Module 8: Probabilistic Reasoning, Forensic Evidence, and the Relevance of Base Rates

The Gileadites seized the fords of the Jordan before the Ephraimites arrived. And when any Ephraimite who escaped said, "Let me cross over," the men of Gilead would say to him, "Are you an Ephraimite?" If he said, "No," then they would say to him, "Then say, 'Shibboleth'!" And he would say, "Sibboleth," for he could not pronounce it right. Then they would take him and kill him at the fords of the Jordan. There fell at that time forty-two thousand Ephraimites.

— Judges 12:5-6

Abstract: The topics developed in this module have at least a tacit connection to Bayes' theorem, and specifically to how base rates operate formally in the use of Bayes' theorem as well as more informally for several legally-related contexts. A number of topic areas are pursued: the general unreliability of eyewitness identification and testimony; polygraph testing; the assessment of blood alcohol level; the legal status and use of base rates; racial and ethnic profiling; false confessions; police interrogations; and the overall dismal state of the forensic "sciences."

An earlier Module 4 discussed the relevance of base rates in the evaluation of diagnostic tests and did so in several important contexts. One involved the Meehl and Rosen (1955) notion of "clinical efficiency" where prediction with a diagnostic test could be shown to outperform prediction using simple base rates. A second was a critique of the area under a Receiver Operating Characteristic curve (the AUC) as the sole mechanism for evaluating how well a particu-

lar diagnostic test performs; in general, the AUC is independent of base rates and fails to assess how well a diagnostic instrument does in specific populations that have relatively low base rates for the characteristic to be detected. When base rates are equal, test sensitivity and the positive predictive value (PPV) are equal (and so are the negative predictive value (NPV) and test specificity). Because of these equivalences, simple functions of the PPV and NPV make sense in communicating just how well or how badly a diagnostic instrument performs.

Contents

1 Bayes' Rule and the Importance of Base Rates

In the formulation of Bayes' rule given in Module 1, the two prior probabilities, $P(A)$ and $P(B)$, are also known as "base rates"; that is, in the absence of other information, how often do the events A and B occur. Base rates are obviously important in the conversion of $P(B|A)$ into $P(A|B)$, but as shown by Tversky and Kahneman, and others (for example, Tversky and Kahneman, 1974), base rates are routinely ignored when using various reasoning heuristics. The example given Module 1 on the importance of base rates in eyewitness identification involved the classic blue and black taxi cab problem; the example was made-up for clarity, but the principle it illustrates

has far-reaching real-world implications.

Some interesting commonalities are present across several forensic and medical domains where a knowledge of Bayes' theorem and the use of prior probabilities (or, base rates) may be crucial to the presentation of science-based recommendations, but which are then subsequently ignored (or discounted) by those very groups to which they are addressed. One area causing a great deal of controversy in the latter part of 2009 was the United States Preventive Services Task Force recommendations on cancer screening in women, particularly regarding when mammograms should start and their frequency. It is clear from the reactions in the media and elsewhere (for example, Congress), that irrespective of what may be reasonable science-based guidelines for women in general, on an individual level they will probably have no force whatsoever, despite recent reassuring results that targeted therapy is just as effective at saving lives without early detection.

Another arena in which Bayes' theorem has a role is in assessing and quantifying in a realistic way the probative (that is, legal-proof) value of eyewitness testimony. The faith the legal system has historically placed in eyewitnesses has been shaken by the advent of forensic DNA testing. In the majority of the numerous DNA exonerations occurring over the last twenty years, mistaken eyewitness identifications have been involved. A 2009 article by Wells, Memon, and Penrod ("Eyewitness Evidence: Improving Its Probative Value," in the series *Psychological Science in the Public Interest*), highlights the place that psychology and probabilistic reasoning have in this endeavor. We quote part of the abstract to give the flavor of the

review:[1]

> Decades before the advent of forensic DNA testing, psychologists were questioning the validity of eyewitness reports. Hugo Münsterberg's writings in the early part of the 20th century made a strong case for the involvement of psychological science in helping the legal system understand the vagaries of eyewitness testimony. But it was not until the mid-to-late 1970s that psychologists began to conduct programmatic experiments aimed at understanding the extent of error and the variables that govern error when eyewitnesses give accounts of crimes they have witnessed. Many of the experiments conducted in the late 1970s and throughout the 1980s resulted in articles by psychologists that contained strong warnings to the legal system that eyewitness evidence was being overvalued by the justice system in the sense that its impact on triers of fact (e.g., juries) exceeded its probative (legal-proof) value. Another message of the research was that the validity of eyewitness reports depends a great deal on the procedures that are used to obtain those reports and that the legal system was not using the best procedures. (p. 45)

A third area in which Bayesian notions are crucial to an understanding of what is possible, is in polygraph examinations and the quality of information that they can or cannot provide. Again, what appears to happen is that people want desperately to believe in some rational mechanism for detecting liars and cheats, and thereby increase one's sense of security and control. So, irrespective of the statistical evidence marshalled, and probably because nothing else is really offered to provide even an illusion of control in identifying prevarication, lie detector tests still get done, and a lot of them. An illuminating tale is Fienberg and Stern's, "In Search of the Magic Lasso: The Truth About the Polygraph," (2005) and the work of the

[1] A very informative *New Yorker* article on eyewitness evidence is by Atul Gawande ("Under Suspicion," January 8, 2001). A more recent news item from *Nature*, concentrates specifically on how lines-ups are (ill)conducted: "Eyewitness Identification: Line-Ups on Trial" (*Nature*, Laura Spinney, May 21, 2008).

National Research Council Committee to Review the Scientific Evidence on the Polygraph (2003). We give the abstract of the Fienberg and Stern article below, followed by three telling paragraphs from their concluding section:[2]

In the wake of controversy over allegations of espionage by Wen Ho Lee, a nuclear scientist at the Department of Energy's Los Alamos National Laboratory, the department ordered that polygraph tests be given to scientists working in similar positions. Soon thereafter, at the request of Congress, the department asked the National Research Council (NRC) to conduct a thorough study of polygraph testing's ability to distinguish accurately between lying and truth-telling across a variety of settings and examinees, even in the face of countermeasures that may be employed to defeat the test. This paper tells some of the story of the work of the Committee to Review the Scientific Evidence on the Polygraph, its report and the reception of that report by the U.S. government and Congress. (p. 249)

At the outset, we explained the seemingly compelling desire for a device that can assist law enforcement and intelligence agencies to identify criminals,

[2]An interesting historical subplot in the development of lie detection involved William Moulton Marston. Marston is usually given credit for promoting the development of an instrument for lie detection based on systolic blood pressure. His doctoral dissertation in experimental psychology at Harvard (1921) was entitled *Systolic Blood Pressure and Reaction-Time Symptoms of Deception and of Constituent Mental States*. It has been suggested (by none other than Marston's son) that it was actually Elizabeth Marston, William's wife, who was the motivation for his work on lie detection and its relation to blood pressure (quoting the son, "when she got mad or excited, her blood pressure seemed to climb"). In any case, Marston lived with two women in a polyamorous relationship—Elizabeth Holloway Marston, his wife, and Olive Byrne. Both these two women served as exemplars and inspirations for Marston's more well-known contribution to American life—the creation of the character and comic strip, *Wonder Women*, in the early 1940s under the pseudonym of Charles Moulton. Supposedly, it was Elizabeth's idea to create a female superhero who could triumph not with fists or firepower, but with love. This character would have a Magic Lasso (or a Golden Lasso, or a Lasso of Truth) that would force anyone captured by it to obey and tell the truth. So, besides introducing Wonder Woman and a lie detection instrument to a United States audience, Marston is credited with several additional cultural introductions. For more detail the reader is referred to the Wikipedia article on Marston.

spies and saboteurs when direct evidence is lacking. The polygraph has long been touted as such a device. In this article and in the NRC report on which it draws, we explain the limited scientific basis for its use, the deep uncertainty about its level of accuracy and the fragility of the evidence supporting claims of accuracy in any realistic application.

How should society, and the courts in particular, react to such a situation? At a minimum they should be wary about the claimed validity of the polygraph and its alternatives for use in the myriad settings in which they are used or proposed for use. This is especially relevant to current forensic uses of the polygraph. We believe that the courts have been justified in casting a skeptical eye on the relevance and suitability of polygraph test results as legal evidence. Generalizing from the available scientific evidence to the circumstances of a particular polygraph examination is fraught with difficulty. Further, the courts should extend their reluctance to rely upon the polygraph to the many quasiforensic uses that are emerging, such as in sex offender management programs. The courts and the legal system should not act as if there is a scientific basis for many, if any, of these uses. They need to hear the truth about lie detection.

As this paper was going to press in January 2005, the Department of Energy finally announced its proposed revised polygraph rules in the Federal Register. They provide a detailed plan for implementing the plan outlined in Deputy Secretary McSlarrow's September 2003 testimony. [Note: This was to only do 4,500 lie detector tests rather than the usual 20,000.] But no other federal agency has stepped forward with a plan to curb the use of polygraphs. All of them have heard the truth about polygraphs as we know it, but they have failed to acknowledge it by action. (p. 259)

We mention one last topic where a knowledge of Bayes' rule might help in arguing within another arena of forensic evidence: the assessment of blood alcohol content (BAC). The United States Supreme Court heard arguments in January of 2010 (Briscoe v. Virginia, 2010) about crime analysts being required to make court appearances, and

to (presumably) testify about the evidence and its reliability that they present now only in written form. The case was spurred in part by a California woman convicted of vehicular manslaughter with a supposed blood alcohol level two hours after the accident above the legal limit of .08. The woman denied being drunk but did admit to taking two shots of tequila (with Sprite chasers).[3]

There are several statistically related questions pertaining to the use of a dichotomous standard for BAC (usually, .08) as a definitive indication of impairment and, presumably, of criminal liability when someone is injured in an accident. Intuitively, it would seem that the same level of BAC might lead to different levels of impairment conditional on individual characteristics. Also, was this value set based on scientifically credible data? A variety of different BAC tests

[3]The woman's name is Virginia Hernandez Lopez; see, for example, Adam Liptak, *New York Times* (December 19, 2009), "Justices Revisit Rule Requiring Lab Testimony." In the actual case being orally argued of *Briscoe v. Virginia* (2010), the Court merely sent it back to a lower court in light of a recently decided case (*Melendez-Diaz v. Massachusetts* (2009)), which held that it is unconstitutional for a prosecutor to submit a chemical drug test report without the testimony of the scientist who conducted the test.

A more recent (5-4) Supreme Court ruling in *Bullcoming v. New Mexico* (2011) reaffirmed the *Melendez-Diaz* decision, saying that "surrogate testimony" would not suffice, and substitutes were not acceptable in crime lab testimony. The first paragraph of the syllabus in the *Bullcoming* opinion follows:

> The Sixth Amendment's Confrontation Clause gives the accused "[i]n all criminal prosecutions ... the right ... to be confronted with the witnesses against him." In Crawford v. Washington ... this Court held that the Clause permits admission of "[t]estimonial statements of witnesses absent from trial ... only where the declarant is unavailable, and only where the defendant has had a prior opportunity to cross-examine." Later, in Melendez-Diaz v. Massachusetts ... the Court declined to create a "forensic evidence" exception to Crawford, holding that a forensic laboratory report, created specifically to serve as evidence in a criminal proceeding, ranked as "testimonial" for Confrontation Clause purposes. Absent stipulation, the Court ruled, the prosecution may not introduce such a report without offering a live witness competent to testify to the truth of the report's statements.

could be used (for example, urine, blood, saliva, breath, hair); thus, there are all the possible interchangeability and differential reliability issues that this multiplicity implies.

The two most common alternatives to the supposedly most accurate blood test are based on urine and breath. Urine tests indicate the presence of alcohol in a person's system, but it takes up to two hours for the alcohol to show up. A positive urine test does not necessarily mean the person was under the influence of alcohol at the time of the test. Rather, it detects and measures usage within the last several days. Breath alcohol does not directly measure BAC but the amount of supposed "alcohol" in one's breath (as well as all chemically similar compounds and extraneous material such as vomit), and can be influenced by many external factors—cell phones, gasoline, blood, exercise, holding one's breath, and so on. We point the reader to an entry, "Blood Alcohol Testing in Drunk Driving Cases," posted by a lawyer, Aaron Larson, on the "expertlaw.com" website (2000).

A knowledge of Bayes' theorem and the way in which sensitivity, specificity, the positive predictive value, and the prior probability all operate together may at times be helpful to you or to others in mitigating the effects that a single test may have on one's assessment of culpability. There are many instances where the error rates associated with an instrument are discounted, and it is implicity assumed that an "observed value" is the "true value." The example of blood alcohol level just discussed seems to be, on the face of it, a particularly egregious example. But there are other tests that could be usefully approached with an understanding of Bayes' rule, such

as drug/steroid/human growth hormone use in athletes, blood doping in bicycle racers, polygraph tests for spying/white collar crime, fingerprint or eyewitness (mis)identification, or laser gun usage for speeding tickets. We are not saying that a savvy statistician armed with a knowledge of how Bayes' theorem works can "beat the rap," but it couldn't hurt. Anytime a judgment is based on a single fallible instrument, the value of the positive predictive value assumes a great

importance in establishing guilt or innocence.[4]

[4]We point to two items regarding lie detection that are relevant to making judgments based on a fallible instrument. One is by Margaret Talbot on using brain scans to detect lying ("Duped: Can Brain Scans Uncover Lies?," *New Yorker*, July 2, 2007); the other debunks voice-based lie detection: "The Truth Hurts: Scientists Question Voice-Based Lie Detection" (Rachel Ehrenberg, *ScienceNews*, June 22, 2010). A more general review devoted to lie detection by Vrij, Granhag, and Porter, appeared in *Psychological Science in the Public Interest* ("Pitfalls and Opportunities in Nonverbal and Verbal Lie Detection," 2010, *11*, 89–121). This article discusses behaviors that are not the best diagnostic indicators of lying. The term "illusory correlation," refers to a false but widely held belief in a relationship between two behaviors, for example, the drawing of big eyes in a Draw-A-Person projective test and a person's paranoia. In lying, there are the two illusory correlations of gaze aversion and nervousness.

The notion that gaze aversion reflects lying appears in our common idiomatic language in phrases such as "he won't look me in the eye." An editorial accompanying the review article cited above (Elizabeth Loftus, "Catching Liars," 2010, *11*, 87–88), comments directly on the cross-racial problem of using gaze aversion to suggest someone is lying:

Using gaze aversion to decide that someone is lying can be dangerous for that someone's health and happiness. And—what was news to me—some cultural or ethnic groups are more likely to show gaze aversion. For example, Blacks are particularly likely to show gaze aversion. So imagine now the problem that might arise when a White police officer interviews a Black suspect and interprets the gaze aversion as evidence of lying. This material needs to be put in the hands of interviewers to prevent this kind of cross-racial misinterpretation. (p. 87)

Coupled with a human tendency to engage in confirmation bias when an illusory correlation is believed, and to look for evidence of some type of "tell" such as "gaze aversion," we might once again remind ourselves to "lawyer up" early and often.

The illusory connection between nervousness and lying is so strong it has been given the name of "the Othello error." A passage from the Vrij et al. (2010) review provides a definition:

A common error in lie detection is to too readily interpret certain behaviors, particularly signs of nervousness, as diagnostic of deception. A common mistake for lie detectors is the failure to consider that truth tellers (e.g., an innocent suspect or defendant) can be as nervous as liars. Truth tellers can be nervous as a result of being accused of wrongdoing or as a result of fear of not being believed, because they too could face negative consequences if they are not believed ... The misinterpretation of signs of nervousness in truth tellers as signs of deceit is referred to as the Othello error by deception researchers ... based on Shakespeare's character. Othello falsely accuses his wife Desdemona of infidelity, and he tells her to confess

1.1 The (Legal) Status of the Use of Base rates

The use of base rates in the context of various legal proceedings, criminal matters, and subject identification has been problematic. The quotation that just opened this chapter shows the historical range for which base rates have come into consideration in a variety of (quasi-)legal settings. This section reviews several of these areas in more detail.

Shibboleth: This word comes directly from the Old Testament Biblical quotation (Judges 12:5-6) regarding the Gileadites and Ephraimites. It refers to any distinguishing practice, usually one of language, associated with social or regional origin that identifies its speaker as being a member of a group. There are a number of famous shibboleths: German spies during World War II mispronounced the initial "sch" in the Dutch port city's name of Scheveningen (and thereby could be "caught"); during the Battle of the Bulge, American soldiers used knowledge of baseball to tell whether there were German infiltrators in American uniforms; United States soldiers in the Pacific used the word "lollapalooza" to identify the Japanese enemy because a repeat of the word would come back with a beginning pronunciation of "rorra."[5]

because he is going to kill her for her treachery. When Desdemona asks Othello to summon Cassio (her alleged lover) so that he can testify her innocence, Othello tells her that he has already murdered Cassio. Realizing that she cannot prove her innocence, Desdemona reacts with an emotional outburst, which Othello misinterprets as a sign of her infidelity. The Othello error is particularly problematic in attempting to identify high-stakes lies because of the observer's sense of urgency and a host of powerful cognitive biases that contribute to tunnel-vision decision making ... (p. 98)

[5]Or, asking a person to say "rabbit" to see if he is Elmer Fudd.

Criminal trials: As noted in Module 1, Rule 403 of the *Federal Rules of Evidence* explicitly disallows the introduction of base rate information that would be more prejudicial than having value as legal proof. For instance, base rate information about which demographic groups commit which crimes and which don't would not be admissible under Rule 403. Although Rule 403 was given in MOdule 1, it is repeated below for completeness of the present discussion.

Rule 403. Exclusion of Relevant Evidence on Grounds of Prejudice, Confusion, or Waste of Time: Although relevant, evidence may be excluded if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury, or by considerations of undue delay, waste of time, or needless presentation of cumulative evidence.

Racial profiling: Although the Arizona governor, Jan Brewer, vehemently denied the label of racial profiling attached to its Senate Bill 1070, her argument comes down to officers knowing an illegal alien when they see one, and this will never depend on racial profiling because that, she says, "is illegal." How an assessment of "reasonable suspicion" would be made is left to the discretion of the officers—possibly a shibboleth will be used, such as speaking perfect English without an accent (or as the then governor of the state adjoining Arizona (Arnold Schwarzenegger) said: "I was also going to go and give a speech in Arizona but with my accent, I was afraid they were going to deport me back to Austria."). The reader is referred to the *New York Times* article by Randal C. Archibold ("Arizona Enacts Stringent Law on Immigration," April 23, 2010) that states succinctly the issues involved in Arizona's "Papers, Please" law.[6]

[6]As discussed in training videos for Arizona law-enforcement personnel, police can consider a variety of characteristics in deciding whether to ask about an individual's immigration

Constitutional protections: Two constitutional amendments protect the rights of individuals residing in the United States. The first amendment discussed is the Fourteenth, with its three operative clauses:

— The Citizenship Clause provides a broad definition of citizenship, overruling the decision in *Scott v. Sandford* (1857), which held that blacks could not be citizens of the United States. Those who follow current politics might note that this clause makes anyone born in the United States a citizen. Calls for its repeal are heard routinely from the political right, with the usual laments about "tourism babies," or those born to illegal immigrants. Irrespective of the citizenship of the parents, a baby born to someone temporarily in the United States is a United States citizen by default, and therefore, under all the protections of its laws.

— The Due Process Clause prohibits state and local governments from depriving persons of life, liberty, or property without steps being taken to insure fairness.

— The Equal Protection Clause requires the States to provide equal protection under the law to all people within its jurisdiction. This was the basis for the unanimous opinion in the famous *Brown v. Board of Education* (1954).

These three clauses are part of only one section of the Fourteenth

status: does the person speak poor English, look nervous, is he traveling in an overcrowded vehicle, wearing several layers of clothing in a hot climate, hanging out in areas where illegal immigrants look for work, does not have identification, does he try to run away, ... See Amanda Lee Myers, "Seventh Lawsuit Filed Over Ariz. Immigration Law" (Associated Press, July 10, 2010). It is difficult to see how any convincing statistical argument could be formulated that the use of behaviors correlated with ethnicity and race does not provide a *prima facie* case for racial profiling.

Amendment, which follows:

Section 1. All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the State wherein they reside. No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.

Although the "due process" and "equal protection" clauses seem rather definitive, the United States judicial system has found ways to circumvent their application when it was viewed necessary. One example discussed fully in Module 3 is the Supreme Court decision in *McCleskey v. Kemp* (1987) on racial disparities in the imposition of the death penalty (in Georgia). But probably the most blatant disregard of "equal protection" was the Japanese-American internment and relocation of about 110,000 individuals living along the United States Pacific coast in the 1940s. These "War Relocation Camps" were authorized by President Roosevelt on February 19, 1942, with the infamous *Executive Order 9066*. The Supreme Court opinion (6 to 3) in *Korematsu v. United States* (1944) upheld the constitutionality of *Executive Order 9066*. The majority opinion written by Hugo Black argued that the need to protect against espionage outweighed Fred Korematsu's individual rights and the rights of Americans of Japanese descent. In dissent, Justices Robert Jackson and Frank Murphy commented about both the bad precedent this opinion set and the racial issues it presented. We quote part of these two dissenting opinions:

Murphy: I dissent, therefore, from this legalization of racism. Racial dis-

crimination in any form and in any degree has no justifiable part whatever in our democratic way of life. It is unattractive in any setting, but it is utterly revolting among a free people who have embraced the principles set forth in the Constitution of the United States. All residents of this nation are kin in some way by blood or culture to a foreign land. Yet they are primarily and necessarily a part of the new and distinct civilization of the United States. They must, accordingly, be treated at all times as the heirs of the American experiment, and as entitled to all the rights and freedoms guaranteed by the Constitution.

Jackson: A military order, however unconstitutional, is not apt to last longer than the military emergency. Even during that period, a succeeding commander may revoke it all. But once a judicial opinion rationalizes such an order to show that it conforms to the Constitution, or rather rationalizes the Constitution to show that the Constitution sanctions such an order, the Court for all time has validated the principle of racial discrimination in criminal procedure and of transplanting American citizens. The principle then lies about like a loaded weapon, ready for the hand of any authority that can bring forward a plausible claim of an urgent need. Every repetition imbeds that principle more deeply in our law and thinking and expands it to new purposes.

. . .

Korematsu was born on our soil, of parents born in Japan. The Constitution makes him a citizen of the United States by nativity and a citizen of California by residence. No claim is made that he is not loyal to this country. There is no suggestion that apart from the matter involved here he is not law abiding and well disposed. Korematsu, however, has been convicted of an act not commonly a crime. It consists merely of being present in the state whereof he is a citizen, near the place where he was born, and where all his life he has lived. . . . [H]is crime would result, not from anything he did, said, or thought, different than they, but only in that he was born of different racial stock. Now, if any fundamental assumption underlies our system, it is that guilt is personal and not inheritable. Even if all of one's antecedents had been convicted of treason, the Constitution forbids its penalties to be visited upon him. But here is an attempt to make an otherwise innocent act

a crime merely because this prisoner is the son of parents as to whom he had no choice, and belongs to a race from which there is no way to resign. If Congress in peace-time legislation should enact such a criminal law, I should suppose this Court would refuse to enforce it.

Congress passed and President Reagan signed legislation in 1988 apologizing for the internment on behalf of the United States government. The legislation noted that the actions were based on "race prejudice, war hysteria, and a failure of political leadership." Over $1.6 billion was eventually dispersed in reparations to the interned Japanese-Americans and their heirs.

The other main amendment that has an explicit rights protection as its focus is the Fourth (from the Bill of Rights); its purpose is to guard against unreasonable searches and seizures, and to require a warrant to be judicially sanctioned and supported by "probable cause." The text of the amendment follows:

The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

Various interpretations of the Fourth Amendment have been made through many Supreme Court opinions. We mention two here that are directly relevant to the issue of law-enforcement application of base rates, and for (racial) profiling: *Terry v. Ohio* (1968) and *Whren v. United States* (1996). The Wikipedia summaries are given in both cases:

Terry v. Ohio ... (1968) was a decision by the United States Supreme Court which held that the Fourth Amendment prohibition on unreasonable searches

and seizures is not violated when a police officer stops a suspect on the street and frisks him without probable cause to arrest, if the police officer has a reasonable suspicion that the person has committed, is committing, or is about to commit a crime and has a reasonable belief that the person "may be armed and presently dangerous." ...

For their own protection, police may perform a quick surface search of the person's outer clothing for weapons if they have reasonable suspicion that the person stopped is armed. This reasonable suspicion must be based on "specific and articulable facts" and not merely upon an officer's hunch. This permitted police action has subsequently been referred to in short as a "stop and frisk," or simply a "Terry stop."

Whren v. United States ... (1996) was a United States Supreme Court decision which "declared that any traffic offense committed by a driver was a legitimate legal basis for a stop," [and] ... "the temporary detention of a motorist upon probable cause to believe that he has violated the traffic laws does not violate the Fourth Amendment's prohibition against unreasonable seizures, even if a reasonable officer would not have stopped the motorist absent some additional law enforcement objective."

In a dissenting opinion in *Terry v. Ohio* (1968), Justice William O. Douglas strongly disagreed with permitting a stop and search without probable cause:

I agree that petitioner was "seized" within the meaning of the Fourth Amendment. I also agree that frisking petitioner and his companions for guns was a "search." But it is a mystery how that "search" and that "seizure" can be constitutional by Fourth Amendment standards, unless there was "probable cause" to believe that (1) a crime had been committed or (2) a crime was in the process of being committed or (3) a crime was about to be committed.

The opinion of the Court disclaims the existence of "probable cause." If loitering were in issue and that was the offense charged, there would be "probable cause" shown. But the crime here is carrying concealed weapons; and there is no basis for concluding that the officer had "probable cause" for believing that that crime was being committed. Had a warrant been sought,

a magistrate would, therefore, have been unauthorized to issue one, for he can act only if there is a showing of "probable cause." We hold today that the police have greater authority to make a "seizure" and conduct a "search" than a judge has to authorize such action. We have said precisely the opposite over and over again.

. . .

There have been powerful hydraulic pressures throughout our history that bear heavily on the Court to water down constitutional guarantees and give the police the upper hand. That hydraulic pressure has probably never been greater than it is today.

Yet if the individual is no longer to be sovereign, if the police can pick him up whenever they do not like the cut of his jib, if they can "seize" and "search" him in their discretion, we enter a new regime. The decision to enter it should be made only after a full debate by the people of this country.

The issues of racial profiling and policies of "stop-question-and-frisk" are ongoing and particularly divisive in big urban areas such as New York City. To get a sense of this continuing controversy, the reader is referred to the *New York Times* article by Al Baker and Ray Rivera (October 26, 2010), "Study Finds Street Stops by N.Y. Police Unjustified." Several excerpts from this article follow that should illustrate well the contentiousness of the "stop-question-and-frisk" policies of the New York City Police Department.

The study was conducted on behalf of the Center for Constitutional Rights, which is suing the New York Police Department for what the center says is a widespread pattern of unprovoked and unnecessary stops and racial profiling in the department's stop-question-and-frisk policy. The department denies the charges.

...

Police Commissioner Raymond W. Kelly has rejected the accusation of racial profiling, and said the racial breakdown of the stops correlated to the racial breakdown of crime suspects. Mr. Kelly has also credited the tactic

with helping to cut crime to low levels in the city and with getting guns off the street.

...

The United States Supreme Court has held that in order for police officers to stop someone, they must be able to articulate a reasonable suspicion of a crime. To frisk them, they must have a reasonable belief that the person is armed and dangerous.

Darius Charney, a lawyer for the Center for Constitutional Rights, said the study crystallized the primary complaints in the lawsuit. "It confirms what we have been saying for the last 10 or 11 years, which is that stop-and-frisk patterns – it is really race, not crime, that is driving this," Mr. Charney said.

Mr. Kelly, responding to the professor's study, said on Tuesday, "I think you have to understand this was an advocacy paper." He also noted that Professor Fagan was paid well to produce the report.

Government institution protections: Although government institutions should protect rights guaranteed by the Constitution, there have been many historical failures. Many of these (unethical) intrusions are statistical at their core, where data are collected on individuals who may be under surveillance only for having unpopular views. To give a particularly salient and egregious example involving the FBI, J. Edgar Hoover, Japanese-American internment, and related topics, we redact the Wikipedia entry on the Custodial Detention Index (under the main heading of "FBI Index") used by the FBI from the 1930s to the 1970s (with various renamed successor indices, such as Rabble-Rouser, Agitator, Security, Communist, Administrative):

The Custodial Detention Index (CDI), or Custodial Detention List was formed in 1939-1941, in the frame of a program called variously the "Custodial Detention Program" or "Alien Enemy Control."

J. Edgar Hoover described it as having come from his resurrected General Intelligence Division in Washington:

"This division has now compiled extensive indices of individuals, groups, and organizations engaged in subversive activities, in espionage activities, or any activities that are possibly detrimental to the internal security of the United States. The Indexes have been arranged not only alphabetically but also geographically, so that at any rate, should we enter into the conflict abroad, we would be able to go into any of these communities and identify individuals or groups who might be a source of grave danger to the security of this country. These indexes will be extremely important and valuable in a grave emergency."

Congressmen Vito Marcantonio called it "terror by index cards." ...

The Custodial Detention Index was a list of suspects and potential subversives, classified as "A," "B," and "C"; the ones classified as "A" were destined to be immediately arrested and interned at the outbreak of war. Category A were leaders of Axis-related organizations, category B were members deemed "less dangerous" and category C were sympathizers. The actual assignment of the categories was, however, based on the perceived individual commitment to the person's native country, rather than the actual potential to cause harm; leaders of cultural organizations could be classified as "A," members of non-Nazi and pro-Fascist organizations.

The program involved creation of individual dossiers from secretly obtained information, including unsubstantiated data and in some cases, even hearsay and unsolicited phone tips, and information acquired without judicial warrants by mail covers and interception of mail, wiretaps and covert searches. While the program targeted primarily Japanese, Italian, and German "enemy aliens," it also included some American citizens. The program was run without Congress-approved legal authority, no judicial oversight and outside of the official legal boundaries of the FBI. A person against which an accusation was made was investigated and eventually placed on the index; it was not removed until the person died. Getting on the list was easy; getting off of it was virtually impossible.

According to the press releases at the beginning of the war, one of the purposes of the program was to demonstrate the diligence and vigilance of the government by following, arresting and isolating a previously identified group of people with allegedly documented sympathies for Axis powers and

potential for espionage or fifth column activities. The list was later used for Japanese-American internment.

Attorney General Francis Biddle, when he found out about the Index, labeled it "dangerous, illegal" and ordered its end. However, J. Edgar Hoover simply renamed it the Security Index, and told his people not to mention it.

USA PATRIOT Act: The attitude present during World War II that resident Japanese-Americans had a proclivity for espionage has now changed after September 11, 2001, to that of Middle Eastern men having a proclivity for committing terrorist acts. The acronym of being arrested because of a DWB ("driving while black") has now been altered to FWM ("flying while Muslim"). Section 412 of the *USA PATRIOT Act* allows the United States Attorney General to detain aliens for up to seven days without bringing charges when the detainees are certified as threats to national security. The grounds for detention are the same "reasonable suspicion" standard of *Terry v. Ohio* (1968). The Attorney General certification must state that there are "reasonable grounds to believe" the detainee will commit espionage or sabotage, commit terrorist acts, try to overthrow the government, or otherwise behave in a way that would endanger national security. After seven days, the detention may continue if the alien is charged with a crime or violation of visa conditions. When circumstances prohibit the repatriation of a person for an immigration offense, the detention may continue indefinitely if recertified by the attorney general every six months. Under the *USA PATRIOT Act*, a person confined for a violation of conditions of United States entry but who cannot be deported to the country of origin, may be indefinitely confined without criminal charges ever being filed.

Profiling, ethnic or otherwise, has been an implicit feature of

United States society for some time. The particular targets change, but the idea that it is permissible to act against specific individuals because of group membership does not. In the 1950s there were popular radio and television programs, such as *The FBI in Peace and War* or *I Led 3 Lives* about the double-agent Herbert Philbrick. These all focused on the Red menace in our midst, bent on overthrowing our form of government. It is instructive to remember our history whenever a new group is targeted for surveillance, and to note that the *Smith Act of 1940* (also known as the *Alien Registration Act*) is still on the books; the enabling "membership clause" and other conditions in the *Smith Act* follow:

> Whoever knowingly or willfully advocates, abets, advises, or teaches the duty, necessity, desirability, or propriety of overthrowing or destroying the government of the United States or the government of any State, Territory, District or Possession thereof, or the government of any political subdivision therein, by force or violence, or by the assassination of any officer of any such government; or
>
> Whoever, with intent to cause the overthrow or destruction of any such government, prints, publishes, edits, issues, circulates, sells, distributes, or publicly displays any written or printed matter advocating, advising, or teaching the duty, necessity, desirability, or propriety of overthrowing or destroying any government in the United States by force or violence, or attempts to do so; or
>
> Whoever organizes or helps or attempts to organize any society, group, or assembly of persons who teach, advocate, or encourage the overthrow or destruction of any such government by force or violence; or becomes or is a member of, or affiliates with, any such society, group, or assembly of persons, knowing the purposes thereof
>
> Shall be fined under this title or imprisoned not more than twenty years, or both, and shall be ineligible for employment by the United States or any department or agency thereof, for the five years next following his conviction.

If two or more persons conspire to commit any offense named in this section, each shall be fined under this title or imprisoned not more than twenty years, or both, and shall be ineligible for employment by the United States or any department or agency thereof, for the five years next following his conviction.

As used in this section, the terms "organizes" and "organize," with respect to any society, group, or assembly of persons, include the recruiting of new members, the forming of new units, and the regrouping or expansion of existing clubs, classes, and other units of such society, group, or assembly of persons.

Eyewitness reliability and false confessions: Several troublesome forensic areas exist in which base rates can come into nefarious play. One is in eyewitness testimony and how base rates are crucial to assessing the reliability of a witness's identification. The criminal case of "In Re As.H (2004)" reported in Module 9 illustrates this point well, particularly as it deals with cross-racial identification, memory lapses, how lineups are done, and so forth. Also, we have the earlier taxicab anecdote of Module 1. One possibly unexpected use that we turn to next involves base rate considerations in "false confessions." False confessions appear more frequently than we might expect and also in some very high profile cases. The most sensationally reported example may be the Central Park jogger incident of 1989, in which five African and Hispanic Americans all falsely confessed. To give a better sense of the problem, a short abstract is given below from an informative review article by Saul Kassin in the *American Psychologist* (2005, *60*, 215–228), entitled "On the Psychology of Confessions: Does Innocence Put Innocents at Risk":

The Central Park jogger case and other recent exonerations highlight the problem of wrongful convictions, 15% to 25% of which have contained con-

fessions in evidence. Recent research suggests that actual innocence does not protect people across a sequence of pivotal decisions: (a) In preinterrogation interviews, investigators commit false-positive errors, presuming innocent suspects guilty; (b) naively believing in the transparency of their innocence, innocent suspects waive their rights; (c) despite or because of their denials, innocent suspects elicit highly confrontational interrogations; (d) certain commonly used techniques lead suspects to confess to crimes they did not commit; and (e) police and others cannot distinguish between uncorroborated true and false confessions. It appears that innocence puts innocents at risk, that consideration should be given to reforming current practices, and that a policy of videotaping interrogations is a necessary means of protection. (p. 215)

To put this issue of false confession into a Bayesian framework, our main interest is in the term, $P(\text{guilty} \mid \text{confess})$. Based on Bayes' rule this probability can be written as

$$\frac{P(\text{confess} \mid \text{guilty})P(\text{guilty})}{P(\text{confess} \mid \text{guilty})P(\text{guilty}) + P(\text{confess} \mid \text{not guilty})P(\text{not guilty})}.$$

The most common interrogation strategy taught to police officers is the 9-step Reid Technique.[7] The proponents of the Reid Technique hold two beliefs: that $P(\text{confess} \mid \text{not guilty})$ is zero, and that they never interrogate innocent people, so the prior probability, $P(\text{guilty})$, is 1.0. Given these assumptions, it follows that if a confession is given, the party must be guilty. There is no room for error in the Reid system; also, training in the Reid system does not increase accuracy of an initial prior assessment of guilt but it does greatly increase confidence in that estimate. We thus have a new wording for an old adage: "never in error and never in doubt."

[7]A discussion of how police interrogation operates was written (and available online) by Julia Layton (May 18, 2006), "How Police Interrogation Works."

A number of psychological concerns are present with how interrogations are done in the United States. Innocent people are more likely to waive their *Miranda* rights (so unfortunately, they can then be subjected to interrogation); but somehow this does not seem to change an interrogator's prior probability of guilt.[8] People have a naive faith in the power of their own innocence to set them free. They maintain a belief in a just world where people get what they deserve and deserve what they get. People are generally under an illusion of transparency where they overestimate the extent that others can see their true thoughts. When in doubt, just remember the simple words—"I want a lawyer." (Or, in the idiom of the *Law & Order* series on TV, always remember to "lawyer-up.") If an interrogation proceeds (against our recommendation), it is a guilt-presumptive process that unfolds (it is assumed from the outset that $P(\text{guilty})$ is 1.0). False incriminating evidence can be presented to you (in contrast to the U.K, which is surprising because the United Kingdom doesn't have a "Bill of Rights"). Some people who are faced with false evidence may even begin to believe they are guilty. The interrogation process is one of social influence, with all the good cards stacked on one side of the table. It does not even have to be videotaped, so any post-confession argument of psychological coercion is hard to make.

As part of our advice to "lawyer up" if you happen to find yourself in a situation where you could be subjected to interrogation (and regardless of whether you believe yourself to be innocent or not),

[8] A minimal statement of a Miranda warning is given in the Supreme Court case of *Miranda v. Arizona* (1966): "You have the right to remain silent. Anything you say can and will be used against you in a court of law. You have the right to speak to an attorney, and to have an attorney present during any questioning. If you cannot afford a lawyer, one will be provided for you at government expense."

there is now a further need to be verbally obvious about invoking one's *Miranda* rights—counterintuitively, you have to be clear and audible in your wish not to talk. The Supreme Court issued the relevant ruling in June 2010. An article reviewing the decision from the *Los Angeles Times* (David G. Savage, "Supreme Court Backs Off Strict Enforcement of Miranda Rights," June 2, 2010) provides a cautionary piece of advice for those of us who might someday fall into the clutches of the criminal system through no fault of our own.

1.2 Forensic Evidence Generally

Most of us learn about forensic evidence and how it is used in criminal cases through shows such as *Law & Order*. Rarely, if ever, do we learn about evidence fallibility and whether it can be evaluated through the various concepts introduced to this point, such as base rates, sensitivity, specificity, prosecutor or defendant fallacy, or the positive predictive value. Contrary to what we may come to believe, evidence based on things such as bite marks, fibers, and voice prints are very dubious. As one example, we give the conclusion of a conference presentation by Jean-François Bonastre and colleagues (2003), entitled "Person Authentication by Voice: A Need for Caution":

> Currently, it is not possible to completely determine whether the similarity between two recordings is due to the speaker or to other factors, especially when: (a) the speaker does not cooperate, (b) there is no control over recording equipment, (c) recording conditions are not known, (d) one does not know whether the voice was disguised and, to a lesser extent, (e) the linguistic content of the message is not controlled. Caution and judgment must be exercised when applying speaker recognition techniques, whether human or automatic, to account for these uncontrolled factors. Under more constrained or calibrated situations, or as an aid for investigative purposes,

judicious application of these techniques may be suitable, provided they are not considered as infallible.

At the present time, there is no scientific process that enables one to uniquely characterize a person's voice or to identify with absolute certainty an individual from his or her voice. (p. 35)

Because of the rather dismal state of forensic science in general, Congress in 2005 authorized "the National Academy of Sciences to conduct a study on forensic science, as described in the Senate report" (H. R. Rep. No. 109-272). The Senate Report (No. 109-88, 2005) states in part: "While a great deal of analysis exists of the requirements in the discipline of DNA, there exists little to no analysis of the remaining needs of the community outside of the area of DNA. Therefore ... the Committee directs the Attorney General to provide [funds] to the National Academy of Sciences to create an independent Forensic Science Committee. This Committee shall include members of the forensics community representing operational crime laboratories, medical examiners, and coroners; legal experts; and other scientists as determined appropriate."[9]

[9]Implications of the NRC study have also appeared in the popular media. For example, an article from the *New York Times* by Clyde Haberman (May 18, 2014), entitled "DNA Analysis Exposes Flaws in an Inexact Forensic Science," emphasizes the fallibility of a heretofore staple of forensic science – microscopic hair analysis. We provide several paragraphs from the article:

This week's offering from Retro Report, a series of video documentaries that re-examine major stories from the past, zeros in on microscopic hair analysis, a staple of forensics for generations. It was long accepted as a virtually unerring technique to prove that this suspect – without a doubt, Your Honor – was the criminal. Wasn't a hair found at the scene?

But with the advent of DNA analysis in the late 1980s, apparent matches of hair samples ultimately proved to be not quite as flawless as people had been led to believe. Instances of wrongful imprisonment make that clear. Retro Report focuses on one such case, that of Kirk Odom, a Washington man who was found guilty of rape in 1981 and spent two decades behind bars. The Federal Bureau of Investigation's vaunted crime lab had asserted that hairs

The results of this National Research Council (NRC) study appeared in book form in 2009 from the National Academies Press (the

---

taken from his head were microscopically like – meaning virtually indistinguishable from – one found on the victim's nightgown. In time, however, DNA testing established that Mr. Odom was not the rapist, as he had asserted all along. Unfortunately for him, that official conclusion came late. By then, he had completed his prison sentence, a man done in by discredited forensic testimony.

Other lab techniques have had their reliability in the courtroom called into question. A 2009 report by a committee of the National Academy of Sciences found "serious problems" with an assortment of methods routinely relied on by prosecutors and the police. They included fingerprinting, blood typing, weapons identification, shoe print comparisons, handwriting, bite marks and – yes – hair testing. DNA was the game changer. The 2009 report said that, with the exception of nuclear DNA analysis, "no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source."

This is not to say that these techniques are no good at all. Indeed, the F.B.I. still affirms its faith in microscopic hair analysis, particularly as a first look. But it now tries to follow that procedure with a deeper and more certain investigation that uses DNA sampling, and it has done so for 18 years. Nonetheless, many forensic methods no longer come wrapped in the shield of invincibility they once widely enjoyed (especially among those prone to take TV shows literally). Fingerprints get blurred, bullets get smashed, blood specimens get tainted, hairs get mischaracterized.

...

The Innocence Project, a nonprofit group based in New York that uses DNA testing to help clear people wrongly convicted of crimes, has played a notable role in casting doubt on how forensic science is applied. Nationwide over the past 25 years, the project says, 316 people sent to prison have been exonerated through DNA analysis; 18 of them served time on death row. Hair comparisons performed by crime labs were factors in nearly one-fourth of those cases.

Even the F.B.I., while asserting the validity of hair analysis, has effectively acknowledged past problems.

In 2012, in an understanding reached with the Innocence Project and the National Association of Criminal Defense Lawyers, the F.B.I. agreed to a more cautious approach to stay squarely within the confines of known science. No absolutes. The bureau would now say, for example, only that a specific person could be included in, or could be excluded from, a "pool of people of unknown size" who might be the source of a specific hair sample. There would also be no statements of statistical probability. In addition, the F.B.I. says it is examining more than 2,500 old cases that lacked DNA evidence, to determine if hair analysis, of itself, played a role in guilty verdicts. It is unclear how far along this review is.

quotations just given are from this source): *Strengthening Forensic Science in the United States: A Path Forward*. The Summary of this NRC report provides most of what we need to know about the state of forensic science in the United States, and what can or should be done. The material that follows is an excerpt from the NRC Summary chapter:

Problems Relating to the Interpretation of Forensic Evidence:

Often in criminal prosecutions and civil litigation, forensic evidence is offered to support conclusions about "individualization" (sometimes referred to as "matching" a specimen to a particular individual or other source) or about classification of the source of the specimen into one of several categories. With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source. *In terms of scientific basis, the analytically based disciplines generally hold a notable edge over disciplines based on expert interpretation.* [italics added for emphasis] But there are important variations among the disciplines relying on expert interpretation. For example, there are more established protocols and available research for fingerprint analysis than for the analysis of bite marks. There also are significant variations within each discipline. For example, not all fingerprint evidence is equally good, because the true value of the evidence is determined by the quality of the latent fingerprint image. These disparities between and within the forensic science disciplines highlight a major problem in the forensic science community: The simple reality is that the interpretation of forensic evidence is not always based on scientific studies to determine its validity. This is a serious problem. Although research has been done in some disciplines, there is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.

The Need for Research to Establish Limits and Measures of Performance:

In evaluating the accuracy of a forensic analysis, it is crucial to clarify the type of question the analysis is called on to address. Thus, although

some techniques may be too imprecise to permit accurate identification of a specific individual, they may still provide useful and accurate information about questions of classification. For example, microscopic hair analysis may provide reliable evidence on some characteristics of the individual from which the specimen was taken, but it may not be able to reliably match the specimen with a specific individual. However, the definition of the appropriate question is only a first step in the evaluation of the performance of a forensic technique. A body of research is required to establish the limits and measures of performance and to address the impact of sources of variability and potential bias. Such research is sorely needed, but it seems to be lacking in most of the forensic disciplines that rely on subjective assessments of matching characteristics. These disciplines need to develop rigorous protocols to guide these subjective interpretations and pursue equally rigorous research and evaluation programs. The development of such research programs can benefit significantly from other areas, notably from the large body of research on the evaluation of observer performance in diagnostic medicine and from the findings of cognitive psychology on the potential for bias and error in human observers.

The Admission of Forensic Science Evidence in Litigation:

Forensic science experts and evidence are used routinely in the service of the criminal justice system. DNA testing may be used to determine whether sperm found on a rape victim came from an accused party; a latent fingerprint found on a gun may be used to determine whether a defendant handled the weapon; drug analysis may be used to determine whether pills found in a person's possession were illicit; and an autopsy may be used to determine the cause and manner of death of a murder victim. ... for qualified forensic science experts to testify competently about forensic evidence, they must first find the evidence in a usable state and properly preserve it. A latent fingerprint that is badly smudged when found cannot be usefully saved, analyzed, or explained. An inadequate drug sample may be insufficient to allow for proper analysis. And, DNA tests performed on a contaminated or otherwise compromised sample cannot be used reliably to identify or eliminate an individual as the perpetrator of a crime. These are important matters involving the proper processing of forensic evidence. The law's greatest dilemma

in its heavy reliance on forensic evidence, however, concerns the question of whether—and to what extent—there is *science* in any given forensic science discipline.

Two very important questions should underlie the law's admission of and reliance upon forensic evidence in criminal trials: (1) the extent to which a particular forensic discipline is founded on a reliable scientific methodology that gives it the capacity to accurately analyze evidence and report findings and (2) the extent to which practitioners in a particular forensic discipline rely on human interpretation that could be tainted by error, the threat of bias, or the absence of sound operational procedures and robust performance standards. These questions are significant. Thus, it matters a great deal whether an expert is qualified to testify about forensic evidence and whether the evidence is sufficiently reliable to merit a fact finder's reliance on the truth that it purports to support. Unfortunately, these important questions do not always produce satisfactory answers in judicial decisions pertaining to the admissibility of forensic science evidence proffered in criminal trials.

A central idea present throughout the collection of modules is that "context counts" and it "counts crucially." It is important both for experts and novices in how a question is asked, how a decision task is framed, and how forensic identification is made. People are primed by context whether as a victim making an eyewitness identification of a perpetrator, or as an expert making a fingerprint match. As an example of the latter, we have the 2006 article by Dror, Charlton, and Péron, "Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications" (*Forensic Science International*, *156*, 74–78). We give their abstract below:

We investigated whether experts can objectively focus on feature information in fingerprints without being misled by extraneous information, such as context. We took fingerprints that have previously been examined and assessed by latent print experts to make positive identification of suspects. Then we presented these same fingerprints again, to the same experts, but gave a con-

text that suggested that they were a no-match, and hence the suspects could not be identified. Within this new context, most of the fingerprint experts made different judgments, thus contradicting their own previous identification decisions. Cognitive aspects involved in biometric identification can explain why experts are vulnerable to make erroneous identifications. (p. 74)

References

[1] Meehl, P., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin*, *52*, 194–215.

[2] Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.