The Statistical Research Group, 1942–1945

W. ALLEN WALLIS*

The organization, work, membership, and influence of the Statistical Research Group, an Office of Scientific Research and Development activity at Columbia University during the Second World War, is described. A detailed account is given of the origins of sequential analysis in 1943, and a brief account of the origins at Stanford University in 1942 of the intensive courses in statistical quality control that were sponsored during the Second World War by the Office of Production Research and Development.

KEY WORDS: Statistical Research Group; Sequential analysis; History of statistics; Quality control.

1. INTRODUCTION

The Statistical Research Group (SRG) was based at Columbia University during the Second World War and supported by the Applied Mathematics Panel (AMP) of the National Defense Research Committee (NDRC), which was part of the Office of Scientific Research and Development (OSRD). An account of the SRG should be history, but this one is reminiscence, albeit fairly well-documented reminiscence. Exploring the history of one's professional field is often a mark of maturity. Reminiscing about it is usually a mark of senility. I have been able, however, to buttress my reminiscences by four bits of research:

1. From the senior staff of the SRG I have collected reminiscences. Collecting and summarizing the reminiscences of others presumably qualifies as research.
2. With the help of Churchill Eisenhart and Charles E. Dewing I have obtained a copy of the *Final Report* of SRG from the National Archives, where it and other SRG documents are kept in the Center for Polar and Scientific Archives.
3. In my own archives I was able to find—more precisely, Kenneth F. Wood was able to find for me—an account I wrote in 1950 of the origins and development of sequential analysis at SRG in 1943.
4. At the last minute I found a needle in a haystack—a brief document in a mound of musty files—that I will describe later.

Note that I refer to *the* Statistical Research Group, or SRG, not to the Statistical Research Group–Columbia, or SRG–C, as some benighted persons do. The situation is analogous to that of *The* Standard Oil Company. It was founded in 1870 in Cleveland by John D. Rockefeller; the other Standard Oil companies were formed later and differentiated themselves from one another and from *The* Standard Oil Company by adding the names of states. Similarly, at least one other organization was later named Statistical Research Group and to differentiate itself added the name of its university, Princeton.

2. QUALITY CONTROL COURSES

My reminiscing can begin appropriately with some activities early in 1942, soon after war was declared against the United States by Japan, Germany, Italy, and their allies. I was then on the faculty at Stanford. The atmosphere there that spring was satirized by a squib in the student paper saying, "It is rumored that in the outside world there is a war and a shortage of Coca-Cola." As one of several statisticians—Holbrook Working, Eugene Grant, Quinn McNemar, Harold Bacon—seeking some way that we at Stanford could contribute to the war effort, I wrote on April 17, 1942, to a friend in Washington, W. Edwards Deming of the Census Bureau, that

> Those of us teaching statistics in various departments here are trying to work out a curriculum adapted to the immediate statistical requirements of the war. It seems probable that a good many students with research training might by training in statistics become more useful for war than in their present work, or might increase their usefulness within their present fields.
>
> It is difficult for us to design such a program because we do not have a picture of the statistical work of the war. You are probably in as good a position as anyone to observe what kinds of statistical training are needed for a wide variety of purposes. It may not be possible to make any generalizations, but if it is we would try to put them to enough use to justify the time and trouble of advising us.
>
> Such problems as the allocation of the available training time among theoretical work, detailed instruction with machines and worksheets, and descriptive or analytical techniques—and what particular techniques to emphasize—are hard to answer in the abstract. What each of us does normally, of course,

* W. Allen Wallis is Chancellor of the University of Rochester, Rochester, NY 14627. He was Director of Research of the Statistical Research Group throughout its existence. He became a fellow of ASA in 1948, was editor of *JASA* from 1950 to 1959 and president of ASA in 1965, received the Chicago chapter's Outstanding Statistician Award in 1968, and delivered the Presidentially Invited Addresses at the ASA annual meetings in 1970 and 1978. This paper is a revision of the Presidentially Invited Address given at the ASA annual meeting in San Diego on August 14, 1978. Useful comments, suggestions, or other assistance have been provided to the author by Kenneth J. Arnold, George A. Barnard, Rollin F. Bennett, Albert H. Bowker, George E.P. Box, James V. Capua, Donald A. Darling, W. Edwards Deming, Charles E. Dewing, Churchill Eisenhart, Harold A. Freeman, Milton Friedman, W. Jackson Hall, Millard Hastay, William H. Kruskal, Robert K. Merton, Frederick Mosteller, Bruce S. Old, Edward Paulson, Mina Rees, Paul A. Samuelson, Herbert Solomon, George J. Stigler, Stephen M. Stigler, John W. Tukey, Jacob Wolfowitz, Kenneth F. Wood, and Holbrook Working.

June 1980, Volume 75, Number 370
Invited Papers Section

is to place emphasis according to his own personal judgment and experience, which unfortunately may not be abreast of war needs.

Deming responded on April 24 with four single-spaced pages on the letterhead of the Chief of Ordnance, War Department. After some explanatory background on the theme that, "The only useful function of a statistician is to make predictions, and thus to provide a basis for action," he wrote:

> Here is my idea. Time and materials are at a premium, and there is no time to be lost. There is no royal short-cut to producing a highly trained statistician, but I do firmly believe that the most important principles of application can be expounded in a very short time to engineers and others. I have done it, and have seen it done. You could accomplish a great deal by holding a school in the Shewhart methods some time in the near future. I would suggest a concentrated effort—a "short" course followed by a "long" course. The short course would be a two day session for executive and industrial people who want to find out some of the main principles and advantages of a statistical program in industry. It would be a sort of popularization, four lectures by noted industrial people who have seen statistical methods used, and can point out some of their advantages. The long course would extend over a period of weeks, or, if given evenings, over a longer period. It would be attended by the people who actually intend to use statistical methods on the job. In many cases they would be delegated by the men who had attended the short course.
>
> I would suggest that both courses be thrown open to engineers, inspectors, and industrial people with or without mathematical or statistical training. Naturally, any person who has had considerable statistical training would be in a position to get much more out of the course, but few would be in this fortunate position . . .
>
> Time is an important factor, and it is highly desirable that programs of this kind should be held in the near future. Late in the summer you could repeat the experiment, with improvements based on the experience of the first one.

On May 1, I was able to write Deming that, "Your letter arrived a few hours ago . . . [T]he specific suggestions struck home so well that Holbrook Working (Chairman of the University Committee on Statistics) has already talked with two or three of the key people and arranged a general meeting of everyone in statistics"; on May 21 the first letter about the course went to firms supplying Army ordnance in the western region; and the first course was given in July 1942 at Stanford.

The program that resulted from Deming's suggestion eventually made a major contribution to the war effort. Its aftermath, in fact, continues to make major contributions not only to the American economy but also to the Japanese economy.

Although I knew nothing about statistical quality control, having scarcely heard of Shewhart, Dodge-Romig, control charts, acceptance sampling, or any of the rest, I was scheduled to teach in the first course, together with Eugene L. Grant of the Stanford Engineering School, a man who knew a great deal about the subject and after the war published a book about it (Grant 1946). Just as I was beginning to wonder how to learn what I was supposed to teach, I was asked by Hildegarde Kneeland to head up an economic research unit in the Office of Price Administration (OPA) in Washington.

Holbrook Working, an outstanding statistician and economist at the Food Research Institute at Stanford, stepped into the quality control course. Before long Working headed a major national program that put on dozens of short, intensive courses in statistical quality control throughout the country and created the nucleus for the American Society for Quality Control (Working 1943, 1945). Working and his colleagues brought about a significant increase in the productivity of American industry, for which he has not yet received proper recognition.

In the meantime, I had begun assembling a staff for the OPA job, even before my own appointment was official; but I received a telegram from Warren Weaver of the Rockefeller Foundation, who was then head of the Fire Control Division of NDRC and later head of the Applied Mathematics Panel, asking if I would take charge of a statistics group that he was assembling with Harold Hotelling as principal investigator.

During the spring of 1942, I hurriedly completed and submitted for publication as much as possible of my work in progress, because it was uncertain when, if ever, I could return to it (Wallis 1942 a,b,c).

3. ORGANIZATION

Thus, on July 1, 1942, I found myself meeting with Weaver, Hotelling, and Jack Wolfowitz, the third charter member of SRG, on the 56th and 64th floors of Rockefeller Center—the first time I had been in an office that high in either altitude or importance. I had been a student of Hotelling's at Columbia in 1935–36, as Wolfowitz had been later. Weaver had been directed to Hotelling by Samuel Wilks, another former student of Hotelling's and something of a protegé of Weaver's, and Hotelling had directed Weaver to Wolfowitz and me. Wolfowitz and I had become friends in the fall of 1939 when both of us were attending late-afternoon lectures at Columbia by Abraham Wald, then a new member of the faculty. We had found a mutual bond not only in statistics but also in our intense concern about the war that had just begun in Europe and about the proper role in it for America. Wolfowitz was teaching on Staten Island while completing his doctorate, and I was at the National Bureau of Economic Research to do economic research on business cycles—which turned out to be statistical research with Geoffrey Moore on runs up and down (Wallis and Moore 1941a,b; Moore and Wallis 1943).

Altogether, 18 principals were to serve on the staff of SRG between July 1, 1942, and its dissolution on September 30, 1945. Only one person left before the end, and he was not a statistician. There were 3 principals in the summer of 1942, 8 by the summer of 1943, 16 by the summer of 1944, and 17 at the end in 1945. The average number for the 39 months was 10. Actually, I stayed on until March 31, 1946, editing and to a large extent writing two books based on SRG's work (SRG 1947, 1948).

In addition to the 18 principals, 60 others served on the staff at one time or another, a total of 78. The maximum at one time was about 50. There were statisticians, mathematicians, computers, typists, secretaries, a switchboard operator, an administrative assistant, a librarian, and a messenger. Lots of locks, safes, and security procedures, but no guards.

The administrative structure of SRG at the beginning consisted of me as Director of Research and Wolfowitz as Associate Director of Research. When Julian Bigelow and Milton Friedman joined, they too were called Associate Directors. Thereafter we ceased giving out that title. While we did not cancel any titles already assigned, in due course Friedman became effectively the Associate or Deputy Director and Albert Bowker the Assistant Director. Bowker's principal responsibility was to organize and manage the computing, which was done by about 30 young women, mostly mathematics graduates of Hunter or Vassar. Some of the basic statistical tables published in *Techniques of Statistical Analysis* (SRG 1948) were computed as backlog projects when there was slack in the computing load.

For offices we rented an apartment at 401 West 118th Street, one block from Columbia and overlooking Morningside Heights and Harlem. Eventually we and two other AMP organizations, the Applied Mathematics Group of Columbia and the strategic bombing section of the Statistical Research Group of Princeton, occupied most of the building. The building adjacent to ours on 118th Street was occupied some of the time by the Columbia section of the atomic bomb project.

SRG spent about $80,000 the first year, $150,000 the second, and $330,000 the third—a growth rate of 100 percent per year. We did not really budget but simply did whatever would advance the cause and submitted the bills. We pinched pennies, however: furniture and equipment, mostly second-hand; no expensive hotels or restaurants; no entertainment funds; salaries set on a no-gain, no-loss basis. One principal staff member still alleges resentfully that the administrative assistant told him to economize by writing his equations on both sides of the paper.

If the SRG merits more than a footnote in the history of American statistics, I suppose it is for one of the 31 reasons that can be composed by using one or more of five points, which I shall list and then discuss.

1. SRG was composed of what surely must be the most extraordinary group of statisticians ever organized, taking into account both number and quality.

2. Sequential analysis, one of the most powerful and seminal statistical ideas of the past third of a century, originated at SRG. Its theory was developed and methods of application were devised there.

3. A variety of useful materials produced at SRG, both theoretical and practical, while not comparable with sequential analysis, have nevertheless become established parts of statistics.

4. The experience of participating in SRG contributed definitively to the subsequent careers of a substantial number of persons who were to be leaders in statistics in the next three decades.

5. SRG was in many respects a model that has not been equaled of an effective statistical consulting group.

4. THE WORK

I discuss the fifth point first, because it gives an opportunity to describe SRG's work.

SRG's sole purpose was to serve the Army, Navy, Air Force, Marines, OSRD, and a few suppliers to these. Although ultimately SRG was accountable to its clients, this accountability was filtered, tempered, modulated, and stabilized by the AMP, which was directed by Warren Weaver, one of the most remarkable, admirable, brilliant, sagacious, and civilized human beings on the American scene in the past half-century. Thus, we were not at the whim of ignorant or arbitrary clients and were under no pressures to shade our findings to a client's prejudices or preferences.

Another advantage was that one of our principal clients, the Navy, had an effective liaison officer, the Coordinator of Research and Development, Admiral Julius A. Furer, who had an excellent staff, which included a remarkable number of young men who were to have outstanding and creative careers (Old 1961). Furer's staff introduced us to persons in the Navy who could benefit from our services and helped us immeasurably in innumerable ways in our relations with the Navy. As a result, SRG's most effective work was with the Navy.

Another advantage was that we were under steady pressure to deliver—there was a war on, as the saying went. Our work, however excellent, was in effect not delivered if it had no influence; so we had to understand the client's viewpoint and needs and be persuasive and accommodating.

Perhaps the strongest discipline resulted from the fact that when we made recommendations, frequently things happened. Fighter planes entered combat with their machine guns loaded according to Jack Wolfowitz's recommendations about mixing types of ammunition, and maybe the pilots came back or maybe they didn't. Navy planes launched rockets whose propellants had been accepted by Abe Girshick's sampling-inspection plans, and maybe the rockets exploded and destroyed our own planes and pilots or maybe they destroyed the target. During the Battle of the Bulge in December 1944, several high-ranking Army officers flew to Washington from the battle, spent a day discussing the best settings on proximity fuzes for air bursts of artillery shells against ground troops, and flew back to the battle to put into effect advice from, among others, Milton Friedman, whose earlier studies of the fuzes had given him extensive and accurate knowledge of the way the fuzes actually performed.[1] We were never wholly responsible for what happened. In fact, we seldom knew whether we were slightly responsible or even knew exactly what happened and to whom. But this kind of responsibility, although rarely spoken of, was always in the atmosphere and exerted a powerful, pervasive, and unremitting pressure.

Sometimes SRG's work evolved by one problem leading to a related problem and sometimes by a technique leading to another application of the technique to an unrelated problem. This is brought out in the SRG *Final Report* (SRG 1945a), which I wrote in September 1945:

> The first project assigned SRG was to evaluate the comparative effectiveness of four 20 millimeter guns on the one hand and eight 50 caliber guns on the other as the armament of a fighter aircraft. This assignment had been preceded by two months of study of air warfare analysis and plane-to-plane fire control. This first assignment involved a study of the geometry and tactics of aerial combat, the probability of hitting, and the vulnerability of aircraft.
>
> The work on the geometry of aerial combat, particularly pursuit curves, led to the assignment to SRG of a problem concerning an automatic dive-bomb sight whose functioning depended upon pursuit curves. This study led to adding a man [Bigelow], and later another [Savage], to the staff to handle such problems. The presence of these men, together with the experience accumulated, led to the assignment of certain problems involving homing torpedoes and in turn to problems involving airborne homing weapons and to work on guided missiles. The presence of one of the men obtained initially for the pursuit curve work [Bigelow] resulted in assignment of a problem of solving linear simultaneous equations electrically. In the work on the geometry of aerial combat some attention was given to the probability distribution of random contacts, and later a study was assigned dealing with random contacts in submarine searches.
>
> The work on the probability of hitting aircraft led to an assignment on the probability of anti-aircraft hits on a directly approaching bomber and this led to a problem concerning the use of shrapnel against directly approaching planes and in turn to a study of the risk to our planes from anti-aircraft. Work on the probability of hitting also led to studies of the dispersion of aircraft machine guns, to work on aircraft turret sights, and to an evaluation of different guns for anti-aircraft use. The work on probability of hitting and on aircraft vulnerability were also the bases for studies of the effectiveness of proximity fuzed rockets in air-to-air combat.
>
> The problem of aircraft vulnerability led SRG to devise a technique for determining vulnerability from damage survived by our own planes, it led to an assignment concerning B-29 vulnerability, and it led to an assignment on the area around standard bombs dangerous to aircraft.
>
> The work on gyroscopic lead computing sights led to an analysis of the use of such sights and to an assignment on Navy anti-aircraft directors. It also led to an analysis of the risk in extending a bombing run and to an analysis of curved flight predictors. It, together with the work on the probability of hitting, led to SRG's participating in studies of trial fire procedures for anti-aircraft and of the value of measuring the muzzle velocity of anti-aircraft fire, and in the preparation of a bibliography on anti-aircraft effectiveness.
>
> The next problem assigned to SRG that was not related to those growing out of the plane-to-plane studies was one concerning the acceptance testing of bomb sights. Although this study did not itself lead to any others, in the course of time a number of similar studies were assigned. One of these involved the comparison of two percentages, which led SRG to consider the problem of analyzing data sequentially. This in turn led to consultative work with the Quartermaster Corps Inspection Service. Another study involved acceptance testing procedures for Naval ordnance and another involved the formulation of specifications and acceptance procedures for rocket propellants. The work on sequential analysis, Naval ordnance, and rocket propellants led to the preparation of standard sampling inspection plans for the Navy.
>
> Early in 1943 SRG was asked to determine lead angles for aircraft torpedo attacks on maneuvering warships. This involved studying torpedo tactics, torpedo performance, the maneuvering characteristics of warships, and combat experience. In order to determine the characteristics of enemy ships, a method was developed and tested (in collaboration with the Applied Mathematics Groups at Columbia University and New York University and the Photo Interpretation Center, Anacostia) for measuring speed and turning radius from the wake shown in aerial photographs. The work on aircraft torpedo tactics led to an assignment on the spread to be used in torpedo salvos fired from surface craft. Another outgrowth of the torpedo work was an assignment on torpedos following zigzagged courses.
>
> A number of studies were assigned to SRG solely because of their dependence on statistical techniques rather than because their subject matter was similar to other SRG work. One of these involved the preparation of artificial test data such as might result from tracking with erratic variations. Another involved reviewing a manual on probability prepared for one of the NDRC divisions. Another, begun in the Applied Mathematics Group, Columbia University, but completed by an SRG member [Eisenhart] as consultant to the Navy, involved a statistical study of the vulnerability of merchant vessels to aircraft torpedoes and bombs. Statistical studies were also made of stereoscopic range finders, food storage data, high temperature alloys, the diffusion of mustard gas, and clothing tests.
>
> A substantial part of the work of SRG consisted of consultation with and informal assistance to the Army, Navy, or NDRC. Much of this consultation was related to the subject matter of AMP Studies assigned to SRG, but a great deal of it grew out of general liaison arrangements and was presented to SRG because of the techniques involved. A large part of this consultative work was done in oral conversations or conferences, although a certain amount is reflected in SRG documents.

To add detail to that general description of SRG's work, I list the titles, classifications, and authors of 13 of the 572 substantive reports, memoranda, and letters that were produced at SRG. I have chosen the first, every 50th one thereafter, and the last, which is numbered 561 because 11 missed serial numbers. (The letters in parentheses after the titles indicate security classifications at the time: Secret, Confidential, Restricted, or Open. The classifications usually reflected the general project to which a report related, not necessarily the content of the specific report.)

1 22 Aug 42	Outline of Cunningham Papers on Mathematical Theory of Air Warfare (C)	Wallis-Wolfowitz
51 18 Jun 43	Conference at Navy Department, 17 June 43 (C)	Friedman
101 28 Oct 43	Minimum Front Necessary for a Torpedo Salvo to Include All Possible Target Maneuvers (S)	Wolfowitz

[1] Proximity fuzes were developed by the OSRD for the Navy. For security reasons they were not made available to the Army until December 16, 1944, during the Battle of the Bulge. Thus, the Army had had no experience in their use. The Navy feared that as soon as the Army had the fuzes, some would fall into enemy hands; this, indeed, did happen during the Battle of the Bulge, but fortunately no harm resulted. (Baxter 1946, pp. 233–236)

151 2 Feb 44	Uses for Aircraft Vulnerability Figures (C)	Wallis
201 25 Apr 44	Number of Trials Necessary for a Given Degree of Assurance That If a New Design Is Said to Produce Fewer Failures Than an Old, Its True Percentage of Failures Is Not More Than Half That of the Old (O)	Hastay
251 12 Jun 44	SRG Work for Division 9 at Bushnell, Florida (S)	Wallis
301 10 Aug 44	Visit to Bureau of Ships, Navy Dept., 3 August 1944 (C)	Friedman
351 7 Nov 44	Sequential Analysis of Double Dichotomies for Grouped Data When the Groups Are of Different Sizes for the Two Methods Being Compared (R)	Paulson
401 9 Jan 45	Surface Damage Criteria for Rocket Propellants (R)	Bigelow
451 1 Mar 45	Flak Risk to Controlled-Missile Bomber (C)	Savage Mosteller Williams
501 19 May 45	Evaluation of Test Data for M-9 Director With and Without Second Derivative Prediction (C)	Wallis
551 28 Aug 45	Relative Effectiveness of Caliber .50, Caliber .60, and 20 mm Guns As Armament for Multiple Anti-Aircraft Machine Gun Turrets (C)	Friedman
561 Oct 45	Pursuit Performance of Pelican (S)	Savage

Four books resulted from SRG's work: *Sequential Analysis* (Wald 1947), *Sequential Analysis of Statistical Data: Applications* (SRG 1945b), *Techniques of Statistical Analysis* (SRG 1947), and *Sampling Inspection* (SRG 1948). All have proved influential—Wald's far more than the others, of course.

5. MEMBERSHIP

Now I turn to the first of the characteristics of SRG that I presume to be of interest, namely, its membership. In order of duration of service, the 18 principal members of the staff were

Allen Wallis, 45 months
Harold Hotelling, 39 months
Jacob Wolfowitz, 39 months
Edward Paulson, 35 months
Julian Bigelow, 31 months
Milton Friedman, 31 months
Abraham Wald, 30 months
Albert Bowker, 27 months
Harold Freeman, 21 months
Rollin Bennett, 20 months
Jimmie Savage, 19 months
Kenneth Arnold, 18 months
Millard Hastay, 18 months
Abraham Girshick, 17 months
Frederick Mosteller, 12 months[2]
Churchill Eisenhart, 11 months
Herbert Solomon, 11 months
George Stigler, 10 months

[2] Mosteller was not formally a member of SRG, since his salary came from another group, a New York branch under John D. Williams of the Statistical Research Group of Princeton University, headed by Samuel S. Wilks. Mosteller worked so closely and extensively with SRG, including being coauthor of two of its books (SRG 1947, 1948) and coeditor of one (SRG 1948), that it will astonish members of SRG, as it did me on reviewing the *Final Report*, to learn that fiscally he was not a member. The 12 months listed is my estimate, to the nearest year, of his full-time equivalence in SRG.

Three of these members were part-time: Hotelling, Wald, and Freeman. In addition, Walter Bartky and Norbert Wiener served briefly as consultants, and Alan Treloar was employed by SRG but assigned full-time to an Army laboratory.

Eight SRG members have been president of the Institute of Mathematical Statistics: Bowker, Girshick, Hotelling, Mosteller, Savage, Solomon, Wald, and Wolfowitz. Four have been president of the American Statistical Association: Bowker, Eisenhart, Mosteller, and Wallis. Two have been president of the American Economic Association: Friedman and Stigler. One is the 1980 president of the American Association for the Advancement of Science: Mosteller. One has received a Nobel Prize: Friedman. At least nine have been chairmen of university departments of statistics: Arnold, Bowker, Girshick, Hotelling, Mosteller, Savage, Solomon, Wald, and Wallis. Two have become heads of major universities: Bowker (of two) and Wallis. I do not know the number of honorary degrees; Guggenheim, Fulbright, and similar fellowships; and special lecturerships, prizes, awards, and other honors; but it must be large.

Of the staff not mentioned in the preceding paragraph, Bennett is an official of one of the largest life insurance companies, Bigelow is a permanent fellow at the Institute for Advanced Study, Freeman is a professor at Massachusetts Institute of Technology, Hastay is a professor at Washington State University, and Paulson is a professor at Queens College (New York).

Members of SRG were in frequent close touch with other statisticians and mathematicians who have had distinguished careers, including Eli Bromberg, Richard Courant, Kurt Friederichs, Saunders MacLane, Jerzy Neyman, Mina Rees, James Stoker, John Tukey, Hasler Whitney, Samuel Wilks, and John Williams. All these had strong influences on members of SRG.

It is perhaps interesting to consider how the major appointments to SRG came about. As I have said, Wilks was responsible for Hotelling and Hotelling for Wolfowitz and me. Hotelling directed my attention to Paulson, a student of his. Bigelow and Savage were suggested to me by Weaver, under whom they were already working, Bigelow with Norbert Wiener at MIT and Savage with R.G.D. Richardson at Brown. Bowker was suggested by Bigelow, and Freeman and Arnold by Bowker. Eisenhart, Mosteller, and Solomon came, at my request, from other applied mathematics panel groups. Hastay had been a graduate student of mine at Stanford in 1938–39. I knew Wald through his lectures at Columbia, and in addition I had become a personal friend in 1939–40. Other SRG members were colleagues from my graduate school days at Chicago and Columbia: Stigler since 1933, Friedman since 1934, and Bennett and Girshick since 1935. Eight members—Bennett, Friedman, Girshick, Paulson, Solomon, Wald, Wallis, and Wolfowitz—had studied or worked with Hotelling before SRG days. Recruiting was by the old-boy network pure and unabashed: no advertising, no competitive examinations, and no attention to race, sex, age, physical handicap, or apparent nationality or surname.[3]

6. SEQUENTIAL ANALYSIS

Turning now to the origins of sequential analysis: The best I can do in 1980, 37 years after sequential analysis was created in April 1943, is to quote a long letter that I dictated to Warren Weaver in March 1950. In 1950 I was 30 years closer to the events than I am now, and, furthermore, I had the use of a 37 year old's memory—something that now I can scarcely recall ever having had.

Dear Warren:

This is in answer to your letter of January 18th asking about the history of sequential analysis. You suggest three headings which I will follow: origin, relation to previous work, and applications; and I'm adding a fourth, miscellaneous.

A. *Origin*

Late in 1942 or early in 1943 you assigned us the task of evaluating an approximation developed by [Navy] Captain Garret L. Schuyler that was supposed to simplify a complicated British formula for calculating the probability of a hit by anti-aircraft fire on a directly approaching dive bomber. Schuyler's approximation was no good. Ed Paulson worked on the problem for us and was able to give rather simple formulas bounding the correct probability. It fell to me to report this to Schuyler during one of my trips to the Navy Department, and before I left you wished me good luck, saying that Schuyler was one of the orneriest characters you had encountered in the Army and Navy; in fact, I believe you said he was the only downright gruff and disagreeable high officer you had to that time encountered. I decided that one of the advantages of not being in uniform was not having to take the lip of any cheeky admirals or generals, so I made my approach brief and direct. I reminded Schuyler of his query and stated that we were now in a position to tell him how good his formula was. He barked, "How good is it?" and I replied "No good at all." He snapped, "What's wrong with it?" This had been anticipated and my reply planned: "What's right with it? We were unable to see any sense in it at all." From then on he and I were good friends, and I still hear from him from time to time.[4] That afternoon we had a long and rambling conversation in the course of which he got around to statistics. He brought up the problem of the necessary sample size for comparing two percentages and gave me a memorandum prepared for him on that subject by a certain John von Neumann. I had never before heard of this von Neumann fellow, but the memorandum was obviously pretty smart for a man who apparently knew little about statistics.

This led Paulson and me to work up the material on comparing two proportions which is now presented in Chapter 7 of TECHNIQUES OF STATISTICAL ANALYSIS. When I presented this result to Schuyler, he was impressed by the largeness of the samples required for the degree of precision and certainty that seemed to him desirable in ordnance testing. Some of these samples ran to many thousands of rounds. He said that when such a test program is set up at Dahlgren it may prove wasteful. If a wise and seasoned ordnance expert like Schuyler were on the premises, he would see after the first few thousand or even few hundred [rounds] that the experiment need not be completed, either because the new method is obviously inferior or because it is obviously superior beyond what was hoped for. He said that you cannot give any leeway to Dahlgren personnel, whom he seemed to think often lack judgment and experience, but he thought it would be nice if there were some mechanical rule which could be specified in advance stating the conditions under which the experiment might be terminated earlier than planned. I told him that I was vaguely aware of the existence of some kind of double sampling procedure used in industrial sampling inspection, though I knew nothing about it personally. Schuyler said that he was familiar with that. Incidentally, Schuyler said that the suggestion of an objective criterion for stopping early had been made by an aviator from Virginia named Gwathmey. Later I talked with Gwathmey, but he seemed to have no idea what it was all about, and no recollection of any such suggestion by him or anyone else.

At any rate, several days after I returned to New York I got to thinking about Schuyler's comment. Just before I had gone to SRG, on 1 July 1942, I had completed a paper on "Compounding Probabilities From Independent Significance Tests," which was published in ECONOMETRICA in 1942. This test procedure, invented like almost all good ideas in statistics by R.A. Fisher, is based on multiplying together the probabilities from the independents tests, the probabilities being calculated on the assumption that the null-hypothesis is correct. One then asks for the probability distribution of the product. This background led me to fool around with Schuyler's idea by multiplying together the successive probabilities under the null-hypothesis of the observations as they arise. I got quite interested in this, and worked out a few numerical examples trying to discover some principle for setting a critical level for the product as a function of sample size.

This was early in 1943, after Milton Friedman had joined SRG but before he had been able to move his family to New York. He was commuting from Washington to New York for two or three days each week. He and I regularly had lunch together, and one day I brought up Schuyler's suggestion. We discussed it at some length, and came to realize that some economy in sampling can be achieved merely by applying an ordinary single-sampling test sequentially. That is, it may become impossible for the full sample to lead to rejection, or for it to lead to acceptance, in which case there is no sense in completing the full sample. The fact that a test designed for its optimum properties with a sample of predetermined size could be still better if that sample size were made variable naturally suggested that it might pay to design a test in order to capitalize on this sequential feature; that is, it might pay to use a test which would not be as efficient as the classical tests if a sample of exactly N were to be taken, but which would more than offset this disadvantage by providing a good chance of terminating early when used sequentially. Milton explored this idea on the train back to Washington one day, and cooked up a rather pretty but simple example involving Student's t-test.

When Milton returned to New York we spent a great deal of time at lunches over this matter, and we began to get so interested that our conversations carried over into the office; and there began to be a noticeable interference with the work we were supposed to be turning out. We finally decided to bring in someone more expert in mathematical statistics than we. This decision was made after rather careful consideration. I recall talking it over with Milton walking down Morningside Drive from the office to our apartment. He said that it was not unlikely, in his opinion, that the idea would prove a bigger one than either of us would hit on again in a lifetime. We also discussed our prospects for being able to work it out ourselves. Milton was pretty confident of our (his?) ability to squeeze the juice out of the idea, but I had doubts and felt that it might go beyond our (my!) depth mathematically. We also discussed the fact that if we gave the idea away, we could never expect much credit, and would have to take our chances on receiving any at all. We definitely decided that even if the credit situation turned out in a way that disappointed us, there would be nothing to do about it,

[3] In spite of this "negative action" recruiting, SRG included a high proportion of women, two blacks, and two persons with severe physical handicaps.

[4] Garret Lansing Schuyler, U.S. Navy Rear Admiral (retired), died at his home in Washington, D.C., at the age of 93, on January 4, 1977. He had won a Navy Cross during the First World War (Eisenhart 1977).

publicly or privately.[5] At the end of the conversation, which ended at our apartment after an hour or two, we decided to turn the whole thing over to Wolfowitz.

The next day we talked with Jack but were totally unable to arouse his interest. Later, either that day or perhaps another day, we talked with him individually. We were still unable to arouse any interest, and there even seemed to be something distasteful about the idea of people so ignorant of mathematics as Milton and I venturing to meddle with such sacred ideas as those of most powerful statistics, etc. No doubt this antipathy was strengthened by our calling the new tests "supercolossal" on the grounds that they are more powerful than "most powerful" tests. Finally, we made a big try with Jack at a lunch at Butler Hall, which ran from twelve until after three in the afternoon. We could not get from him any criticism of our idea that made sense to us, and yet we could not make a dent on him. We finally asked what he would think of our talking with Wald about the problem; he was somewhat cool to this proposal, apparently considering Wald's time too valuable to be wasted.

We got Wald over the next morning and explained the idea to him. Wolfowitz was present, I believe, and I know that Hotelling and Paulson were. In discussing the idea with Paulson, he had said that Hotelling had once mentioned something along the same lines, pointing out that if one is to toss a coin one hundred times to test its bias, and has established the criterion that a 60–40 division or worse will cast doubt on the coin, then there is no use continuing to 100 if earlier 60 heads or 60 tails are reached. Hotelling, according to Paulson, had further suggested that he had some notions about how to go about looking for a solution to the problem of stopping such sampling earlier, saying that the ideas were related to those of heat flow in the presence of an absorptive barrier. Hotelling contributed the term "sequential" to describe the method of analysis, though I believe that this was at a later conference when Wald presented his results. We presented the problem to Wald in general terms for its basic theoretical interest, and as a practical example cited the problem of comparing two fire control devices with a hit or miss classification of each round.

At this first meeting Wald was not enthusiastic and was completely noncommital. I am inclined to attribute this to the fact that Wolfowitz had spoken to him the preceding evening, after our appointment had been arranged.

The next day Wald phoned that he had thought some about our idea and was prepared to admit that there was sense in it. That is, he admitted that our idea was logical and worth investigating. He added, however, that he thought nothing would come of it; his hunch was that tests of a sequential nature might exist but would be found less powerful than existing tests. On the second day, however, he phoned that he had found that such tests do exist and are more powerful, and furthermore he could tell us how to make them. He came over to the office and outlined his sequential probability ratio to us. This is the ratio of the probability under the null-hypothesis, with which I had been puttering around, to the probability under the alternative hypothesis—or rather, the reciprocal of this ratio. He found the critical levels by an inverse probability argument, showing that the same critical levels result no matter what assumption is made about the a priori distribution. After some further work on the general method, he proposed the procedure for testing double dichotomies sequentially that is presented in Section 3 of SEQUENTIAL ANALYSIS OF STATISTICAL DATA: APPLICATIONS.

It was quite a while before Wald worked out all of the theoretical justifications of his initial results. Actually it was a very short time in view of the amount of work he achieved.

B. *Relation to Previous Work*

While it later developed that there had been previous work related to sequential analysis, you can see from the foregoing account that Wald's development did not actually grow out of preceding work. Though you will note that at my first conversation with Schuyler, I had mentioned Dodge's double sampling developed at the Bell Laboratories, it was not until applications of sequential analysis had progressed rather far that the relations to double sampling received further attention. After Deming heard about sequential analysis at that dinner you and I had with him at the Cosmos Club, he wrote me that sequential analysis seemed to be mathematically similar to some work by Lord Rayleigh on the random walk problem in physics. In talking with Walter Bartky about Wald's work when he and I were both consultants for Francis Bitter, he told me that he had done similar work some years before at Western Electric. Apparently this had been refused for publication earlier, but it was now speeded through and appeared in the *Annals of Mathematical Statistics*. Later we heard of British work along the same lines. In the interest of international amity we used to refer to this as if it were somehow on a par with Wald's sequential analysis, but my impression is that in point of fact it was not really comparable at all except for some vague similarity of purpose.[6] It has been pointed out by Polya, and no doubt others, that sequential analysis is very similar to the seventeenth or eighteenth century (which is it?—you will know) problem of the gambler's ruin.

The real relationship of Wald's work to previous work, however, is to the Neyman-Pearson theory of testing statistical hypotheses. In a sense, Wald's development of the sequential probability ratio test is a straight-forward application of the principles of testing hypotheses developed by Neyman and Pearson. They showed that the probability ratio (they call it the likelihood ratio) is the appropriate basis for a test of significance. Wald simply sequentialized it.

C. *Applications*

On the whole SRG did not have much direct influence on or responsibility for applications, but I suppose most of the major applications were indirectly influenced by our work. While Wald was still preparing his monograph on the theory, we started to work on a book on applications. We were understaffed at that time, and other work had higher priority. Finally, we arranged with Harold Freeman of MIT to take on the job as a special assignment. He wrote the first version of SEQUENTIAL ANALYSIS OF STATISTICAL DATA: APPLICATIONS. While he was working on this, he was called in by the Boston office of the Quartermaster Corps for advice on acceptance inspection, and it seemed to him that sequential analysis was eminently suitable for their problem. He therefore gave a series of lectures to the staff, including the top officer, a Colonel Rogow who had come to the Quartermaster Corps from Sears Roebuck and who after the war became president of Eversharp. Rogow was apparently extremely energetic and dynamic (I never actually met him myself), and shortly thereafter was put in charge of the entire Inspection Service of the Quartermaster Corps. At that time the QMIS was in pretty bad shape all around, and Rogow shook it up thoroughly. They made innumerable changes, one being the introduction of sequential analysis. Freeman worked with them closely at this time, in developing their sampling procedures and in lecturing to their staff. It was in these initial training courses that the demonstration involving a field cook stove (gasoline burning) was put on. I have forgotten the details of this, except that the sequential method reached the correct decision much more often than did the method previously in use and with much less inspection. Rogow encountered considerable opposition in introducing sequential analysis, particularly from the Army Ordnance Department which had thought that the QMC would use its double sampling plans and had sent Dodge on a tour of QMC depots to explain the system. Rogow caused a good deal of resent-

[5] As it turned out, we were satisfied with the acknowledgments expressed by Wald (Wald 1945a, p. 120; 1947, p. 2).

[6] This is not, of course, the last word in relation to the British contribution to sequential analysis, but it does describe the attitude that those of us at SRG had at the time, which was essentially to be so absorbed in our work as to pay only passing attention to similar work by others. Communication was difficult and we had very little information about the work going on in England except that some was going on.

ment by his high-handed and extremely forceful methods, but he achieved an amazingly quick revolution in the QMIS. Actually, sequential analysis deserves only a small part of the credit for the total improvement achieved. Much of the improvement was due simply to better methods of inspecting given items, better methods of reporting, etc. Nevertheless, sequential analysis became the opposite of a scapegoat: something to which all the credit could be attached, so that it would not be necessary to say that they were simply doing what could have been done twenty years sooner.

The Navy interests in sequential analysis came first from John Curtiss. I gave him Wald's basic formulas at lunch one day at the Water Gate Inn. He was quick to perceive the usefulness of sequential analysis in sampling-inspection work. Indeed, my impression is that this may have preceded Freeman's work. Curtiss was the first to suggest to me that the decision criteria be transformed from levels of the likelihood ratio to levels for the actual count of defectives, to be shown as a function of sample size. This was an adaptation of the standard tables of acceptance and rejection numbers used by Army Ordnance and taken by them from the Bell Laboratories. At SRG we later thought of the graphical presentation of these acceptance and rejection numbers.

As far as I know, sequential analysis was never actually used on the rocket propellant study. Rocket propellant production was under Army Ordnance, even when done for the Navy, and they would have nothing to do with sequential analysis. Some of the big savings there were made by increasing lot size, since the product was essentially homogeneous, and by using variables instead of attributes as a basis of decision.

D. *Miscellaneous Comments*

So much for your three headings. A few additional comments occur to me that might have been put in earlier. One of the big theoretical developments after Wald's initial results was the discovery of methods of deriving power functions or operating characteristic curves. This came about in the following manner. In most of the simple, standard tests four parameters are specified to derive the test, usually described as specifications of good and of bad quality (there must be a gap between them) and of the two risks of error. When the test is constructed, however, it turns out to involve two parallel straight lines. Thus, you put in four parameters and you get out three. Friedman was the first to point this out, and he naturally asked, "Where is the vanished parameter?" More specifically he pointed out that there must be many combinations of the original four parameters which lead to the same final set of three. Now, finding all of the sets of four that lead to a given set of three amounts to finding what the probabilities of acceptance are for all possible lot qualities, which is what is known as the operating characteristic curve. Its complement is called the power function. Friedman found this by a rather ingenious device for the binomial distribution. Wald overcame a technical obstacle to extending Friedman's method to continuous distributions. Later we had a memo from England in which a Miss Stockman had derived the operating characteristic curve for the binomial case. Wald says that George Brown had derived it, too, though neither Milton nor I knows anything of this directly. At any rate, this opened up a whole new range of activity for Wald's theoretical abilities, and a whole new spurt in theoretical developments.

Another point which received considerable attention was the problem of making an estimate from a sample collected for sequential test purposes. Such samples lead to biased estimates if one applies the estimating procedures that could be used with ordinary random samples. A great deal of high powered work was put on this during the war, most of it by Wald, quite a bit by Girshick, and some by Wolfowitz, Tukey, Savage, and others. Just after the war you sent me a 3×5 slip saying "See *Nature*, no. 3924, page xxx." We saw this, and it was essentially an estimate from a highly special type of sequential sample.[7] It started Savage and Mosteller off, and eventually they discovered how to form unbiased estimates from sequential samples. At about the same time, Girshick, who had by then returned to Washington, discovered the same results. The three of them got together and published the paper.[8] This opened up another large flurry which is still under way, and dragged in several mathematicians not previously very familiar to statisticians. Attention shifted to the problem of sequential estimation proper, that is, the problem of collecting a sample sequentially for maximum efficiency in estimation. I had raised this problem with Wald right after he worked out his test procedure, and he had gone so far as to show that in the case of a normal distribution with known variance there is no sequential estimate as efficient as a single sample estimate. More recently, Stein has shown that when the variance is unknown, a sequential estimate can be obtained which is practically as efficient as a single sample estimate with known variance. Anscombe published a note on sequential estimation in BIOMETRIKA for 1949.

We had a little incident with Thornton Fry after he became Acting Chief of the Panel. This is an incident which I look back on with some satisfaction, arising partly from the fact that I won out and got the book published over his determined opposition, and partly from the amusing fact that this was not due to my astuteness but to his outfoxing himself. One of the highlights of the Fry affair was a conference of Mina Rees, Sam Wilks, Fry, and myself to decide what to do about the book, Fry having promised to support the decision of the group. After some discussion he ruled Mina voteless and called for Sam's vote. It took him a good hour to wring a vote out of Sam, who wouldn't say yes or no about publication but always thought of an alternative or a compromise. Finally Fry said, "Sam, you're the hardest man to pin down that I ever encountered. Just say yes or no." Sam said that if it was put that way, and if it really wasn't possible to do this, that, or the other, then he really wouldn't want to see the book not published. Fry said he had given his word, so would support the decision, but he wanted us to know that it was the wrongest decision he ever participated in. He then wrote to Ward Davidson recommending publication but explaining that this recommendation went much against his grain, was thoroughly against both NDRC and SRG interests, and was made because he had promised to make it. Davidson turned it down. Fry insisted then that we drop the whole matter of publication and immediately order 200 copies that had been requested by the Air Force. As soon as the order was irrevocably placed, I got Columbia to write NDRC, by-passing Fry and Davidson, an offer to provide 200 copies free, thus saving the government $2,025, if allowed to publish the book commercially. Since no government official could risk refusing so clear-cut a direct cash saving, we had no further troubles.

Somewhere around the office I have various documents that would fill in this story in more detail. For one thing, shortly after Wald started work on sequential analysis in April 1943, he asked me about the origin of the idea and I gave him a written memo on it. This is surely somewhere in my files. I would hate to hunt anything in those files, but if it really interests you I will.

Sincerely yours,
Allen Wallis

On August 7, 1978, after futile searches for that April 1943 memo had been made for me at the National Archives, the Library of Congress, and the Defense Department, I poked into a veritable haystack of old files in my basement. The second drawer I chose started with letters to Albert Bowker in the spring of 1946 offering him an appointment at Stanford and

[7] See Haldane (1945).

[8] Mosteller informs me (1978) that, "Girshick had not discovered the same results. He had some asymptotic work he had done. It turned out that his work was not at all close to ours. We published with him primarily because of your concerns about sending two groups off to do the same job."

ended with the needle I was seeking, the 1943 memorandum to Wald. I quote the memorandum complete, not so much because it adds to or subtracts from the account in the letter to Weaver, as for the interest in comparing an account written within a week or two of the events with an account written after seven years. To the best of my recollection, I had not seen the 1943 memorandum again until August 1978. It is dated April 3, 1943.

Memorandum to: A. Wald
From: W. A. Wallis
Subject: HISTORY OF SEQUENTIAL TEST

Your inquiry last night about the history of the problem you are working on set me to "reminiscing"—if that term can be applied to events only a few weeks old.

The idea that sequential tests would have far reaching consequences for the theoretical foundations of statistics, quite beyond their immediate practical applications, and that in a sense they improve the power of "most powerful" tests, originated jointly with Milton Friedman and myself.

I can perhaps claim to have supplied the initiative in the first stages by devising an elementary sequential test which demonstrated that in circumstances where the test criterion used was "most powerful" exactly the same power and exactly the same region of rejection could be secured with a smaller expected number of trials (my memorandum 3–23–3A [March 23, 1943] presents the test). In other words, this demonstrated that the classical tests become more efficient when the observations are treated sequentially instead of *en masse* as in the classical procedure.

I presented this idea to Milton during the period March 23 to 26. (I had discussed with him the practical merits of sequential tests during the preceding week but neither of us had developed any definite ideas about how to construct such a test, nor had we seen any theoretical interest in the problem.) He was greatly interested and we discussed the implications at some length. During these conversations we developed the idea of modifying the region of rejection in such a way that it would become most powerful relative to a sequential analysis of the data—i.e., sacrificing some of the power relative to a fixed expectation of N. I should say that it was at this second stage that we attained a fairly clear-cut formulation of the problem and its potentialities, with perhaps some vague notions about the kind of solution required. (At this point we talked the idea over with Jack Wolfowitz, but he, being somewhat less impetuous and considerably more mathematically sophisticated than we, was inclined to doubt the existence of a sequential test which would have for given power and size of region an expected number of trials less than the number required for the same power and size of region with classical tests.)

Milton then assumed the initiative and actually constructed a simple example of what we were then calling a colossal test (in contrast with those merely "most powerful"). He showed this to me during the period March 30 to April 2. Though it left much to be desired, it seemed to us pretty definite confirmation of our ideas of the preceding week that "most powerful" tests can be in a certain sense more powerful if the region of rejection is altered to take sequence into account. It also became clear that the idea deserved much more thorough exploitation than Milton and I were able to provide, and besides constituted by its fascination a menace to Milton's, Jack's, and my obligations to the NDRC; so we decided to try to foist it on you. Our pangs at parting with so promising a child are offset by our pride and joy at its amazing development in its present opulent home.

The idea of a sequential test was first mentioned to me by Ed Paulson in February when we were working on our memorandum 3–6–3AE [March 6, 1943] on "Number of Trials Necessary for Comparing Two Percentages." Ed said that Professor Hotelling had mentioned that he had thought of setting some kind of a bound, a function N, the crossing of which would lead to rejection of the null hypothesis. This would avoid the necessity of determining N in advance. Ed said that Hotelling had said he had some ideas on how to go about constructing such a test. Ed and I agreed that such a solution would be highly practical but it slipped into the back of my mind.

However, on March 13 I talked with Captain G.L. Schuyler about our memo 3–6–3AE. Captain Schuyler said that Lt. E.W. Gwathmey had suggested devising a chart on which results would be plotted continuously, with two lines drawn on the chart in advance. Crossing one would mean rejection, crossing the other acceptance. I talked with Gwathmey for a couple of hours or more that afternoon. (Incidentally, he gave no indication of having heard of the control line idea before, which puzzled me.) It seemed to me that Gwathmey was pretty badly confused on the whole subject, and since he is unquestionably a first-rate man that worried me. He made two points: that he couldn't understand what we did, and anyway didn't think we faced the real problem. All in all, I felt that while what we had said was undoubtedly "true," and the best that could be done, Gwathmey's confusion was not unjustified. Consequently, I began thinking about the problem of a sequential test. At that stage, I was viewing it purely as a specific practical problem. And as we explained to you, it is for the double dichotomy that we hope to have the first application of the procedure you have worked out.

In 1974, Donald A. Darling gave a paper on "The Birth, Growth, and Blossoming of Sequential Analysis" (Darling 1976). In several essential respects, it is not consistent with my accounts of 1943, 1950, and now 1980.

The impact of sequential analysis on statistical theory and practice can be judged by others better than by me, for it is 25 years since I last participated in technical statistics. Harold Freeman suggested using as an index the number of articles with the term *sequential analysis* in their titles listed in *An Author and Permuted Title Index to Selected Statistical Journals* (Joiner et al. 1970). This volume covers six journals for varying periods, mostly in the 1960's. I counted 195 such titles. Since 75 journal-years are covered, this comes to 2.6 articles on sequential analysis per year per journal. Each of the 1975 and 1976 volumes of the *Current Index to Statistics* lists between 50 and 55 articles including the term *sequential analysis* in their titles. Sequential analysis even today, 37 years after its discovery, continues to exert a strong influence on statistical research.

My last two headings must be disposed of briefly.

7. INFLUENCE

The continuing influence of other work done at SRG can be documented by the continuing citations to *Sampling Inspection* and to *Techniques of Statistical Analysis*. The heart of *Sampling Inspection* survives in revision after revision of military acceptance sampling specifications. Requests for permission to reprint parts of both these books continue to arrive occasionally—especially *Techniques of Statistical Analysis*, which one publisher asked to reprint complete. (Both books have been in the public domain for more than 20 years.) If I were researching rather than reminiscing, I would count entries in citation indexes.

The effect of SRG on the subsequent careers of its members was probably greatest on Jimmie Savage, who came to SRG as an applied mathematician, innocent of statistics. The fact that 7 of the 18 principals were primarily or secondarily economists—Bennett, Friedman, Hastay, Hotelling, Stigler, Wald, and Wallis—unquestionably exerted a profound influence on the later statistical work of all members of SRG, but especially on Savage's. It is, of course, hard to separate the impact of the economists in SRG from the great influence on all statistical theory immediately after the war of game theory and decision theory, both of which are essentially economic theories.

The only members of SRG whose subsequent careers were not much affected by the SRG experience, as far as I can tell, were Bennett, Bigelow, and Stigler—an electrical engineer and two economists.

Of the 18 principals in SRG, 4 have died: Girshick, Hotelling, Savage, and Wald; their biographies appear in the *International Encyclopedia of Statistics* (Kruskal and Tanur 1978). All those who are living, except Bigelow, responded to my request for reminiscences. Several themes that appear in most of the responses were well articulated by Kenneth Arnold:

> It is a pleasure to recall the days of SRG. I had not seen such concentrated cooperation before nor have I since. The joy of working with others toward a goal we all considered of paramount importance was enhanced by the urgency we felt. The importance of the problems, association with highly intelligent individuals, the fact that what we did made a difference, all contributed to satisfaction with being a part of the group. The individuals composing SRG had a common core of knowledge of statistics but differed greatly in their approach to problems.
>
> My recollections have less to do with specific accomplishments by specific individuals than with the feeling that the interchange of ideas among members of SRG in their attempts to solve practical problems brought forth a new balance among the concepts which give statistics its vitality. Each member of SRG had his own version of this balance, but there was much in common. It could be that the development of statistics in the United States for the next twenty years was foreordained but I prefer to believe that SRG was responsible for the fact that many a discussion at a meeting of statisticians sounded like a continuation of a discussion begun at SRG.

Several of the younger men expressed a view that Bowker put this way:

> I was able to meet and talk with many people who were or were to be major leaders in the field. I could drop in on most of them. Doors were frequently left open and I must say I found it very exciting intellectually to be treated as a colleague and friend rather than as a student. I had many conversations with people who like to talk and explain. I tried to pattern the department at Stanford on SRG.

Several letters mentioned that SRG functioned during what Hastay called "one of the great creative periods in statistical science," a time when, in Freeman's words,

> Statistical interest and relevance were high. Testing hypotheses, estimation theory were alive. The Fisher-Neyman controversies were unsettled. Sequential theory, decision theory, game theory, Bayesian notions, all exciting innovations, were underway. Those were effective years for statistical theory and application.

Then Freeman added a sentiment expressed in several letters and felt by all of us: "Ghastly that such progress should have involved, even depended on, the deaths of millions of people thousands of miles away."

Nearly everyone mentioned at least one person who had influenced him particularly strongly. It is no surprise that the two mentioned most frequently by far were Friedman and Wald. Others mentioned by a substantial fraction of the respondents were Bigelow, Bowker, Eisenhart, Girshick, Hotelling, Savage, and Wolfowitz, and nearly everyone was mentioned at least once. (I have ignored references to me for reasons obvious to any statistician.)

Bennett reminded me of a "mathematical recreation" that consumed a large amount of lunchtime attention for several weeks:

> Given 12 coins, all of the same weight except one, and using only a two-pan balance, find the odd coin and determine whether it is heavier or lighter than the others, making only three weighings.

Others remembered what an elegant general solution to this whole class of problem had been produced by Bennett.

Friedman reported an important contribution of his work at SRG to his later work in economics:

> One article of mine that in a very important sense grew almost entirely out of the work of SRG was the article which I wrote in 1953 on "Choice, Chance, and the Personal Distribution of Income." It traced directly to our work on the proximity fuze. One element in the work on the proximity fuze was the attempt to approximate the time distribution of bursts. The proximity fuze had two impulses, one forward and one backward. As a result there were generally two modes in the distribution. We treated this as the sum of two separate distributions and ultimately decided that it could best be treated as the sum of two exponential distributions. The resemblance of those distributions of bursts to income distributions got me started to thinking about whether the same method could not be used to describe income distributions, and that result is directly and immediately reflected in the article I referred to. (Friedman 1953)

Friedman's inspiration to fit the fuze data by joining a positive and a negative exponential distribution, thus generating a sharp-pointed mode, followed an evening of conversation about economics with Arthur F. Burns, who questioned the assumption, which is made a priori by economists, that no economic series or distribution ever has a discontinuity in the first derivative—or probably any other derivative either.

There is a legend about SRG that I should mention. It is said that as Wald worked on sequential analysis his pages were snatched away and given a security classification. Being still an "enemy alien," he did not have a security clearance so, the story has it, he was not allowed to know of his results. This story justifies a remark with which Jimmie Savage often responded to astonishing news: "I wouldn't believe it if I had made it up myself." As a matter of fact, I did make this one up myself. We were irritated and amused by the fact that Wald's work on sequential analysis was classified while he had no clearance. In a casual conversation one day, I exclaimed in irritation, "What do they expect us to do, snatch the pages away as each is completed and keep it secret from him?" This petulant joke has become part of the history of statistics. The reason sequential analysis was classified, and the reason it could not be declassified, even though it was requested by the top people in SRG, AMP, NDRC, and OSRD, was that the funds supporting SRG had been appropriated by Congress for classified work. Therefore, anything substantive that we did was classified. About six months before the war ended, the Army got sequential analysis declassified. They argued that its wide dissemination throughout industry was important to the war effort and that classification impeded full and speedy dissemination. A pleasant by-product of the classification of sequential analysis was that Wald's citizenship was granted expeditiously, making him one of the few enemy aliens to receive citizenship during the war.

8. CONCLUSION

The invitation to prepare this paper referred also to my other organizational activities in statistics. I had parts, ranging from minor to major, in the establishment of five university departments of statistics: Columbia, Stanford, Chicago, Harvard, and Rochester, all of which I believe to be strong enough to justify pride in even small roles. I edited the *Journal of the American Statistical Association* for almost the whole decade of the 1950's; what I would point to about my editorship is the group of associate editors who worked with me[9] and the way we organized the editing and refereeing. As chairman of the editorial advisory board and of the executive committee of the *International Encyclopedia of the Social Sciences*, which was prepared between 1962 and 1968, I was largely responsible for selecting the editors for statistics and economics.[10] I was chairman of the President's Commission on Federal Statistics in 1970–71, and since I was allowed a major part in selecting its members, there too I can take pride in its membership and staff[11] and in its continuing influence.

REFERENCES

Baxter, James Phinney III (1946), *Scientists Against Time*, Boston: Little, Brown. (This is the brief official history of the Office of Scientific Research and Development.)

Darling, Donald A. (1976), "The Birth, Growth, and Blossoming of Sequential Analysis," in *On the History of Statistics and Probability*, ed. D.B. Owen, New York and Basel: Marcel Dekker, 369–375.

Eisenhart, Churchill (1977), "The Birth of Sequential Analysis" (Obituary note on retired Rear Admiral Garret Lansing Schuyler), *Amstat News*, 33, 3.

Freeman, Harold A. (1944), *Sequential Analysis of Statistical Data: Applications*, New York: Statistical Research Group, Columbia University. (Ring binder of seven booklets. Also see Statistical Research Group 1944.)

Friedman, Milton (1953), "Choice, Chance, and the Personal Distribution of Income," *Journal of Political Economy*, 61, 277–290.

Grant, Eugene L. (1946), *Statistical Quality Control*, New York: McGraw-Hill Book Co.

Haldane, J.B.S. (1945), "A Labour-Saving Method of Sampling," *Nature*, 155, 49–50.

International Encyclopedia of the Social Sciences (1968), ed. David L. Sills, New York: The Free Press.

International Encyclopedia of Statistics (1978), eds. William B. Kruskal and Judith M. Tanur, New York: The Free Press.

Joiner, Brian L., Laubscher, N.F., Brown, Eleanor S., and Levy, Bert (1970), *An Author and Permuted Title Index to Selected Statistical Journals*, Washington, D.C.: U.S. National Bureau of Standards Special Publication No. 321.

Moore, Geoffrey H., and Wallis, W. Allen (1943), "Time Series Significance Tests Based on Signs of Differences," *Journal of the American Statistical Association*, 38, 153–164.

Old, Bruce S. (1961), "The Evolution of the Office of Naval Research," *Physics Today*, 14, 30–35.

SRG, Columbia University (1945a), *Final Report* (submitted to the Applied Mathematics Panel, National Defense Research Committee in completion of research under Contract OEMsr-618 between the trustees of Columbia University and the Office of Scientific Research and Development. Available at the U.S. National Archives and Record Service, Center for Polar and Scientific Archives, Washington, D.C.).

——— (1945b), eds. Harold A. Freeman, M.A. Girshick, and W. Allen Wallis, *Sequential Analysis of Statistical Data: Applications*, New York: Columbia University Press. (Ring binder of eight booklets. A revision and expansion of Freeman 1944.)

——— (1947), eds. Churchill Eisenhart, Millard W. Hastay, and W. Allen Wallis, *Selected Techniques of Statistical Analysis for Scientific and Industrial Research and Production and Management Engineering*, New York and London: McGraw-Hill Book Co.

——— (1948), eds. Harold A. Freeman, Milton Friedman, Frederick Mosteller, and W. Allen Wallis, *Sampling Inspection: Principles, Procedures, and Tables for Single, Double, and Sequential Sampling in Acceptance Inspection and Quality Control Based on Percent Defective*.

Wald, Abraham (1945a), "Sequential Tests of Statistical Hypotheses," *Annals of Mathematical Statistics*, 16, 117–186.

——— (1945b), "Sequential Method of Sampling for Deciding Between Two Courses of Action," *Journal of the American Statistical Association*, 40, 277–306.

——— (1947), *Sequential Analysis*, New York: John Wiley & Sons.

Wallis, W. Allen (1941), "A Significance Test for Time Series Analysis," *Journal of the American Statistical Association*, 36, 401–409.

——— (1942a), "The Temporal Stability of Consumption Patterns," *Review of Economic Statistics*, 24, 177–183.

——— (1942b), "Compounding Probabilities From Independent Significance Tests," *Econometrica*, 10, 229–248.

——— (1942c), "How to Ration Consumers' Goods and Control Their Prices," *American Economic Review*, 32, 501–512.

Wallis, W. Allen, and Moore, Geoffrey H. (1941), *A Significance Test for Time Series and Other Ordered Observations*, New York: National Bureau of Economic Research.

Working, Holbrook (1943), "Making Statistical Methods Contribute to the War Effort," *American Statistical Association Bulletin*, 3, 2–4.

——— (1945), "Statistical Quality Control in War Production," *Journal of the American Statistical Association*, 40, 425–447.

[9] The associate editors of *JASA*, usually six at a time, while I was editor were Sidney S. Alexander, Albert H. Bowker, Churchill Eisenhart, Solomon Fabricant, Harold A. Freeman, Millard Hastay, Clifford Hildreth, Howard L. Jones, Lester S. Kellogg, Nathan Keyfitz, George M. Kuznets, William G. Madow, Philip J. McCarthy, Frederick Mosteller, I. Richard Savage, and C. Ashley Wright.

[10] The editor for statistics was William H. Kruskal, who was assisted by Judith M. Tanur. They have compiled the statistics articles from *IESS*, with additions, to form the *International Encyclopedia of Statistics*, published in 1978.

[11] Members of the President's Commission on Federal Statistics of 1970–71 were Ansley J. Coale, Paul M. Densen, Solomon Fabricant, Robert D. Fisher, W. Braddock Hickman, William H. Kruskal, Stanley Lebergott, Frederick Mosteller (vice chairman), Richard M. Scammon, William H. Shaw, Frank D. Stella, James A. Suffridge, John W. Tukey, and W. Allen Wallis (chairman). Daniel B. Rathbun was executive director of the staff, Paul Feldman was deputy executive director, Norman V. Breckner was assistant director, and there were seven other professionals on the staff.