Statistical Science
2023, Vol. 38, No. 2, 285–302
https://doi.org/10.1214/22-STS870

The Secret Life of I. J. Good

Sandy Zabell

Abstract. I. J. ("Jack") Good was a leading Bayesian statistician for more than half a century after World War II, playing an important role in the post-war Bayesian revival. But his graduate training had been in pure mathematics rather than statistics (one of his doctoral advisors at Cambridge had been the famous G. H. Hardy). What was responsible for this metamorphosis from pure mathematician to applied and theoretical statistician? As Good himself only revealed in 1976, during the war he had initially served as an assistant to Alan Turing at Bletchley Park, working on the cryptanalysis of the German Naval Enigma, and it was from Turing that he acquired his life-long Bayesian philosophy. Declassified and other documents now permit us to understand in some detail how this came about, and indeed how many of the ideas Good explored and papers he wrote in the initial decades after the war, in fact, gave in sanitized form, results that had their origins in his wartime work. Drawing on these sources, this paper discusses the daily and very real use of Bayesian methods Turing and Good employed, and how this was gradually revealed by Good over the course of his life (including his return to classified work in the 1950s).

Key words and phrases: I. J. Good, Alan Turing, Bayesian statistics, Bletchley Park, cryptanalysis, Enigma machine, Banburismus, deciban, weight of evidence, Tunny.

Irving John (I. J., "Jack") Good (December 9, 1916–April 5, 2009) was one of the most prominent Bayesian statisticians during the second half of the 20th century.[1] His many contributions to statistics include a seminal book, *Probability and the Weighing of Evidence* (1950), the "Turing-Good estimator" for the sampling of species (Good, 1953), many papers in statistical journals such as the *Annals of Statistics*, *Statistical Science*, the *Journal of the American Statistical Association*, the *Journal of the Royal Statistical Society* and *Biometrika*, as well as hundreds of additional papers in the philosophical, computing, and scientific literatures. But he became a Bayesian precisely when it was in a state of near total eclipse thanks to attacks by statistical giants such as R. A. Fisher and Jerzy Neyman. Why?

The answer is closely connected with his "secret life," the work he did on cryptanalysis at Bletchley Park during World War II (and, as it turns out, afterwards as well). At Bletchley, Good initially worked under the famous Alan Turing (1912–1954) and this had a decisive impact on his career: his Bayesian approach to statistics, his interest in data, his ready adoption of the use of computers. Good only began to reveal this three decades later, and even then he only gradually divulged over a period of several decades specific details about exactly what he did. The structure of this paper parallels this: the first part discusses what Good did and said up to 1976, and how this might have appeared to someone on the outside; the second part what Good and others revealed starting in 1976.

1. THE CAREER OF I. J. GOOD

Here is a short version of what Good's *curriculum vitae* might have looked like in the early 1970s:

Good, Irving John (Born London, December 9, 1916)

1938:	BA Cambridge (Jesus College)
1941:	PhD University of Cambridge (Mathematics)
1941–45:	Foreign Office
1945–48:	Lecturer, University of Manchester
1948–59:	Government Communications Headquarters
1959–62:	Admiralty Research Laboratory
1962–64:	Consultant, Institute for Defense Analyses
1964–67:	Senior Research Fellow, Oxford (Trinity)
1967–:	Professor, Virginia Polytechnic Institute

As can be seen, there are several intriguing aspects of Good's career that are already evident even in so simple

Sandy Zabell is Professor, Departments of Mathematics and Statistics, Northwestern University, Evanston, Illinois, 60208, USA (e-mail: zabell@math.northwestern.edu).

[1] Born Isadore Jacob Gudak, he began to use the anglicized version of his name by the time he published his first paper (Good, 1940).

a document. Good received his doctorate in mathematics at Cambridge (under the joint supervision of G. H. Hardy and A. S. Besicovitch, two of the purest of pure mathematicians); he did no graduate work at all in statistics. Nor did he start out as a typical academic: 20 of the first 26 years after he received his PhD were spent outside of academia, in a succession of mostly defense-related establishments. It was only in 1967, when he was 50, that he settled down at Virginia Polytechnic Institute (now known as Virginia Tech), where he remained for the rest of his life.

2. BLETCHLEY PARK YEARS

Good was an undergraduate at Jesus College, Cambridge from 1934 to 1938. Continuing on at Cambridge as a graduate student under Hardy and Besicovitch, he received the prestigious Smith's Prize in mathematics in 1940, and was awarded his PhD in 1941.

When Good received his PhD, Britain was at war with Germany. What did he do during the war? If asked, he would have said he worked in the Foreign Office. This was his first secret.

2.1 "I Worked in the Foreign Office . . . "

Shortly before Good began his his wartime work in 1941, he met Stuart Milner-Barry at a chess match and indiscreetly asked him if he was working on German ciphers. Milner-Barry replied "No, my address is Room 47, Foreign Office" (Banks, 1996, p. 8). Milner-Barry was in fact then working at Bletchley Park on the cryptanalysis of German Army and Luftwaffe messages.

There was a reason for this evasion. During the war (and long after), anyone who worked at Bletchley Park, then the center of British wartime signals intelligence, was strictly forbidden to say anything about the nature of the highly sensitive work done there. The officially approved euphemism instead was that they "worked in the Foreign Office."[2] When Sarah Turing, Alan Turing's mother, wrote a biography of her son in 1959 after his death, all she knew of her son's wartime work was that he was a "Temporary Civil Servant in the Foreign Office, in the Department of Communications" (Turing, 1959, p. 67).[3] Of course to the initiated, such a bland statement is itself a strong hint. [4]

2.2 Good Reports to Bletchley

Good reported to Bletchley Park on May 27, 1941 (coincidentally, the day the *Bismarck* was sunk). He was a talented chess player (the 1939 Cambridgeshire chess champion) and this may have played some role in his ending up at Bletchley—one of the three people who interviewed him as a potential recruit was C. H. O'D. (Conel Hugh O'Donel, "Hugh") Alexander (1909–1974), twice British Chess Champion and the best player in England before the war. Alexander at that point was working under Turing at Bletchley Park and played a key role in the cryptanalysis of the Naval Enigma. Good and Alexander knew each other from the chess world, and Alexander met Good at the train station the day Good first reported to work.

Good was fortunate in his wartime assignment: he spent his first two years (1941–1943) at Bletchley Park working in Hut 8 (Naval cryptanalysis) under Alan Turing, from whom he learned the practical Bayesian approach to statistics;[5] and his last two years (1943–1945) working in the Newmanry (one of two sections devoted to cryptanalysis of the SZ 40/42, an online teleprinter system), working under the Cambridge mathematician M. H. A. ("Max") Newman (1897–1984), using an attack centered on the use of the "Colossus." These were very different experiences.

2.2.1 *Hut* 8. Hut 8 was the section at Bletchley Park devoted to the cryptanalysis of German Naval Enigma messages. Employed throughout the *Kriegsmarine* (German Navy), the Naval Enigma was used to communicate with the Kriegsmarine's ships and U-boat (submarine) fleet. The ability of the Allies to read Naval Enigma traffic starting in the summer of 1941 was ultimately a significant factor in the Battle of the Atlantic, enabling the Allies to locate and sink many U-boats, as well as steer Allied convoys and other ships away from them.[6]

[2]Strictly speaking, this was actually true: Bletchley Park was part of GC&CS, the Government Code and Cypher School, which since the end of World War I had been administratively housed in the Foreign Office. For the arcane reason this came to be, see Budiansky (2000, pp. 51–52).

[3]She did know his work was of considerable importance, that at one point he supervised 100 women who worked under him, had traveled to the US during the war and was afterwards awarded the OBE (Order of the British Empire) in recognition of his wartime work.

[4]Two examples are J. H. Plumb's 1950 *England in the Eighteenth Century*, and A. T. Hatto's 1965 translation of the German epic saga *The Nibelungenlied*: in both cases. the short author biographies at the beginning say that the authors worked in the Foreign Office during the war. In fact, both spent the war at Bletchley Park. Plumb worked there from December 1940 to 1945: he was Head of German, Italian and (later) Japanese Signals Intelligence, and later Head of the Japanese Naval Traffic Analysis subsection. Hatto was at Bletchley Park from September 1939 to 1945; he worked on *Abwehr* (German military intelligence), Gestapo and weather ciphers; see Flood (2011, pp. 177–178).

[5]Earlier influences, more mathematical and philosophical rather than practical, included Hall and Knight's *Higher Algebra* (1891), John Maynard Keynes, Frank Plumpton Ramsey and Harold Jeffreys; see Good (1983, p. x). My thanks to Glenn Shafer for pointing out this passage to me.

[6]First documented in Patrick Beesly's *Very Special Intelligence* (Beesly (1977)). During World War II, Beesly worked in the Operational Intelligence Centre of the Admiralty's Naval Intelligence Division, and so had first-hand knowledge of the utility of naval signals intelligence. In 1977, the technical aspects of this were not yet public knowledge; Ralph Erskine's "Afterword" in the 2000 edition brought

The section was headed by the brilliant Alan Turing. Turing had arrived at Bletchley Park on September 4, 1939, one day after England declared war on Germany. Although the Naval Enigma was then regarded as unbreakable, within a just few months (December 1939) Turing had worked out the theoretical basis of a statistical attack on the Naval Enigma; as a result, Hut 8 was set up in January 1940, initially staffed by Turing, Peter Twinn (in Naval Intelligence) and two clerical assistants. By the summer of 1941, Hut 8 had become fully operational and able to read some messages quickly enough (within days, sometimes less) to provide actionable intelligence. It eventually employed more than a hundred people, including the mathematicians Peter Hilton and Shaun Wylie (later the authors of a well-known book on homology, Hilton and Wylie, 1967), and the famous English historian, J. H. Plumb.

The key to Turing's attack was *Banburismus*, a statistical method for determining the right-most wheel out of three used in the encrypting of Naval Enigma messages.[7] These were selected from a set of eight (for a total of $8 \cdot 7 \cdot 6 = 336$ orders), making impractical an attack by the *Bombe* (a special purpose electromechanical device constructed to attack Army and Luftwaffe traffic, which only selected the three wheels of the Enigma from a set of five rather than eight, and which therefore only had 60 wheel orders to search, a factor of more than five less). Banburismus was the basic method used to attack Naval Enigma traffic for more than two years.

When Good joined Hut 8 the basic attack on the Naval Enigma had already been worked out, and was just then being employed operationally. Despite this, Good soon made several important contributions to improving the efficiency of the attack, and was "responsible for a considerable amount of the most valuable statistical work done in the hut" (Alexander, 1945, p. 63); some specifics of this are discussed below in Section 5. It was only half a century later that detailed information about the work in Hut 8 began to be released.

2.2.2 *Transfer to the Newmanry.* In the spring of 1943, the United States and Britain started producing new and more powerful Bombes, which could be used in the attack against the Naval Enigma without the need for the shortcut of Banburismus, and so the use of Banburismus was discontinued in September 1943.[8] As a result, Good was transferred in September from Hut 8 to the Newmanry, one of two sections at Bletchley Park devoted to the cryptanalysis of the SZ 40/42.[9]

The SZ 40/42 was a German Army machine used to encrypt teleprinter traffic. It was an online device (meaning that it could simultaneously encrypt and send a message, unlike the Enigma, in which these two operations had to be performed separately), and was used to send messages typically much longer than those of the Enigma (thousands of characters as opposed to the imposed maximum of 200 for Enigma messages). Its contents were often strategic in nature rather than the tactical content found in many Army and Luftwaffe Enigma messages, and often gave valuable insights into German intentions, order of battle and so on. (So, e.g., it is possible Hitler's infamous message to the Commandant of Paris shortly before the city fell to the Allies in August 1944, instructing it be destroyed rather than let it fall into enemy hands, was sent in encrypted form using this device.)

Messages encrypted by the SZ 40/42 were transmitted using the Baudot code, in which characters were represented by a sequence of five impulses (think of these as 0s and 1s), which were then subjected by the machine to two successive layers of encryption. Thanks to a defect in the encryption process, Bletchley Park discovered these two layers of encryption could in fact be successively stripped off. The *Newmanry*, named after its head, Max Newman, removed one layer of encryption primarily by statistical means, using the *Colossus*, a complex device capable of compiling various statistical summaries at very high speed. The result was then passed on to the *Testery*, named after its head, Major Ralph Tester (1902–1998), which then removed the second layer of encryption, primarily by classical linguistic means.

Good and Wylie, being mathematicians, were naturally assigned to the Newmanry. Good stayed there for the remainder of the war, leaving in September 1945. Just as in Hut 8, he made many practical and theoretical contributions to the operation of the section, including bringing over Turing's Bayesian approach. He also gained valuable experience in machine computation, experience which was to serve him well after the war. The details of the attack on *Tunny* (the codename for the SZ 40/42) will be discussed in Section 6.

the book up to date in this regard as well as updating the book's bibliography.

[7]Named after the English town of Banbury, where the sheets of paper used to carry out part of the process were printed.

[8]See Alexander (1945, Chapter 6, Section 10) and Mahon (1945, Chapter 10, "The abandonment of Banburismus"). The fact that the US Bombes were in the US did not present a problem for the British. Bletchley Park and OP-20-G (the US Navy's cryptologic organization) were able to communicate rapidly and securely by sending enciphered messages via cable. If urgent, a message "would take under an hour from the time we began to write the signal out in Hut 8 to the completion of its decyphering in Op. 20 G. As a result of this, we were able to use the Op. 20 G bombes almost as conveniently as if they had been at one of our outstations 20 or 30 miles away" (Alexander, 1945, p. 90).

[9]For the month of Good's transfer, see Alexander (1945, p. 63). Good, looking back decades later, first gave the month as October but afterwards changed his mind, saying it was April (1993, p. 160, 2006, p. 208), but Alexander's 1945 internal history, written by the head of the section at the time of the transfer, is clearly to be preferred.

3. POSTWAR YEARS UP TO 1976

After the war Good spent several years at the University of Manchester (1945–1948), working on statistics and computing,[10] after which he returned to GCHQ (Government Communications Headquarters, the postwar successor to GC&CS), where he remained for the next 11 years (1948–1959). Much later, he said he returned to intelligence work because (apart from a dislike for teaching) "the Cold War was heating up and I thought I could do more good in government service" (Banks, 1996, p. 13).

Even now there is essentially no information about what Good did during his GCHQ years. When asked by David Banks in a 1993 interview "can you give me any sense as to what types of mathematics or what types of things you were thinking about" at GCHQ, Good dodged the question, saying "I think it's better that I don't say anything" and immediately changed the topic (Banks, 1996, p. 13). The only hint appears to be that at one point he worked on VENONA, a top-secret US–UK project devoted to deciphering Soviet messages (Budiansky, 2006, p. 61).

Good resigned from GCHQ in the late 1950s in order to accept a full professorship at the University of Chicago, but changed his mind at the last moment for "personal reasons" and stayed in England.[11] (This may have been due to the illness of his mother.) After visiting several institutions for two- to three-year stints (Admiralty Research Laboratory, 1959–1962; Institute for Defense Analyses, 1962–1964; Trinity College, Oxford, 1964–1967), he became a Professor at Virginia Polytechnic Institute, where he remained for the rest of his life.

With the benefit of hindsight, the influence of his wartime work on much of his published postwar statistical research is apparent and easily traced.

3.1 Probability and the Weighing of Evidence (1950)

Immediately after the war, Good wrote his classic book *Probability and the Weighing of Evidence* (1950), espousing the subjective, Bayesian viewpoint (but with a strong pragmatic streak running throughout).[12] In retrospect, it is clear the book advances a view of the subject shaped by his Bletchley Park experiences. In his preface Good wrote (p. vi):

> Dr. A. M. Turing, Professor M. H. A. Newman and Mr. D. Michie were good enough to read the first draft (written in 1946) and I am most grateful for their numerous suggestions.

Apart from the date of its first draft, and the fact that Turing, Newman and Donald Michie had been respectively his two bosses at Bletchley and his closest collaborator in the Newmanry, there is also direct internal evidence in the book of the Bletchley Park influence.

3.1.1 *Likelihood ratios and the weight of evidence*. In the Bayesian attack on a cryptosystem (as envisaged by Turing), the evidence E at one's disposal typically consisted of the contents of an encrypted message, and the goal was to identify the correct setting of the machine (or system) H, versus one or more alternative settings $\bar{H}$, by computing their likelihood ratio. Central to Turing's approach was the use of the *Bayes factor*

$$\frac{P(E \mid H)}{P(E \mid \bar{H})}.$$

To ease computing, Turing introduced the **deciban**:

$$10 \cdot \log_{10} \frac{P(E \mid H)}{P(E \mid \bar{H})}.$$

There were two reasons for this:

- the log converted products into sums (which were easier to compute);
- the factor of 10 was used to simplify the arithmetic.

In his 1950 book, Good devoted an entire chapter (Chapter 6, "Weighing evidence") to exploring the log-likelihood, describing it as "the *weight of evidence* or amount of information for H given E." This was a topic he returned to many times in his later work, for example, Good (1960, 1968a, 1975) and Good and Toulmin (1968). Good was always careful to acknowledge Turing in this, albeit cautiously. The deciban was a multiple of the *ban*, so called because it was used in Banburismus; but in his book Good (1950, p. 63) avoided this sensitive connection, writing instead:

> Turing suggested further that it would be convenient to take over from acoustics and electrical engineering the notation of bels and decibels (db).

3.1.2 *Sequential analysis*. Turing's Banburismus was part of a sequential cryptanalytic process using decibans as inputs in a Baysian sequential probability ratio test he had developed (Banks, 1996, pp. 9–11), independently of both Abraham Wald and George Barnard (who had also

[10] Good was hired by Newman, who became head of the Manchester mathematics department at the end of the war. For further information on Newman, see Adams (1985).

[11] Some of the details of this episode are documented in the archives of the Department of Statistics at the University of Chicago.

[12] Ironically, the book was rejected when initially submitted to the Cambridge University Press. Discouraged, Good did not pursue publication further until urged to do so by his Bletchley colleague, Donald Michie, in 1948.

come up with the idea for wartime applications).[13] Good's 1950 book treats sequential analysis in Section 6.2 (pp. 64–66). He did not have to worry about having this particular discussion in his book cleared, because he could simply point to both Barnard (1946) and Wald's papers (1945a and 1945b) and 1947 book, but he could not resist ending however with the cryptic comment:

> The sequential technique is clearly not restricted to the quality control of goods. It can be used for deciding between any two "simple statistical hypotheses",

clearly having Banburismus in mind.

3.1.3 *The theorem of the weighted average of (partial) factors*. In considering a message, one usually knew the sender and recipient. These might influence the statistical characteristics of a message in a known way. In addition, based on prior experience one might also know that a certain fraction of the time one of several different distributions occurred (depending, e.g., on the operator and the type of message being sent, such as a weather report) although one did not know which type beforehand. (For example, two-thirds of the time one might encounter one type of message and one-third of the time another.) Such information was incorporated into an attack using the *theorem of the weighted average of (partial) factors*.

Good's 1950 book gives a clear statement and proof of this result (p. 68). Suppose $H = H_1 \cup \cdots \cup H_n$ is a composite hypothesis (so that the H_i are mutually exclusive), $\overline{H}$ the negation of H and E evidence. Then if

$$p_i = P(H_i \mid H), \qquad f_i = \frac{P(E \mid H_i)}{P(E \mid \overline{H})},$$

the theorem states that the factor in favor of H given E is

$$\frac{P(E \mid H)}{P(E \mid \overline{H})} = \sum_i P(H_i \mid H) \frac{P(E \mid H_i)}{P(E \mid \overline{H})} = \sum_i p_i f_i.$$

3.2 Postwar Papers

But Good's Bletchley Park-inspired contributions to statistics in the years immediately after the war were not confined to just a general advocacy of the Bayesian viewpoint. He proceeded to publish (always carefully crediting Turing) refinements of a number of technical advances Turing had developed during the war. As Good later explained:

> Turing did not publish these wartime statistical ideas because after the war he was too busy working on the ground floor of computer science and artificial intelligence. I was impressed by the importance of his statistical ideas, for other applications, and developed and published some of them in various places. Much of my delay was caused by the wartime attitude that everything was classified, from Hollerith cards to sequential statistics, to empirical Bayes, to Markov chains, to decision theory, to electronic computers. These extreme standards of secrecy only gradually abated after the war. [Good, 1992a, p. 211]

Good instead waited "until it was clear that Turing's interests lay elsewhere," and "statistics was no longer regarded as a classified topic" (Good, 2000, p. 106). The resulting papers touched on a variety of topics:

- the sampling of species problem (Good, 1953 and 1956), used in the attack on the Naval Enigma;
- the variance of the weight of evidence (Good, 1961), which Turing had analyzed in the normal case, and which Good extended to other cases;
- a scoring method for repeats (Good, 1973), which extended the method Turing used in Banburismus;
- the discrete Fourier transform (Good, 1951, 1958, 1962), which Good had learned about from Turing during the war.

A frequent clue throughout these papers is an acknowledgement to Turing. See also Banks (1996, pp. 10–11).

3.2.1 *The sampling of species problem*. As part of the process of encryption when using the Naval Enigma, a three letter trigram was chosen from a book and enciphered using one of nine tables, which were eventually known to the British. Determining which table was in use on a given day was an important step in the process of decryption.

Different users had different copies of the book, and experience over time revealed the trigrams were not being chosen at random.[14] This provided the basis for an attack: each of the nine possible tables were used to determine an underlying candidate trigram, and these were scored on the basis of whether they were more or less common. Inasmuch as there were $26^3 = 17{,}576$ possible trigrams, this presented a statistical challenge: estimating the probabilities of a large number of "species" (trigrams), each of which necessarily had a small probability, based on relatively limited data.

[13] Good briefly discussed the method with Barnard at some point in 1941–1942, not mentioning its cryptanalytic application, a conversation he remembered vividly decades later because at the time he was concerned about avoiding a potential breach of security (Good, 1992a, pp. 219–220).

[14] "The popular ones turned out to be at the top of the blocks of 25, particularly those on the central pages, as captures observed." Letter from Joan Clark Murray to I. J. Good, September 26, 1993; quoted in Good (2000, p. 110).

Turing's solution to this challenge was to devise what is today called the "Turing-Good estimator." Suppose a particular trigram has been observed a total of r times in a sample of N trigrams. Turing's insight was that in order to estimate the probability of seeing the trigram again in the future there was, in addition to r and N, valuable information to be gleaned from the "frequencies of the frequencies"; that is, for each r, $1 \le r \le N$, the number n_r of distinct trigrams each occurring r times in the sample. Evidently,

$$\sum_{r=1}^{N} r n_r = N.$$

Turing's proposed estimator for the frequency of the trigram was

$$\frac{n_{r+1}}{n_r}\left(\frac{r+1}{N}\right).$$

It is apparent that the interest and utility of such an estimator goes far beyond its original cryptanalytic use, and—with Turing's permission—Good wrote a paper in *Biometrika* (Good, 1953) which, using Turing's formula as a starting point, considerably expanded on its theory, discussed improvements in its practical application involving smoothing of the observed n_r and gave many illustrations of its utility. Good's paper was the starting for a now considerable body of literature about this estimator, which performs well in a broad variety of circumstances and which typically represents a considerable improvement on the MLE r/N.

Good was careful to hide the cryptanalytic origins of Turing's estimator. Contrary to the usual practice of *Biometrika*, the author's affiliation (GCHQ) was not given. And regarding Turing, Good carefully says (p, 237):

> The formula was first suggested to me, together with an intuitive demonstration, by Dr. A. M. Turing several years ago. Hence. a very large part of the credit for the present paper should be given to him, and I am most grateful to him for allowing me to publish this work.

Good returned to this subject later in a follow-up paper, Good and Toulmin (1956).[15]

3.2.2 *The discrete Fourier transform*. The DFT (the discrete Fourier transform) was used in the Newmanry to calculate discrete convolutions (Reeds et al., 2015, p. 583). Turing first drew Good's attention to it during the war; after the war, Good made good use of it, employing it in some 20 publications, covering at least 10 distinct areas; see Banks (1996, p. 10), Reeds et al. (2015, pp. 583–584). These postwar papers sometimes reflected earlier wartime conversations.[16]

4. THE LIFTING OF THE EMBARGO

Up until 1976, Good remained entirely silent about his actual wartime work. It can be hard today to appreciate just how complete the silence was regarding Allied successes in attacking German encryption devices. One instructive example is provided by David Kahn's pathbreaking book *The Codebreakers* (Kahn (1967)): although it contains an entire chapter on the US success in reading the Japanese "Purple" cipher, and several chapters on German signals intelligence, it is entirely silent about Bletchley Park and Ultra.[17]

All this changed in 1973, when General Gustave Bertrand (1896–1976) wrote *Enigma, ou la plus grande énigme de la guerre 1939–1945* ("Enigma, or the Greatest Enigma of the War of 1939–1945"). This revealed that since 1932 the Poles had been reading the Enigma, as well as the Polish-French collaboration in the years leading up to and after the outbreak of the war. The publication of Bertrand's book apparently served as an inducement to the British to lift their embargo a year later (1974) on any discussion of their cryptologic successes during the war; the first beneficiary of this change in policy was F. W. Winterbotham's *The Ultra Secret* (1974). After this, the floodgates open and an ever-increasing succession of books and papers appeared, including notably Good (1976 and 1979), Hinsley (1979–1990), Rejewski (1981), Welchman (1982), Hinsley and Stripp (1993), Good (2000), Copeland (2006) and Reeds et al. (2015).

Good's silence prior to 1976 was certainly well advised. The history of signals intelligence contains a number of celebrated instances of old hands feeling free to publish without prior approval accounts of their wartime successes, only to suffer serious consequences after. A cautionary tale here is that of Gordon Welchman (1906–1985), one highly relevant to Good. Welchman had headed Hut 6 (Army and Luftwaffe cryptanalysis) at Bletchley Park, and was responsible for many important advances during the war. (He was also the moving

[15] Also published in *Biometrika*. Once again, no affiliation is given, but in a later paper on a different subject, Good and Toulmin (1968), submitted February 17, 1967, Toulmin's affiliation is listed as "Government Communications Headquarters, Cheltenham, Gloucestershire, England."

[16] In a paper using the discrete Fourier transform to derive the Poisson summation formula, Good wrote: "I am indebted to Dr. S. Wylie, Professor D. Rees and Professor M. H. A. Newman for stimulating discussions sixteen years ago" (Good, 1962, p. 259). Presumably submitted in 1961, 16 years earlier would be 1945, when Good was still at Bletchley Park. (Besides Newman and Wylie, David Rees, 1918–2013, was another a colleague of Good in the Newmanry. Later a distinguished mathematician, he is well known for the "Artin-Rees theorem" in algebra.)

[17] As Kahn (2010, p. 16) himself wryly noted several decades later.

force behind a famous letter to Prime Minister Winston Churchill, personally delivered to Churchill's principal private secretary on October 21, 1942, complaining about a lack of sufficient resources, as a result of which Hut 6 and Hut 8 were immediately given virtual carte blanche in obtaining personnel and materiel.) Welchman emigrated to the US in 1948, and spent the rest of his life working primarily for the US defense establishment. He kept scrupulously quiet about his outstanding contributions to the Allied war effort for more than 35 years, but in the late 1970s, as revelations about the work at Bletchley Park began to emerge, Welchman concluded silence was no longer required. And so he came to write his highly informative book *The Hut Six Story* (1982), which detailed the many successes in the attack on the Enigma, the devices (such as the Bombe and diagonal board) used in its attack, and Alan Turing's crucial role in all this. But he made a fatal error, not submitting his book beforehand for prepublication review. He was promptly stripped of his security clearance, forbidden to speak to the press and remained under a cloud for the remainder of his life.[18]

4.1 The Dance of the Seven Veils

Good first publicly disclosed he had worked at Bletchley Park in a lecture at the National Physical Laboratory, Teddington, on April 28, 1976 (Good, 1976). The resulting NPL report foreshadows just how carefully Good was to approach the subject over the next several decades: a footnote on the first page notes "This paper was cleared by the British Cabinet Office in June 1976" (p. 31 of the 1980 reprint). As its title ("Early Work on Computers at Bletchley") suggests, the paper discusses the role of hardware but says nothing whatever about cryptanalysis, for, as Good explains, "I have not been told that I can refer to the cryptanalytic techniques" (p. 38). Good obviously thought the paper important for, besides its initial release as an NPL report, in 1979–80 it was reprinted in two journals and a book with a much wider circulation. Good also discussed computing at Bletchley Park in a few other places at this time, but without adding anything further of substance (see, e.g., Good, 1982, pp. 53–59).

The paper also recounts a number of Good's personal experiences at Bletchley, including some of the most interesting people he met there, starting with Turing. (This may in fact have been the first public mention of the nature of Turing's work during the war.) Among those mentioned were the chess masters Hugh Alexander, Stuart Milner-Barry and Harry Golombek, who worked in Huts 6 and 8; his mathematicial colleagues, M. H. A. Newman, J. H. C. Whitehead, David Rees, Shaun Wylie and Peter Hilton, who worked in the Newmanry; and public figures such as Roy Jenkins (later a prominent member of the Labor Party and Chancellor of the Exchequer) and Peter Benenson (the founder of Amnesty International), both of whom worked in the Testery.

It was only later that Good first began to reveal (in a very limited way) some of the technical statistical aspects of the attacks on the Enigma and Tunny. One might describe this as the "the dance of the seven veils," some of the highlights include:

- 1976: Lecture at the National Physical Laboratory
- 1979: *Biometrika* paper on Turing's wartime work
- 1993: "Enigma and Fish" chapter in *Codebreakers*
- 1996: David Banks *Statistical Science* interview
- 2000: Use of the Turing–Good estimator in attack on the Naval Enigma revealed (Good, 2000)
- 2006: "From Hut 8 to the Newmanry" (in Copeland volume)
- 2015: *The General Report on Tunny* (written in 1945, 500 pages, declassified in 2004, scholarly edition published in 2015)

4.2 A. M. Turing's Statistical Work in World War II

This was the title of a brief paper Good published in *Biometrika* in 1979 describing Turing's "unpublished contributions to statistics" at Bletchley Park during WWII. At first glance, the paper seems like a curiosity: a jumble of simple results and techniques in statistical inference. In fact, it is clear in retrospect that what the paper actually does is lay out the sequence of steps in the statistical attack on the Enigma, each step being an integral part in that attack:

- Bayes factors
- Sequential analysis and log factors
- The deciban
- Weighted averages of Bayes factors
- Design of experiments and expected weight of evidence
- The variance of the weight of evidence
- Expected values of Bayes factors
- Search trees
- Repeat rates
- Empirical Bayes

Here, the *link* with Good's book and some of his papers (Good, 1950, 1953, 1956, 1961, 1969, and Good and Toulmin, 1956 and 1968), not merely to Turing but to Bletchley Park and cryptanalysis, was revealed, but no detail given. The deciban, for example, is described as being used as part of "an important classified process called Banburismus" (but we are not actually told what Banburismus is), and that the main application of the deciban "was to sequential analysis, not for quality control but for discriminating between hypotheses" (but we are not told what those hypotheses were).

Good's 1979 *Biometrika* paper was later reprinted in the volume on pure mathematics in the *Collected Works*

[18] For an excellent biography of Welchman, see Greenberg (2014).

of A. M. Turing (Britton, 1992), together with an accompanying commentary (Good, 1992a). That commentary, in addition to providing background on the Enigma, gives a detailed description of the linkages between Turing's wartime results and no fewer than seventeen papers written by Good between 1960 and 1989 (several of them having to do with the philosophical implications of the weight of evidence).

Even more interesting is a paper Good wrote 14 years earlier (Good, 1965a), which appeared in the *NSA Technical Journal*. The paper, "A list of properties of Bayes-Turing factors", although marked "Unclassified", was not publicly available and only approved for release by the NSA in 2011.[19] It provides a useful complement to Good's 1979 paper because it is more specific and gives more mathematical detail about how Good's later work expanded on Turing's earlier efforts.

4.2.1 *The central role of Bayes factors*. The Bayes factor played a central role in Turing's overall philosophy of cryptanalysis. This was not confined to the special case of the attack on the Enigma, as became clear several decades later in 2012 (the centenary of Turing's birth), when GCHQ declassified and released a paper Turing had written during the war, "The Applications of Probability to Cryptography" (Turing, 2012a). In this remarkable document (intended for newcomers to Bletchley or at least the Bayesian approach), Turing illustrated how Bayesian methods could be applied to four distinct classical cryptanalytic problems (the Vigenère cipher, a letter subtractor problem, the theory of repeats and transposition ciphers). In each case, an attack based on Bayes factors was described, together with a discussion of how the computations needed for the attack could be carried out in practice. (This wedding of the theoretical and the practical was a key reason for Turing's spectacular success at Bletchley.) For a commentary on the essay, see Zabell (2012). One of the reasons for Good's transfer to the Newmanry in 1943 appears to have been in part to bring over the Bayesian approach.

4.2.2 *The cryptanalytic use of the weight of evidence*. In several sections of his 1979 paper (3, 6–9), Good discussed a number of properties of the Bayes factor, weight of evidence and expected weight of evidence. In a paper from this same period, while discussing Harold Jeffreys's contributions to statistics Good (1980, p. 27) elaborated on the wartime use of the concept of weight of evidence:

> Turing pointed out that, when expected weight of evidence per observation is small:
>
> $$\int f(x)\log\frac{f(x)}{g(x)}\,dx \quad \text{and}$$
>
> $$\int g(x)\log\frac{g(x)}{f(x)}\,dx$$
>
> are approximately equal and opposite. I therefore found it natural to use as a quasi-utility, in some classified applications during World War II, their sum
>
> $$\int \big[f(x)-g(x)\big]\log\big[f(x)/g(x)\big]\,dx,$$
>
> which is now often called the *divergence* between two probability densities.

Good added that Solomon Kullback (of Kullback–Leibler fame, who worked in the US Army's Signal Intelligence Service during the war and retired as a senior member of the NSA in 1962), was also an "intensive user" of both expected weight of evidence and the divergence.[20]

In later years, Good wrote extensively in the outside literature about the weight of evidence as a central element in his statistical philosophy.

In the next two sections, we discuss some of the details of the specific statistical contributions Jack Good made in the attacks on the Enigma and the SZ 40/42. In each case, this requires a preliminary discussion of the operation of the machine.

5. THE ENIGMA

The *Enigma* was a commercial German cryptographic device invented by Arthur Scherbius in 1918. The core of

[19] Curiously, even then it was only released after a FOIA appeal, despite the fact that the technical results in it appear to have been in the public domain for more than three decades.

[20] People working on classified research sometimes encounter the frustration that an important result published by someone else in the outside literature had in fact been discovered by themselves earlier, but in a classified setting, and for precisely this reason were unable to publish it. (One well-known example is the discovery of the RSA and Diffie–Hellman public key encryption methods at GCHQ several years before their outside publication; see Singh, 1999, pp. 279–292.) It is natural to speculate that this comment by Good was intended to stake a claim to his priority in recognizing the divergence as a statistically useful quantity. It was important for him to state this occurred during the war because, as he notes, Jeffreys had made use of the divergence shortly after in 1946. Indeed, Good even goes so far as to speculate that Bletchley Park may have been the indirect source for Jeffreys!

> It seems possible that [Jeffreys] thought of this integral through familiarity with Gibb's work, or perhaps its unsymmetrical form was mentioned to him by Turing. I know that John Wishart and G.H. Hardy, also both Cambridge men, were told top secret facts about the work at Bletchley, so certainly Jeffreys *should* have been initiated.

the machine involved three wheels, each effecting a permutation of the alphabet; as part of initially setting up the device the three wheels were arranged in one of $3! = 6$ possible orders (the *Walzenlage*). Each wheel had a set of electrical contacts along its rim numbered from 1 to 26; prior to each message being sent the three wheels were rotated so that a specific contact was in the upmost position (the *message setting*). Each time a letter of the message was encrypted one or more of the wheels moved, resulting in a new permutation of the alphabet, thus avoiding the weaknesses inherent in a classical monoalphabetic substitution cipher such as those discussed in Edgar Allan Poe's *The Gold-Bug*, or Sir Arthur Conan Doyle's *The Adventure of the Dancing Men*.[21]

5.1 The German Army Enigma

The Enigma was adopted by the German military in the late 1920s (by the Navy in 1926, the Army in 1928), but modified to increase its security. The Germans added a plugboard (the *Steckerbrett*) through which the electric current generated during encryption first entered and later exited the machine. The plugboard interchanged several pairs of letters and left others unchanged. By 1939, ten pairs of letters were selected each day, giving rise to a total of 150, 738, 274, 937, 250 possible settings for the plugboard connections.

Showing truly remarkable foresight, in 1932 Polish intelligence hired three young mathematicians—Marian Rejewski (1905–1980), Jerzy Różycki (1909–1942) and Henryk Zygalski (1908–1978)—to attack the Enigma. Although in principle relatively secure, thanks to both a blunder in its use (a three-letter message setting was sent to the message recipient by encrypting it *twice*—the blunder—using a *Grundstellung* (an initial setting used throughout the day), as well as brilliant cryptanalysis by the Polish mathematicians, the Poles were in fact able to read a majority of the Enigma messages (some 75%, closer to 90% had they been given additional personnel) for much of the decade. Over time, the Germans made successive improvements in their use of the machine, but until the end of 1938 the Poles were always able to meet these new challenges; see Rejewski (1981), Turing (2021).

This happy state of affairs changed in the winter of 1938. On December 15, 1938, two additional wheels were added to the original set of three, and the three wheels in the machine were now selected from this set of five, so that there were now $5 \cdot 4 \cdot 3 = 60$ different possible arrangements, resulting in a ten-fold increase in the work necessary for decrypting messages, well beyond the capacity of Rejewski's team. In July 1939, the Poles, sensing the inevitable, met with their French and English counterparts in Pyry, just south of Warsaw, and passed on to them the fruits of their nearly decade-long exploitation of the device, including replicas of the machine together with the internal wiring of the five wheels; see Turing (2021).

When Turing arrived at Bletchley Park in September 1939, he set to work constructing a special purpose device, the *Bombe*, which substantially improved on the mechanical devices employed by the Poles. This, together with a further modification due to Gordon Welchman (the *diagonal board*) permitted the British to begin reading some Army and Luftwaffe Enigma traffic by the beginning of 1940. This is a well-known story; excellent accounts include Welchman (1982) and Budiansky (2000). The cryptanalysis of Army and Luftwaffe Enigma traffic was the responsibility of Hut 6, headed by Welchman.

5.2 The Naval Enigma

The Enigma used by the *Kriegsmarine* (German Navy) was a much more secure device than the one used by the *Heer* (Army) and Luftwaffe. As noted earlier, instead of five wheels, the three wheels of the Naval Enigma were chosen from a set of eight, increasing the number of wheel orders from 60 to 336. Not even Turing's 1940 Bombe could handle this.[22] Furthermore, the Kriegsmarine used a much more complex (and initially unknown) method of sending the three-letter message setting using a set of bigram tables. So when Turing arrived at Bletchley the Naval Enigma was considered unbreakable, and no one was working on it. But (characteristically) Turing viewed this as a challenge and an opportunity rather than a deterrent (he later said he started to work on it because he could "have it all to myself"). By the beginning of 1940, Turing had invented *Banburismus*, a paper-and-pencil method of determining (under favorable circumstances) the right-most of the three wheels, which meant the work of determining the wheel order would be narrowed down to $7 \cdot 6 = 42$ possible orders, well within the capability of the Bombes then being constructed. The same evening Turing was also able to deduce the method by which the message setting was being encrypted. When a set of Enigma

[21] Each time a letter was encrypted the wheels advanced in approximate odometric fashion: the right-most wheel advanced every step, the middle wheel advanced once every twenty-six steps, and each time the middle wheel advanced to a particular ("turnover") position, both it and the left wheel advanced in unison on the next step. (This meant the machine cycled through all possible settings of the wheels in $25 \cdot 26^2 = 16{,}900$ steps, not $26^3 = 17{,}576$ steps.) A lettered ring attached to the left side of a wheel established a correspondence between letters and contacts. In order to use the machine one needed to know both the order of the wheels (the *Walzenlage*) and how the rings were attached (the *Ringstellungen*). These were part of the daily setting, shared by all machines in a particular network.

[22] This was a practical rather than theoretical issue. A single wheel order took about 20 minutes to test on the 3-wheel Bombe, so testing all 336 wheel orders would take about 112 hours on one machine, or 11 hours on ten. (And if the crib—the conjectured plaintext the process required—was wrong one had to start over again from scratch.)

keys were captured from a German trawler (the *Krebs*) on March 4, 1941, it was possible to reconstruct the bigram tables being used to perform the encryption and this, together with a statistical attack on their use devised by Turing, enabled one to decrypt the message setting. Within a few months Bletchley Park was able to read Naval Enigma traffic with regularity; see Kahn (1991).[23] It is hard to overstate the central role that Turing played in all this. Shortly after the war Hugh Alexander, who had succeeded Turing as the head of the section, wrote in his classified internal history of Hut 8:

> There should be no question in anyone's mind that Turing's work was the biggest factor in Hut 8's success. In the early days, he was the only cryptographer who thought the problem worth tackling and not only was he primarily responsible for the main theoretical work within the Hut (particularly the developing of a satisfactory scoring technique for dealing with Banburismus) but he also shared with Welchman and Keen the chief credit for the invention of the Bombe. It is always difficult to say that anyone is absolutely indispensable but if anyone was indispensable to Hut 8 it was Turing. The pioneer work always tends to be forgotten when experience and routine later make everything seem easy and many of us in Hut 8 felt that the magnitude of Turing's contribution was never fully realized by the outside world. [Alexander, 1945, pp. 42–43.]

When Good joined Hut 8 in May 1941, a systematic attack was up and running, and so there was no question of him making fundamental contributions to the basic attack. Nevertheless, he almost immediately played an important role in increasing the efficiency of the attack in several ways, both practical and theoretical. Several of these will now be discussed.

5.2.1 *Refining the deciban*. As noted earlier, to facilitate computing, Turing had introduced the *deciban*:

$$10 \cdot \log_{10} \frac{P(E \mid H)}{P(E \mid \bar{H})},$$

which in Banburismus was computed to one decimal place of accuracy. Soon after arriving at Bletchley, Good advocated using a factor of 20 instead of 10, and rounding to the nearest integer. He made this suggestion both because it turned out most of the individual scores would then be single digits (and so easier to add), as well as having computed how much information (in terms of expected weight of evidence) would lost by this additional rounding and found it was relatively minor. Although this change sounds relatively minor, it ended up saving half the time needed for Banburismus (Good, 1993, p. 158, Banks, 1996, p. 9).

This contribution illustrates both Good's practical turn of mind (despite his training being exclusively in pure mathematics), one which made him such a useful a member of Bletchley Park, as well as the sometimes surprising utility of having a fresh pair of eyes take a second look at a problem.

5.2.2 *Attacking the bigram tables*. In addition to the general daily setting (*Walzenlage, Ringstellungen*, and *Steckerverbindungen*) known ahead of time to everyone in a network, the recipient of a message also had to know the specific message setting (in terms of rotating the wheels so that specific contacts were pointing up) in order to decrypt it. In the case of the Naval Enigma, this was done by choosing a trigram from a book, encrypting it using one of nine fixed *bigram tables*, and then appending the resulting encrypted trigram to the beginning of the message. The recipient would then reverse this process and read off the trigram.[24]

Here is an example. The *Kenngruppenbuch* was a book containing all 17,576 possible trigrams in a scrambled order. The sender would:

- choose two trigrams, say *LQR*, *CPY*, from the *Kenngruppenbuch*;
- choose an additional pair of "haphazard" letters, say G and O;
- Use these to form a rectangle and then encrypt each column of the rectangle using the bigram table in force for that day. For example:

$$\begin{array}{cc} G\ L\ Q\ R & T\ A\ L\ I \\ \rightarrow & \\ C\ P\ Y\ O & U\ H\ S\ U \end{array}$$

(So the bigram table told the sender to replace GC by TU, and so on.) The resulting eight letters (the *message indicator*) were then appended to the start of the message. The receiver would then reverse this process, and use CPY to decrypt the message.

[23] This oversimplifies a more complex reality. One complication was that various materials such as key lists and short signal books (useful as a source of cribs) were changed from time to time, and it considerably simplified matters if these could be "pinched" from a captured German vessel; see Budiansky (2000, pp. 191–196, 283–285). This could be a dangerous business: on one occasion two men (Lieutenant Anthony Fasson and Able Seaman Colin Grazier) were lost when they were unable to get off a scuttled U-boat in time.

[24] Strictly speaking there was another step: both the sender and receiver would encrypt this trigram using a common daily setting—the *Grundstellung*—and it was this encrypted trigram that was actually used as the message setting. Successfully dealing with this additional complication was the genius behind Turing's method of Banburismus.

This apparently impressive procedure however had two fatal weaknesses, due to the presence of two sources of supposed randomness introduced by the sender. The first was that the trigrams were not selected in genuinely random fashion by the operators from the *Kenngruppenbuch*: there was a tendency, for example, to pick trigrams from the tops of pages. When an operator chose a trigram from the *Kenngruppenbuch*, he crossed it out, but since different operators used different copies of the book, as time went on knowledge of which trigrams had previously been used by the operators gave information about which trigrams were more likely to be used in the future by other operators. Turing devised an attack that exploited this (the "sampling of species" approach described earlier). This was the attack in use when Good came on the scene. It was effective, but it required a fair amount of data and became less and less useful as more and more trigrams were crossed out.

The other weakness was that humans are also very poor at selecting individual letters of the alphabet in a genuinely random way. As Good later related:

> I noticed on one night shift that about twenty messages were enough to identify which digraph table was in use, because the "haphazard" letters (*G* and *O* in the example) were not "flat-random." This discovery then provided the routine method for identifying the table. [Good, 2000, p. 109]

Curiously, looking back half a century later Good initially did not think "this discovery was of much importance" (Hinsley and Stripp, 1993, p. 156), but later that year Joan Murray, a colleague in Hut 8, wrote to tell him "your discovery ... was more valuable than you said, providing the regular method of quickly identifying which digraph table was then in use. Originally, it had been possible to determine that by the occurrence of popular starting positions, but that method had soon become unsatisfactory" since the popular trigrams had been used up by the German operators (Good, 2000, p. 110).

5.2.3 *"Depth-finding" and the repeat rate*. Turing's Banburismus was a hand method for finding the right-wheel in favorable situations, cutting down the number of wheels orders from 336 to 42, which made use of the Bombe practical. (More generally, it could sometimes identify both the right and middle wheels, although not with certainty.) It became fully operational in the summer of 1941, was in use for the next two years, and was a key element in the attack on the Naval Enigma during this time.

Banburismus required finding *depths*, sections of two or more Enigma messages enciphered using the same setting on the Enigma. Enigma depths were identified using a strengthened version of a classical cryptanalytic tool, the *repeat rate* (or *index of coincidence*, Friedman, 1922), the strengthened version initially due to Turing and later refined by Good. (Alexander jokingly referred to Good's improvement as "ROMSING," the Resources Of Modern Science.) How depth-finding was used in Banburismus is outlined below in Section 5.2.4; here, we discuss the narrower statistical task of finding depths.

Suppose two strings of letters are juxtaposed, one above the other. The *empirical repeat rate* is the fraction of pairs of letters, one above the other, that coincide. Consider, for example,

```
I M E T A T R A V E L L E R F R O M A
N A N T I Q U E L A N D W
W H E N A P R I L W I T H I T S S H O
W E R S S W E E T T H E D
```

The empirical repeat rate is $4/32 = 1/8$, since there is a pair of E's in both the 3rd and 27th places, a pair of A's in the 5th place and a pair of R's in the 7th place.

If letters are output in uniform random fashion, then the probability two letters in two juxtaposed strings match at a given place is $26(1/26)^2 = 1/26$. If however the letters consist of meaningful text (e.g., journalistic English or German naval communications) then they do not occur with equal frequency. If p_j, $1 \le j \le 26$, is the frequency of the jth letter in the language under consideration, then the *theoretical repeat rate*, the frequency of two letters matching, is

$$\rho = \sum_{j=1}^{26} p_j^2.$$

The repeat rate of English is about 1/15 and in the case of Naval Enigma messages, about 1/17.

Note *the repeat rate is invariant under permutations of the alphabet*. That is, if one had two messages encrypted by the Naval Enigma *using the same setting*, then the two plaintexts would not only exhibit a repeat rate of 1/17 in a sufficiently long stretch of text, but so would their corresponding ciphertexts. This is because if at a given step in the message the Enigma permutation is $k = \sigma(j)$ (the jth letter in the alphabet of the plaintext is encrypted by the kth letter in the ciphertext), and q_k is the resulting frequency of the kth letter of ciphertext, then the repeat rate for the ciphertext is

$$\rho_{\text{cipher}} = \sum_{k=1}^{26} q_k^2 = \sum_{j=1}^{26} q_{\sigma(j)^2} = \sum_{j=1}^{26} p_j^2 = \rho_{\text{plain}}.$$

This provides a test for when two messages—suitably aligned—are in depth: one slides one message relative to the other and at each offset counts the number of repeated letter pairs. When the offset messages are not in depth, the expected repeat rate is that of random text (1/26); but when they are properly offset so as to be in depth, then

the observed repeat rate will increase from 1/26 to to 1/17 provided the message is long enough.

Enigma messages were short, however, no more than 200 letters, too short for the crude classical method. The key to overcoming this difficulty was, as Turing realized, to recognize that in actual text one not only expects individual letters to match at a higher rate than flat random, but also multiletter strings. For example, if a common word like "the" occurs at the same place in a pair of aligned messages, then a three letter repeat will be observed, even though our oversimplified model of language assigns this a probability of only $(1/17)^3$, or about 1 in 4913. To exploit this, Turing developed a method of scoring multiletter repeats which was highly effective in identifying depths resulting from correctly offsetting two messages. For a discussion of the details of Turing's scoring system, see Turing (2012a, Section 2.3), Turing (2012b), Zabell (2012, pp. 202–207).[25]

5.2.4 *Banburismus*. The use of depths in Banburismus exploited the process by which the wheels in the machine turned. The ring of each wheel had a turnover notch (and in the case of the Naval Enigma, sometimes two).[26] For example, the ring for wheel 1 had a notch between letters Q and R. If wheel 1 was set to Q (its "turnover position") and there was a wheel to its left (i.e., it was in either the right or middle position), then when the next letter was encrypted both wheel 1 as well as the one to its left would simultaneously step. For example, if the initial message setting were ADL and the right wheel was wheel 1, then the successive settings of the machine as the letters were encrypted would be

ADL
ADM
ADN
ADO
ADP
ADQ
AER
AES
AET

and so on. The turnover process for the wheels can be summarized thus:

- The right wheel advanced each time.
- If either the right or middle wheel was in turnover position, then on the next step both it and the wheel to its left advanced by one.
- Because there was no wheel to the left of the left wheel, whether or not the left wheel was in turnover position had no effect on the turning of the wheels.

The turnover positions for wheels 1 to 5 were in different locations, corresponding to the letters Q, E, V, J, Z. (So after the turnover, the wheels were in positions R, F, W, K, A; the Bletchley Park mnemonic for this was "**R**oyal **F**lags **W**ave **K**ings **A**bove.") Because of this, finding the turnover pattern for a wheel meant one knew which wheel was in use. This was the weakness that Banburismus exploited.[27]

But how did Banburismus determine when a turnover occurred? Suppose a pair of intercepted messages have message settings AAA and AAH. After typing the first seven letters of the AAA message, the right wheel will advance to the H position and the setting will then be either ABH or AAH, depending on whether or not the middle wheel has advanced. This in turn will depend on which wheel is in the right-most position. If, for example, the right wheel is wheel 2 (whose turnover position is E), then a turnover will occur and the setting will be ABH; while if the right wheel were any of the others (1, 3, 4 or 5), then a turnover will not occur during the passage from A to H, and the setting will be AAH.

How can one distinguish between these two cases? The answer is simple: if the right wheel is not wheel 2, then the machine starting out at the AAA setting will be at the AAH setting seven letters later, so from this point on the setting of the first machine will be the same as the setting of the other machine (starting out at AAH) at the start of its message; *appropriately shifted, the two messages will be in depth*. On the other hand, because a turnover has to occur somewhere during a full circuit of the wheel, if one starts at AAH in the second machine and types 19 letters (so that the right wheel advances from H to A), a turnover will necessarily occur at some point in the transition from H to A, and the machine setting will be ABA, not AAA; the two messages will not be in depth. So all you had to do was look at two "slides."

For the Poles, this was a relatively straightforward matter, because—thanks to the German blunder of doubly encrypting the message setting using a Grundstellung—they were able to easily determine the message setting given enough messages (for a clear description of why, see Singh, 1999, Chapter 4). For Turing and Hut 8, though,

[25] Alexander (1945, p. 54) also briefly mentions a success by Good shortly before he left Hut 8 that although of limited operational value was technically impressive. ("The other coup was Good's success in breaking into a recalcitrant Limpet (Shark Offizier) day. All our previous successes in discovering Offizier steckers had been when the message set-up was known; by use of Hollerith machinery Good succeeded in making a break with over 20 possible alternative set-ups from which to choose. This was not an intrinsically important result but it was an achievement of some technical interest.")

[26] The three additional wheels employed by the Naval Enigma, wheels 6, 7 and 8, had two notches, at Z and M.

[27] There was some irony here: presumably the five different turnover positions were chosen under the mistaken impression this would improve the security of the device. In reality, this was a *Schlimmbesserung* (German for "bad improvement").

matters were much less straightforward: in the case of the Naval Enigma, after stripping off the bigram level of encryption of the message indicator to obtain a trigram, you still did not know the message setting: you still had to encrypt the trigram using the (still unknown) Grundstellung in order to obtain the message setting. If you only knew at this stage that the trigrams for the two messages were AAA and AAH then, although you knew the left and middle wheels had started out in the same position for both messages, all you knew about the right wheels was that they were in different positions, but not how far apart the third letters in the two message settings were. To determine this, you had to look at 50 slides, $\pm k$ for $1 \leq k \leq 25$, not just two, in order to determine when a turnover occurred. This was why a much more accurate scoring system was essential.

Given enough trigrams, this enabled one to work out the *relative* positions of the letters in the cipher alphabet. In our AAA, AAH example, for instance, if the two messages were in depth four letters into the AAA message then A and H, when enciphered using the Grundstellung, were four apart: either A and D, or B and E, or C and F and so on.

Finally, given enough pairs of messages in depth, one could decide between the twenty-six possible alignments by exploiting the special characteristics of the Enigma. One of these was basic, the reciprocal nature of the Enigma permutation: the same setting used to encrypt the message was also used to decrypt it. If A was sent to Q, then Q was sent to A, since if the electrical path followed by the current led from A to Q, then equally it led from Q to A. (Mathematically, the Enigma permutation was a product of two cycles.) So you could rule out any candidate alignment in which both A was encrypted by Q and Q by K, say. For similar reasons, the Enigma permutation also enjoyed a second important property: no letter could encrypt to itself (because otherwise the entering and exiting current would collide). So in our example above, you could immediately rule out the first alignment (A and H with A and D), because this would mean that A had been enciphered by A.

There are several accounts that can be consulted for the technical details of Banburismus. First, there are the two internal classified histories written immediately after the war, Alexander (1945, pp. 94–109) and Mahon (1945, pp. 16–20). Alexander was Turing's immediate successor as the head of Hut 8, and Mahon was in turn Alexander's successor, serving until the end of war; these two accounts are the primary source documents for the wartime history of the organization. But Budiansky (2006, Appendix B) may be a gentler place to start.

6. 1943: GOOD MOVES TO THE NEWMANRY

Good left Hut 8 in September 1943, to join Hut F (the Newmanry), the section at Bletchley Park working on a machine attack on Tunny, the codename for the SZ 40/42.

The Poles were familiar with the basic structure of the Enigma, in part because it was a modified version of a commercially available device, and thanks in part to an informant inside the Reichswehr's Cipher Bureau, Hans Thilo Schmidt (see Sebag-Montefiore, 2000, Kahn, 2010). The SZ 40/42, in contrast, had been designed by the German military, and so there was no corresponding commercial model to work from.

Despite this, in a remarkable tour-de-force of cryptanalytical skill the cryptanalysts of Bletchley Park (in particular the mathematician William Tutte, 1917–2002) were able to deduce *the entire structure of the machine* thanks to a single operational slip on the part of the German operators, who had once sent two long and almost identical messages at the same setting.[28]

6.1 The Design of Tunny

This was no mean feat, given the complexity of the device. Letters were represented in it using the then standard five impulse *Baudot code* (so, e.g., A = 00011, B = 11001, ...). The encryption used

- five "chi" wheels (employing regular motion)
- five "psi" wheels (employing irregular motion)
- two "mu" wheels (determining when irregular motion occurs)

(Here, "irregular" means that sometimes the wheels moved, and sometimes did not.)

Despite its apparent complexity, the process of encryption may be simply and schematically represented as

$$P \to P + \psi \to P + \psi + \chi = C$$

(P denoting "plaintext," C "ciphertext"). There were two stages of the attack, handled by two separate organizations:

- The *Newmanry*: striped off the χ layer of encryption using primarily statistical methods;
- The *Testery*: which then stripped off the ψ layer of encryption using primarily linguistic methods.

6.2 The Attack on Tunny

Despite its impressive appearance, Tunny suffered from a serious design flaw: when the five psi (irregularly moving) wheels did move, they did so simultaneously. As a result, a crafty combination of the output of a pair of wheels (in the initial stage of the attack, the ψ_1 and ψ_2 wheels) resulted in a biased stream of 0-1 bits.[29] This could be

[28] Lieutenant Colonel, later Brigadier John Tiltman, 1894–1982, played a particularly vital role at the start in exploiting the gaffe, using it to strip off the encryption, and determine the key the machine had generated in encrypting the two messages.

[29] The fact that this is the case is referred to by Friedrich Bauer as "Newman's theorem" (Copeland, 2006, pp. 393–395).

used as a test for the correct setting of the chi wheels for the given message. Because there were $41 \cdot 31 = 1271$ possible settings for the χ_1 and χ_2 wheels, if the correct setting for these was used to remove the chi layer of encryption from the crafty combination, then the resulting decrypted stream of bits would reveal the underlying *biased* sequence of 0s and 1s, whereas if one of the other 1270 (and incorrect settings) were used to decrypt the test stream, the resulting 0-1 stream would remain and appear as *unbiased*. Thus, the task of setting the first two chi wheels was converted into the purely statistical task of finding a single biased stream among a total of 1271.

This required a vast amount of computing, and for this the *Colossus* was constructed (see Copeland, 2006 for a detailed discussion of this device from a variety of viewpoints). It has been argued that in many ways the Colossus was the first programmable computer, not because it could store a program in memory, but because it could be rewired to perform other tasks. (This was in contrast with the Bombe, which was a special purpose device, constructed for the sole task of attacking the Enigma in a very specific way.)

Once a pair of chi wheels had been set in this way, other pairs of chi wheels could then be set by a similar process, eventually resulting in setting all five. This work was performed in the Newmanry and after the chi wheels had been set there, the message and settings were then sent on to the Testery to strip off the additional psi layer of encryption.

The intelligence supplied by reading Tunny intercepts was of enormous importance to the Allies, but qualitatively different from that of the Enigma. The Enigma often contributed immediate, actionable intelligence relating primarily to the Kriegsmarine and Luftwaffe, whereas Tunny supplied vital strategic information regarding the Army, throwing light "on the intentions and the condition of the German Army and on the thinking and planning of the whole of the German High Command" (Hinsley and Stripp, 1993, p. 144). Thus, the attack on Tunny combined technology, mathematics and linguistics, and rooms of complex equipment, but was most certainly worth it; see Reeds et al. (2015) and Zabell (2015).[30]

We know a great deal about the work in the Newmanry because in the months immediately after Germany surrendered to the Allies on May 7/8, 1945, Good wrote a long (500+ pages) classified report on the attack on the SZ 40/42, *The General Report on Tunny*.[31] During his time in the Newmanry, Good made many contributions to its successful operation, but because the *General Report on Tunny* does not in general attribute specific advances or techniques in the Newmanry to specific individuals, it is difficult to determine which of these were due to Good (although it seems likely that many of the more advanced statistical innovations in the Newmanry were in fact due to him).[32]

In the following, we single out two statistical issues of interest to which Good will unquestionably have contributed substantially.

6.2.1 *The need for Bayesian statistics*. Let us go back to the setting of a pair of wheels, say the χ_1 and χ_2 wheels. At this point, an interesting statistical issue arose. If a message was short enough, then there would not be enough information to identify the right setting from the competing 1270 other, incorrect settings. Conversely, if a message were very long, then the right setting would be obvious. But in the intermediate case, although there might be a clear candidate for the right setting, this might not be certain.

Recall the model is that the stream of 0-1 bits (resulting from a crafty combination of the five impulses) is biased if the chi wheels are correctly set (the probability of a 1 is $p \neq 1/2$), or unbiased if the wheels are not correctly set ($p = 1/2$). This was scored by just adding up the number of 1s. The *General Report on Tunny* describes the statistical approach initially used:

> Suppose the best score is 4σ (i.e., four standard deviations) without serious rivals. 4σ or better occurs at random once in 30, 000 experiments so it would be natural to imagine that the odds of the setting given are 30, 000 divided by 1271 or 23 : 1 on.
> In fact they are more like 3 : 1 on, (i.e., even after a factor has been set against all the other settings due to the existence of no serious rival), though the odds depend to a reasonable

[30] The importance of the section may be gauged in part by its size at the end of the war: 26 cryptographers, 28 engineers and 273 "Wrens" (WRNS, Women's Royal Naval Service, women who performed a variety of clerical functions); see Copeland (2006, p. 158).

[31] Newman left Bletchley Park in May 1945, immediately after the war in Europe ended, but directed Good, who remained on for several months, to prepare a comprehensive report on the activities of the section. A similar report on the operations of the Testery was prepared at the same time by Michie, who after completing the Testery report joined Good to work on the Newmanry report. (It appears not to be known when a third author, Geoffrey Timms, joined this effort.) Declassified in June 2000, the Newmanry report is now available in a reset, scholarly edition, Reeds et al. (2015). Curiously, the Testery report was only declassified much later ("opened" on August 3, 2018).

[32] It is only occasionally the current record permits us to point to specific instances. One of these is that in February 1944 Good and Michie "found a way of using Colossus to discover the Tunny wheel patterns" (Copeland, 2006, p. 76). Even after the war Good's experience with Colossus proved useful: at some point the NSA planned to construct a special purpose device for some unidentified purpose but decided not to do so when Good demonstrated that a surviving Colossus (most had been destroyed after the war at the direction of Churchill) could be reprogrammed to accomplish the same task (Copeland, 2006, p. 173).

> extent on the particular link and length of tape and d. [Reeds et al., 2015, p. 49]

This was important because, even given a promising candidate for the chi wheels, you did not know if you were right, because after stripping off the chi layer of encryption, you were still looking at $P + \psi$ (i.e., encrypted text). This could give rise to problems:

> In the very early days of the section, there was a tendency to continue with a message for some time if it gave a 4σ, since it was not believed that the odds could be much below 20 : 1 on.
> This was before the deciban had been brought over from Hut 8 [by Jack Good of course...].

What is going on here? The question is not just how unlikely it is for a random stream of 0s and 1s to result in a 4σ result, but *how strong the evidence is in favor of a particular setting*. Suppose:

- $H_j, 1 \le j \le 1271$ denote the 1271 possible settings of χ_1 and χ_2,
- X_j the score for H_j (the sum of 1s if H_j is used to strip off χ),
- p is the probability of getting a 1 in the 0-1 stream, $q = 1 - p$.

Then, assuming all 1271 possible settings are equally likely to occur (a not unreasonable assumption here), applying Bayes's theorem gives

$$P(H_i \mid X_1, \ldots, X_{1271}) = \frac{(p/q)^{X_i}}{\sum_{j=1}^{1271} (p/q)^{X_j}}.$$

(Note this depends not only on p, the bias, but also on N, the message length, via the X_j.)

Was the *General Report* right that the odds were "more like 3 : 1 on"? To take a specific example: suppose $N = 1080$ and $p = 0.55$ (very natural values in this situation). Then a "4σ" result would be a score of 606 or greater. Suppose H_i is a setting, and $X_i = 606$. The posterior probability for H_i will depend on all 1271 observed X_j scores, and is therefore random. Monte Carlo is our friend here: in an experiment involving 1,000 trials, on average the posterior odds in favor of H_i based on a score of $X_i = 606$, came out as 2.97—Good's comment is spot on.

6.2.2 *Wheel breaking*: *Rectangling and flagging*. The patterns of 0s and 1s on the wheels were changed regularly, and so part of the attack on Tunny required finding these new wheels configurations.[33] This was called *wheel-breaking*, and in turn involved *rectangling*, a method related to the modern iterated proportional fitting procedure.

To understand this, we need to say something about the craft combination. Let ψ_i^t denote the output of wheel ψ_i at time t. If the wheel does not step, then $\psi_i^t = \psi_i^{t+1}$, and their mod two sum (or difference) is

$$\Delta\psi_i^t := \psi_i^t + \psi_i^{t+1} = 0.$$

So, the irregular stepping introduced an inherent bias in the first differences in favor of 0s, and it was natural for the Newmanry to work at the level of these differences. (This introduces a small ambiguity in terms of finding ψ_i, since $\Delta\psi_i$ only determines ψ_i up to parity, depending on the value of its first component.)

The German designers of the machine were aware of this issue, and took steps to ensure that not only the output of ψ_i, but also $\Delta\psi_i$ would generate an apparently unbiased stream of 0s and 1s. What they overlooked, however, was that because when the psi wheels did step, they stepped in unison, their outputs were correlated and as a result the output of a sum of first differences

$$\begin{aligned}\Delta\psi_{i,j}^t &:= \Delta\psi_i^t + \Delta\psi_j^t \\ &= \psi_i^t + \psi_i^{t+1} + \psi_j^t + \psi_j^{t+1}\end{aligned}$$

would be biased. (It is easy to see this by a straightforward calculation.)

This was fatal, permitting one to not only set the wheels (as seen above), but to also "break" the wheels to determine new wheel patterns. As we have seen, the lengths of chi wheels 1 and 2 (χ_1 and χ_2) were 41 and 31, respectively. Because both chi wheels advanced each time a character was encrypted, and their lengths were relatively prime, the pair had an overall period of $1271 = 41 \cdot 31$: so after 1271 steps the two chi wheels would again be in the same position. The bias of the craft combination of $P + \psi$, the output of encrypting the plaintext by just the psi wheels, persisted if instead of looking at all outputs, you restricted yourself to the 1/1271 fraction of the time when χ_1 and χ_2 were in a common position such as i, j. If the two chi wheels had the same parity (either both 0 or both 1) at this point (so that their sum was 0), then when the layer of chi encryption was added to that of the psi, the bias in the crafty combination at the ψ level would persist; while if the two chi wheels had opposite parity (one 0 and one 1, so that the sum was 1), then the bias in the crafty combination would reverse. Exploiting this fact was the key to breaking the wheels.

Given a message, let a_{ij} record the *difference* in the number of 1's and 0's that occur in the crafty combination

[33] Initially, the patterns on the μ wheels were changed daily, those on the χ wheels monthly, and those on the ψ wheels first quarterly and later (October 1942) also monthly (Copeland, 2006, pp. 48 and 381). Beginning on August 1, 1944, however, the patterns on all three sets of wheels were changed every day, and so efficient wheel breaking became a matter of urgency.

when χ_1, χ_2 are in position i, j. (So, e.g., if the message length is $3 \times 1271 = 3813$, then every position i, j occurs exactly 3 times, and $-3 \le a_{ij} \le 3$.) The mission is to recover the two wheel configurations from (a_{ij}), and the attack exploits the fact that the bias in the crafty combination permits one to determine (with sufficient data) the parity of the sum of the two chi wheel contributions when χ_1 and χ_2 are in the i, j position.

Let $x = (x_i)$ and $y = (y_j)$ denote the configurations of the two wheels, where (for reasons that will become immediately apparent), we will now denote the sequence of 1s and 0s on each wheel by a sequence of $+1$s and -1s. Suppose that (x_i) represents our current guess regarding the configuration of the first wheel, and we are trying to guess (y_j), the configuration of the second wheel. Then each position i on the first wheel has a "vote" regarding the sign of y_j: namely $a_{ij}x_i$. (For example, if $a_{ij} < 0$ and $x_i < 0$, then $a_{ij}x_i > 0$; so the vote is for $y_j > 0$, consistent with the expectation that x_i and y_j have opposite parity since $a_{ij} < 0$.) It is then natural to add up the votes and decide on the basis of the sign of the sum:

$$y_j = \begin{cases} +1 & \sum_i a_{ij}x_i > 0; \\ 0 & \sum_i a_{ij}x_i = 0; \\ -1 & \sum_i a_{ij}x_i < 0. \end{cases}$$

Now we iterate: using the computed values of y_j, we in turn "push them through the rectangle," obtaining new values for the x_i, and then continue going back and forth in "ping-pong" fashion. It can be shown this iterative procedure eventually terminates in an unchanging (x_i), (y_j) configuration (as opposed to endlessly cycling through different configurations), but in general the limiting configuration will depend on the initial start or guess. So, it was important to start off with a good initial guess in order to arrive at the right answer; this was accomplished by a separate, preliminary technique called *flagging*. For rectangling and flagging, see Reeds et al. (2015, pp. 125–127, 135–136, 185–186, 215–216), Zabell (2015, p. xcv), and Copeland (2006, pp. 396–405).

Such iterative procedures were not unknown in statistical practice in 1940, but they were certainly uncommon.[34] Rectangling is very close to a variant of the power method in linear algebra used to find the eigenvectors and eigenvalues of a matrix, the variant using the singular value decomposition of a matrix to find the left and right singular vectors of the matrix. The variant is discussed by Good in his 1965 book *The Estimation of Probabilities: An Essay on Modern Bayesian Methods* (Good, 1965b), yet another instance of his wartime activities informing his postwar research and writings.

[34]For one example, see Deming and Stephan (1940).

7. I. J. GOOD, PHILOSOPHER OF SCIENCE

Good's advocacy of the Bayesian position was not limited to the statistical profession. He had a life-long interest in philosophy, and championed the use of subjective probability and the Bayesian approach in this literature, publishing extensively in it as well. This included twenty-two papers in the *British Journal for the Philosophy of Science*, seven papers in *Philosophy of Science*, three in *Synthese* and contributions to many conference proceedings, as well as speaking at numerous conferences. Twenty-three of these papers were later reprinted in a book, *Good Thinking* (1983), together with an introduction drawing the different threads together. It was, as Good wrote in the introduction (p. ix), "a book about applicable philosophy, and most of the articles contain all four of the ingredients philosophy, probability, statistics, and mathematics." For a later paper specifically crafted for a statistical audience, summarizing his work on the interconnections between statistics and the philosophy of science; see Good (1988), a paper in this journal.

One example of Good's approach was a paper on Rudolf Carnap's *principle of total evidence*. Carnap had constructed a formal logic of inductive inference, but as A. J. Ayer pointed out, it was unclear in Carnap's system why there was any value in acquiring new information. Proceeding from the Savage axioms for utilities, Good was able to show in a short paper (Good, 1967) that under natural assumptions the expected utility of acquiring cost-free information is monotonically increasing thanks to a simple minimax theorem. (This casting of a question in philosophy in quantitative terms, in order to exhibit a simple and often elegant answer to it, was characteristic of much of Good's work in the philosophy of inductive inference.)

But Good's advocacy of the Bayesian position had a strong practical streak to it, and was never dogmatic. He once wrote a short letter to the *American Statistician* (Good, 1971) titled "46656 varieties of Bayesians," the point being there was no unique Bayesian position, but a large number, depending on your take as to what kinds of judgements were possible, how precise they had to be, if you used utilities, whether or not you believed in the existence of physical ("objective") probabilities, and so on. Elsewhere (Good, 1992b) he wrote about a Bayes–non-Bayes compromise, reflecting his Bletchley Park mindset that in the end, if it worked and was efficient, you used it.

It must have been frustrating for him in the extreme during the pre-1976 period to be unable to rebut attacks on Bayesian methods as being merely theoretical in nature, lacking genuine practical application. (This author once witnessed such an attack during a conference in 1971, when Good was asked a hostile question of this nature by a member of the audience.)

8. EPILOGUE

By now it will be abundantly clear the decisive impact I. J. Good's "secret life"—his wartime experiences—played in his becoming a major advocate for the postwar Bayesian approach. Let's end by giving Jack Good the last word. Many years later, looking back on his time at Bletchley Park, he wrote (Good, 2006, p. 222):

> Most of the cryptanalysts in the Newmanry dispersed into the various universities and most of us achieved some measure of success in our unclassified work. But the success of our efforts during the war, and the feeling that we were helping substantially, and perhaps critically, to save much of the world (including Germany) from heinous tyranny, was a hard act to follow.

ACKNOWLEDGMENTS

My thanks to Steve Stigler for information about Good's proposed move to the University of Chicago, and to Skip Garibaldi, Al Hales, Jim Reeds, Glenn Shafer, Frode Weierud and an anonymous referee for helpful comments on a draft of the paper.

REFERENCES

ADAMS, J. F. (1985). Maxwell Herman Alexander Newman, 7 February 1897–22 February 1984. *Biogr. Mem. Fellows R. Soc.* **31** 437–452.

ALEXANDER, C. H. O'D. (1945). Cryptographic History of Work on German Naval ENIGMA. HW 25/1, UK National Archives, Kew, London.

BANKS, D. L. (1996). A conversation with I. J. Good. *Statist. Sci.* **11** 1–19. MR1437124 https://doi.org/10.1214/ss/1032209661

BARNARD, G. A. (1946). Sequential tests in industrial statistics. *Suppl. J. R. Stat. Soc.* **8** 1–21. MR0019287

BEESLY, P. (1977). *Very Special Intelligence*: *The Story of the Admiralty's Operational Intelligence Centre*, 1939-1945. Hamish Hamilton, London. Reprinted in 2000 by Greenhill Books, London, with a new Introduction by W. J. R. Gardner, and a new Afterword by Ralph Erskine, and an updated Bibliography by Gardner and Erskine.

BRITTON, J. L., ed. (1992) *Collected Works of A. M. Turing*. *Pure Mathematics* North-Holland, Amsterdam.

BUDIANSKY, S. (2000). *Battle of Wits*: *The Complete Story of Codebreaking in World War II*. The Free Press, New York.

BUDIANSKY, S. (2006). Colossus, codebreaking, and the digital age. In *Colossus*: *The Secrets of Bletchley Park's Codebreaking Computers* (B. J. Copeland, ed.) 52–63. Oxford Univ. Press, Oxford.

COPELAND, B. J., ed. (2006). *Colossus*: *The Secrets of Bletchley Park's Codebreaking Computers* Oxford Univ. Press, Oxford.

DEMING, W. E. and STEPHAN, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.* **11** 427–444. MR0003527 https://doi.org/10.1214/aoms/1177731829

FLOOD, J. L. (2011). Arthur Thomas Hatto: 1910–2010. In *Biographical Memoirs of Fellows of the British Academy* (R. Johnson, ed.) **10** 172–198. Oxford Univ. Press, Oxford.

FRIEDMAN, W. F. (1922). *The Index of Coincidence and Its Applications in Cryptography*. Riverbank Laboratories, Geneva, IL. Department of Ciphers, Publication 22. Second ed., 1935, War Department, Office of the Chief Signal Officer, Government Printing Office, Washington, DC. Declassified and approved for release by NSA on 01-06-2014.

GOOD, I. J. (1950). *Probability and the Weighing of Evidence*. Charles Griffin & Company, Limited, London.

GOOD, I. J. (1951). Random motion on a finite Abelian group. *Proc. Camb. Philos. Soc.* **47** 756–762. MR0044061 https://doi.org/10.1017/s0305004100027201

GOOD, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* **40** 237–264. MR0061330 https://doi.org/10.1093/biomet/40.3-4.237

GOOD, I. J. (1956). Some terminology and notation in information theory. *Proc. Inst. Elec. Engrs. C.* **103** 200–204. MR0076232

GOOD, I. J. (1958). The interaction algorithm and practical Fourier analysis. *J. Roy. Statist. Soc. Ser. B* **20** 361–372. MR0102888

GOOD, I. J. (1960). Weight of evidence, corroboration, explanatory power, information and the utility of experiments. *J. Roy. Statist. Soc. Ser. B* **22** 319–331. MR0116394

GOOD, I. J. (1961). Weight of evidence, causality and false-alarm probabilities. In *Information Theory* (*Symposium*, *London*, 1960) 125–136. Butterworths, Washington, DC. MR0131907

GOOD, I. J. (1962). Analogues of Poisson's summation formula. *Amer. Math. Monthly* **69** 259–266. MR0184006 https://doi.org/10.2307/2312938

GOOD, I. J. (1965a). A list of properties of Bayes-Turing factors. *NSA Tech. J.* **10** 1–6. Issue 2 for 1965. Approved for Release by NSA on Appeal on 03-9-2011, FOIA Case #58820.

GOOD, I. J. (1965b). *The Estimation of Probabilities. An Essay on Modern Bayesian Methods*. MIT Press, Cambridge, MA. MR0185724

GOOD, I. J. (1967). On the principle of total evidence. *British J. Philos. Sci.* **17** 319–321.

GOOD, I. J. (1968a). Corroboration, explanation, evolving probability, simplicity and a sharpened razor. *British J. Philos. Sci.* **19** 123–143. MR0242205 https://doi.org/10.1093/bjps/19.2.123

GOOD, I. J. (1969). Statistics of language. In *Encyclopedia of Linguistics*, *Information and Control* (A. R. Meetham, ed.) 567–581. Pergamon, London.

GOOD, I. J. (1971). 46656 varieties of Bayesians. *Amer. Statist.* **25** 62–63. Letter in the December 1971 issue.

GOOD, I. J. (1973). The joint probability generating function for run-lengths in regenerative binary Markov chains, with applications. *Ann. Statist.* **1** 933–939. MR0341612

GOOD, I. J. (1975). Explicativity, corroboration, and the relative odds of hypotheses. *Synthese* **30** 39–73.

GOOD, I. J. (1976). Early work on computers at Bletchley. *Natl. Phys. Lab. Rep. Com. Sci.* **82**. Updated and reprinted in *Annals of the History of Computing* **1** (1979), 38–48.

GOOD, I. J. (1979). Studies in the history of probability and statistics. XXXVII. A. M. Turing's statistical work in World War II. *Biometrika* **66** 393–396. MR0548210 https://doi.org/10.1093/biomet/66.2.393

GOOD, I. J. (1980). The contributions of Jeffreys to Bayesian statistics. In *Bayesian Analysis in Econometrics and Statistics*: *Essays in Honor of Harold Jeffreys* (A. Zellner, ed.) 21–34. North-Holland, Amsterdam.

GOOD, I. J. (1982). A report on a lecture by Tom Flowers on the design of Colossus. *Ann. Hist. Comput.* **4** 53–59.

GOOD, I. J. (1983). *Good Thinking*: *The Foundations of Probability and Its Applications*. Univ. Minnesota Press, Minneapolis, MN. MR0723501

GOOD, I. J. (1988). The interface between statistics and philosophy of science. *Statist. Sci.* **3** 386–412. MR0996411

GOOD, I. J. (1992a). Introductory remarks for the article in *Biometrika* 66 (1979). In *The Collected Works of A. M. Turing. Pure Mathematics* (J. L. Britton, ed.) 211–223. North-Holland, Amsterdam.

GOOD, I. J. (1992b). The Bayes/non-Bayes compromise: A brief review. *J. Amer. Statist. Assoc.* **87** 597–606. MR1185188

GOOD, I. J. (1993). Enigma and Fish. In *Codebreakers*: *The Inside Story of Bletchley Park* (F. H. Hinsley and A. Stripp, eds.) 149–166. Oxford Univ. Press, Oxford. Corrected paperback edition, 1994.

GOOD, I. J. (2000). Turing's anticipation of empirical Bayes in connection with the cryptanalysis of the Naval Enigma. *J. Stat. Comput. Simul.* **66** 101–111.

GOOD, I. J. (2006). From Hut 8 to the Newmanry. In *Colossus*: *The Secrets of Bletchley Park's Codebreaking Computers* (B. J. Copeland, ed.) 204–222. Oxford Univ. Press, Oxford.

GOOD, I. J. and TOULMIN, G. H. (1956). The number of new species, and the increase in population coverage, when a sample is increased. *Biometrika* **43** 45–63. MR0077039 https://doi.org/10.1093/biomet/43.1-2.45

GOOD, I. J. and TOULMIN, G. H. (1968). Coding theorems and weight of evidence. *IMA J. Appl. Math.* **4** 94–105.

GREENBERG, J. (2014). *Gordon Welchman*: *Bletchley Park's Architect of Ultra Intelligence*. Frontline Books, London.

HALL, H. S. and KNIGHT, S. R. (1891). *Higher Algebra*: *A Sequel to Elementary Algebra for Schools*, 4th ed. Macmillan, London.

HATTO, A. T. (1965). *The Nibelungenlied*. Penguin Books Inc, New York.

HILTON, P. J. and WYLIE, S. (1967). *Homology Theory*: *An Introduction to Algebraic Topology*. Cambridge Univ. Press, Cambridge.

HINSLEY, F. H. (1979–1990). *British Intelligence in the Second World War*. Her Majesty's Stationery's Office, London. Four volumes, Volume 3 issued in two parts.

HINSLEY, F. H. and STRIPP, A. (1993). *Codebreakers*: *The Inside Story of Bletchley Park*. Oxford Univ. Press, Oxford. Corrected paperback reprinting, 1994.

KAHN, D. (1967). *The Codebreakers*: *The Story of Secret Writing*. Macmillan, New York. Second edition, 1996, Macmillan, New York.

KAHN, D. (1991). *Seizing the Enigma*: *The Race to Break the German U-Boat Codes* 1939–1943. Houghton Mifflin, Boston, MA. Second edition, 1998, Barnes and Noble, New York; 2012 edition, Naval Institute Press, Annapolis.

KAHN, D. (2010). How I discovered World War II's greatest spy. *Cryptologia* **34** 12–21. Reprinted in *How I Discovered World War II's Greatest Spy, and Other Stories of Intelligence and Code*, CRC Press, Boca Raton, FL, 2014, Chapter 1.

MAHON, A. P. (1945). The History of Hut 8. HW 25/2, UK National Archives, Kew, London.

PLUMB, J. H. (1950). *England in the Eighteenth Century*. Pelican Books, London.

REEDS, J. A. et al. (2015). *Breaking Teleprinter Ciphers at Bletchley Park*. Wiley–IEEE Press, Piscataway, NJ.

REJEWSKI, M. (1981). How Polish mathematicians deciphered the Enigma. *Ann. Hist. Comput.* **3** 213–234. Includes discussion at the end by Cipher A. Deavours and I. J. Good.

SEBAG-MONTEFIORE, H. (2000). *Enigma*: *The Battle for the Code*. Wiley, Hoboken, NJ.

SINGH, S. (1999). *The Code Book*: *The Science of Secrecy from Ancient Egypt to Quantum Cryptography*. Doubleday, New York. Very readable, but to be consulted with caution; see the review by Jim Reeds in the *Notices of the American Mathematical Society* **47** (2000) 369–372.

TURING, S. (1959). *Alan M. Turing*. Heffer, Cambridge. MR0106802

TURING, A. M. (2012a). The applications of probability to cryptography. Unpublished paper, c. 1941, UK National Archives, HW 25/37, released 2012 on the centenary of Turing's birth.

TURING, A. M. (2012b). Paper on the statistics of repetitions. Unpublished paper, c. 1941, UK National Archives, HW 25/38, released 2012 on the centenary of Turing's birth.

TURING, D. (2021). *X*, *Y & Z*, 2nd ed. The History Press, Cheltenham, UK. First edition, 2018. The author, Sir John Dermot Turing, is a nephew of Alan Turing, and has—in addition to a very successful career in his own right—recently written several books on the history of cryptography in WWII..

WALD, A. (1945a). Sequential method of sampling for deciding between two courses of action. *J. Amer. Statist. Assoc.* **40** 277–306. MR0013276

WALD, A. (1945b). Sequential tests of statistical hypotheses. *Ann. Math. Stat.* **16** 117–186. MR0013275 https://doi.org/10.1214/aoms/1177731118

WALD, A. (1947). *Sequential Analysis*. Wiley, New York. MR0020764

WELCHMAN, G. (1982). *The Hut Six Story*: *Breaking the Enigma Codes*. McGraw-Hill, New York.

WINTERBOTHAM, F. W. (1974). *The Ultra Secret*. Weidenfeld and Nicolson, London.

ZABELL, S. (2012). Commentary on Alan M. Turing: The applications of probability to cryptography. *Cryptologia* **36** 191–214.

ZABELL, S. (2015). Statistics at Bletchley Park. In *Breaking Teleprinter Ciphers at Bletchley Park* (J. Reeds et al., eds.) lxxv–ci. Wiley–IEEE Press, Piscataway, NJ.