KIM KLEINMAN*

Why Edgar Anderson Visited Math Departments: Natural History, Statistics, and Applied Mathematics

ABSTRACT

Edgar Anderson of the Missouri Botanical Garden had long and rich collaborations with such mathematicians and mathematically inclined biologists as R. A. Fisher, Sewall Wright, and John Tukey. It was Anderson's *Iris* data that Fisher used to develop his linear discriminant function to capture multiple variations. A sabbatical with Wright in 1933 helped hone Anderson's mathematical skills while helping him understand what mathematics could and could not do. He and Tukey shared an interest in conveying data graphically. This long-standing commitment to applying mathematics to natural history problems informed his scientific career as he sought to capture the variations he recognized in the natural populations. He used graphical tools to examine hybridization as an evolutionary mechanism and to use the taxonomic data from these variations to study the underlying genetic forces at work in evolution. In important synthesis articles in the mid-1950s, he summarized his conclusions about applied mathematics and natural history. They were not mere technical tools, but reflected a commitment to observation and pattern recognition as the basis of his science. Understanding these views more fully deepens an appreciation of this constantly independent-minded contributor to evolutionary theory in the twentieth century.

KEY WORDS: Edgar Anderson, R. A. Fisher, Sewall Wright, John Tukey, applied mathematics, statistics, natural history

*Webster University, 470 East Lockwood Avenue, St. Louis, MO 63119, Kleinman@webster.edu

The following abbreviations are used: Anderson Papers *for* Edgar Anderson Papers, Missouri Botanical Garden Archives; Fisher Papers *for* RA Fisher Correspondence, Barr Smith Library, University of Adelaide; MBG *for* Missouri Botanical Garden; Tukey Papers *for* John Tukey Papers, Courtesy of Frank Anscombe; Wright Papers *for* Sewall Wright Papers, Correspondence File "Edgar Anderson, 1932–1960," American Philosophical Society.

Historical Studies in the Natural Sciences, Vol. 49, Number 1, pps. 41–69. ISSN 1939-1811, electronic ISSN 1939-182X. DOI: https://doi.org/10.1525/HSNS.2019.49.1.41.

INTRODUCTION

In December 1956, Edgar Anderson gave readers of *The Missouri Botanical Garden Bulletin* a popular overview of his efforts to capture graphically and mathematically his natural history observations of species, their variability, and the evolutionary forces shaping them. This project began with his first major paper where he used "ideographs" to capture visually multiple variations in *Iris*[1] and continued through sabbaticals and collaborations with such preeminent population geneticists R. A. Fisher and Sewall Wright. In 1956, he was excited to explain "Why Botanists Visit Math Departments,"[2] as he had a grant from the Guggenheim Foundation to work at Princeton University with John Tukey, who would later dedicate his *Exploratory Data Analysis*[3] in part to Anderson.

Anderson was a leading participant in the "Evolutionary Synthesis" discussions of the 1930s–1950s, having presented with Ernst Mayr on "Systematics and the Origin of Species" from their respective disciplinary perspectives for the 1941 Jesup Lectures. I have argued that he did not complete a botanical companion to Mayr's landmark work[4] because he pursued "his own synthesis," using maize to explore plant evolution under human interactions.[5] Further, given the complexity of plant reproduction, he was less sure that a fully integrated view of evolution had been reached, despite the significant progress that had been achieved. In one contribution to the broader discussion, he said, though we have solid knowledge about "how evolution is proceeding" in particular taxa, "Evolution in general, evolution with a big E, that is another matter."[6]

He came to the Missouri Botanical Garden (MBG) in 1922 from Harvard's Bussey Institution where he had studied self-sterility in *Nicotiana* with E. M. East.[7] At the MBG, he was "Geneticist to the Garden," contributing to the

1. Edgar Anderson, "The Problem of Species in the Northern Blue Flags, *Iris versicolor* L. and *Iris virginica* L.," *Annals of the Missouri Botanical Garden* 15 (1928): 241–332.

2. Edgar Anderson, "Why Botanists Visit Math Departments," *Missouri Botanical Garden Bulletin* 44 (1956): 148–51.

3. John W. Tukey, *Exploratory Data Analysis* (Boston: Addison-Wesley, 1977).

4. Ernst Mayr, *Systematics and the Origin of Species: From the Viewpoint of a Zoologist.* (New York: Columbia University Press, 1942).

5. Kim Kleinman, "His Own Synthesis: Corn, Edgar Anderson, and Evolutionary Theory in the 1940s," *Journal of the History of Biology* 32 (1999): 293–320.

6. Edgar Anderson, "Supra-Specific Variation in Nature and in Classification: From the View-Point of Botany," *The American Naturalist* 71 (1937): 223–35, on 223.

7. Edgar Anderson, "Studies on Self-Sterility VI. The Genetic Basis of Cross-Sterility in *Nicotiana*." *Genetics* 9 (1924): 13–39.

education of both graduate students in botany at Washington University and students in the Garden's program for landscape gardeners. Even as he taught genetics, he joined his systematist colleagues and students in the field benefitting from their perspectives.

His graphical and mathematical work was part of his dedication to creating visual tools for both observation and analysis that could accurately capture the subtle variation in populations in nature without washing out significant details. He was, as always, committed to getting the biology right first before forming conclusions. Anderson, though sympathetic to the earlier work of Karl Pearson and the Biometricians,[8] found their work inadequate for biological reasons (organisms do not vary randomly, so they must be analyzed accordingly[9]). He felt the need to create and contribute to such mathematical tools that were up to the biological tasks at hand. To that end, this paper is a chronicle of these mathematical collaborations and their published results to capture this key element of Anderson's evolutionary views.

This story then includes Anderson's initial *Iris* work in 1928; his 1929–1930 sabbatical at The John Innes Horticultural Institute, where he worked with Fisher; a 1933 term spent with Sewall Wright at the University of Chicago; his collaboration with Fisher wherein his *Iris* data set resulted in Fisher's linear discriminant analysis[10]; Anderson's continued work on *Tradescantia*, maize, and other taxa, which contributed to his appreciation of introgressive hybridization as an evolutionary mechanism[11]; his collaboration with Tukey; and finally, his own major mathematical papers ("Efficient and Inefficient Methods of Measuring Specific Differences,"[12] "Natural History, Statistics, and Applied Mathematics,"[13] and "A Semigraphical Method for the Analysis of Complex

8. William B. Provine, *Origins of Theoretical Population Genetics* (Chicago: University of Chicago Press, 1971).

9. Edgar Anderson, "Introgressive Hybridization," *Biological Review of the Cambridge Philosophical Society* 28 (1953): 280–307.

10. Ronald A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics* 7, no. 2 (1936): 179–88.

11. Kim Kleinman, "'Bringing Taxonomy to the Service of Genetics': Edgar Anderson and Introgressive Hybridization," *Journal of the History of Biology* 49, no. 4 (2016): 603–24.

12. Edgar Anderson, "Efficient and Inefficient Methods of Measuring Specific Differences," in *Statistics and Mathematics in Biology*, eds. Oscar Kempthorne, Theodore A. Bancroft, John W. Gowen, and Jay L. Lush (Ames: The Iowa State College Press, 1954).

13. Edgar Anderson, "Natural History, Statistics, and Applied Mathematics," *American Journal of Botany* 43 (1956): 882–89.

Problems"[14]). Importantly, Anderson made a point of discussing these graphical techniques in his two books, devoting whole chapters: the more soberly titled "Special Techniques for the Study of Introgression" in *Introgressive Hybridization*[15] and "How To Measure an Avocado" in *Plants, Man, and Life.*[16]

More broadly, I aspire to complement Joel Hagen's examination of a "statistical frame of mind in systematic botany."[17] There he notes the marked increase in the use of statistics in biology over the course of the twentieth century, with George Gaylord Simpson and Anderson[18] representing an older generation—albeit as particularly mathematically curious and adept representatives of that generation—than others like Robert Sokal, who embraced statistics without reservation. Simpson and Anderson saw mathematics as a tool, but saw a role for judgment, experience, and pattern. The ready availability of sophisticated computer power and tools has only accelerated the consolidation of a "statistical frame of mind." Anderson contributed to important statistical tools like R. A. Fisher's linear discriminant function while recognizing that many scientists, himself included, were visual-minded. For him and them, graphical depictions of data were a complement to the patterns they recognized in the field and so were a bridge between older natural history methods and the application of statistics.

Recent work by Thomas L. Hankins[19] and by David Sepkoski and Mario Tamborini[20] capture other attempts to use graphical and visual depictions of the natural world to advance science grounded in the nuanced visual judgments of seasoned observers. In these colleagues' valuable contributions, I see a complementary approach among Anderson, John Herschel, and H. G. Bronn across their respective disciplines.

14. Edgar Anderson, "A Semigraphical Method for the Analysis of Complex Problems," *Proceedings of the National Academy* 43 (1957): 923–27.

15. Edgar Anderson, *Introgressive Hybridization* (New York: John Wiley and Sons, 1949).

16. Edgar Anderson, *Plants, Man, and Life* (Boston: Little, Brown, and Company, 1952).

17. Joel Hagen, "A Statistical Frame of Mind in Systematic Biology from *Quantitative Biology* to *Biometry*," *Journal of the History of Biology* 36 (2003): 353–84.

18. Ibid., 360–64.

19. Thomas L. Hankins, "A 'Large and Graceful Sinuosity': John Herschel's Graphical Method," *Isis* 97, no. 4 (2006): 605–633.

20. David Sepkoski and Mario Tamborini, "'An Image of Science': Cameralism, Statistics, and the Visual Language of Natural History in the Nineteenth Century," *Studies in the History of the Natural Sciences* 48, no. 1 (2018): 56–109.

But even more broadly still, this discussion can be seen in relation to Lorraine Daston's and Peter Gallison's view that objectivity has a history.[21] As such our standards of what constitutes objective knowledge develops and changes over time as tools, like statistics and applied mathematics in this case, become more refined. But this development is not linear and reflects shifting historical judgments. Even as our tools, techniques, and judgments shift, Anderson's commitment to depicting variation, using pattern recognition, and the natural history method in general remain, both as synthesized in newer practices and as an important antithesis still to be considered.

Anderson's work in this area reflects dedicated commitment to creating biologically sound techniques, useful for both observation in the field and analysis back in the laboratory. They were graphical and visual because, as he told Sewall Wright, "My chief difficulty in reading ["General, Group and Special Size Factors"[22]] is that like many naturalists I am visually minded. A string of numbers means very little to me until I have turned it into a diagram."[23] But, as his protege Charles B. Heiser remembered him, his visual talents were particularly acute:

> I became aware that he could see things in plants that others couldn't, or at least I couldn't . . . he could look at a field of plants and actually see things that escaped others. In trying to analyze this ability I have decided it involved two things: an ability to correlate a large number of independent observations in a very short time . . . and an ability to analyze patterns.[24]

Anderson's aim for these techniques was simply to correlate those many and varied observations in a way that would help him find the patterns in nature that he sought to explain. Their relevance to fields as disparate as archaeology, psychology, even linguistics in proto-digital humanities projects is testimony to their power and Anderson's breadth of interests. So, they are not mere technical tools but are integral to understanding the view of science of a constantly independent minded contributor to evolutionary theory in the twentieth century.[25]

21. Lorraine Daston and Peter Gallison, *Objectivity* (Cambridge, MA: MIT Press, 2007).

22. Sewall Wright, "General, Group, and Special Size Factors," *Genetics* 17 (1932): 603–19.

23. Anderson to Wright, 11 Mar 1932, in Wright Papers.

24. Charles B. Heiser Jr., "Student Days with Edgar Anderson or How I Came to Study Sunflowers," *Annals of the Missouri Botanical Garden* 59 (1972): 362–372, on 365.

25. Kim Kleinman, "His Own Synthesis" (ref. 5); "How Graphical Innovations Assisted Edgar Anderson's Discoveries in Evolutionary Biology," *Chance* 15, no. 3 (2002): 17–21; "Biosystematics and the Origin of Species: Edgar Anderson, W. H. Camp, and the Evolutionary Synthesis," in

THE *IRIS* PROJECT IN 1928

Edgar Anderson joined the MBG staff in 1922 as Geneticist to the Garden, having just received his doctorate from Harvard University, where he studied at the Bussey Institution with E. M. East. His initial responsibilities included reorganizing the Garden's practical School for Gardeners and teaching in the Henry Shaw School of Botany at Washington University in St. Louis. Late in his career he reflected on the impact of working with colleagues and students who focused on taxonomy, morphology, and physiology. "I taught genetics but I explored the Ozarks with my students. They learned genetics from me, and they convinced me that I should take a serious interest in taxonomy." They studied the "wide field between genetics and taxonomy," focusing on the species problem just as much as the Ozarks.[26]

He brought that approach into his later work, including on the evolution of cultivated plants, which itself contributed to his broader insights into the impact of back-crossing hybridization on evolution in general. He valued "the taxonomic method."

> What is the taxonomic method? Well, it is the method of natural history . . . . It is easiest to define by example or by contrasting it with approaches more suitable for other types of problems. It customarily does not deal with exact units, it uses a minimum of mathematics, and until recently has had only a grudging respect from my fellow geneticists. Though I am not a taxonomist, though I am by nature fascinated with mathematics and have used mathematical methods in nearly all my scientific papers, I was fortunate fairly early in my scientific career to have acquired a real respect for the taxonomic method. It is probably no accident that Dr. Alfred C. Kinsey and I, who were graduate students together, should both be using the taxonomic method in places where it has not recently been generally used, he in the study of sex and I in the study of cultivated plants.[27]

Descended from Darwin: Insights into the History of Evolutionary Studies, 1900–1970, eds. Joe Cain and Michael Ruse (Philadelphia: American Philosophical Society, 2009): 73–91; "Systematics and the Origin of Species from the Viewpoint of a Botanist: Edgar Anderson Prepares the 1941 Jesup Lectures with Ernst Mayr," *Journal of the History of Biology* 46 (2013): 73–101; "'Bringing Taxonomy to the Service of Genetics'" (ref. 11).

26. Edgar Anderson, "What We Do Not Know about *Zea mays*," *Transactions of the Kansas Academy of Science* 71 (1968): 373–78 on 373.

27. Anderson, *Plants, Man, and Life* (ref. 16).

It was in his first major scientific paper on *Iris* that he combined the taxonomic method with the use of mathematical tools to address the species problem through an intensive and extensive look at its manifestation in the Northern Blue Flags. His study was "intensive" in that he strove to look comprehensively at the variation across the entire range of the species; it was "extensive" in that it contributed to the underlying question of how species originate.[28] That is, he practiced the serious interest in taxonomy urged on him by his MBG colleagues and students by studying the Northern Blue Flags, both on their own terms and with an evolutionist's broader curiosity.

What he found was that *Iris versicolor* is made up of two species—a northern one with short lanceolate petals and short ovaries, which extends from New England west to northern Michigan, and a more southern one with obviate-spatulate petals and long ovaries, which ranges from the Great Lakes to the Gulf of Mexico and up the southern Atlantic seaboard. By examining Linnaeus's own specimens and other contemporary collections, Anderson determined that the second, a more southern population, was in fact *Iris virginica.*

To explicate those differences, including making sense of the differences between the two kinds of petal, he prepared a table showing petal lengths, widths, taper as well as taper/width ratios for five plants from each species. No single measure sufficed to separate the two populations, and only developing a complex ratio could do the job mathematically. In the face of this, Anderson cited Charles Sedgwick Minot, whom he found via D'Arcy Wentworth Thompson's *On Growth and Form* (1917). Minot wrote:

> The fact that men of genius have evolved wonderful methods of dealing with numerical relations should not blind us to another fact, namely, that the observational basis of mathematics is, psychologically speaking, very minute compared with the observational basis of even a single minor branch of biology. . . . While therefore here and there the mathematical methods may aid us, we need a kind and degree of accuracy of which mathematics is completely incapable.[29]

There is a deeper reason why mathematics, particularly before modern computers, is of limited use for taxonomic problems. Not only is it "cumbersome in analyzing differences in form" and "swift and efficient only in recording differences in number,"[30] it is essentially useless in handling the multiple

28. Anderson, "The Problem of Species" (ref. 1), 242–43.

29. Cited in ibid., 283–84.

30. Ibid.

variations a taxonomist deals with. One could perhaps develop a measure that could distinguish between apples and oranges by shape, but such an effort would be both difficult and ludicrous.[31] In this paper, he wrote more soberly, "Mathematics can deal with such problems through the study of correlation but it is slow and laborious work and though it may be useful in the analysis of some particular problem it is not adapted to general taxonomic use."[32]

As an alternative, Anderson developed "ideographs" as "a new method of presenting biometric data which combines the good points of the methods of mathematics and comparative morphology. Like mathematics, it is accurate and objective. Like morphology, it leaves something to the trained eye"[33] (Fig. 1).

These ideographs allowed Anderson to record four variables at once and show their relationships in pictorial form. His convention was to construct a white rectangle with the dimensions of the petal and then to superimpose it on a black rectangle with the dimensions of the sepal. With ideographs for each plant measured as in Figure 1, he had a ready pictorial method for comparing variation among populations. As he explained later in *Introgressive Hybridization*:

> Though they are laborious to construct, the importance of the ideograph lies in the fact that they show so many things at once. For the iris ideographs, each shows fifteen separate facts. That is, if the ideographs were to be replaced with statistics, it would be necessary to employ fifteen separate measurements and ratios for each ideograph. There are first of all the four original measurements—sepal length, sepal width, petal length, and petal width; then there are the six proportions between these four, taken two at a time (the length of the petal in proportion to its width, the width of the petal in proportion to the width of the sepal, etc.); then there are four three-way relationships (such as the length-width of the petal in relation to the length of the sepal); and finally there is the relationship of all four measurements taken at once.[34]

Finding such a graphical tool as the ideograph was an integral part of Anderson's application of the taxonomic method to the species problem. He sought ever better graphical and mathematical tools to capture the reality and complexity of nature.

31. Anderson, "Efficient and Inefficient Methods" (ref. 12), 103–04.
32. Anderson (ref. 1), 284.
33. Ibid.
34. Anderson, *Introgressive Hybridization* (ref. 15), 86–88.

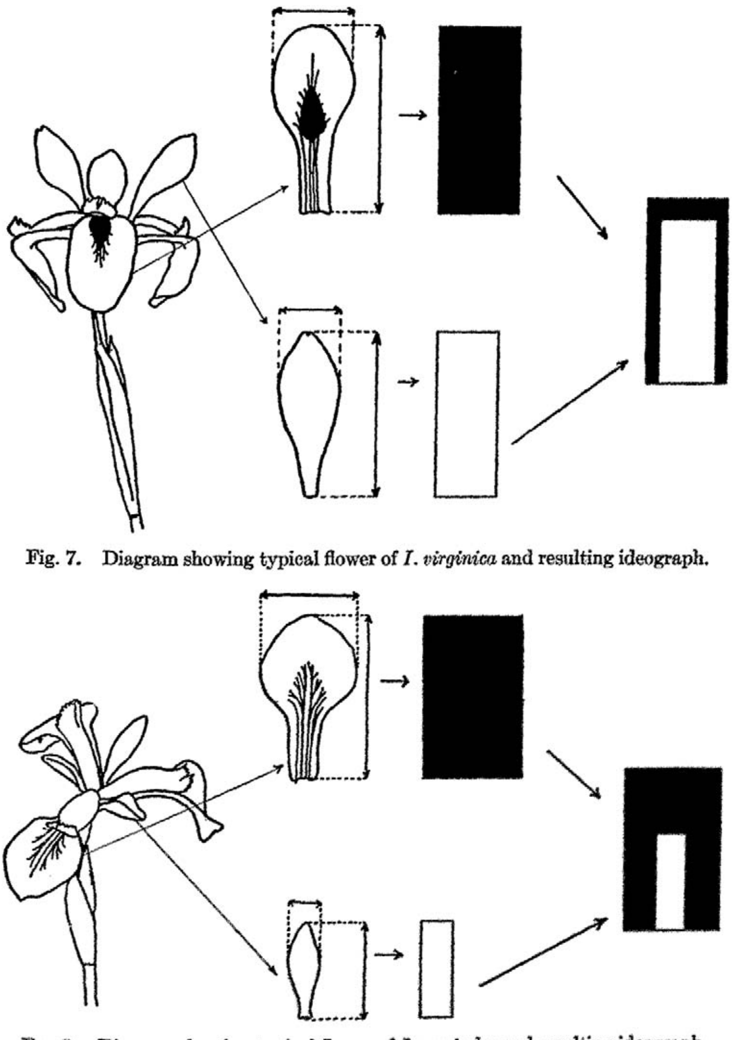

FIG. 1. Ideographs of *I. versicolor* and *I. virginica*. *Source:* Anderson, *Introgressive Hybridization* (ref. 15), 87.

IRIS AND R. A. FISHER

As Anderson's *Iris* project deepened, he collaborated with the noted mathematical geneticist R. A. Fisher. It was Anderson's *Iris* data[35] from which Fisher

35. Edgar Anderson, "The Species Problem in *Iris*," *Annals of the Missouri Botanical Garden* 23 (1936): 457–509.

derived the linear discriminant function[36] as a way of rendering precisely the type of multiple variations that Anderson saw as essential in understanding the differences between species.

Anderson worked with Fisher during his 1929–1930 National Research Fellowship, which he spent at the John Innes Horticultural Institute. During this time in England, he was able to work with J.B.S. Haldane and C. D. Darlington as well. Late in his career, he recounted a meeting with Fisher that underscored the importance of heterosis (hybrid vigor) as an object of study. He presented Fisher with six *Acquilegias* (columbines), three very different species and their three F1 hybrids. Fisher probed the material with interest and intelligence, but was unable to sort the plants. However, their colleague Dorthea DeWinton was able to discover the relationship very quickly by noting the increased size of the hybrids, which she attributed to hybrid vigor; that is, DeWinton recognized, as had plant and animal breeders even before Mendel, that hybrids often had qualities that were enhanced even beyond either parent. For Anderson, her observations meant "that there might be various subtle aspects of heterosis which . . . might eventually, to an able, experienced observer, be recognized (as a variable complex) almost instantly."[37]

In a May 28, 1930, letter to Fisher before the meeting discussed above, Anderson outlined the problem, noting, "What has particularly interested me has been the comparative morphology of the F1s?"[38] Assuming a random distribution of dominants for several characters, the combinations were not as expected. Anderson looked to Fisher for assistance because "[t] here seem to be several ways of getting at the problem but with my limited mathematical equipment I am naturally not getting very far." Ultimately, Anderson resolved the problem by exploring the role of linkage of genes in tending to retain certain characters and that such linkage therefore limited genetic recombination.

After Anderson returned to the United States (and a job at the Arnold Arboretum between 1930 and 1936), he and Fisher corresponded sporadically before their intense collaboration on the *Iris* data in late-1935 through mid-1936. Fisher took particular interest in Anderson's quantitative comparison of specific and generic differences,[39] saying "construct[ing] a three-dimensional

36. Fisher, "The Use of Multiple Measurements" (ref. 10).

37. Anderson, "What We Do Not Know" (ref. 26).

38. Anderson to Fisher, 28 May1930, Fisher Papers.

39. Edgar Anderson and E. C. Abbe, "A Quantitative Comparison of Specific and Generic Differences in the *Betulaceae*," *Journal of the Arnold Arboretum* 15 (1934): 43–49.

model seems to me an extremely good idea."[40] He did suggest that Anderson in the future orient such models in the same three-dimensional space for more ready comparison, but nonetheless valued Anderson's visual and graphical approach to such data.

Anderson reopened the *Iris* discussion with a December 19, 1935, letter to Fisher. On the Gaspe Peninsula in Canada he found two related species, *Iris setosa* var. *canadensis* and *I. versicolor*, growing intermingled together over an acre. The former seemed to be a preglacial remnant population that had 36 chromosomes, whereas the more prevalent *I. versicolor* was a polyploid with n+108, meaning that it had three sets of the chromosomes from *I. setosa* var. *canadensis*. While in London, they had discussed examining inter-specific and intra-specific variation using *Prunus*, but Anderson thought that "this will serve your purposes even better." He concluded the letter with mention of his biometrical work on oak leaves: "It is simple sort of stuff but I think it will show that the human biometricians would have gotten further and faster if they had decided biologically what they were going to measure before they started into a tough mathematical maze."[41] Time and again, Anderson underscored, literally in this case with his own pencil, the centrality of the biology in any such study, and that the mathematics must be strictly subordinated to that reality.

Fisher's answer on January 22, 1936, reflected both his general enthusiasm for the problem and his understanding that the elegant truths of mathematics had to fit the biological facts being studied. Discussing sepal lengths, he pointed out that whereas *setosa* sepals are shorter than those of *versicolor* in absolute terms, they are longer relative to other measurements when compared to *versicolor*; thus, his conclusion was: "This is a useful warning against thinking that one could build up a discriminating compound, using coefficients merely based on the mean differences and their variances without taking into account the whole system of co-variances between the measurements of each species."[42]

These results pleased Anderson tremendously as he was coming to the conclusion that "*Iris versicolor* is a pre-glacial or inter-glacial amphidiploid derivative of *I. virginica* plus *I. setosa*. . . . The relationship between the two species, in other words, is like the relationship between coffee and coffee with

40. Fisher to Anderson, 16 May 1934, Fisher Papers.
41. Anderson to Fisher, 19 Dec 1935, Fisher Papers.
42. Fisher to Anderson, 22 Jan 1936, Fisher Papers.

cream in it."[43] Recall that in *I. virginica*, n = 18, so with n = 108 *I. versicolor* is a hexaploid (six sets of chromosomes) of it with some *I. setosa* germplasm included. And that is precisely what Fisher's mathematics ultimately confirmed.

Anderson sent Fisher more data on March 17, 1936, and received a draft of the paper in a letter dated April 15, 1936. Anderson offered just a few small corrections while commenting "I wish I understood it mathematically as well as biologically." He returned to that persistent tension between the subject and the tool in his postscript: "Of course I now feel that such purely quantitative measures as petal lengths and widths are not very efficient discriminators compared to many qualitative characters. Were I to start work over at the beginning I should be tempted to use the latter, at least in part."[44] Although Anderson was very pleased with Fisher's "precise and elegant Index," seeing it as "a tremendous step" in the direction of advancing biometrics, he remained cautious about its over-application. He knew from the field that there were few among the enormous number of differences between any two species that would help very much in discriminating between them. Further, "the chances of hitting upon these few in a random collection of measurements or scoring are very slight," and thus, "an unguided biometrician would certainly use a large number of [characters], and he would end, I feel sure, with a discriminatory index which would be inferior to my own unaided eye."[45]

Fisher agreed immediately: "In fact, the point of my method is rather the development of discriminants specifically designed for the special purpose immediately in view than for anything like the production of a taxonomic panacea, against which you, I think rightly, offer some words of warning."[46] Still, he viewed his linear discriminant function as "a step towards objectivity in judgments based on multiple observations."[47]

Anderson shared this view and in a later letter pointed to Fisher's own discussion of "the naive assumptions of the founders of biometry and on the relation of biometry to the underlying taxonomic problems" as "a perfect expression of what I have felt, and tried to say. It comes as a relief therefore to have these thoughts and feeling made articulate in such a beautiful fashion."[48]

43. Anderson to Fisher, 6 Feb 1936, Fisher Papers.
44. Anderson to Fisher, 11 May 1936, Fisher Papers.
45. Anderson to Fisher, 28 Jan 1937, Fisher Papers.
46. Fisher to Anderson, 10 Feb 1937, Fisher Papers.
47. Ibid.
48. Anderson to Fisher, 7 Jun 7, 1937, Fisher Papers.

Fisher had recently written on a different subject:

> It will have occurred to the reader who has followed this article so far that the science of craniometry must be in a very primitive condition, if it is still concerned with clarifying its fundamental notions at the stage we have been discussing. It seems, indeed, undoubtedly true that the theoretical concepts developed in the subject have lagged far behind the sheer mass of observational material which has been accumulated.

Whether this was attributable, as Fisher offered, to the "sheer magnitude" of the task, a misplaced optimism that the real difficulties could be overcome, or "to an unconscious minimizing of those difficulties," he acknowledged that even statistically reliable measures "does not [*sic*] in itself bring us nearer the stage of recognising which, if any, among our measurements are of the greatest and which are of the least value as indicators of racial affinity," in this case.[49]

Whatever caveats he might offer about getting the biology right, Anderson nonetheless accepted as valid the assumptions behind the search for racial differences by physical anthropologists. But, with an eye toward honing the tools for such an inquiry, he wrote Fisher "to propose, however cheeky it may seem to suggest it, is that the data [on the leaves of two different oak species] might be worked up jointly as an example of the way in which species and races differ." Anderson's view was that understanding just how species differ was the starting point if anthropology or any science was to get anywhere, and so that "appropriate mathematical tools can be adjusted to deal with the problem. Your index is one such tool, but it still does not take in every kind of difference which exists in the data it uses."[50] Once again, for Anderson the limitations of even as powerful a tool as Fisher's linear discriminant function were at the fore. Fisher was interested in Anderson's proposal, and though it came to naught, he pushed Anderson to clarify the problem for both of them.

Anderson wrote back on December 31, 1938, to acknowledge that he was heading into unexplored territory and so was necessarily vague about what he was after. He saw the problem as having two components. The first was "How do species differ?," by which he meant both theoretically (as genotypes influence phenotypes in general) and practically (in terms of such diagnostic specifics of the oak leaves). Finding the most efficient mathematical tools for

49. Ronald A. Fisher, "The Coefficient of Racial Likeness' and the Future of Craniometry," *Journal of the Royal Anthropological Institute* 66 (1936): 57–63.

50. Anderson to Fisher, 30 Nov 1938, Fisher Papers.

treating these measurements is the second component of the project. The scope was vast—no less than remaking biometry. As Anderson wrote,

> In other words I propose to start at the very beginning of biometric problems; where biometry should have started had it been set going by a logical mind. So far as I can see Pearson (and the Pearsonians) started out on the naive assumption that if they just made enough accurate measurements and treated them neatly it would all make sense. Your methods have been a tremendous step forward because you have asked what questions were to be answered with the measurements and then have devised the most efficient means of using them . . . [I]n my opinion even you are attacking the problem from the wrong end (It may be of course that when we work it out logically from the front that no better means of selecting what to measure can be devised . . . but I don't think so).[51]

They collaborated no further on this project as the advent of World War II and other scientific projects are much more likely explanations for their mutual shift of attention than Anderson's self-described cheekiness. Still, Anderson's attempt here to create a non-Pearsonian biometrics on a firm biological footing captures his caution about mathematical tools and their application to his scientific problems.

At the same time, he worked closely with the finest mathematically minded biologists he could. His collaboration with Fisher unfolded virtually side-by-side with his discussion with Sewall Wright.

STUDYING WITH SEWALL WRIGHT IN 1933

While Anderson was at the Arnold Arboretum in the early 1930s, William E. Castle shared a draft of Sewall Wright's paper, "General, Group, and Special Size Factors."[52] Anderson was "very much interested by it," but told Wright:

> My chief difficulty in reading it is that like many naturalists I am visually minded. A string of numbers means very little to me until I have turned it into a diagram. I inclose [*sic*], on a green sheet, one of the diagrams I made when reading your paper. You will see that the connecting lines are roughly proportional in width to the values of A′. I send it on to you, not because I think it would help you, or any other statistician, but because I am sure that

51. Anderson to Fisher, 31 Dec 1938, Fisher Papers.
52. Wright, "General, Group" (ref. 22).

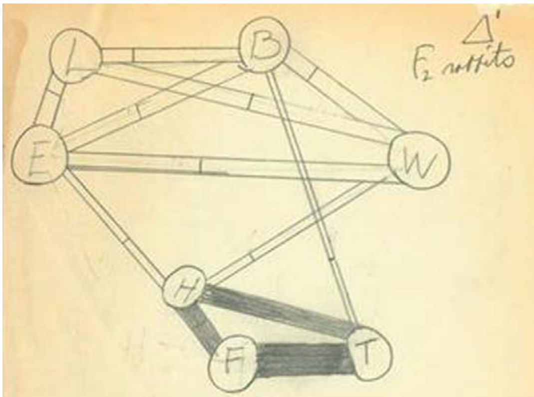

FIG. 2. Anderson's drawing, visualizing findings of Wright (ref. 22). *Source:* Anderson to Wright, 11 Mar 1932, Wright Papers. Used with permission of the American Philosophical Society.

> something of the sort would be a help to those biologists who think in pictures.
>
> Is there any mathematically ethical fashion by which the variance of the F-1 and F-2, in your last tables could be directly compared to the proportion measured?[53]

In this, his first contact with Wright, Anderson identified himself as a naturalist for whom images are more meaningful than tables of numbers. His figure (Fig. 2) was as much a visual metaphor as a precise rendering; indeed he worried to Wright that "this has been written very hurriedly and as I read it over it sounds positively impertinent. I assure you it is not meant in that way."[54]

Indeed, here Anderson was perfectly earnest while at the same time seeking to make the profound tools of mathematics applicable to his own work.

Wright responded on March 31, 1932, commenting, "I appreciate the advantage of a graphical presentation of the results and spent a good deal of time trying to work one out. There seems to be no accurate way of doing this, however."[55] Wright pointed out numerous mathematical objections to Anderson's (or anyone's) attempt to do so because he was committed to finding "an *accurate* way of doing this." Anderson, however, at an early stage

53. Anderson to Wright, 11 Mar 1932, Wright Papers.
54. Ibid.
55. Wright to Anderson, 31 Mar 1932, Wright Papers.

of his mathematical education, sought something he saw as slightly different, "a mathematically *ethical* fashion" by which Wright's data could be "directly compared and the proportion measured."[56]

He spent January and February 1933 at the University of Chicago, learning from Wright and others that, for mathematicians, accuracy defined the ethical. Still, Anderson's aims were never quite the same as the mathematicians'; he sought tools, like the ideographs for *Iris*, that were objective and accurate while capturing the taxonomist's concern with morphological form.

Returning home after a very productive term, Anderson wrote appreciatively back to Wright:

> As I take up again some of the problems I was working with before I went to Chicago I begin to realize how much I learned in my short stay with you. My primary purpose in going there was to bridge the gap between my fragmentary mathematics and the useful statistical tools which I knew about but did not understand. That objective was more than realized. The sight of prf-1 or the mystical x2, or "e" with a mystical exponent now greet me on the page with a happy flash of recognition.[57]

It had been a very useful apprenticeship that advanced Anderson's science. Still, he remained committed to his visual, natural history/taxonomic approach.

He applied it to his discussion of "A Quantitative Comparison of Specific and Generic Differences in the *Betulaceae*"[58] using another iteration of the graph he had sent Wright in that first letter. Because the criteria for morphological and taxonomical distinctions are disadvantaged for being neither objective or commensurable, he offered this paper as "a crude first attempt" to resolve things by "carry[ing] on consciously and mathematically the same sort of process which with a good naturalist is subconscious and unmathematical," by distinguishing a species by "the total impression of a very large number of variables."[59]

In this exercise, Anderson scored and indexed six characteristics important in distinguishing among members of the birch family. By aggregating the differences between taxa he established "the magnitude of difference [as] the length of a line between two points each of which is defined in six dimensions."[60] That index is mathematically sufficient, as he learned from Wright, but he still offered

56. Anderson to Wright, 11 Mar 1932, Wright Papers.
57. Anderson to Wright, 7 Mar 1933, Wright Papers.
58. Anderson and Abbe, "A Quantitative Comparison" (ref. 39).
59. Ibid., 45.
60. Ibid., 47.

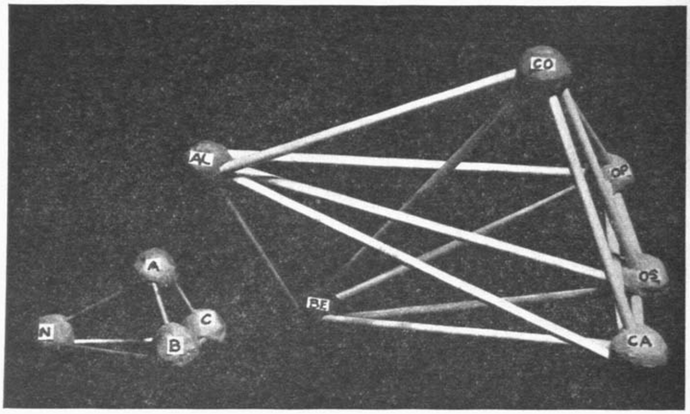

FIGURE 1. Models showing the comparative distances between the genera of the Betulaceae and the sub-sections of the genus *Betula*. For the larger model, AL, *Alnus;* CA, *Carpinus;* CO, *Corylus;* OP, *Ostryopsis;* OS, *Ostrya*. For the smaller model, N, §§*Nanae;* A, §§*Albae;* B, §§*Acuminatae;* C, §§*Costatae*.

FIG. 3. Spatial model of specific and generic differences as distances. *Source:* Anderson and Abbe, "A Quantitative Comparison" (ref. 39), 48.

a three-dimensional geometric model[61] that nonetheless does not reflect phylogenetic development. He offered this visual metaphor as a bridge between the old qualitative methods of taxonomy and these emerging quantitative ones (Fig. 3).

When he reopened his correspondence with Wright in 1938 with a discussion of genetic linkage and hybridization, he thanked Wright for "[y]our interesting computations [which] I accept as I do the pleasant sunshine this afternoon; something which I appreciate and value even though I do not understand it."[62] Anderson had sent Wright a draft of his paper, "The hindrance to gene recombination imposed by linkage: an estimate of its total magnitude."[63] He considered this as "my most important and best written technical paper [which] has won over every mathematical geneticist who set out to disprove my conclusions."[64] The upshot of the paper was "that the total cohesive effect of linkage [of genes on the same chromosome] in the germplasm comes out to figures of interstellar magnitude when you try to estimate them."

61. Ibid., Figure 1, 48.

62. Anderson to Wright, 14Oct 1938, Wright Papers.

63. Edgar Anderson, "The Hindrance to Gene Recombination Imposed by Linkage: An Estimate of Its Total Magnitude," *American Naturalist* 73 (1939): 185–88.

64. Anderson, 1967 typescript, Anderson Papers.

Wright's initial response was limited. He did think Anderson was correct about linkage as a hindrance to recombination at least as far as the F2 generation, but it "falls off rapidly after a certain number of generation and I doubt whether linkage (by itself) has very much significance in terms of geological time."[65]

Although Anderson demurred that he was not trying to work out the problem beyond the F2 generation, nonetheless his framework for raising linkage and recombination is part of his examination into the role of hybridization as an evolutionary mechanism. Thus, he hoped to draw Wright into looking at his problem:

> There are so many ifs and and's and but's about the whole matter that I wish you would work out the evolutionary mathematics of hybridization as thoroughly as you have done that of mutation. As has doubtless occurred to you, evolution by hybridization is much like that by mutation, hybridization rate replacing mutation rate. The chief difference is that whereas mutation was at random from gene to gene, hybridization brings in whole blocks of genes. Furthermore, these blocks have been selected to go together.[66]

Wright was skeptical about the role of hybridization, noting the strong selection pressures against survival of such crosses. He did envision the theoretical possibility of a stable zone of contact between species (or better, he thought, subspecies) that would result in a 100% hybrid population. In those cases of geographic separation leading to very occasional crossing, such hybridization is "very closely allied to mutation in this case."[67]

Two years, to the day, Wright told Anderson that, yes, hybridization (or even migration) could be substituted for mutation in his coefficient. Thus, "The principal difference between my views and Fisher's is that while he bases evolution exclusively on selection pressure with substantially panmictic groups I have stressed the importance (but not exclusive importance) of breeding structure."[68]

Anderson was heartened by this letter, "For once you are not talking over my head." He continued to push for "something much more far-reaching" because "[t]here are certainly groups of species in which hybrid segments are much more important than gene mutations as material for selection." He was

65. Wright to Anderson, 13 Oct 1938, Wright Papers.
66. Anderson to Wright, 14 Oct 1938, Wright Papers.
67. Wright to Anderson, 17 Oct 1938, Wright Papers.
68. Wright to Anderson, 17 Oct 1940, Wright Papers.

referring to the then recent findings of Mangelsdorf and Reeves on the relationships between *Tripsacum* and maize,[69] as well as his opinion that "most of your guinea pig genes look that way to me." He concluded his letter by explaining what he saw as happening in nature:

> As for the formulae what I would like to see is the logical deductions from the hypothesis that a population receives its diversity from hybridization. This means that all the genes come in together in one jump and gradually break up. In a few generations they would be following regular Mendelian rules. I suspect that the chief effect of the other isolators (besides linkage) would be to release them more slowly (inversions) and to preserve some of them in the heterozygous condition.[70]

Wright ended this phase of the discussion applauding Anderson's call for "a more far-reaching mathematical theory dealing with the consequences of hybridization and with the behavior of chromosomes and chromosome segments rather than genes." But he cautioned that there are limitations, particularly from the point of view of statistics.

> A statistical theory can only be constructed for recurrent phenomena. I suspect that our evolutionary theory will always be a combination of rather heterogeneous elements—statistical treatment of the recurrent aspects, verbal treatment of those aspects that depend on unique events—allopolyploidy, for example. The sort of thing that you are writing about seems to come too close to this latter category to look very promising from the statistical standpoint. However, experimental study may bring to light general principles that give better premises for mathematical elaboration than are visible now.[71]

Anderson was undaunted though and continued to develop his ideas about hybridization as an evolutionary mechanism. The point was to study nature first and then see if the necessary mathematical tools could be developed. It was through this specific approach that Anderson expanded and honed his mathematical and graphical thinking.

69. Paul Mangelsdorf and R. G. Reeves, "Hybridization of Maize, Tripsacum, and Euchlaena," *Journal of Heredity* 22 (1931): 329–43.

70. Anderson to Wright, 18 Oct 1940, Wright Papers.

71. Wright to Anderson, 31 Oct 1940, Wright Papers.

INTROGRESSIVE HYBRIDIZATION (1949): CHAPTER 6, SPECIAL TOOLS FOR THE STUDY OF INTROGRESSION

Introgressive hybridization—the repeated backcrossing of hybrid populations to parental lineages, and thereby introducing linked traits for natural selection to work on—was Anderson's signature contribution to the rich discussions on evolutionary theory in the second quarter of the twentieth century. It was based on the concepts of genetic linkage and recombination he discussed with Sewall Wright. Whereas, in his view, the majority of his colleagues (Fisher, Wright, and Theodosius Dobzahnsky) used the principles of genetics as a starting point to explain variation in nature, he sought the opposite, putting "taxonomy in the service of genetics," asking what observed variation posed for the understanding of genetic process.[72] He built on the approach of Karl Pearson and the biometricians "to study species differences and population differences so effectively and so exactly that the effects of the evolutionary forces actually at work could be measured in natural populations." To do so, though, he had to recognize and correct the flawed assumption of the biometricians, that "species vary at random." His graphical tools were developed for both observation and analysis of how species actually vary. Powerful as he deemed them to be, they were subordinated to the biological reality they were designed to study.

Anderson devoted Chapter 6 of his technical book *Introgressive Hybridization*[73] to "Special Tools for the Study of Introgression." The chapter's aim was to outline "the special techniques that have been developed for apprehending introgression in the field." He meant them to be used in conjunction with more traditional techniques (transplant experiments, cytology, experimental repetition of suspected crosses), but those techniques were limited to cases when the parental species were known or suspected. They could not be used predictively. These techniques could "apprehend introgression in the field."[74]

Since we have difficulty judging more than one variable at a time, he sought methods that would allow him to capture several variations at once. We have already seen how the ideographs for *Iris* were derived from the same data that Fisher's linear discriminant function rendered numerically, making graphically

72. Anderson, "Introgressive Hybridization" (ref. 9). Kleinman, "Bringing Taxonomy to the Service of Genetics" (ref. 11).

73. Anderson, *Introgressive Hybridization* (ref. 15), 81–101.

74. Ibid., 81–82.

visible the variation that both tools captured. But, particularly, before Fisher, "the methods of conventional biometry are laborious and inefficient," especially "for exploring relationships between groups of characters, particularly when we do not know in advance the general nature of the relationship."[75]

Anderson predicated these tools of "*polygraphic analysis*" (his emphasis) on "exact, objective, and verifiably accurate measures." Only with such data first could he suggest which tool was right for the job. Besides the ideograph discussed above, he used hybrid indices (based on a scoring of traits), radiate indicators (to capture traits in populations), and standardized photographs that could be incorporated into inclusive herbaria.[76]

But Anderson focused on and used adaptations of the Cartesian grid, the scatter diagram (arraying two factors, one for each axis, on the grid, rendering correlation coefficients visible), and even more, the pictorialized scatter diagram that could convey even more information by making the points on the grid convey even more information.[77] His students called this pictorialized version a "whisker diagram," but he gave it a more formal name, "metroglyph," when he used these scatter diagrams to analyze trends in British poetry.[78]

The power of the scatter diagrams, demonstrating, in Anderson's view, the validity of introgressive hybridization as a concept, was clearest when he used it to "apprehend introgression in the field" with *Oxytropsis* (locoweed) species in Dory Hill, Colorado (Fig. 4).

They varied in many ways at once—leaf width, plant height, "hairs" on their stems, flower color, and petal shape. He then plotted leaf width on the y-axis of a Cartesian grid and plant height on the x-axis. If the plant was red, its plot point had a long line or "whisker" at 10 o'clock; if pink, the "whisker" at that position was shorter; and if white, there was no whisker. The hairiness of stems were depicted with "whiskers" at the 12 o'clock position and petal notches similarly at 2 o'clock. Arraying these data on a grid helped Anderson see that shorter, narrow-leafed white plants without hairs or notches would cross with taller red plants with wider leaves, hairier stems, and notches to create mostly pink, medium-height plants with hairiness, leaf width, and notches variably intermediate. The data that captured these variations were made visible and so could be approached in a way that was familiar and

75. Ibid., 82.

76. Ibid., 81–101.

77. Ibid., 83–86.

78. Edgar Anderson, "A Botanist Looks at Poetry," *University of Michigan Quarterly Review* 4 (1965): 177–84.

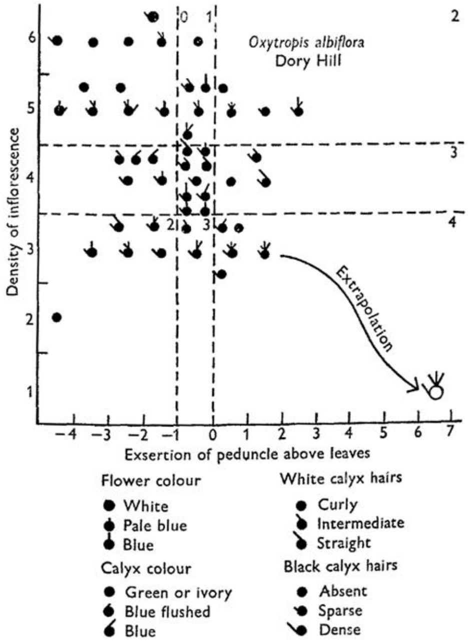

Fig. 1. Pictorial scatter diagram showing interrelationships of six variables in a population of fifty plants of *Oxytropis*.

FIG. 4. Scatter diagram, also called "whisker diagram." *Source:* Anderson, "Introgressive Hybridization" (ref. 9), 285.

comfortable for a skilled field observer to use. Just as he could in the field with the plants themselves, Anderson could look at one of these graphs and see a generalized version of the dynamics of the variation in the population he had observed.

Anderson came to be able to map the variation in a population and infer a possible hybridizer that would bring such variation to the species he was studying. In the locoweed case, he inferred sight unseen the hybridizing source of variation in the primary population. From the introgressed traits, he predicted the kind of plant that could be responsible, found it in the local flora and guidebooks, and only then found it nearby.[79]

This tool was important enough that he included a generalized explanation in his popular book, *Plants, Man, and Life* with a chapter called "How to

79. Anderson, "Introgressive Hybridization" (ref. 9); "The Role of Hybridization in Evolution," in *This is Life: Essays in Modern Biology*, eds. Willis H. Johnson and William C. Steere (New York: Holt, Rinehart, and Winston, 1962) 287–314.

Measure an Avocado."[80] There variations in *Planta alba* and *P. rubra* were built up and arrayed on the grid.

SUMMARIZING THE GRAPHICAL LESSONS: PAPERS FROM 1954 AND 1956

The fundamentals of Anderson's visual tools merited popular discussion in *Plants, Man, and Life* and in his report to *MBG Bulletin* readers on visiting mathematicians. But another, different sign of their importance was in his reflections on this work for professional colleagues.

Along with a subsequent paper published after his visit to Princeton University to work with John W. Tukey,[81] Anderson captured his fundamental conclusions about the graphical depiction of mathematical data from natural history in "Efficient and Inefficient Methods of Measuring Specific Differences"[82] and "Natural History, Statistics, and Applied Mathematics."[83] Together these papers recapitulate an overall methodology while applying and describing his specific techniques.

He summarized his approach to readers of the *MBG Bulletin* in his 1956 article on his interactions with mathematics departments:

> If the basic facts in any particular problem are numbers (let us say the yield in bushels of a series of wheat varieties) then statistics does a good job. If the basic facts are patterns (as for instance the intricate and complicated kinds of differences between two species), then at best statistics is inefficient; at the worst it gives the wrong answer but veils its failure in impressive technical language.[84]

He mentioned the two professional articles named above in this popular summary.

In the 1954 article, "Efficient and Inefficient Methods of Measuring Specific Differences," Anderson rated the efficiency of measuring species difference methods by the number of sense impressions in play. Interestingly, he noted that early biometricians started out with his kind of problem, those like species

80. Anderson, *Plants, Man, and Life* (ref. 16), 89–106.

81. Anderson, "A Semigraphical Method" (ref. 14); reprinted with a new note in *Technometrics* 2 (1960): 387–91.

82. Anderson, "Efficient and Inefficient Methods" (ref. 12).

83. Anderson, "Natural History" (ref. 13).

84. Anderson, "Why Botanists Visit" (ref. 2), 150.

differences with many sense impressions, but "it was not until efficient methods for dealing with the former type [fewer sense impressions as with agronomic problems] were developed that biometry produced really practical results or received any general recognition."[85]

His contribution then was to provide an approach to deal with multiple sense impressions "to return to Pearson's original task of analyzing evolution." Unlike wheat yields, these problems require both a biological grounding in what is relevant and important to measure and a technique to deal with "multiple attribute statistics."[86] Returning to the dictum from his 1953 review article on "Introgressive Hybridization," he underscored that species do not vary at random, which he demonstrated using ideographs of *Uvularia* species to create first a hybrid index and then a pictorialized scatter diagram.[87] In essence, he recapitulated here in practical application the "special tools for the study of introgression" that he had presented in Chapter 6 of *Introgressive Hybridization.*

But he saw theoretical implications and explicated them by the seemingly mundane problem of distinguishing between apples and oranges, that paradigm of what we know enough about not to compare. Whereas we could elaborately seek measures (diameter, perhaps? height? width? indentation at stem end?) that would distinguish them biometrically, Anderson proposed that, even if one had never seen either fruit, one could use one's multiple sense impressions and minimal guidance (a standardized photograph, perhaps) to perform the task. Further, though, he argued that such a judgment was more than subjective. If multiple operators using the same technique come up with the same judgment, the aggregate of those subjective judgments reflects an objective standard.[88] Anderson offered his techniques as a way of providing an objective and efficient method of measuring species differences.

His 1956 article "Natural History, Statistics, and Applied Mathematics" stands as a general statement not only on the topics in the title but indeed on his approach to science. Above all, he advanced a perspective and approach to capturing "pattern data," the very substance of natural history. Large, complex problems with multiple factors are what scientists confront initially; "Everything looks chaotic at first but we do not live in a chaotic universe.

85. Anderson, "Efficient and Inefficient Methods" (ref. 12), 93.
86. Ibid., 93.
87. Ibid., 94–99.
88. Ibid., 103–04.

There may be confusion in our minds but there is no chaos in the way the world is running." A training in natural history can arm an investigator with an approach that bounces between observation and hypothesis in an effort to discern patterns. We may need to begin with separate individual readings ("pointer data"), but Anderson cautioned against too much reliance on them "until we know what kind of problem we are up against. . . . Precision has little advantage until we have enough understanding to use precise analysis."[89]

Natural history already had a diminished reputation when Anderson was writing, yet he well knew its importance, including in the field he trained, genetics. "Those who have not actually participated in such routine genetic chores as the making of linkage maps, seldom realize how qualitative the basic data of Genetics can be." Indeed, he attributed rapid advance in the field because "it was using pattern data with their broad observational basis, but using them with great precision." His example was eye mutants in *Drosophila*, which were in fact separated by pattern observations and not "on a truly objective, statistical basis!"[90]

Anderson defended keeping natural history in the curriculum precisely because that instinct cannot be replaced mechanically. Still, he knew there was important work to be done: "If however, we are to bring Natural History to its full stature as a science, we need to learn how to use its pattern data with greater precision; and even more to learn how to think about them analytically. Greater precision is no great benefit unless it leads to sharper analysis."[91]

Yet, Anderson was skeptical of the siren call of statistical precision and invoked the South African statistician B. De Loor, who noted, "Statistics in not interested in the individual," whereas the applied mathematicians Anderson tried to influence were interested in "the significant individual." That distinction was significant for Anderson, and it was this reorientation he argued for so forcefully for in this essay.[92]

An important part of the argument made by the long arc of this project was the techniques that could capture the patterns of natural history, not just for observation but for analysis, and aggregate those subjective observations into something usefully objective.

89. Anderson, "Natural History" (ref. 13), 883.
90. Ibid., 883–84.
91. Ibid., 884.
92. Ibid., 886 and 887.

WHY EDGAR ANDERSON VISITED JOHN TUKEY AT PRINCETON IN 1957

As Edgar Anderson ended a short (1954–1956), frustrating tenure as MGB Director, he had a Guggenheim Fellowship that allowed him to return to his research projects, including mathematics, while allowing the Garden leadership to reorganize in his absence. He was actually quite sanguine about giving up the Directorship, telling his friend Paul Allen of the Escuela Agricola Panamericana, "This will begin in January when I am going as a guest of the Mathematics Department at Princeton. My methods of studying evolution are beginning to be widely used in fields as remote as economics and psychology, and they raise some pretty fundamental issues in the fields of applied mathematics."[93]

In his host John W. Tukey, Anderson found a kindred spirit. Tukey was intensely curious about many subjects, commenting, "The best thing about being a statistician is that you get to play in everyone's backyard."[94] Throughout his career, he offered his expertise on such subjects as the statistics of the Kinsey sex surveys, the United States Census, and the impact of aerosol sprays on the ozone layer. In the 1940s, he coined the computer term "bit" to stand for binary digit, and a decade later, contrasted computer programs as "software" to the hardware of the actual machines. Besides his work at Princeton, he maintained a separate office at American Telephone and Telegraph's Bell Laboratories, where he worked as a researcher.

Tukey was a legendary character around Princeton. He typically listened to seminar presentations while attending to his own paperwork with slide rule in hand. Yet, his comments were unerringly acute. One colleague celebrated a seminar as a particular success because Tukey put down his work and listened.[95] Tukey also handled room scheduling for the school, fitting class sizes and times together in his head in the days before computers.

Tukey dedicated his 1977 book *Exploratory Data Analysis*, in part, to Anderson as one "from whom the author learned much that could not have been learned elsewhere."[96] They shared a commitment to finding ways of displaying data that could suggest new ways of interpreting them, rather than merely fitting them into existing theories and theorems.

93. Anderson to Allen, 20 Nov 1956, Anderson Papers.

94. Cited in Tukey's *New York Times* obituary, 28 Jul 2000, http://www.nytimes.com/2000/07/28/us/john-tukey-85-statistician-coined-the-word-software.html

95. Author's interview with John Tyler Bonner, 5 Apr 2001, Princeton, NJ.

96. Tukey, *Exploratory Data Analysis* (ref. 3).

So, with Tukey, Anderson adapted the pictorialized scatter diagram he used to explain introgression into what he called a "metroglyph." The basic diagram had been refined for use not only in scatter diagrams but also in maps and indices to show sequences in time, space, or development. In the upper left hand corner of such a figure, four individuals could be categorized on the basis of five different qualities. The upper right hand corner of the figure then would be the template for each glyph, so that the four individual glyphs are slightly below the center of the right side. The numbers underneath these glyphs serve as index numbers based on assigning no points for a low ranking for each quality, one point for medium, and two points for a high ranking. Thus, the data can be arranged in the index in the lower right hand quality or in a pictorialized scatter diagram as in the lower left hand comer of the figure.

As Anderson pointed out, "In working with these glyphs, it has gradually been realized that they are a device for helping the eye to aid the mind. For maximum efficiency it is best to make some concessions to the eye. The mind can be trained to adjust itself more readily than can the eye."[97] Among the concessions to the eye he recommended were limiting the number of rays, keeping them in the upper half of each glyph, and making the length of the medium and high rays distinct.

During this sabbatical, Anderson gave formal seminars to the mathematics, biology, psychology, and geology departments as well as to a special Population Research Group. He consulted with students and staff in archaeology, linguistics, psychology, biology, and horticulture. Finally, according to John Tyler Bonner, Fisher took advantage of Anderson's presence in Princeton to visit for a three-way conversation with him and Tukey.

Anderson enjoyed the opportunity Tukey and Princeton afforded him to do sustained productive work after the frustrations of administrative responsibilities at the MBG. Whatever bitterness that might have existed was quickly washed away in stimulating activity. His enthusiasm was reflected back to him in a letter the physicist and mathematician Edwin Bidwell Wilson wrote to him at Princeton. After reviewing some recent literature touching on statistics, Wilson ended his letter with this comment: "I am glad you are having a good time. I believe you are working at an important problem which our modem mechanistic or quantum-mechanical philosophers have overlooked. The point

97. Anderson, "A Semigraphical Method" (ref. 14); reprinted with a new note in *Technometrics* 2 (1960): 387–91 on 389.

of view of the naturalist must be kept alive and to keep it alive we have to have a philosophy of it which may sometimes incorporated into philosophy."[98]

Anderson developed this approach and these techniques over a career that synthesized careful field work with a rigorous interaction with the leading workers who took a mathematical approach to biological problems—R. A. Fisher, Sewall Wright, and John Tukey. Its elements were to focus on the specific facts and patterns of biology with visual tools that would not be elevated above and therefore alienated from those facts and patterns. That was the point of view of the naturalist, which Wilson complimented Anderson for keeping alive and incorporating into a rounded philosophy of science.

CONCLUSION

Edgar Anderson's graphical techniques were a thoroughly integral part of an approach to science that manifested itself in his important evolutionary work on introgressive hybridization and the more general commitment to the power of the pattern data of Natural History.

It was of broader applicability, as demonstrated in his seminars at Princeton in 1957 with colleagues in mathematics, biology, psychology, and geology, as well as with an interdisciplinary group interested in population research.

He left to them the task of knowing the substance of their fields so that the techniques could illuminate and not obscure the work at hand, just as he carefully applied them in his own work on particular taxa (*Iris*) or particular evolutionary mechanisms (introgressive hybridization) in his efforts to understand "Evolution with a big E."

His circumspection about these techniques is perhaps best shown in the conclusion of his charming essay on applying metroglyphs to British poetry.

> If, by diagramming several things at once, one could get at some of the essential differences in style, then from the surface of English poetry one might study some of the important surges of thought and feeling which lie beneath the surface. I should wish, if such studies were ever to be made, that they might be done by men of imagination and insight, with a real affection for English poetry. I have an uncomfortable feeling that this might not always be so. I have at times been distressed to see my methods for analyzing species-differences in the hands of unimaginative men, who had little feeling

98. Wilson to Anderson, 22 Feb 1957, Tukey Papers.

> for plants. I might be even more uncomfortable to find such minds starting out after William Blake or John Donne, armed with my sharp new techniques. But as Professor Miles remarked to me when I shared my concern with her, "Plants and poems are tough."[99]

Tough as they are—poems, plants, indeed the natural world as a whole—each has an integrity that Anderson approached with graphical techniques born of an understanding grounded in a deep appreciation of each phenomenon that he sought to study in all its individual fullness and on its own terms.

ACKNOWLEDGEMENTS

Throughout this long-term project on Edgar Anderson, the support and encouragement of Gar Allen, Betty Smocovitis, and Peter Hoch have been indispensable. I am glad that I have more recently cajoled Nuala Caomhanach into this group. I deeply appreciate the insights of my reviewers whose suggestions I have benefitted from greatly. Terrence Cotton of the Webster University Library staff at the time helped with formatting and copy editing.

99. Anderson, "A Botanist Looks at Poetry" (ref. 78), 184.