

no consideration of the differences among them. The trial court observed that "much of Plaintiffs' testimony was not even contested." It noted that USB "failed to put on any of the immediate supervisors of 7 of the 21 RWGs" or "all of the immediate supervisors for 13 of the 21 RWGs."[47] While these omissions may be relevant to the claims of particular members of the RWG, Plaintiff A generally is not entitled to prevail because there is a dearth of evidence against Plaintiff B. The strength or weakness of the defense against others is generally irrelevant to the case of a given class member.[48] That it was deemed material by the trial court strongly indicates that its decisions were not independent.

Nevertheless, by the court's accounting, USB controverted two-thirds of the claims of the RWG to non-exempt status through the testimony of their immediate supervisors. Yet, the trial court discounted this testimony by finding that these witnesses "lacked any knowledge and/or foundation of the hours worked or work activities performed by the RWGs." However, Michael Lewis testified that he supervised RWG member Matt Gediman for a little over a year and estimated that during the first quarter Gediman spent 60-70 percent of his time outside the office, and

---

[47] *Duran v. U.S. Bank National Association*, Case No. 2001-035537, Alameda Country, Northern Division, September 22, 2008, Statement of Decision for Phase I, at 31-32.

[48] *Cf. Ginty v. Ocean S. R. Co.*, 172 Cal. 31, 43 (1916) (excluding evidence under the doctrine of *res inter alios acta*).

55-60 percent during the second quarter of the year.[49] Once again, the court ignored the details pertaining to individual claimants in favor of more global findings regarding the class in general.

"Independence" was further compromised because the court permitted the non-randomly selected Fitzsimmons' case to be tried first, which likely affected the court's receptivity to the subsequent claims of the RWG through the mechanism of "priming."[50] It is a well-established principle of psychology that the perceptions of individuals are influenced by the sequence in which other stimuli are perceived. Numerous laboratory studies have documented the tendency of a subject's own behavior to be influenced by the behavior they observe in other subjects, photographs, videos and other stimuli prior to their own actions. This phenomenon

---

[49] Although Lewis conceded that he did not track the time of his supervisees, the trial court placed too much weight on this fact in the mistaken belief that this issue is controlled by *Anderson v. Mt. Clemens Pottery Co.*, 328 U.S. 680 (1946). *See* Statement of Decision at 32. It is not. *Mt. Clemens* concerned the amount of overtime pay due employees who concededly were non-exempt. The issue in dispute in that case was how long it took these employees to walk from the time-clock to their workstations—a question of damages. Under those circumstances, the Supreme Court announced the rule that because the employer in that case failed to keep accurate records, the burden shifts to the employer to come forward with evidence of the precise amount of work performed ..." *Id.* at 687-88. However, this case turns on the exempt status of the class members—a question of liability—on which the employer already has the burden of proof. The fact that precise records do not exist regarding how much time an employee spends in or out of the bank does not affect the employer's burden of proof whatsoever and makes it no more or less likely that the employee is non-exempt

[50] D. J. Watts, Everything Is Obvious—Once You Know the Answer (2012), p. 52 n. 6.

should be neutral in its impact if a sample of trials is truly random, because "randomness" requires that both the *sequence* of selections, as well as the selections themselves, be random.[51] By ceding to class counsel the right to select the order in which cases were tried, the trial court departed from a random methodology and gave the class an unfair advantage.

B. Absent independence, the trial results cannot be extrapolated to other class members with any degree of confidence

This Court has recognized that estimates of hours of work, or the probability that an employee is exempt, obtained from any sample inevitably are distributed around the "true" value for the class or population of interest. For example, if one repeatedly estimates the weekly hours of a company's workforce by means of random samples, the average for the sample might range between, say, 38 and 42, although the true average is 40. The degree of inaccuracy of this average technically can be characterized by a "confidence interval," which is similar to the more commonly-used "margin of error."[52]

---

[51] *Cf. Griffith v. Consumer Portfolio Serv.*, 838 F. Supp. 2d 723, 725 (N.D. Ill. 2011) ("As we understand these terms, 'random number generation' means random sequences"); and Kaye and Freedman, Reference Guide on Statistics at 230 ("With a randomized controlled experiment, subjects are assigned to treatment or control at random in the strict sense—by.... using a random number generator on a computer").

[52] *Bell v. Farmers Insurance Exchange*, 115 Cal. App. 4th at 752 ("This measure of possible inaccuracy consisted of a 'confidence interval,' based on a 95-percent degree of confidence. The term 'interval' can be

The confidence interval, in turn, depends on whether the sampled cases are independent, in the terms described previously. Because the assumption of independence is unproven and untenable, the estimates relied upon by the trial court greatly understate the size of the true confidence interval and the proportion of class members who may be exempt. Thus, Dr. Drogin testified:

> "the 95 percent confidence interval ranges from 87 percent to 100 percent for the percentage of class members who are misclassified." When asked if, "statistically speaking," would not this mean that up to 13 percent of the class could possibly be properly classified, Drogin responded: "That's true."[53]

However, that conclusion rests squarely on the assumed independence of the verdicts.[54]

The confidence interval described by Dr. Drogin, because it *assumes* independence, has no necessary connection to the more relevant confidence interval that applies in this case, when sampled cases are *dependent*. That is, suppose the result of the first trial is so powerful that it influences the verdict in subsequent trials. Now permit class counsel to select the first case to be tried and assume that this class member is found to be non-exempt. That makes it more likely that the second trial will result in a

---

expressed in more familiar language in terms of a margin of error, which is one-half of the interval.").

[53] Slip op. at 33.

[54] *See* note 40, *supra* (demonstrating that the calculation of the 13 percent confidence interval depends on the "multiplication rule" of *Collins*, which rests on the unwarranted assumption of independence).

finding of "non-exempt," which in turn will influence the third trial in that same direction, and so on. At the end of the day, the number of verdicts in that direction will be more extreme than they would be otherwise, which makes the confidence interval much larger than the one calculated by Dr. Drogin.

Dr. Drogin testified that a 95-percent confidence interval ranges from 87 to 100 percent of the entire class being non-exempt.[55] If true, the implication would be that it is very unlikely for more than 13 percent of the class to be exempt.[56] However, that testimony depends critically on the unproven assumption that each trial is independent of all others. If that assumption is false, which seems probable given the trial plan, the true confidence interval is larger than Dr. Drogin estimated, and the likelihood that more than 13 percent of the class is exempt will be greater as well. Accordingly, Dr. Drogin's conclusions rest on a foundation that is unsupported by any evidence and they therefore are entitled to no weight.[57]

V. CONCLUSION

The trial court purported to base its trial plan on a statistical methodology that is flawed both conceptually and as applied in this case.

---

55 Slip op. at 33.

56 Dr. Drogin testified that "up to 13 percent of the class could possibly be correctly classified." *Id.*

57 *See People v. Collins*, 68 Cal. 2d at 327 (disregarding probabilistic evidence because, among other defects, "the inadequate proof of statistical independence").

As noted in the Federal Judicial Center's *Reference Manual on Scientific Evidence*, "It is randomness in the technical sense that provides assurance of unbiased estimates from a randomized controlled experiment or a probability sample. Randomness in the technical sense also justifies calculations of standard errors, confidence intervals, and p-values. *Looser definitions of randomness are inadequate for statistical purposes.*"[58]

This brief has explained the numerous ways in which the methods of the trial court, and its *ad hoc* decisions, favored expedience over the technical requirements of random sampling, and raise serious doubts about whether any trial procedure can be faithful to the rigors of statistical analysis.[59] In any event, there can be no doubt that here the trial court erred in designing and implementing its trial plan in this case, and for that reason the decision of the Court of Appeal should be affirmed.

---

[58] Kaye and Freedman, Reference Guide on Statistics, at 230 (emphasis added).

[59] As noted at the outset, the concept of litigating class actions based on evidence pertaining only to a sampling of the class members is no longer viable in light of *Dukes* and *Comcast*. In any event, this case amply illustrates the numerous practical problems attending the theoretical solution proposed in A.G. King and M. S. Muraco, "The False Dichotomy Posed by Sav-On and a Suggested Solution," 21 *Lab. Lawyer* 257, 269 (2006).

Dated: May 3, 2013

LITTLER MENDELSON, P.C.
501 W. Broadway, Suite 900
San Diego, CA 92101-3577

By: _____
ALLAN G. KING
JULIE A. DUNNE
MARGARET HART EDWARDS

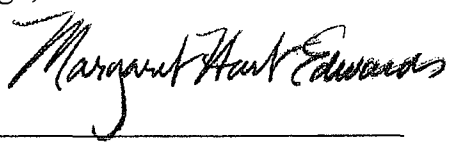
Attorneys for AMICUS CURIAE
THE GALLUP ORGANIZATION

CERTIFICATE OF COMPLIANCE

Pursuant to Rule 8.204(c)(1) of the California Rules of Court and in reliance on the word count of the computer program used to prepare this brief, counsel certifies that this Brief of Amicus Curiae The Gallup Organization in Support of Position of U.S. Bank National Association was produced using 13-point, Roman-type font and contains 8,857 words.

Dated: May 3, 2013

LITTLER MENDELSON, P.C.
501 W. Broadway, Suite 900
San Diego, CA 92101-3577

By: 

ALLAN G. KING
JULIE A. DUNNE
MARGARET HART EDWARDS
Attorneys for AMICUS CURIAE
THE GALLUP ORGANIZATION

PROOF OF SERVICE BY MAIL

I am employed in San Francisco County, California. I am over the age of eighteen years and not a party to the within-entitled action. My business address is 650 California Street, 20th Floor, San Francisco, CA 94108. I am readily familiar with this firm's practice for collection and processing of correspondence for mailing with the United States Postal Service. On May 3, 2013, I placed with this firm at the above address for deposit with the U.S. Postal Service a true and correct copy of the within documents:

BRIEF OF AMICUS CURIAE THE GALLUP ORGANIZATION IN SUPPORT OF POSITION OF U.S. BANK NATIONAL ASSOCIATION

in a sealed envelope, postage fully paid, addressed as follows:

SEE ATTACHED SERVICE LIST

Following ordinary business practices, the envelope was sealed and placed for collection and mailing on this date, and would, in the ordinary course of business, be deposited with the United States Postal Service on this date.

I declare under penalty of perjury under the laws of the State of California that the above is true and correct. Executed on May 3, 2013, at San Francisco, California.

JEFF SMITH

SERVICE LIST
Duran v. U.S. Bank National Association
CA Supreme Court Case No. S200923

Ellen Lake
The Law Offices of Ellen Lake
4230 Lakeshore Avenue
Oakland, CA 94610

Counsel for Plaintiffs & Respondents, Sam Duran and Matt Fitzsimmons

Edward Joseph Wynne
Wynne Law Firm
100 Drakes Landing Road, Suite 275
Greenbrae, CA 94904

Brad S. Seligman
Lewis Feinberg Lee Renaker
and Jackson P.C.
476 Ninth Street
Oakland, CA 94607

Alison L. Tsao
Timothy Mark Freudenberger
Kent Joseph Sprinkle
Carothers DiSante & Freudenberger LLP
601 Montgomery Street, Suite 350
San Francisco, CA 94111

Counsel for Defendant & Appellant, U.S. Bank National Association

Kimberly Ann Kralowec
Kralowec Law Group
188 The Embarcadero, Suite 800
San Francisco, CA 94105

Counsel for California Employment Lawyers Association : Pub/Depublication Requestor

Michael D. Singer
Cohelan Khoury and Singer
605 "C" Street, Suite 200
San Diego, CA 92101

Firmwide:119816602.2 026429.2000

ATTACHMENT

DAVID FREEDMAN
ROBERT PISANI
ROGER PURVES

Statistics

Fourth Edition

W • W • NORTON & COMPANY

NEW YORK • LONDON

the remaining 97 measurements come closer to the normal curve. In sum, most of the data have an SD of about 4 micrograms. But a few of the measurements are quite a bit further away from the average than the SD would suggest. The overall SD of 6 micrograms is a compromise between the SD of the main part of the histogram—4 micrograms—and the outliers.

In careful measurement work, a small percentage of outliers is expected. The only unusual aspect of the NB 10 data is that the outliers are reported. Here is what the Bureau has to say about *not* reporting outliers.[4] For official prose, the tone is quite stern.

> A major difficulty in the application of statistical methods to the analysis of measurement data is that of obtaining suitable collections of data. The problem is more often associated with conscious, or perhaps unconscious, attempts to make a particular process perform as one would like it to perform rather than accepting the actual performance .... Rejection of data on the basis of arbitrary performance limits severely distorts the estimate of real process variability. Such procedures defeat the purpose of the... program. Realistic performance parameters require the acceptance of all data that cannot be rejected for cause.

There is a hard choice to make when investigators see an outlier. Either they ignore it, or they have to concede that their measurements don't follow the normal curve. The prestige of the curve is so high that the first choice is the usual one—a triumph of theory over experience.

4. BIAS

Suppose a butcher weighs a steak with his thumb on the scale. That causes an error in the measurement, but little has been left to chance. Take another example. Suppose a fabric store uses a cloth tape measure which has stretched from 36 inches to 37 inches in length. Every "yard" of cloth they sell to a customer has an extra inch tacked onto it. This isn't a chance error, because it always works for the customer. The butcher's thumb and the stretched tape are two examples of *bias*, or *systematic error*.

> Bias affects all measurements the same way, pushing them in the same direction. Chance errors change from measurement to measurement, sometimes up and sometimes down.

The basic equation has to be modified when each measurement is thrown off by bias as well as chance error:

$$\text{individual measurement} = \text{exact value} + \text{bias} + \text{chance error}.$$

If there is no bias in a measurement procedure, the long-run average of repeated measurements should give the exact value of the thing being measured: the chance

The method in example 3 is called the multiplication rule.

> *Multiplication Rule.* The chance that two things will both happen equals the chance that the first will happen, multiplied by the chance that the second will happen given the first has happened.

Example 4. Two cards will be dealt off the top of a well-shuffled deck. What is the chance that the first card will be the seven of clubs and the second card will be the queen of hearts?

Solution. This is like example 3, with a much bigger box. The chance that the first card will be the seven of clubs is 1/52. Given that the first card was the seven of clubs, the chance that the second card will be the queen of hearts is 1/51. The chance of getting both cards is

$$\frac{1}{52} \times \frac{1}{51} = \frac{1}{2{,}652}.$$

This is a small chance: about 4 in 10,000, or 0.04 of 1%.

Example 5. A deck of cards is shuffled, and two cards are dealt. What is the chance that both are aces?

Solution. The chance that the first card is an ace equals 4/52. Given that the first card is an ace, there are 3 aces among the 51 remaining cards. So the chance of a second ace equals 3/51. The chance that both cards are aces equals

$$\frac{4}{52} \times \frac{3}{51} = \frac{12}{2{,}652}.$$

This is about 1 in 200, or 1/2 of 1%.

Example 6. A coin is tossed twice. What is the chance of a head followed by a tail?

Solution. The chance of a head on the first toss equals 1/2. No matter how the first toss turns out, the chance of tails on the second toss equals 1/2. So the chance of heads followed by tails equals

$$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

Exercise Set C

1. A deck is shuffled and two cards are dealt.
   (a) Find the chance that the second card is a heart given the first card is a heart.
   (b) Find the chance that the first card is a heart and the second card is a heart.

2. A die is rolled three times.
   (a) Find the chance that the first roll is an ace ⚀.
   (b) Find the chance that the first roll is an ace ⚀, the second roll is a deuce ⚁, and the third roll is a trey ⚂.

3. A deck is shuffled and three cards are dealt.
   (a) Find the chance that the first card is a king.
   (b) Find the chance that the first card is a king, the second is a queen, and the third is a jack.

4. A die will be rolled six times. You have a choice—
   (i) to win $1 if at least one ace shows
   (ii) to win $1 if an ace shows on all the rolls

   Which option offers the better chance of winning? Or are they the same? Explain.

5. Someone works example 2(a) on p. 226 this way:

   > For me to win, the queen can't be the first card dealt (51 chances in 52) and she must be the second card (1 chance in 51), so the answer is
   >
   > $$\frac{51}{52} \times \frac{1}{51} = \frac{1}{52}.$$

   Is the multiplication legitimate? Why?

6. "A cat-o'nine-_____ can be used to punish _____ of state, but this is seldom done." A coin is tossed twice, to fill in the blanks. What is the chance of the coin getting it right?

7. A coin is tossed 3 times.
   (a) What is the chance of getting 3 heads?
   (b) What is the chance of not getting 3 heads?
   (c) What is the chance of getting at least 1 tail?
   (d) What is the chance of getting at least 1 head?

The answers to these exercises are on p. A67.

4. INDEPENDENCE

This section introduces the idea of independence, which will be used many times in the rest of the book.

> Two things are *independent* if the chances for the second given the first are the same, no matter how the first one turns out. Otherwise, the two things are *dependent*.

Example 7. Someone is going to toss a coin twice. If the coin lands heads on the second toss, you win a dollar.

(a) If the first toss is heads, what is your chance of winning the dollar?
(b) If the first toss is tails, what is your chance of winning the dollar?

Assistant. Can you be more specific?

Kerrich. Let me write an equation:

number of heads = half the number of tosses + chance error.

This error is likely to be large in absolute terms, but small compared to the number of tosses. Look at figure 2. That's the law of averages, right there.

Assistant. Hmmm. But what would happen if you tossed the coin another 10,000 times. Then you'd have 20,000 tosses to work with.

Kerrich. The chance error would go up, but not by a factor of two. In absolute terms, the chance error gets bigger.[2] But as a percentage of the number of tosses, it gets smaller.

Assistant. Tell me again what the law of averages says.

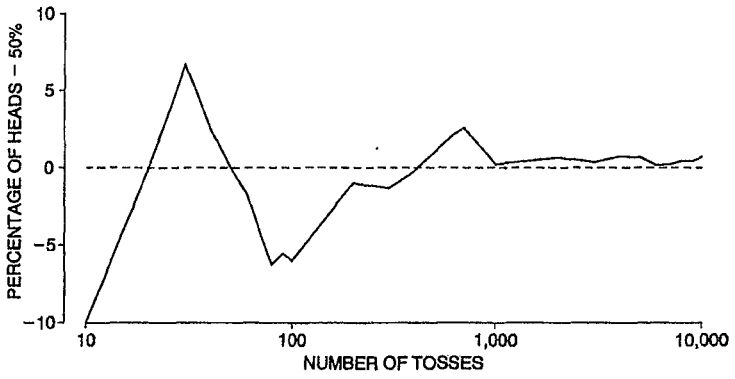
Kerrich. The number of heads will be around half the number of tosses, but it will be off by some amount—chance error. As the number of tosses goes up, the chance error gets bigger in absolute terms. Compared to the number of tosses, it gets smaller.

Assistant. Can you give me some idea of how big the chance error is likely to be?

Kerrich. Well, with 100 tosses, the chance error is likely to be around 5 in size. With 10,000 tosses, the chance error is likely to be around 50 in size. Multiplying the number of tosses by 100 only multiplies the likely size of the chance error by $\sqrt{100} = 10$.

Assistant. What you're saying is that as the number of tosses goes up, the difference between the number of heads and half the number of tosses gets

Figure 2. The chance error expressed as a percentage of the number of tosses. When the number of tosses goes up, this percentage goes down: the chance error gets smaller relative to the number of tosses. The horizontal axis is not to scale and the curve is drawn by linear interpolation.



bigger; but the difference between the percentage of heads and 50% gets smaller.

Kerrich. That's it.

Exercise Set A

1. A machine has been designed to toss a coin automatically and keep track of the number of heads. After 1,000 tosses, it has 550 heads. Express the chance error both in absolute terms and as a percentage of the number of tosses.
2. After 1,000,000 tosses, the machine in exercise 1 has 501,000 heads. Express the chance error in the same two ways.
3. A coin is tossed 100 times, landing heads 53 times. However, the last seven tosses are all heads. True or false: the chance that the next toss will be heads is somewhat less than 50%. Explain.
4. (a) A coin is tossed, and you win a dollar if there are more than 60% heads. Which is better: 10 tosses or 100? Explain.
   (b) As in (a), but you win the dollar if there are more than 40% heads.
   (c) As in (a), but you win the dollar if there are between 40% and 60% heads.
   (d) As in (a), but you win the dollar if there are exactly 50% heads.
5. With a Nevada roulette wheel, there are 18 chances in 38 that the ball will land in a red pocket. A wheel is going to be spun many times. There are two choices:
   (i) 38 spins, and you win a dollar if the ball lands in a red pocket 20 or more times.
   (ii) 76 spins, and you win a dollar if the ball lands in a red pocket 40 or more times.

   Which is better? Or are they the same? Explain.

The magnitude of the *Digest*'s error is staggering. It is the largest ever made by a major poll. Where did it come from? The number of replies was more than big enough. In fact, George Gallup was just setting up his survey organization.[4] Using his own methods, he drew a sample of 3,000 people and predicted what the *Digest* predictions were going to be—well in advance of their publication—with an error of only one percentage point. Using another sample of about 50,000 people, he correctly forecast the Roosevelt victory, although his prediction of Roosevelt's share of the vote was off by quite a bit. Gallup forecast 56% for Roosevelt; the actual percentage was 62%, so the error was 62% − 56% = 6 percentage points. (Survey organizations use "percentage points" as the units for the difference between actual and predicted percents.) The results are summarized in table 1.

Table 1. The election of 1936.

	Roosevelt's percentage
The election result	62
The *Digest* prediction of the election result	43
Gallup's prediction of the *Digest* prediction	44
Gallup's prediction of the election result	56

Note: Percentages are of the major-party vote. In the election, about 2% of the ballots went to minor-party candidates.
Source: George Gallup, *The Sophisticated Poll-Watcher's Guide* (1972).

To find out where the *Digest* went wrong, you have to ask how they picked their sample. A sampling procedure should be fair, selecting people for inclusion in the sample in an impartial way, so as to get a representative cross section of the public. A systematic tendency on the part of the sampling procedure to exclude one kind of person or another from the sample is called *selection bias*. The *Digest*'s procedure was to mail questionnaires to 10 million people. The names and addresses of these 10 million people came from sources like telephone books and club membership lists. That tended to screen out the poor, who were unlikely to belong to clubs or have telephones. (At the time, for example, only one household in four had a telephone.) So there was a very strong bias against the poor in the *Digest*'s sampling procedure. Prior to 1936, this bias may not have affected the predictions very much, because rich and poor voted along similar lines. But in 1936, the political split followed economic lines more closely. The poor voted overwhelmingly for Roosevelt, the rich were for Landon. One reason for the magnitude of the *Digest*'s error was selection bias.

> When a selection procedure is biased, taking a large sample does not help. This just repeats the basic mistake on a larger scale.

The *Digest* did very badly at the first step in sampling. But there is also a second step. After deciding which people ought to be in the sample, a survey