Project TALENT:
A Short History of a
Long Project

MARION F. SHAYCOFT

AMERICAN INSTITUTES FOR RESEARCH
Post Office Box 1113 / Palo Alto, California 94302
Project TALENT: A Short History of a L-o-o-n-g Project

MARION F. SHAYCOFT

This work was done under Grant No. NIE-G-74-0003 from the National Institute of Education, Department of Health, Education, and Welfare. The content does not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement should be inferred.

AMERICAN INSTITUTES FOR RESEARCH
PALO ALTO, CALIFORNIA
1977
A few words about the nature and purpose of this book

This is an "internal" history of Project TALENT. In other words, it is not about the nearly half a million people "out there" to whom we administered two days worth of tests when they were high school students in 1960 and who are now in their early 30's. No, it is not about them and what has happened to them in the years since 1960; there are already a great many massive research reports telling about that. What the "internal history" focuses on primarily is what happened at places like:

- 410 Amberson Avenue, Pittsburgh
- 6135 Kansas Avenue, Washington
- 1808 Adams Mill Road, Washington
- 200 South Craig Street, Pittsburgh
- 135 North Bellefield Avenue, Pittsburgh
- 1791 Arastradero Road, Palo Alto

(where Project TALENT has had its headquarters at various times), and at places in Washington like:

- 400 Maryland Avenue S.W.
- 1200 19th Street N.W.

(where the Office of Education and the National Institute of Education respectively have their headquarters).

In the 20 years that have elapsed since the project was just a gleam in John Flanagan's eye, a lot has happened and we have learned a lot. If it is true that one learns by experience, then one should learn a lot in a long-term longitudinal study--just for the very reason that the study does last so long and has so many facets. One of the ideas behind this History is that it would be nice if future researchers--particularly researchers doing large-scale longitudinal studies--could learn from our experience.
In a period as long as 20 years, it is inevitable that we would have made some mistakes. But we must have done many things right, too; otherwise the project would have collapsed and died long ago. In this History then, we are trying to take a long retrospective look, so that we can tell our readers about our errors and our successes, in the hopes that future researchers will avoid the former and may see fit to emulate the latter.

A literary genre known facetiously as the "HIBK murder mystery" was popular at one time. This was a murder story written in the first person singular, in which the narrator, usually an amateur detective, frequently started statements with "Had I but known [HIBK] then what I know now..." I mention this because the present History, though not a murder mystery, also partakes, to some extent, of the HIBK feature, at least by implication, every time it contains a discussion of errors made, or things done in less than the optimal way.

Inevitably, in discussing what was done right and what could have been done better, a lot of personal opinion must enter the picture. That is an inherent feature of a History such as this. Whenever opinions are expressed in this volume, they are my opinions unless explicitly attributed to someone else. This is, as someone who read the manuscript commented, a "personal document." Given the nature of the assignment and the function of the History, I think that that is inevitable. And some of these opinions expressed in this "personal document" may be wrong; almost certainly some of them are wrong, since nobody is infallible. But since I wouldn't want my errors attributed to someone else, I repeat: the opinions I have expressed here are mine. Other people may or may not share them. Some readers may be bothered by the fact that some of the opinions expressed are not exactly dispassionate, but I know of no way that it could be avoided.

And if there is an apparent undue emphasis on my own role in the project, I apologize for that. I certainly have tried not to slight anyone else's role, or to give the impression that my role has been more important than it
really was. Thus if there seems to be an undue emphasis on my role it is because that is what I know best; and if there is an undue emphasis on my own viewpoint, it is because that is what I understand best.

Nevertheless it is at least a little embarrassing that the Technical Appendices turned out to consist almost exclusively of my own writings. This should not suggest that the author of this History is the only staff member to have written papers or presented papers at professional meetings. Quite the contrary! As indicated in Chapter 21, over the years more than 100 papers have been presented by staff members at professional meetings and a very large number of staff members have been among their authors. But the author of this History has specialized in methodological problems to a somewhat greater degree than most of the other staff members and therefore has had more occasion to write papers with methodological implications. The Technical Appendices are intended largely as a technical guide to various rather special aspects of research with TALENT data, sometimes requiring special methodological twists. For a long time I was virtually the only one on the project staff working on problems of that sort rather than almost exclusively on substantive problems (as were most of the other members of the research staff). And that is how it happens that the author of this History is also the sole author of most of the Appendices.

The period covered

A History needs a cut-off point. This History covers the events occurring from the inception of TALENT in 1957 through the final processing of the 11-year follow-up data in early 1977. But readers (and authors) may find a cut-off date frustrating when they know that things must have happened after that date. Therefore a "Postscript" has been appended after the last chapter, to bring things up to date.

How to find your way around in this book

I suspect that this volume will tell some readers more about Project
TALENT than they are interested in learning. Since I have no wish to burden the reader with unwanted information, there is a fairly detailed Table of Contents. This, together with the usual "List of Tables" and "List of Figures" will, it is hoped, let the reader with specific information needs find what he is particularly interested in at a particular time.

There is also a glossary of acronyms and abbreviations which the reader may find helpful if he stumbles on a mysterious jumble of letters, such as AFPTRC or NCERD or T.E.A.R.S. The glossary is the last thing in this volume.

Most of the appendices are bound together in a separate volume.

A few words on nomenclature

In writing this book the author encountered a couple of slight problems of nomenclature. The first one concerned how she should refer to present and former colleagues on Project TALENT. Five options clamored for consideration: (1) full name as used officially in a professional context (e.g., John C. Flanagan, A. Carp, William V. Clemans, Dorothy S. Edwards); (2) first and last names as used conversationally to refer to the person named (e.g., John Flanagan, Al Carp, Bill Clemans, Jo Edwards); (3) last name with honorific title attached (e.g., Dr. Flanagan, Dr. Carp, Dr. Clemans, Dr. Edwards); (4) last name only (e.g., Flanagan, Carp, Clemans, Edwards); (5) first name only (e.g., John, Al, Bill, Jo). Options 2 and 5 have the advantage of informality—which fits in with the fact that this is a somewhat informal history. Option 5 corresponds to the prevalent mode of address but it can lead to ambiguity except where context helps. (Does "Bill" mean Bill Gorham? Bill Cooley? Bill Clemans? And does "Jo" really mean "Dorothy"?) Option 3 is somewhat more formal and certainly less subject to ambiguity. Option 4 corresponds to the standard mode of reference used in professional articles. If the problem posed above were a five-option multiple-choice item it would be a rotten one; it doesn't clearly have one and only one correct answer—nor even one and only one best answer.
The solution chosen was to use whichever style of reference seemed most appropriate or most convenient in the context in which it appeared. This means that almost every one of the five options was used at one point or another. The author takes what comfort she can from Ralph Waldo Emerson's well-known comment on consistency.*

That brings us to the second nomenclature problem. How should the author refer to herself? As "the author"? Or by name? Or in the first person plural? Or first person singular? Again I invoke Emerson's comment on consistency!

Marion F. Shaycoft

September 1977

* "A foolish consistency is the hobgoblin of little minds."
ACKNOWLEDGMENTS

There are two people whose contributions to this effort were such that without them this book couldn't have been—or wouldn't have been—written. Chronologically Dr. Andrew C. Porter is the first, since it was he who really stimulated the idea, when he was directing the Basic Skills Group at the National Institute of Education. (He is now on the faculty of Michigan State University.) As a result of a proposal that the author of this History submitted to the Basic Skills Group, which bore on the topic of the desirability of disseminating what we had learned in Project TALENT about running a large-scale longitudinal study and processing and analyzing the data from such a study, Andy had the idea that all of that could be incorporated in a history of Project TALENT as viewed from the inside. That idea of Andy's led, eventually, to a grant, and that grant resulted in this History. Thus it is clear that without Andy's interest and action this History would not have been written.

After the grant was awarded and I started working on the History, Andy maintained close contact and a continuing interest. He reviewed the outline for the History and made suggestions on it; I discussed new developments with him; and his interest in it has continued to this day.

In an early conference, Andy urged me to write in an informal style rather than in the style of a typical research report. I don't know whether this History has turned out in line with his hopes, but I am absolutely certain that it doesn't sound like a typical research report!

The other person whose contribution to this History has been of critical importance is John Flanagan. Certainly without the information he provided about events that occurred before the author of this History joined the Project TALENT staff and about other events that were going on soon after that, most of Chapter 2 couldn't have been written.

Furthermore throughout the writing process he has been a sounding board. I have checked recollections with him, tried out ideas on him, and given him draft copy to review. That review has proven most helpful—as were the
numerous conferences at which he functioned as an informal in-house consultant. Thus it is clear that without John's very substantial contribution this History probably couldn't have been written; or if it had been, it would have looked quite different. It would have been a lot shorter, being truncated at the front end. And the errors that John caught in reviewing the draft wouldn't have been caught by anybody.

John Dailey, who was the Project Director from 1958 to 1964 is another person who has been extremely helpful. I have conferred with him frequently, both in person and by phone, and he too has helped to fill the gaps in my information about some of the things that happened in the early days of the Project. He was kind enough to review the first half of the manuscript at my request.

Jo Edwards of the AIR staff in the Washington Office, who was also on the TALENT staff in the early days of the project, was also kind enough to answer lots of questions I threw at her about her recollections, and to review the first half of the manuscript. She provided many helpful suggestions.

One other person whom I was able to consult who was associated with the project in its early days, and later too, was Alice Scates, of the Office of Education. She was the project's first monitor, and like all the other people I have mentioned, she was very helpful.

It must be clear from the foregoing that while this History is based to a substantial degree on documented evidence, there are many parts that are dependent on memory—mine and others'. Thus it wouldn't be at all unlikely that there are some errors of fact—but I imagine there are very few since we have tried to check everything checkable, and since, reassuringly, my consultants' memories agree with each other in general.

Jack Schwille, who until recently was at the National Institute of Education, continued Andy Porter's interest in the developing "History" after Andy left NIE.

I cannot end this Acknowledgments section without saying thank you to the two staff members—Rachel Holmen and Nancy Hull—who have done the
most to convert this history from a completely handwritten partly illegible scrawl into a finished report. Rachel has been associated with this effort for about a year, and in addition to her typing skills she has provided very competent assistance in other aspects, by functioning also as a research assistant and administrative assistant. Nancy Hull entered the picture later, but she too has provided help and facilitated the effort in many ways besides typing.

My thanks to all the people mentioned above.

Of course I take responsibility for any errors of fact in this manuscript, and as I hope the Preface has made clear, if there are any errors of opinion those errors are mine alone.

Marion F. Shaycoft

September 1977
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vii</td>
</tr>
<tr>
<td>Table of contents</td>
<td>x</td>
</tr>
<tr>
<td>List of tables</td>
<td>xviii</td>
</tr>
<tr>
<td>List of figures</td>
<td>xix</td>
</tr>
</tbody>
</table>

Chapter

1. Prologue                                          | 1-1  |
   A. What Project TALENT is                         | 1-1  |
   B. Frame of reference                             | 1-2  |
   C. Current status of the Project                  | 1-3  |
   D. In the public eye!                             | 1-3  |

2. The Prenatal Stage:
   (Project TALENT Before It Was Project TALENT)     | 2-1  |
   A. Roots of the idea                              | 2-1  |
   B. The start of the search for funding            | 2-4  |
   C. The planning and the plans                     | 2-10 |
   D. Finding funding                                 | 2-12 |
   E. Bonfires and brouhahas                          | 2-13 |
       1. The Baltimore brouhaha                       | 2-14 |
       2. The Texas bonfire                            | 2-15 |

3. Issues and Decisions                              | 3-1  |
   A. The sample                                      | 3-1  |
       1. The Sampling Advisory Panel                 | 3-2  |
       2. After the Sampling Advisory Panel meeting   | 3-7  |
       3. The 16-year-olds and the 15-year-olds        | 3-9  |
       4. The 10 subsamples: an interim feature        | 3-10 |
   B. The design                                      | 3-11 |
   C. To build or buy: The TALENT Battery             | 3-12 |

4. Naming the Baby                                   | 4-1  |

5. Building the Battery                              | 5-1  |
   A. Schedule                                        | 5-1  |
   B. "State-of-the-art"                              | 5-1  |
   C. Content and organization of the battery         | 5-2  |
   D. The battery-building crew                       | 5-5  |
       1. The test builders                            | 5-5  |
       2. Questionnaire for students                   | 5-8  |
       3. School questionnaire                         | 5-8  |
   E. Tryout and item analysis                        | 5-9  |
   F. The end of the process                          | 5-9  |
   G. What we ended up with                           | 5-13 |
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Pre-&quot;Criterion-Referenced&quot; Criterion-Referenced Tests (The TALENT Domain Tests)</td>
<td>6- 1</td>
</tr>
<tr>
<td>A. The concepts</td>
<td>6- 1</td>
</tr>
<tr>
<td>B. The tests</td>
<td>6- 2</td>
</tr>
<tr>
<td>7. Execution of the Plan</td>
<td>7- 1</td>
</tr>
<tr>
<td>A. Selecting the sample</td>
<td>7- 1</td>
</tr>
<tr>
<td>B. Getting the cooperation of the selected schools</td>
<td>7- 3</td>
</tr>
<tr>
<td>C. Administering the tests</td>
<td>7-12</td>
</tr>
<tr>
<td>D. E. F. Lindquist and Measurement Research Center</td>
<td>7-19</td>
</tr>
<tr>
<td>E. Implementation</td>
<td>7-25</td>
</tr>
<tr>
<td>1. Before MRC (The &quot;TALENT Laundry&quot;)</td>
<td>7-25</td>
</tr>
<tr>
<td>2. At MRC</td>
<td>7-27</td>
</tr>
<tr>
<td>3. After MRC</td>
<td>7-27</td>
</tr>
<tr>
<td>8. With a Cast of Thousands!</td>
<td>8- 1</td>
</tr>
<tr>
<td>A. The staff</td>
<td>8- 1</td>
</tr>
<tr>
<td>1. The Principal Investigator</td>
<td>8- 5</td>
</tr>
<tr>
<td>2. Project Directors</td>
<td>8- 5</td>
</tr>
<tr>
<td>3. Associate Directors</td>
<td>8-11</td>
</tr>
<tr>
<td>4. Data Bank Directors</td>
<td>8-12</td>
</tr>
<tr>
<td>5. Other recent senior staff members</td>
<td>8-13</td>
</tr>
<tr>
<td>6. Personnel in charge of computer operations</td>
<td>8-14</td>
</tr>
<tr>
<td>7. Computer trouble-shooter task forces</td>
<td>8-16</td>
</tr>
<tr>
<td>8. Other research staff members</td>
<td>8-18</td>
</tr>
<tr>
<td>a. Washington phase</td>
<td>8-18</td>
</tr>
<tr>
<td>b. Second Pittsburgh phase</td>
<td>8-20</td>
</tr>
<tr>
<td>c. Palo Alto phase</td>
<td>8-21</td>
</tr>
<tr>
<td>d. Washington, Pittsburgh, and Palo Alto</td>
<td>8-22</td>
</tr>
<tr>
<td>B. Early advisory panels</td>
<td>8-23</td>
</tr>
<tr>
<td>C. Regional Coordinators</td>
<td>8-23</td>
</tr>
<tr>
<td>D. OE and NIE personnel</td>
<td>8-23</td>
</tr>
<tr>
<td>E. TALENT post-doctoral fellows</td>
<td>8-26</td>
</tr>
<tr>
<td>9. The Geography of the Project</td>
<td>9- 1</td>
</tr>
</tbody>
</table>
Chapter
10. Early Reports
   A. Basic reports
   B. Equating studies
   C. Subsequent studies applying the results of
      the equating

11. Returning to the Scene
   A. Designing the retest study
      1. Who
      2. Where
      3. What
      4. When
      5. How much
      6. How
   B. Processing the answer sheets
   C. Data analyses and final report
   D. Test reliability data

12. The 4% Sample

13. Tapes: Trauma, Troubles, Travail, Trials, Tribulations, and T.E.A.R.S.
   A. TALENT's tapes: Ancient history
   B. 1. Iowa City (1960)
      2. Coping with geography
      3. Zeros and blanks
      4. Building a tape master file
      5. Design of the tape master file
      6. Other instances of data loss
   B. Interim Alphabetic Directory
   C. The concept of follow-back studies
   D. Tape editing and reformatting system (T.E.A.R.S.)
      1. The NIE panel
      2. The proposal
      3. September 1975 - April 1976
      4. April 1976 - May 1977

14. Computers and Computer Programs
   A. Problems with programs
   B. Keeping up with technology
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>15. Struggling with Weights</strong>&lt;br&gt;(Not a Chapter About Weight-Lifting)</td>
<td>15-1</td>
</tr>
<tr>
<td>A. Sets of weights that have been developed</td>
<td>15-1</td>
</tr>
<tr>
<td>1. Weights for 1960 data (where no follow-up data are being used)</td>
<td>15-1</td>
</tr>
<tr>
<td>2. Weights for follow-up data</td>
<td>15-3</td>
</tr>
<tr>
<td>B. Whether and how to weight</td>
<td>15-5</td>
</tr>
<tr>
<td>C. Self-weighting: A solution to the trade-off dilemma</td>
<td>15-6</td>
</tr>
<tr>
<td><strong>16. Other Statistical Problems Encountered and Solved</strong></td>
<td>16-1</td>
</tr>
<tr>
<td>A. Sampling</td>
<td>16-1</td>
</tr>
<tr>
<td>1. How to handle schools that are in the project on a specially invited or volunteer basis</td>
<td>16-1</td>
</tr>
<tr>
<td>2. The &quot;feeder junior high school&quot; problem</td>
<td>16-2</td>
</tr>
<tr>
<td>3. Coping with complexly structured samples in computing standard errors</td>
<td>16-2</td>
</tr>
<tr>
<td>4. Is the Special Sample of nonrespondents necessary? (An empirical validation of follow-up Weight C)</td>
<td>16-2</td>
</tr>
<tr>
<td>5. Can the selection of the Special Sample be improved?</td>
<td>16-3</td>
</tr>
<tr>
<td>B. Reducing sampling error—both random and systematic</td>
<td>16-3</td>
</tr>
<tr>
<td>C. Reliability coefficients and related statistics</td>
<td>16-4</td>
</tr>
<tr>
<td>1. Reliability coefficients</td>
<td>16-4</td>
</tr>
<tr>
<td>2. Coefficients of internal precision</td>
<td>16-5</td>
</tr>
<tr>
<td>D. Inferring score distributions for an unselected age group</td>
<td>16-6</td>
</tr>
<tr>
<td>E. Problems in methodology related to career guidance</td>
<td>16-6</td>
</tr>
<tr>
<td><strong>17. Management Problems</strong></td>
<td>17-1</td>
</tr>
<tr>
<td>A. Record-keeping problems</td>
<td>17-1</td>
</tr>
<tr>
<td>1. Identification numbers and other codes</td>
<td>17-1</td>
</tr>
<tr>
<td>2. Some comments about redundancy</td>
<td>17-1</td>
</tr>
<tr>
<td>3. Confidentiality</td>
<td>17-3</td>
</tr>
<tr>
<td>4. Retrievability</td>
<td>17-3</td>
</tr>
<tr>
<td>B. Securing and maintaining cooperation</td>
<td>17-3</td>
</tr>
<tr>
<td>C. Getting funding (the cliff-hanger serial)</td>
<td>17-5</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>17. Management Problems (continued)</td>
<td></td>
</tr>
<tr>
<td>D. Problems</td>
<td>17-9</td>
</tr>
<tr>
<td>1. Turnover</td>
<td>17-11</td>
</tr>
<tr>
<td>2. Personnel selection</td>
<td>17-11</td>
</tr>
<tr>
<td>E. Management problems related to computer and computer-related operations</td>
<td>17-12</td>
</tr>
<tr>
<td>1. Management of keypunching operations</td>
<td>17-12</td>
</tr>
<tr>
<td>2. Coordination between the research staff and the computer staff</td>
<td>17-13</td>
</tr>
<tr>
<td>3. Backup tapes</td>
<td>17-14</td>
</tr>
<tr>
<td>F. Pitt leaves TALENT before TALENT leaves Pittsburgh leaves Pittsburgh!</td>
<td>17-14</td>
</tr>
<tr>
<td>18. The follow-ups</td>
<td>18-1</td>
</tr>
<tr>
<td>A. Mailing of the questionnaires</td>
<td>18-1</td>
</tr>
<tr>
<td>B. Processing the returned questionnaires</td>
<td>18-2</td>
</tr>
<tr>
<td>1. Keypunching vs. optical scanning</td>
<td>18-2</td>
</tr>
<tr>
<td>2. Coding</td>
<td>18-4</td>
</tr>
<tr>
<td>3. Name and address changes</td>
<td>18-4</td>
</tr>
<tr>
<td>C. Mail response rates</td>
<td>18-5</td>
</tr>
<tr>
<td>D. Special samples of nonrespondents</td>
<td>18-6</td>
</tr>
<tr>
<td>E. Efforts to maintain and improve response rates</td>
<td>18-10</td>
</tr>
<tr>
<td>1. For mailed questionnaire</td>
<td>18-10</td>
</tr>
<tr>
<td>2. For special sample</td>
<td>18-11</td>
</tr>
<tr>
<td>F. Content of special-sample questionnaire</td>
<td>18-13</td>
</tr>
<tr>
<td>G. Ten becomes eleven; twenty equals seventeen?</td>
<td>18-14</td>
</tr>
<tr>
<td>19. TALENT Post-Doctoral Fellowships</td>
<td>19-1</td>
</tr>
<tr>
<td>20. Later Studies</td>
<td>20-1</td>
</tr>
<tr>
<td>A. The one-year follow-up studies</td>
<td>20-1</td>
</tr>
<tr>
<td>B. The MAP studies</td>
<td>20-1</td>
</tr>
<tr>
<td>C. The 1963 retest study</td>
<td>20-1</td>
</tr>
<tr>
<td>D. The five-year follow-up studies</td>
<td>20-3</td>
</tr>
<tr>
<td>E. Ted Cureton's two TALENT studies</td>
<td>20-4</td>
</tr>
<tr>
<td>F. The Career Data Book and the supplement to it</td>
<td>20-4</td>
</tr>
<tr>
<td>G. The 11-year follow-up studies</td>
<td>20-5</td>
</tr>
</tbody>
</table>
21. Some Other Uses of TALENT Data
   A. Equating studies
      1. For the Air Force
      2. For the Navy
      3. Knoxville equating
      4. Inter-American Tests
      5. PCG equating
   B. Knoxville follow-up studies
   C. Other studies of marriage, the family, and parenthood
   D. Changes in norms
   E. Quality-of-life studies
   F. Follow-back studies
   G. A guidance study for NSF
   H. Sex differences in achievement
   I. Volunteer military forces
   J. Project PLAN
   K. Data Bank studies
   L. Miscellany
   M. Undone studies

22. The Data Bank
   A. History of the Data Bank
   B. Extent of use
   C. Charging for Data Bank services
   D. Policy on use of TALENT Battery
   E. Policy on use of Data Bank services
      1. Maintaining confidentiality
      2. Raw data tapes versus data analysis printouts
   F. Handbooks, brochures, etc.

23. Bridging the Gap:
    The PCG (Planning Career Goals) Battery
    A. Resolution of the dilemma
    B. Planning the battery
    D. Developing the PCG Battery and equating it to TALENT

Postscript
References . . . . . . . . . . . . . . . . . . . . r- 1

Appendix A. The TALENT Battery . . . . . . . . . . . . A- 1

Appendix A1. Composition of the Project TALENT Battery . . . . . . . . . . . . A1- 1

Appendix A2. Thumbnail descriptions of the Tests . A2- 1

Appendix B. Choosing Which 1960 Variables to Use . . . B- 1

Glossary of Acronyms and Abbreviations . . . . . . . . g- 1
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Project TALENT original follow-up schedule</td>
<td>1-4</td>
</tr>
<tr>
<td>3-1</td>
<td>Sampling ratios and senior class size strata</td>
<td>3-9</td>
</tr>
<tr>
<td>15-1</td>
<td>Taxonomy and etiology of statistical error</td>
<td>15-2</td>
</tr>
<tr>
<td>17-1</td>
<td>TALENT funding for first 20 years (1957-76)</td>
<td>17-10</td>
</tr>
<tr>
<td>18-1</td>
<td>Mailing schedule and procedure for 11-year follow-up of grade 10 cohort, and TALENT News</td>
<td>18-12</td>
</tr>
<tr>
<td>23-1</td>
<td>List of the occupations included in each of the the 12 career groups</td>
<td>23-4</td>
</tr>
<tr>
<td>23-2</td>
<td>PCG and corresponding TALENT instruments</td>
<td>23-6</td>
</tr>
</tbody>
</table>

# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure No.</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Remarks by Senator Hubert H. Humphrey in the &quot;Congressional Record&quot; of January 20, 1960</td>
<td>1-6</td>
</tr>
<tr>
<td>1-2</td>
<td>Letter from Vice President Lyndon B. Johnson to John T. Dailey, March 19, 1962</td>
<td>1-9</td>
</tr>
<tr>
<td>2-1</td>
<td>Advisory Panels</td>
<td>2-8</td>
</tr>
<tr>
<td>2-2</td>
<td>Research Advisory Committee of the Cooperative Research Program (1958-59)</td>
<td>2-9</td>
</tr>
<tr>
<td>2-3</td>
<td>Advisory Council</td>
<td>2-11</td>
</tr>
<tr>
<td>3-1</td>
<td>Sampling Advisory Panel</td>
<td>3-3</td>
</tr>
<tr>
<td>5-1a</td>
<td>Record Form Z (side 1)</td>
<td>5-10</td>
</tr>
<tr>
<td>5-1b</td>
<td>Record Form Z (side 2)</td>
<td>5-11</td>
</tr>
<tr>
<td>5-2</td>
<td>Extract from Local Coordinators' Manual</td>
<td>5-12</td>
</tr>
<tr>
<td>7-1</td>
<td>List of schools in sample as of 22 June 1960</td>
<td>7-13</td>
</tr>
<tr>
<td>7-2</td>
<td>The national sample of schools [Map]</td>
<td>7-17</td>
</tr>
<tr>
<td>8-1</td>
<td>Project professional, technical, and supervisory personnel: 1957-76</td>
<td>8-2</td>
</tr>
<tr>
<td>8-2</td>
<td>The changing cast: 1957-76</td>
<td>8-4</td>
</tr>
<tr>
<td>8-3</td>
<td>Regional Coordinators in 1960</td>
<td>8-24</td>
</tr>
<tr>
<td>9-1</td>
<td>The geography of Project TALENT (Who and what were where when)</td>
<td>9-3</td>
</tr>
<tr>
<td>13-1</td>
<td>Priority chart for each additional step</td>
<td>13-13</td>
</tr>
<tr>
<td>13-2</td>
<td>Description of Part A of tape documentation</td>
<td>13-15</td>
</tr>
<tr>
<td>13-3a</td>
<td>A Part B form intended primarily for test scores</td>
<td>13-17</td>
</tr>
<tr>
<td>13-3b</td>
<td>A Part B form intended for other kinds of information including questionnaire items</td>
<td>13-18</td>
</tr>
<tr>
<td>13-4</td>
<td>Project TALENT data file format</td>
<td>13-21</td>
</tr>
</tbody>
</table>
TECHNICAL APPENDICES
(separately bound)

Appendix C. The Coefficient of Internal Precision:
A New Psychometric Statistic for
Heritability Research . . . . . . . . C- 1

Appendix D. Sampling
Appendix D1. Must All Members of a Probability
Sample Be Randomly Selected? . . . . D1- 1

Appendix D2. The "Closed School-Cluster" Method
of Selecting a Probability Sample . D2- 1

Appendix E. Weighting
Appendix E1. When is Bias Better? . . . . . . E1- 1

Appendix E2. A Stratified Procedure to Compensate for Nonrespondents . . . . E2- 1

Appendix E3. Minimizing Total Error:
A Trade-Off Problem . . . . . . E3- 1

Appendix F. Statistical Procedure for Inferring Score
Distribution of Unselected Group of
18-Year-Olds from Score Distributions of
Unselected Group of 15-Year-Olds and
Twelfth-Grade Students . . . . . . . F- 1

Appendix G. Career Guidance
Appendix G1. Propinquity Analysis
A New Multivariate Index for Use
in Educational Planning . . . . . . . . G1- 1

Propinquity Analysis:
A New Guidance Tool . . . . . . . . G1-20

Appendix G2. The Project TALENT Approach to
the Career-Grouping Problem:
Cutting the Gordian Knot . . . . G2- 1
Chapter 1. PROLOGUE

If you were born between 1942 and 1945 and went to high school in the United States, there is roughly one chance in 25 that you were a participant in Project TALENT; maybe you were a participant even though you were born a little earlier or a little later—say in 1941 or 1946.

But even if you weren't a participant it is quite likely you know someone who was (although he or she may never have had occasion to mention it to you). And if you don't, the chances are still pretty good that you know someone who is or was on the staff of the project or who was or is associated with one of its sponsors. The fact that you are reading this book suggests that you are a member of the educational research community. Over the years there have been approximately 54 professional staff members, and in the course of the nearly 20 years since the beginnings of the project those 54 people have scattered widely. There have been innumerable college undergraduates and graduate students in the Palo Alto, Pittsburgh, and Washington, D.C. areas who have worked part-time as coding clerks or other junior staff members, many of whom eventually sought careers in educational research. They too have scattered widely. Therefore, to reiterate, perhaps you know someone who is or was on the Project TALENT staff. Chapter 8 (which has the somewhat hyperbolic title "With a Cast of Thousands!") tells a bit about who some of those people are.

A. What Project TALENT is

Project TALENT is a very-large-scale jointly cross-sectional and longitudinal educational research study that some people may think has been going on forever. It hasn't. But it was going on as far back as 1957, if one counts the planning phases that were carried out before the project really began to take shape in anything resembling its present form. And it had its start even further back than that, if one traces it back to when the idea was born. Actually in Chapter 2 we trace the idea even further back than that, to its "pre-natal stage."
To condense the description of Project TALENT into a single sentence: In 1959, primarily under the sponsorship of the U.S. Office of Education, the American Institutes for Research (AIR) and the University of Pittsburgh jointly undertook a massive project, which involved scientifically selecting a stratified random probability sample of all secondary schools (public, parochial, and private) in the United States, testing all their students in grades 9-12 (nearly half a million students) with a two-day battery of aptitude tests, information tests, other tests of ability and achievement, and inventories probing the examinees' background, interests, and personality traits, and then following up the examinees by mailed questionnaire periodically over a very long period of time (which hasn't ended yet). There is no point in expanding that very condensed description of Project TALENT at this point, because that is what most of the rest of this book is about.

B. Frame of reference

A frame of reference for the project is provided by its original purposes and its original schedule. The purposes of the project, as originally stated, were fivefold:

a. To provide an inventory of human resources.

b. To provide a set of standards for educational and psychological measurement.

c. To provide a comprehensive counseling guide indicating the patterns of aptitude and ability which are predictive of success in various careers.

d. To result in a better understanding of how young people choose their life work.

e. To result in a better understanding of the educational experiences which prepare students for their life work.
The original plan was to follow up each of the four grade cohorts four times—once a year, five years, ten years, and twenty years after they were scheduled to graduate from high school. Table 1-1 shows the schedule.

C. Current status of the Project

The one-year and five-year follow-ups were completed on schedule. As things turned out, however, the 10-year follow-up did not occur. Because of a one-year delay in funding, we conducted an 11-year follow-up instead.

The 10-year follow-up had been scheduled to start in 1970. Instead, the 12th-grade cohort was followed up in 1971. It was about this time that plans for the creation of the National Institute of Education (NIE) were being worked out. Project TALENT was among the activities scheduled to be transferred to NIE whenever it came officially into existence. It is probable that uncertainties associated with the imminent creation of NIE and the planned transfer of some of the functions of the Office of Education (OE) to the new Institute had something to do with the delay in funding.

Table 1-2 shows the schedule on which follow-ups have been completed thus far.

It perhaps should be mentioned at this point, although to do so involves breaching the definition of a "history" by talking about the future, that the possibility of another major change in the schedule has been broached recently (by NIE personnel). It has been suggested that the 20-year follow-up be changed to a 17-year follow-up.

D. In the public eye!

Over the years Project TALENT has managed to attract a fair amount of public attention and public interest, most of it favorable, and some of it from very eminent persons. For instance, the Congressional Record of January 20,
### TABLE 1-1. Project TALENT original follow-up schedule

<table>
<thead>
<tr>
<th>Grade When Tested in 1960</th>
<th>Years for Follow-up Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-Year</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
</tr>
<tr>
<td>9</td>
<td>1964</td>
</tr>
</tbody>
</table>

### TABLE 1-2. Follow-ups completed thus far

<table>
<thead>
<tr>
<th>Grade When Tested in 1960</th>
<th>Years for Follow-up Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-year</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
</tr>
<tr>
<td>12</td>
<td>1961</td>
</tr>
<tr>
<td>11</td>
<td>1962</td>
</tr>
<tr>
<td>10</td>
<td>1963</td>
</tr>
<tr>
<td>9</td>
<td>1964</td>
</tr>
</tbody>
</table>
1960, carries some very kind remarks by Senator Hubert H. Humphrey of Minnesota (later to become Vice-President Humphrey). The Congressional Record pages carrying Senator Humphrey's remarks are reproduced as Figure 1-1. Two years later, in March 1962, another Vice-President, this time the future President, Lyndon Johnson, wrote a very complimentary letter to the Project Director, John Dailey. Johnson's letter is reproduced as Figure 1-2. It may be of some interest that Lyndon Johnson and John Dailey were "noncontemporaneous schoolmates" in Southwest Texas State Teachers' College, the college in which they both earned their bachelor's degrees.

Project TALENT has also attracted its share—possibly more than its share—of unfavorable attention. In the early stages of the project it was under attack by various extremist groups, mostly at the far right edge of the political spectrum. It also attracted its share of interest in Congress. After the test battery was developed there were requests by various senators and congressmen for an opportunity to examine it. These requests, which presumably arose as a result of complaints by one or two of their constituents, were of course granted. After they examined the materials they were in every instance satisfied that whatever dire complaint had been made was without foundation, and we heard no more about it. The furthest anything like this ever got was a bill introduced in Congress early in 1959 to compel the Office of Education to cancel the contract for Project TALENT. The bill was introduced by a Representative from Michigan (who had not been among the small number of members of Congress who actually contacted us and examined the materials). Obviously the bill got nowhere. If it had been passed there would have been no occasion to write this book, seventeen years later! As Project TALENT has become better known over the years, hostility seems to have died down and support to have grown.
New Honors for Ellsworth Tompkins

EXTENSION OF REMARKS

OF HON. JOHN A. BLATNIK
OF MINNESOTA
IN THE HOUSE OF REPRESENTATIVES
Wednesday, January 20, 1960

Mr. BLATNIK. Mr. Speaker, I think all my colleagues agree with me when I say that one of the major, if not the major, issues facing the 86th Congress is revision of the social security law. Researches and on-the-scene surveys conducted throughout 1959 have demonstrated beyond question that our elderly citizens have suffered a steady deterioration so far as their income positions are concerned, and that this trend is certain to continue unless this Congress takes remedial action.

In this connection, I should like to call my colleagues’ attention to an anniversary that is being remembered and celebrated this month by millions of elderly citizens throughout the vast reaches of this great nation.

Dr. Francis E. Townsend who pioneered the pension movement in America, and still heads the organization he founded, was 93 years of age on January 15. Still hale and hearty despite his advanced years, the doctor, even at this moment engaged in a vigorous speaking tour on the west coast, urging upon his audiences the acceptance of the plan he advocated.

That plan is before this Congress in the form of my bill, H.R. 4000. It calls for universal retirement benefits of $140 monthly for older citizens age 65 and older, and to certain other groups including the disabled and widows with dependent children.

This program, to be financed from the proceeds of a modest 2 per cent tax on gross incomes, would yield about $140,000 in benefits of about $140, and would be paid as a matter of right H.R. 4000, unlike the social security program, calls for genuine pay-as-you-go financing.

The genius of Mr. Townsend lies in the administration of the agencies which he started. His imagination of millions of his fellow citizens some 25 years ago and today commands respect, not only among the aged people of this country, but now, too, among the students of the social security problem and among those who sit in this Chamber who have been elected to serve the best interests of our constituents.

It is fitting, therefore, that we pay tribute to Mr. Townsend on the occasion of his 97th birthday. It is given to few men to live so long and accomplish so much and earn the devotion of so many different followers. He may live to celebrate his 100th birthday—and to realize at long last the fruits of his efforts in behalf of his fellow Americans.

The Story of Project Talent

EXTENSION OF REMARKS

OF HON. HUBERT H. HUMPHREY
OF MINNESOTA
IN THE SENATE OF THE UNITED STATES
Wednesday, January 20, 1960

Mr. HUMPHREY. Mr. President, I was pleased to learn recently of "The Story of Project Talent," a survey of our aptitudes and abilities as a Nation to be undertaken by the University of Pittsburgh with the support of our administration. The bulletin of the project talent office describes the program that is being set up.

Certainly, at this time we must use all our available abilities and talents to aid our national security. I applaud this university scientific survey that will take a census of our scientific resources.

Mr. President, I ask unanimous consent that the bulletin explaining "The Story of Project Talent," be printed in the Appendix of the Record.

There being no objection, the bulletin was ordered to be printed in the Record, as follows:


Our country's continued growth and development require the identification, development, and utilization of all available talents. We cannot use all of these talents unless we know what is available. A national inventory of aptitudes and abilities will enable our personnel to assist their students in developing and making the most of their
If we are to plan properly for meeting the increased educational needs of the present day, we must know, for example, how many educators or physicians we may expect to educate and train in the years ahead—we need a better understanding of how many of our people have the aptitudes, the inclination, and the education for entrance into various careers. These are among the basic questions that Project Talent must be able to answer. But how many persons fail to receive proper encouragement and education? There are many tests in standard subjects, and for some of them it has been possible to test one person or a thousand persons for general intelligence, musical ability, or capacity in language or mathematics. But Project Talent is to be an enterprise that will provide the first scientifically planned national inventory of human talents: the aptitudes and abilities of a people.

NINETEEN SIXTY IS A CENSUS YEAR—PROJECT TALENT IS A SCIENTIFIC CENSUS

This inventory will be conducted in March of 1960, in order to coincide with the 1960 population census. When the census taker comes to the door of a home, the household earns not feel the census is an examination which they are subjected to. Thus Project Talent is true of Project Talent. No special preparation is necessary for this examination. Other occupations have become an integral part of national existence; in the eyes of students, their superiors, and their professors, and in the eyes of students, their superiors, and their professors, and in the eyes of students, their superiors, and their professors.

Thus, leading educational research institutions, agencies of the Federal Government, and the major research universities and educational associations have joined together to assist in guiding the study and interpreting the results from Project Talent.

FINANCIAL SUPPORT IS BEING PROVIDED BY THE OFFICE OF EDUCATION WITH ASSISTANCE FROM OTHER FEDERAL AGENCIES

Project Talent is financed largely by the Office of Education through Commissioner of Education Lawrence G. Derthick. These funds came from a cooperative project directed by Assistant Commissioner of Education Roy M. Hall. The plan for the project, as developed by the University of Pittsburgh and the American Institute for Research, was reviewed and recommended for approval by the advisory committee of the cooperative research program.

The chairman of this committee was Dr. Ralph W. Tyler, Center for Advanced Study in the Behavioral Sciences, Stanford, Calif. Other members of the committee were: Dr. Philip L. Engleman, executive secretary, American Association of School Administrators; Dr. Charles S. Arneson, chancellor, University of Wisconsin; Dr. E. B. Lindman, professor of school administration, George Peabody College for Teachers; Dr. Henry J. Otto, professor of education, University of Texas; Dr. J. C. Gage, assistant commissioner for research, New York State Education Department (retired); Dr. H. H. Remmers, director, division of educational research, Purdue University; Dr. Donald B. Watts, dean, College of Arts, State University of Iowa; and Dr. Dorothy Adkins Wood, chairman, department of psychology, University of North Carolina.

Supplementary support is provided also by the National Institute of Mental Health and the Office of Naval Research, with aid in the planning stages from the National Science Foundation.

In 1959, Dr. John C. Flanagan, professor at the University of Pittsburgh and director of the American Institute for Research, prepared the report for the initial survey of the study and is the responsible investigator.

The Planning and Administration of Project Talent is under the guidance of a National Board of Advisors and a National Advisory Panel.

Four panels of the Nation's leading specialists worked with the staff for more than a year planning the study. This panel included six full-time research workers headed by Dr. John T. Dailey. The overall chairman of the advisory panels during the planning phases was Dr. John H. Fischer, then superintendent of the Norfolk (Va.) Schools.

The present chairman is Dr. Kenneth E. Oberholtzer, superintendent of the Denver (Colo.) Public Schools. The other panelists include: Mr. Walter A. Chapman, executive director, Carnegie Foundation; Mr. Stanley J. Chater, dean of the College of Humanities, Pennsylvania State University; Dr. Edward Landy, director, division of counseling services, Newton (Mass.) public schools; Mr. Robert D. Keller, director, University High School, University of Minnesota. The chairman of the testing problems panel is Dr. Robert L. Thorndike, head, department of psychological foundations and services, Teachers College, Columbia University. The chairman of the morale and counseling panel is Dr. Samuel A. Stouffer, director, laboratory of social research, Harvard University.

The plans developed through these groups were reviewed by an advisory council composed principally of presidents and executive secretaries of the professional associations most interested in the findings from this study. These associations included the American Association of School Administrators, National Association of Secondary-School Principals, Council of Chief of School Officers, National Association of Secondary-School Boards, American Association for the Advancement of Science, American Personnel and Guidance Association, and American Psychological Association.

Regional coordinators will administer the project locally. They will arrange for cooperation of schools, and serve as advisers on procedures. These local officials, familiar with the educational needs and programs of the schools in their region, will be the key representatives of one of the largest educational studies ever undertaken.

Reports from the survey studies which follow will be published and will be freely available to educators and research specialists. Prior to publication, these studies will be reviewed by panels of experts and the advisory council. Many educational and research groups have expressed keen interest in these studies and will be kept fully informed. A series of reports will be made to educational agencies throughout the country. These will go to school systems, superintendents of departments of education, and to private educational associations as well as to the support agencies. The findings will also be made available to parents, to industry, to the military services, and to anyone concerned with planning for proper future use of manpower.

The responsibility for administering the tests and collecting the data about the students and their past experiences is placed directly in the hands of the selected schools. The scientific analysis of the original data and the collection of the follow-up data will be done under the direction of the project's staff at the University of Pittsburgh with special assistance from the American Institute for Research.

The success of the survey phase depends upon the cooperation and good will of local coordinators, principals, teachers, and others.

PROJECT TALENT IS A SCIENTIFIC STUDY USING PROVEN TECHNIQUES

It is expected that this historic national census and follow-up will give the Nation a great many new facts. Some will be surprising. Also, it is expected that administration of such a large program will produce new methods and new ideas in the field of educational research itself. This should be one of several important by-products from the study.

"Project Talent" is not an experiment to try out new theories. All the major components of this project have been tested and proven in previous studies. Systematic measurement and description of students is es-
FIGURE 1-1 (continued)

INTRODUCTION

In this kind of scientific study, it is possible to know in advance some of the major types of information which will be produced.

Some of the important results of this Nation's Aptitude and Ability Census will be:

- An inventory of human resources: Project Talent will be a tremendous inventory or stocktaking in which we find out the capabilities of our youth.
- We also plan to study the relationships between one kind of ability and another, between one type of school course and another, and between personal hobbies and the development of many types of competence.

A set of standards for educational and psychological measurement: When reliable measurement methods are obtained by studying the strengths of persons in a systematically selected sample, it will be possible to provide a more accurate and meaningful test of students for college admissions committees, for state examiners, for test authors in standardizing tests so that scores indicate comparable levels of ability.

This may be roughly compared to the basic standards such as the marked bars for measuring length, or the standard weights which the National Bureau of Standards maintains.

A comprehensive counseling guide indicating the patterns of aptitude and ability which will lead to success in various careers: In the follow-up after the national examinations and analysis, students who took the tests will be located and asked to report on educational and vocational experiences. A young girl may have become a receptionist, or a housewife, or she may have gone to college. By studying thousands of student aptitude, interest, and ability patterns, and finding out the person's later activities, this will be a great deal. This will help students by predicting more precisely what kinds ofaptitudes and abilities, what kinds of courses, and what kinds of interests constitute the best basis for various kinds of careers. This kind of information is needed for success in a career. Motivation is a necessary ingredient, but the best use of the student's special talents requires that he identify this talent early and obtain the educational essential for the full development and effective use of his powers.

It is anticipated that Project Talent will make a significant contribution toward meeting this need: better prediction, based on actual follow-ups, of a young person's chances for success in a given field. To some extent, this may be done by using aptitude tests as an enormous incentive to teachers, young people and parents to be told, for example, "This boy unquestionably has talent in a given direction and should continue his education in that direction, he may reasonably expect to master his chosen trade or profession." A beginning of how young people choose their life work: Many people follow their family trade or profession. They tend to think that people know quite early what their life work will be. Other people drift into an occupation and they tend to think that everyone else more or less drifts into a particular trade, business, or profession. Many people feel that they have very little choice.

However, many people today do have a choice, and are willing to pay the price and the need for special training continue to increase. We have begun to learn something about the processes by which a young person decides whether to be a designer, a lawyer, or an apprentice for a trade. This study and others can help us understand what ages certain lifetime careers tend to be chosen.

A better understanding of the educational experiences which prepare students for their life work: American education is noted for its diversity. Only through the analysis of detailed information about students, their educational experiences, and their subsequent successes or failures can we hope to make educational systems as flexible and responsive to the individual needs of its students as it must be if our Nation is to continue to develop and grow.

Project Talent has been carefully designed to fill an important national need for facts regarding the identification, development, and growth of this kind of information. This information is intended as a basis for manpower policies and as a basic resource for the many individuals responsible for the education of our children.

extension of REMARKS

OF HON. THOMAS J. LANE
OF MASSACHUSETTS
IN THE HOUSE OF REPRESENTATIVES
Wednesday, January 20, 1960

Mr. LANE. Mr. Speaker, under leave to extend a congratulation, I wish to include a statement by Daniel J. Murphy, Assistant Director, Bureau of Litigation of the Federal Trade Commission, designed to protect the consumer from exploitation by salesmen who misrepresent their products.

The statement follows:

FEDERAL TRADE COMMISSION CONFERENCE ON PUBLIC DECEPTION

extension of REMARKS

OF HON. THOMAS J. LANE
OF MASSACHUSETTS
IN THE HOUSE OF REPRESENTATIVES
Wednesday, January 20, 1960

Mr. LANE. Mr. Speaker, under leave to extend a congratulation, I wish to include a statement by Daniel J. Murphy, Assistant Director, Bureau of Litigation of the Federal Trade Commission, designed to protect the consumer from exploitation by salesmen who misrepresent their products.

The statement follows:

FEDERAL TRADE COMMISSION CONFERENCE ON PUBLIC DECEPTION

Mr. Speaker, today, I have the pleasure of presenting a statement on the subject of public deception. The Federal Trade Commission has been working on this problem for some time and has developed a comprehensive program to combat this type of business practice.

Mr. Speaker, the problem of public deception is one that affects all of us. It is a matter of great concern to the Federal Trade Commission, which is responsible for enforcing the Federal Trade Commission Act, the purpose of which is to prevent unfair methods of competition and restraints of trade that are injurious to commerce.

The Commission has found that public deception is often used by unscrupulous salesmen to sell products that are not as good as they are represented to be. This type of conduct not only harms the consumer, but also damages the reputation of the honest businessman who is trying to sell high-quality products.

The Commission has been working closely with industry and law enforcement agencies to develop a comprehensive program to combat public deception. This program includes enforcement actions, education, and research.

Enforcement actions are taken against businesses that engage in unfair or deceptive trade practices. These actions may include the imposition of fines, the revocation of licenses, and the referral of cases to the Department of Justice for criminal prosecution.

Education is also an important part of the Commission's program. The Commission has developed educational materials for use by schools, businesses, and consumers to help them understand the nature of public deception and how to avoid it.

Research is another key component of the Commission's program. The Commission has conducted research on the nature of public deception and the best ways to combat it. The results of this research are used to guide the Commission's enforcement actions and to develop new educational materials.

The Federal Trade Commission is committed to working closely with industry and law enforcement agencies to combat public deception. We believe that by working together, we can effectively prevent this type of conduct and protect the interests of consumers.

We are confident that with the support of Congress, the Federal Trade Commission will be able to continue its work in this important area.
March 19, 1962

Dear John:

Thank you so much for sending me a copy of the first volume in your study of the talents of American Youth.

It would be difficult to think of a more worthwhile undertaking than your efforts to assure that the best use is made of our nation's most valuable resource.

Congratulations on being a part of this wonderful project, and I hope you will continue to keep me advised on the progress of your work.

Kindest regards and all good wishes.

Sincerely yours,

Lyndon B. Johnson

Dr. John T. Dailey
1808 Adams Mill Road, N.W.
Washington, D.C.
Chapter 2: THE PRENATAL STAGE:
(Project TALENT Before It Was Project TALENT)

A. Roots of the idea

The initial idea for Project TALENT goes back as far as December 1956. At that time John Flanagan was working on the technical manual for the Flanagan Aptitude Classification Tests (FACT Battery) and was concerned about the problem of norms. Data were available on about 3,000 or 4,000 cases but he didn't regard this as a really adequate sample for the norms. Pondering the thought that it was very difficult (perhaps impossible) for the individual test-developer to arrange for a really satisfactory norm group, he decided that the only feasible solution would be for the Federal government to do it. What would be needed would be a very large and truly representative sample of the population to which the norms were to apply. If a fairly comprehensive "anchor battery" of tests were administered to such a sample it would be to the advantage of all test authors and all test publishers, since through the medium of equipercentile equating it would then be possible to develop excellent norms for any test that was comparable (in level and type of content) to a test in the anchor battery.

Checking back in the literature to see whether anything like this idea had been suggested he found that over a quarter of a century earlier Herbert Toops had suggested something similar. Toops had proposed that a million-member sample be selected, representative of the entire American population; he dubbed this proposed sample the "Standard Million". This sample would be used for the standardization of all sorts of tests.

Because he was focusing primarily on the high school level John Flanagan was able to modify the Toops idea by making it a bit less general. The sample would still be representative of a huge segment of the American population, but not of the entire population. The population represented would consist of everybody in high school at a given time. Because of this restriction, the size of the sample could
be scaled down, from a million to a mere (?) half-million. That meant that roughly 5 percent of all high school students would be in the sample. A considerably larger proportion of the corresponding population than Toops' "Standard Million" would be represented; that would have been only about three-quarters of a percent.

If Toops' Standard Million was one root of the idea for Project TALENT, what were the others? The answer to that question lies in large part in the history of aptitude and achievement testing and of research with the development and validation of methods for predicting success in various careers. John Flanagan had played a major role in these areas; in his capacity as Associate Director of the Cooperative Test Service (CTS) of the American Council on Education he had played a leading role in the development and scaling of achievement tests. And in his World War II role as the Director of the Army Air Force Aviation Psychology Program he blazed some trails in the development of aptitude tests and in differential prediction (the basic problem in personnel classification).

In developing plans for the project, John was of course fully cognizant at all stages not only of the tremendous possibilities the project would offer but also of the difficult problems and potential pitfalls that would be encountered. For instance he knew that it would be necessary to cope with the criterion problem, a stumbling block which had been the downfall of so many previous research projects. In an address given at Lake Washington High School for the Counseling and Guidance Training Institute, University of Washington, on April 21, 1960, he said:

"...How is it planned to determine 'the specific patterns of aptitudes and abilities and interests which provide the best basis for college courses and careers'? We plan to do this by comparing the thousand items of information about these people with their subsequent success in school courses, college courses and occupational experiences. We
will do this by following these students up 1 year, 5 years, 10 years, and 20 years after they graduate from high school. We do not plan to retest them all at this time but we will use follow-up questionnaires to get information as to what courses they're taking, what grades they are getting in their college courses, what employment they have been in, what salaries, how satisfied, how successful they are, and similar types of information. We hope that before 20 years have gone by we will be able to make some improvements on methods for assessing the success of people in various occupations such as teaching so that we will be able to do some comparison within groups."

In that last sentence Flanagan was referring to the by then well-known fact that it is much easier to validate predictive tests in terms of how well they differentiate among groups—in other words validating against the group-membership criterion—than in terms of how well they correlate with levels of success within group. The last sentence in the quotation above mentions a 20-year goal for solving this difficult problem. If that goal is to be met there are only about 4 years left. In the 16 years that have elapsed since then some thought has been given to the problem and perhaps some progress has been made, but the problem is by no means solved.

Flanagan's speech continues as follows, pointing out some additional areas where problems could be anticipated, and indicating how Project TALENT might bypass them:

"Studies such as 'Ten Thousand Careers' and similar follow-up studies have not been very successful in differentiating the people who are most successful in occupations from the people who are less successful. But many of these studies have had special problems. For example, in Thorndike's 'Ten Thousand Careers', the group was quite heterogeneous in the amounts of education and age. This group of people came into

---

* At about the time that the plans for what was to become Project TALENT were first being developed (1957) Robert L. Thorndike and Elizabeth Hagen were engaged in a monumental follow-up study of about 10,000 men who had been in the Air Force during World War II. The results are reported in a most enlightening book, entitled "Ten Thousand Careers" (1959).
the Air Force, some after completing college and some right out of high school. When Thorndike followed them up 12 years later, some of them had just recently finished their schooling while others had ten or more years of experience in a field after college. We hope that members of our group are sufficiently similar with respect to age and to educational experience that we will be able to make comparisons without some of the disturbing factors which made it difficult to establish relations in the Thorndike data. We still recognize, however, that the factors that cause one lawyer to earn $100,000 and another lawyer to earn only $5,000* a year may not be predictable from any of the present kinds of tests."

In this section we have discussed, on a somewhat limited basis, the roots of the idea for Project TALENT. From those roots the full-scale idea began to burgeon, including the principal characteristics that would define its scope: large-scale testing, on an enormous sample, with a very comprehensive battery of objective aptitude and achievement tests and questionnaires, long-term follow-up, and predictive validation.

B. The start of the search for funding

Clearly the idea that was eventually to develop into "Project TALENT" was a very ambitious one—indeed, an audacious one—and it would take a very large amount of money, over a long period of time, to put it into effect.

How does one go about getting funding for an educational research project that one foresees is likely to cost seven or eight million dollars before it is through? (Actually the project total cost may go substantially higher than that, because of the inflation that has occurred since 1957.) And in particular, how can one hope to accomplish this in a milieu not geared to large-scale costly educational research projects? Remember, back in the late 1950's the National Institute of Education didn't exist, and the Office of Education

* In 1960, levels of earnings and cost of living were far lower than now!
was a much smaller-scale operation than it is now. Educational research projects much more limited in scope but costing nearly as much (or even, in some cases, more) are now almost commonplace, but in 1957 the idea was daring in its magnitude. To repeat the question posed at the beginning of this paragraph: How, in 1957, would one go about getting funding for such a project? The answer, if the person seeking the funding happens to be John Clemans Flanagan, can probably be summed up in two words: indefatigability and thoroughness.

He immediately wrote a prospectus for the project and discussed it informally with the heads of many Federal government agencies—including, among others, the National Science Foundation (NSF), National Institutes of Health (NIH), Office of Education (OE), National Institute of Mental Health (NIMH), Office of Naval Research (ONR), Research Division of the Bureau of Naval Personnel (BuPers), Personnel Research Branch (PRB) of the Adjutant General's Office of the Department of the Army, and the Personnel Laboratory of the Air Force Personnel and Training Research Center (AFPTRC). The response was generally favorable in the sense that the government personnel with whom the idea was discussed thought it was a great idea and that someone ought to fund it. None of them, however, perceived his own agency as being the appropriate one to finance the project.

At this early stage the plan for the project was almost identical with the plan that ultimately developed. Most of the main features were there: the large national sample consisting of all students in grades 9-12 in a random sample of all high schools, the aptitude testing with a comprehensive battery, the comprehensive biographical inventory blank, and the follow-ups by questionnaire 1, 5, 10, and 20 years after high school graduation. There was one major difference, however. Since it was the problem of developing norms for the FACT Battery that had been the source of the idea for the project in the first place, and since the FACT battery comprised a very comprehensive set of aptitude tests which would be eminently suitable for the purposes of the project, naturally the use of that battery was an inherent part of the original plan. (As Chapter 3 indicates, however, it turned out that there were excellent reasons for changing that part of the plan in favor of developing a new battery of tests designed especially for the project; eventually, therefore the part of the plan that involved use of the FACT Battery was completely changed.)
After the groundwork had been laid by innumerable informal conferences at which the preliminary prospectus was discussed, the next step was to organize a planning conference, the purpose of which would be to give agencies planning to participate in the program or having a direct interest in its outcomes an opportunity to discuss the design of the study and contribute their ideas prior to the preparation of a formal proposal. The organizations invited to send representatives to the planning conference, which was held late in February 1957, included government agencies, foundations, professional organizations such as the American Psychological Association (APA), the American Association for the Advancement of Science (AAAS), and the National Education Association (NEA), and test publishers. It was hoped that if project plans could be developed in such a way that the needs of all these various organizations would be served, there might result a groundswell of support for the project.

Among those who attended the meeting was Alice Scates of the Office of Education. She was later to become Project TALENT's first monitor. The meeting was generally successful, in that a favorable attitude toward the prospective project was generated—but there was still not one cent of funding in hand.

To permit plans to be developed, the Cooperative Research Program of the Office of Education agreed to provide a five-year grant (July 1957 - June 1962) of about $114,000 to cover project planning, "get ready costs," and some other miscellaneous expenses. Because the Cooperative Research Program was by law not permitted to make grants to any kind of organization except state departments of education and institutions of higher learning, the grant could not go directly to AIR. Instead, the proposal written by John Flanagan was submitted by the University of Pittsburgh, where he was Professor of Psychology; and it was to the University of Pittsburgh that the Cooperative Research Program awarded the grant.

The Principal Investigator, who was of course John Flanagan, appointed as the incipient project's first Project Director Robert C. Craig, who was at that time a member of AIR's research staff. (A little over a year later, in November 1958, the Cooperative Research Program awarded another grant to Pitt. This one was for $150,000 and was to run for 2 years; it was for a supplementary sample of 16-year-olds not in grade 9, 10, 11, or 12.) Meanwhile early in 1958 NIMH provided a supplementary grant ($46,000) to the University of Pittsburgh, and ONR and NSF awarded small contracts to AIR ($35,000 and $15,000 respectively) to provide supplementary support for the project.
Thus at this point a total of $360,000 altogether had been obtained (mostly from the Office of Education, but partly from three other agencies). This $360,000 welcome though it was, was certainly not enough to assure the future of the project. The efforts to get funding for going beyond the planning stage, by actually carrying the project out, could not and did not abate. Prospectuses and proposals continued to be written, discussed, revised, submitted.

One very important step that took place early in 1958 was the appointment of four Advisory Panels. They were in the following areas and had the following chairmen:

1. Guidance and Counseling
   Edward Landy, Director, Division of Counseling Services, Newton (Massachusetts) Public Schools

2. Educational Research
   Robert J. Keller, Director, University High School, University of Minnesota

3. Testing Problems
   Robert L. Thorndike, Head, Department of Psychological Foundations and Services, Teachers College, Columbia University

4. Manpower and Sociology
   Samuel A. Stouffer, Director, Laboratory of Social Relations, Harvard University

Membership of the Panels is shown in Figure 2-1.

It will be noted that the first overall chairman of the Panels was John Fischer, at that time Baltimore's School Superintendent. Originally Ralph Tyler, the Director of the Center for Advance Study in the Behavioral Sciences in Palo Alto, had been slated to be Chairman, but he had resigned from this position before any meetings were held, because he perceived that in view of his appointment as Chairman of the OE Cooperative Research Program's Research Advisory Committee for 1958-59, there might be a conflict of interest. The Research Advisory Committee is shown in Figure 2-2. When Dr. Tyler therefore resigned as Chairman of the Advisory Panels, Dr. Fischer agreed to become his replacement. In the spring of 1959, Dr. Fischer, too, resigned, because he was leaving his position as Baltimore Superintendent of Schools to become the Dean of Teachers College, Columbia University. (He is now the President of that institution and has been since 1962.) Dr. Kenneth Oberholtzer, Denver's School Superintendent, replaced Dr. Fischer as Chairman.
Chairmen (seriatim):

John H. Fischer, Superintendent, Baltimore Public Schools
Kenneth Oberholtzer, Superintendent of the Denver Public Schools

GUILDANCE AND COUNSELING PANEL

Chairman, Edward Landy, Director, Division of Counseling Services, Newton (Mass.)
Public Schools
Ralph F. Berdie, Director, Student Counseling Bureau, University of Minnesota
Bruce E. Shear, Chief, Bureau of Guidance, New York State Education Department
John M. Stalnaker, President, National Merit Scholarship Corporation
David V. Tiedeman, Associate Professor of Education, Harvard University
Arthur E. Traxler, Executive Director, Educational Records Bureau
Leona E. Tyler, Professor of Psychology, University of Oregon

EDUCATIONAL RESEARCH PANEL

Chairman, Robert J. Keller, Director, University High School, University of Minnesota
Reverend O'Neil C. D'Amour, Assistant to the Executive Secretary, National Catholic
Education Association
Paul L. Banfield, Headmaster, Landon School (Maryland)
Warren G. Findley, Assistant Superintendent for Pupil Personnel Services, Atlanta
Public Schools
Earl J. McGrath, Professor of Education, Teachers College, Columbia University
Paul T. Rankin, Assistant Superintendent of Schools, Detroit Public Schools
James W. Reynolds, Professor of Education, University of Texas
J. Wayne Wrightstone, Director, Bureau of Educational Research, New York City Board
of Education

TESTING PROBLEMS PANEL

Chairman, Robert L. Thorndike, Head, Department of Psychological Foundations and
Services, Teachers College, Columbia University
Henry Chauncey, President, Educational Testing Service
Wayne H. Holtzman, Associate Director, Hogg Foundation, University of Texas
A. Paul Horst, Professor of Psychology, University of Washington
Lloyd G. Humphreys, Chairman, Department of Psychology, University of Illinois
E. Lowell Kelly, Director, Bureau of Psychological Services, University of Michigan
Joseph Zubin, Principal Research Scientist, Biometrics Research Unit, New York State
Department of Mental Hygiene

MANPOWER AND SOCIOLOGY PANEL

Chairman, Samuel A. Stouffer, Director, Laboratory of Social Relations, Harvard University
E. Franklin Frazier, Head, Department of Sociology, Howard University
Seymour E. Harris, Chairman, Department of Economics, Harvard University
Donald G. Marquis, Massachusetts Institute of Technology
Irving D. Lorge, Executive Officer, Institute of Psychological Research, Teachers College,
Columbia University
C. Joseph Nuesse, Dean, School of Social Science, Catholic University of America
Fred L. Strodtebeck, Associate Professor of Sociology, University of Chicago
FIGURE 2-2. Research Advisory Committee of The Cooperative Research Program (1958-59)

U.S. Office of Education
Commissioner Lawrence G. Derthick
Assistant Commissioner Roy M. Hall

Committee Members
Chairman, Ralph W. Tyler, Director
    Center for Advanced Study in the Behavioral Sciences
    Stanford University

Finis E. Engleman, Executive Secretary
    American Association of School Administrators

Chester W. Harris, Professor of Education
    University of Wisconsin

Erick L. Lindman, Professor of School Administration
    George Peabody College for Teachers

Henry J. Otto, Professor of Education
    University of Texas

J. Cayce Morrison, Assistant Commissioner for Research
    New York State Education Department (retired)

H. H. Remmers, Director
    Division of Educational Reference, Purdue University

Dewey B. Stuit, Dean
    College of Liberal Arts, State University of Iowa

Dorothy Adkins Wood, Chairman
    Department of Psychology, University of North Carolina
After the four Advisory Panels had met quite a few times, an Advisory Council was set up; its principal function was to review plans developed by the Advisory Panels. The Advisory Council was composed mostly of Presidents and Executive Secretaries of the professional associations most interested in the project and its findings. Membership of the Council is shown in Figure 2-3.

One result of the existence of this Council was that in effect it provided a channel of communication between the Principal Investigator and the organizations represented by the Council members. Factions in some of these organizations had been opposed to the project because they were afraid it would result in unfair and inappropriate comparisons of different schools, or that unwarranted inferences might be made about the relative merits of the school systems in different states, merely on the basis of a comparison of the results for a handful of schools. For the most part it was possible to allay this uneasiness by making a firm commitment that not only would no data be released concerning individuals but that we would go far beyond that, in that we would not release data about any participating school, or school system, or city, or county, or state. In fact the smallest geographical unit about which our commitment would permit us to release data would be region of the country. Furthermore we made a commitment not to release data that would permit comparisons among the three major categories of school in terms of type of control (i.e., public, parochial, and other private schools).

C. The planning and the plans

As the basic plans took shape ever more firmly, it became apparent that the eventual design of the project would be very similar to John Flanagan's early idea, which had first popped into his head in December 1956. The following features were agreed upon early by all concerned (although not necessarily in the following words):

1. It would be a survey study, not an experimental study.
2. The sample would be very large, and it would be a true probability sample, selected by random means.

3. The study would be both cross-sectional and longitudinal.

4. Its longitudinal feature would be a long-term one, with several follow-ups, over as long a period as 20 years.

The panels met quite a few times. There were two two-day meetings in the spring of 1958, one the following fall, and another the following spring. There was also considerable consultation by mail. Thus the panels were very active ones. (Further details about what was decided on the basis of their recommendations are provided in the next chapter.)
D. Finding funding

Recognizing that it was in the public interest to get the project funded, John Wilson, Assistant Director of the National Science Foundation, took the lead in convening a Federal Inter-Agency Committee to consider the problem. The committee met late in 1958. The organizations represented included the Army, Navy, Air Force, National Science Foundation, Atomic Energy Commission, National Institutes of Health, and, of course, the Office of Education. John Flanagan of course attended, too. Though all the participating agencies were interested in the project and wanted to see it get under way, none had been ready to fund it. In the course of the committee deliberations, however, a consensus was reached that in view of the nature of the proposed project, by far the most appropriate agency to fund it would be the organization that was already providing the principal support for the planning phase—the Office of Education. Roy M. Hall, the Assistant Commissioner of Education, who was a participant in the Interagency Committee meeting, agreed. And since he happened to be Director of the Cooperative Research Program, he was in a position to recommend that that agency finance the project. Fortunately the Commissioner of Education, Lawrence G. Derthick, agreed with him. Thus some very important hurdles had been passed in the search for funds. The proposal was formally submitted to the Office of Education in December 1958. But despite the backing of Dr. Derthick and Dr. Hall, one major hurdle yet remained: all projects to be funded by the Cooperative Research Program had to be approved first by the Research Advisory Committee. (See Figure 2-2.) Fortunately, under the chairmanship of Ralph Tyler, who understood the importance of the project and its potential value, the Committee gave its blessing! That was in the spring of 1959. In April 1959 the Cooperative Research Program awarded a $480,000 grant to the University of Pittsburgh, and Project TALENT was at last on its way! The prenatal stage was over. It had been about the same length (28 months) from conception in December 1956 to birth in April 1959 as the gestation period of an elephant—a coincidence that probably has no significance whatever.

The University of Pittsburgh, not AIR, was of course the prime grantee for the same reason that applied in the case of the planning grants. However, it was fully understood by all concerned that the project was to be carried out as a cooperative venture of the University and AIR, with the University's
contribution being mainly to provide administrative services. All AIR staff members who were to be on the Project TALENT staff received University appointments (while continuing also to be associated with AIR). These arrangements, while they caused a little awkwardness and inconvenience, nevertheless worked quite well.

E. Bonfires and brouhahas

We did not start to develop the Project TALENT Battery until about the beginning of 1959. But in the meantime we were busy with experimental tryouts of pseudo-batteries composed of parts of commercial tests; in some instances, after preliminary forms of our own instruments had been developed, we arranged to incorporate them in other researchers' batteries for a single administration. The purposes of these early tryouts varied. In some cases the aim was to work out the logistical and other details of handling the administration of a very large test battery to all students in a school. Another purpose was to get some preliminary data (intercorrelations, for instance) that might prove helpful in our test development efforts. A third purpose, at least later in the 1958-59 school year, when development of preliminary forms of most of our tests and questionnaires was well under way, was to get some empirical data on the new items—some preliminary item analysis data. In two of the communities where activities of this sort were undertaken, personality inventories were being used that had not been constructed by TALENT personnel, were not intended to be part of the TALENT Battery, and were in both instances somebody else's instruments—NOT OURS! But there was sufficient confusion that when a few outraged parents took umbrage at some of the personality inventory questions that they regarded as an offensive invasion of privacy, Project TALENT was blamed, and for a while the press was enlivened by reports of public protests against the project, defense of the project by school personnel, and liberal quoting of the items which aroused the ire of those parents and other members of the public who were most vociferous. For convenience, we may name these two episodes after the sites where they occurred—calling them respectively "the Baltimore brouhaha" and "the Texas bonfire." Some readers of this
History may remember reading in their local newspapers something about one
or both of these episodes at the time they occurred. What they read is very
likely to have been wrong—either in toto or in large part. Let us try to
set the record straight.

1. The Baltimore brouhaha

Arrangements were made with a school in Reistertown, Maryland,
a small community in Baltimore County, to administer a "pseudo-battery"* to
the students. Among the instruments included was the Minnesota
Counseling Inventory, a widely used personality questionnaire published
by the Psychological Corporation. Three items in this questionnaire
angered some parents and others who heard about them and unfortunately
attributed them to Project TALENT. The items that led to the resulting
brouhaha were:

1. Do you dream about sex?
2. I am worried about sex.
3. Have you ever stolen anything?

The Baltimore newspapers had articles on it for several days.
The matter was discussed at PTA meetings and at faculty meetings
in the high school in question. School authorities stalwartly defended
the Minnesota Counseling Inventory and Project TALENT. (They did not
distinguish between the two.) The school faculty passed the following
resolution:

"Resolved, that the faculty go on record as firmly
believing that in spite of uninformed and misguided
criticism of this pilot program, the broad overall
purposes merit and should receive full support from
educators, from our community, and from the public
at large."

Despite this praise, it would appear to the author of this History,
in the bright light of hindsight, that the representatives of TALENT
who made the arrangements with the school should have made clearer to
them that this preliminary testing did not involve the Project TALENT
Battery at all—or indeed any tests that would eventually be included
in it.

* "Pseudo-battery" is the term I am using here to distinguish the set of
instruments used from a formally organized battery available commercially
or still in the developmental stage. The term was not used in 1958.

- Author
2. The Texas bonfire

Wayne Holtzman, who was a member of the Advisory Panel on Testing Problems, and was associated with the Hogg Foundation for Mental Health, in Houston, informed us that the Hogg Foundation was planning a county-wide testing program called the Houston Youth Study, which would include every 9th-grader not only in Houston but in many schools in the surrounding county, Harris County. Only noncognitive instruments (personality inventories, attitude scales, etc.) were being included. He offered to include in the battery any sets of noncognitive items (e.g. background information items, interest items) that we would like tried out. The offer was accepted, and history (the Baltimore brouhaha) almost repeated itself. Some of the Hogg Foundation instruments were personality inventories and attitude scales that were regarded by some of the local citizenry as undue invasion of privacy. There was at that time (and there may still be) an organization in Houston called "Minute Women." The Minute Women objected vociferously to the questionnaire items and demanded that all the answer sheets be burned. The Houston Board of Education, three members of which were reported to be active Minute Women, angrily denounced the study, and with only one dissenting vote, ordered that the answer sheets be destroyed "at once." Early the next morning, with reporters breathing down her neck, the Assistant Superintendent, who also was the chairman of the study committee, burned every answer sheet from Houston. An application by a concerned citizen for a temporary restraining order to prevent destruction of the data came several hours too late. The data were already ashes.

The episode resulted in headlines all over the country. It was generally referred to as "book-burning," and was widely deplored. In Houston, however, the press was largely hostile to the testing program. The Houston Chronicle editorialized: "It seems that no family skeleton is safe from the sociologists anymore."

Among the personality inventory items that caused resentment were the following:
- I enjoy soaking in the bathtub.
- A girl who gets into trouble on a date has no one to blame but herself.
- If you don't drink in our gang, they make you feel like a sissy.
- Sometimes I tell dirty jokes when I would rather not.
- Dad always seems too busy to pal around with me.

But the Youth Study had its defenders, too. The Child Welfare Section of the Community Council called on the Council's board to ask school board officials to cooperate with research organizations in the future in psychological and sociological studies involving field tests in the schools. The Community Council, however, included in its membership some school board members who managed to keep motions of that sort from passing.

Meanwhile the opposition to the Youth Study in Houston had stimulated similar opposition in surrounding suburbs where the Houston Youth Study battery had also been administered. However, there was no test-burning outside of Houston, and most or all of the county data reached the Hogg Foundation intact.

Again, however, as in the Baltimore Brouhaha, the instruments and items that were criticized were in no case Project TALENT's.
Chapter 3. ISSUES AND DECISIONS

This chapter is concerned with issues, mostly but not exclusively of a technical nature, on which decisions had to be made before the "main event", the 1960 data collection, could get under way—and with the decisions that were made concerning those issues. The issues fall into three categories: the sample, the design, and the battery.

A. The sample

As was indicated in Chapter 2, certain features of the sample were inherent in the basic idea of the project. The sample would have to be very large and it would have to be a probability sample. The probability sample feature meant that we could not settle for just a very large sample consisting of whoever wanted to participate (a variety of the "available cases" method); nor for a very large sample achieved by inviting schools selected in some nonrandom way (most likely because they were convenient) to participate, and cutting off the rolls whenever the desired number of acceptances had been received (another variety of the "available cases" method); nor for the "quota method". (The quota method is a modification of the available cases method whereby instead of just setting an overall sample size to be sought, the size of the sample is set separately for various categories of schools, but within those limitations the "available cases" approach is used.)

But the decision that the sample should be large and that it should be a true probability sample didn't answer all the questions. How large should the sample be? And just how should it be structured? How many grade levels should be included? Which ones?

Then, of course, there was the whole area of highly technical questions that needed answers. For instance, should the primary
sampling unit be the student? the classroom? the school? the school system? the county? If the sampling unit is something larger than the individual student how many sampling stages should there be? If the sampling unit is the school, who should we cope with the problems of the junior high school? And in any event, should the sample be stratified? If so, on what? Should a uniform sampling ratio be used for all strata?

1. The Sampling Advisory Panel

The sheer magnitude and complexity of the sampling problem, combined with the fact that none of us, at that time, had had experience with a sampling problem of that magnitude, and none of us considered himself (or herself) an expert on the fine points of probability sampling, meant that we needed help from experts—a fact that was underlined when a member of the Cooperative Research Program's Research Advisory Committee raised a question on our sampling plans that seemed to require an answer from a sampling expert.

Because of that, and also because of all of our own unanswered questions, an advisory panel on sampling was established. Its four members are listed in Figure 3-1.

The panel met twice—first on 12 March 1959 and then on 26 September 1959, both times in AIR's Washington Office. The Office of Education was represented at the first meeting by Dr. J. William Asher (now on the faculty at Purdue, and a loyal user of TALENT data, via the Data Bank) and at the second meeting by Bill Asher again, and also by Howard Hjelm (who is still at the Office of Education, where he is now the Associate Commissioner for Adult Vocational, Technical, and Manpower Education). At the meetings, most of the issues mentioned above were considered, along with a few others.
FIGURE 3-1. Sampling Advisory Panel

Morris H. Hansen
Assistant Commissioner for Research and Development
Bureau of the Census

William G. Cochran
Professor of Statistics
Harvard University

Phillip J. Rulon
Professor of Education
Harvard University

Frederick F. Stephan
Professor of Social Statistics
Princeton University
Considerable attention was devoted to the problem of sample size. Some members of the Cooperative Research Program's Research Advisory Panel had wondered whether we really needed as large a sample as had originally been planned (close to half a million students in over a thousand schools). Wouldn't a carefully chosen sample of, say, 50,000 serve the purpose at least as well? Some members of the sampling panel initially had the same doubt. But their doubts vanished when John Flanagan explained to them why the very large sample was central to his idea. He made it clear that for some of the major purposes of the project it would be necessary to analyze the data in terms of sample members' career fields (as identified in the follow-ups). For some career fields, such as physicist, anthropologist, meteorologist, the number of cases would be barely large enough to give useful results if we started with half a million students. The sampling panel members were won over to this viewpoint.

When the question of choice of sampling unit was tackled consensus was quickly reached in favor of our original plan—to select a sample of schools, and test all students in the selected grade range who were in the chosen schools. (In other words there would be no second-stage sampling of students or classrooms within school.) Admittedly, using students rather than schools as the primary sampling unit would have had some advantages from a theoretical viewpoint but it would have been utterly out of the question from any practical viewpoint.

As for the grade range to be sampled, it was quickly agreed that grades 9 through 12 would provide a useful range.

The next problem tackled was stratification. We (the project staff) knew it would be advantageous to stratify but we weren't sure what combination of stratification variables to use. There
were a lot of possible stratification variables that we considered. Some had to be eliminated because it wouldn't be feasible to acquire the necessary information prior to sample selection; others would have to be eliminated just to cut down the number of stratification variables. Having too many would create practical problems as well as theoretical ones. (Classifying the universe of secondary schools into an enormous number of tiny cells would mean that there would be a lot of cells with so few schools in them that in the sample selection those cells would have to be very heavily oversampled to insure even one sample member.)

Among the stratification variables that we (the project staff) had been considering as possibilities were school control (3 categories--public, parochial, other private), region of the country, state, population of community, kind of community (urban, suburban, other small town, rural), socioeconomic level of the neighborhood served by the school, school size, grade range of the school, "school quality" (whatever that might mean), kind of high school (e.g., academic, comprehensive, commercial, vocational). Clearly, we weren't suffering from a paucity of ideas as to what to stratify on; the problem was an overabundance of ideas.

The Sampling Advisory Panel quickly confirmed that it would be highly desirable to use a stratification procedure in sampling; and they quickly endorsed our tentative plan to stratify on the public-parochial-private categorization.

As for geographical location they suggested geographical region, with the states arranged geographically within region, so that any "remnant" sets of schools (e.g., a remnant set consisting of 16 schools left over after one school had been selected at
random from each set of 20) would "spill over" into an adjacent state.

A further refinement of the plan was introduced when a member of the panel pointed out that the very largest cities (cities with populations much in excess of a million) were more like each other than they were like the rest of their states; Chicago, for instance, resembled Philadelphia more than it resembled the rest of Illinois. That being the case, it was suggested that each of these largest cities be treated, for purposes of sampling, as if it were a state; and that together they should form a "quasi-region".

We pointed out to the panel that 5 percent of the schools had 50 percent of the students and 50 percent of the schools had 5 percent of the students. This led to the consideration of the idea of stratifying on school size. The idea was endorsed, and in connection with it, Dr. Hansen pointed out the importance of "oversampling" the very large schools (in other words, using a higher sampling ratio for such schools than for ones with fewer students). This would help reduce sampling error on student statistics (though not on school statistics). For similar reasons he recommended undersampling the very tiniest schools.

He also pointed out that to be able to get an empirical check on the magnitude of sampling errors it might be advantageous to select two parallel subsamples and compute statistics for them separately, before combining the subsamples.

The panel also agreed that it would be desirable, if possible, to stratify on some measure of "school quality" or of socioeconomic status (which were of course recognized not to be synonymous)—but nobody present ventured a suggestion as to what that measure might be.

One further topic discussed by the panel was what to do about schools that declined to participate. Should we select a replace-
ment school? Dr. Hansen answered this question firmly in the negative. He pointed out that it would be every bit as satisfactory, from a theoretical viewpoint, and considerably less trouble from a practical viewpoint, just to increase the weight given to other schools in the same stratification cell to compensate, insofar as possible, for the missing school. He also stressed the fact that no participating school, whether it was a replacement school or just a more heavily weighted school from the originally selected sample, can wholly substitute for a school that refuses to participate, and that therefore it was imperative that we make every effort to get a near-perfect participation rate, for the initially selected sample.*

One important sampling problem, unfortunately, had been overlooked in preparing the agenda for the panel meeting. That was the problem of how to select the sample of junior high schools. I have no idea as to whether the panel would have come up with a good solution to this troublesome and nearly (though not quite) intractable problem if the question had been raised at the meeting. It wasn't, and so we shall never know. (To find out how we eventually handled the problem, see Chapter 7, section B.)

2. After the Sampling Advisory Panel meeting

The meeting of the Sampling Advisory Panel had been extremely helpful. It had confirmed the initial ideas on sample size and it had provided us with the outlines of a sampling plan, and with a strategy for coping with the problem of schools in the sample that might decline to participate in the project. But the details of the sampling plan remained to be worked out.

We were going to try to select our sample in such a way as to include about 400,000 to 500,000 students. The grades tested would

* We took Hansen's advice to concentrate our efforts on getting the initially selected sample to cooperate; and those efforts proved very successful. (See Chapter 7, section B.)
be grades 9 through 12. The primary sampling unit would be the school—and there would be no secondary sampling. There would be three, or perhaps four, stratification variables:

1) Public-parochial-private

2) Geographical location
   (Region, and state within region, with the very largest cities treated as "quasi-states" combining to form a "quasi-region". A strategy for handling "remnant" groups of schools in a state was also suggested.)

3) School size

4) "School quality" or socioeconomic level of the school, or some combination of the two

There would be differential sampling ratios, primarily on the basis of school size, with the highest ratio being used for the largest schools.

But what cutting points would be used to define the school size strata? And what sampling ratios would be used? (Whatever ratios were selected, they would have to be ones that would yield the right sample size, in terms of both number of students and number of schools.) And what on earth were we going to use for a direct or surrogate "school quality" or SES measure that would be available before sample selection?

On the basis of a review of what kinds of information were available to us on the 90,000 IBM cards (3 per school) provided by the National Center of Educational Statistics (NCES) and containing data from their 1958-59 National Survey of Public Secondary Day Schools (see Chapter 7, section A), and some preliminary computer runs of a sample of these cards, John Flanagan made decisions on these basic unanswered questions. For the public schools four strata would be established on the basis of senior
class size, with sampling ratios as shown in Table 3-1. For the nonpublic schools there would be stratification on only two variables: (1) Control (2 categories: "parochial" and "other private"), and (2) location (region and state, including quasi-states and a quasi-region, as in the case of the public schools). From the preliminary computer runs it could be correctly inferred that the ratios shown in Table 3-1 would yield about the right number of schools and students.

<table>
<thead>
<tr>
<th>School Control</th>
<th>No. of seniors</th>
<th>Sampling ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>400 or more</td>
<td>1:13</td>
</tr>
<tr>
<td></td>
<td>100-399</td>
<td>1:20</td>
</tr>
<tr>
<td></td>
<td>25-99</td>
<td>1:20</td>
</tr>
<tr>
<td></td>
<td>Under 25</td>
<td>1:50</td>
</tr>
<tr>
<td>Parochial</td>
<td></td>
<td>1:20</td>
</tr>
<tr>
<td>Other private</td>
<td></td>
<td>1:20</td>
</tr>
</tbody>
</table>

After studying the types of data available on secondary schools, John (Flanagan) decided that the remaining stratification variable would be "retention ratio"—defined as the ratio of graduates to 10th-graders. This could be expected to be correlated with socio-economic status, and as a measure of the school's "holding power" it could be regarded as, in one sense, an indication of "school quality" (still undefined).

3. The 16-year-olds and the 15-year-olds

It will be recalled from Chapter 2 section B that funding had been secured from the Cooperative Research Program for supplementing the grade 9-12 probability sample with a probability sample of 16-year-olds not in grade 9, 10, 11, or 12. These were to be

* The term "parochial" is used here as a kind of shorthand for what would more precisely be described as "parochial or diocesan".
all the 16-year-olds in the school districts corresponding to
about one-tenth (a representative tenth) of the public schools
in the regular sample. These 16-year-olds, properly weighted
and combined with the 16-year-olds in the grade 9-12 probability
sample, would constitute a probability sample of the age group.

The necessity for getting this representative tenth of the
schools led to the decision by TALENT's Principal Investigator
to adopt, in a considerably modified form, the suggestion Dr.
Hansen had made at the sampling panel meeting, that we have two
parallel samples. The modification consisted in splitting the
schools in the sample into 10 subsamples ("subsamples 0-9")
parallel in terms of the geographical stratification variable and
the public-parochial-private stratification variable. One of
these 10 subsamples (specifically "subsample 0") would define the
school districts for the 16-year-old sample.

Shortly after this decision was made it was decided that for
practical reasons it would be better to have a probability sample
of 15-year-olds than of 16-year-olds. The principal reason for
this change was that it became apparent that there would be too
many 16-year-olds not in school to permit a good job of locating
and testing them to be done within budget. Because of the com-
 pulsory education laws which in every state required school atten-
dance until at least age 16, there would be far fewer 15-year-old
dropouts. Thus a probability sample of 15-year-olds could be
established more easily, and would serve the purpose at least as
well as a probability sample of 16-year-olds would. Therefore
permission was secured from the Office of Education to change the
sampling plan in this one respect.

4. The 10 subsamples: an interim feature

The 10 subsamples had served one of their two functions when
one of the subsamples was used to define the school districts for
the sample of 15-year-olds not in grades 9-12. The second important function was to provide on a preliminary basis a sample one-tenth the size of the total sample, for use in analyses not requiring the full sample. But for that function the school subsamples were useful only at the beginning to make it possible to secure preliminary results, based on subsample 0, when only that subsample had been put on tape in a form suitable for data analysis. When all 10 school subsamples had been put on tape in master file form the need for the subsamples as such vanished, since we could get a much better 10% sample by selecting a representative tenth of students; this was easy to do on the basis of last digit of student number. Thus the ten school subsamples have hardly been used since good 10% student subsamples became available.

B. The design

Among the problems that were considered by the four advisory panels (see Figure 2-1) at one time or another were some that were concerned with the basic design of the study. For instance:

1. Should the testing be a one-shot event or should there be retesting in the same schools in subsequent years?

2. And if there is to be retesting, should it be of the same cohort at a higher grade level level? Or a different cohort, and the same grade level? Or both?

The recommendations of the Advisory Panels were uniformly in agreement with the original plans. There would be one-shot testing. Follow-ups would be by questionnaire only; there would be four—one, five, ten, and twenty years after the cohort's graduation from high school. There would be no routine retesting.*

* The special retesting described in Chapter 11 was not decided upon until the year it occurred (1963).
Obviously retesting would have had some advantages. But the decision against it was really the only possible one in view of the long-range purposes of the study. It will be recalled that the near-half-million sample that had been settled on was about the smallest that would give an adequate number of cases in some important but relatively infrequent occupations. But if there were to be any retesting, the per-capita costs would inevitably be increased substantially, and the sample size would have to be cut down accordingly.

C. To build or buy: The TALENT Battery

The principal problem to be solved in connection with the battery was where it should come from. Should it be a commercially distributed battery available in toto from some test publisher? Or should it be put together from a bunch of commercially distributed separate tests, available from one or more publishers? OR SHOULD A BRAND-NEW BATTERY BE DEVELOPED BY THE PROJECT STAFF, SPECIFICALLY FOR USE IN PROJECT TALENT?

The answer to these questions is telegraphed by the capitalization at the end of the preceding paragraph. The decision was—and had to be—that we would develop our own battery. That was really the only sensible decision—for a multiplicity of reasons, which are listed below:

1. In the first place it would avoid giving a tremendous commercial advantage to any one test publisher. Buying an intact battery from a single publisher would have given that publisher an enormous commercial advantage—not just in terms of the profits from selling half a million copies but also because of the increase in the value of the property as a result of all the normative and follow-up data that would be developed by Project TALENT, at no cost at all to the publisher. All the test publishers except the one whose battery was selected would have objected strenuously (and with good reason).
2. A specially built battery would insure that the tests would be new to all students alike. Thus distortion due to differential practice effects and to the fact that some examinees would already be quite familiar with a test while it would be brand new to others would be avoided. (A commercial test to which none of the sample had been exposed would not have been available.)

3. We needed a lot of very short tests rather than a smaller number of longer tests, since the battery would have to cover a very wide spectrum of abilities. Although with a battery of the sort we needed, some of the tests would be too short to give highly reliable results for individuals, a battery of this sort would still be better for research purposes than would a smaller number of longer more reliable tests. (Furthermore even though short individual tests may not yield highly reliable scores, they can be combined to produce very reliable composites.)

4. The final battery would have to be very tightly tailored not only physically but also in terms of test length, in order to make optimum use of the limited amount of testing time (2 days) that would be available in which to administer a very large number of instruments. The chances that each of a set of commercially distributed instruments would be precisely the optimum length would be virtually infinitesimal.

5. Unlike most users of commercial tests, we would not be dependent on the publisher for norms. (There goes one more argument for commercial tests, right out of the window!) With a probability sample consisting of nearly half a million students we felt confident that we could build our own norms, and that they would be better than the ones then generally available for commercial tests, both because they could be based on an enormous sample (nearly 100,000 cases per grade) and because it would be a genuine probability sample. It should be realized that in 1960 it was not really customary for test publishers to secure probability samples for norming. Samples were far more likely to consist of "available cases", and the number of
available cases was more likely to be counted in thousands than in hundreds of thousands.

Even after the decision had been made that we would build our own battery, and thus not bestow an enormous advantage on a single test publisher, to the detriment of its competitors, some of the publishers were still a little uneasy about what effect Project TALENT might have on their business. Reassurance was provided by a commitment on Project TALENT's part, that the TALENT battery would be used only for research—and that even then it would be made available only if the research study were such as to require the use of TALENT tests specifically, in other words if there were no tests commercially available that would do the job. The TALENT Battery would of course be made available at cost for use in equating commercial tests, so that they could benefit from TALENT's normative (and follow-up) data. These commitments allayed the uneasiness of those test publishers that had been uneasy, and we secured their support.
Chapter 4. NAMING THE BABY

Lots of people wonder where the name "Project TALENT" came from and what it means. They assume, because TALENT, used as the name of the project, is always spelled all in capitals, that it is an acronym. It isn't.

To begin with, the name of the project didn't come with the idea for the project. That idea goes a long way back, as explained in Chapter 2. The beginnings of John Flanagan's idea that AIR ought to do something along those lines took shape in 1956. But the developing idea was not dubbed Project TALENT until sometime in 1959. Originally it had a much more cumbersome name, "Research into the Identification, Development, and Utilization of Human Talents".

The Project Director, John Dailey, may have been the first person to realize how badly a short easily remembered name was needed to identify the project and make it stick in the public consciousness. After giving the matter some thought he announced one day that he had thought of the title: Project TALENT. He explained that the reason this name was appropriate was that everyone has some sort of talent--not necessarily one related to intellectual or artistic ability, but possibly a talent for assembly-line work or for tasks requiring physical strength or manual dexterity--or perhaps a talent for friendship. This broad (and humane) definition of "talent" makes the name of the project compatible with the character of the sample, representing virtually everybody, rather than a highly selected segment of the population.

The title has served its purpose well. The name Project TALENT rolls off many tongues all over the country. Inevitably, however, there has been some confusion among people who have heard the project name but know nothing whatever about the project. It has been confused occasionally with enterprises designed to identify high school
graduates in the top levels of intellectual ability—for instance the Westinghouse Science Talent Search, or the National Merit Scholarship competition. But this kind of confusion is of little consequence, since it is so easily cleared up.

Somewhat more prevalent than the misconception about the implication of the word "talent" as used in the project name is the misconception mentioned earlier, that "TALENT" is an acronym. And numerous attempts have been made, mostly by uninvolved bystanders, to develop an a posteriori acronym. The author of this History offers the following demonstration that such efforts are to no avail:

<table>
<thead>
<tr>
<th>T</th>
<th>Totally</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Acronymless</td>
</tr>
<tr>
<td>L</td>
<td>Large-Scale</td>
</tr>
<tr>
<td>E</td>
<td>Educational</td>
</tr>
<tr>
<td>N</td>
<td>Nationwide</td>
</tr>
<tr>
<td>T</td>
<td>Testing</td>
</tr>
</tbody>
</table>
Chapter 5. BUILDING THE BATTERY

A. Schedule

The schedule for getting from the preliminary phase to a published battery in final form was going to be an extraordinarily tight one. That much became apparent as the delays in securing funding continued. It was clear that if the goal of doing the testing phase in the spring of 1960 were to be met, test development would have to be begun before funding was received, and so that is what was done. It was a gamble, but it was one that paid off; it made it possible to meet the schedule for a spring 1960 testing. (It was felt that 1960 was a particularly advantageous base year because it would coincide with the Decennial Census, thus facilitating all sorts of comparisons.)

B. "State-of-the-art"

It had been decided early that Project TALENT would develop a "state-of-the-art" battery. This meant that it was to consist primarily of well-established types of tests and inventories rather than novel or experimental ones about which little was known. But the instruments we were to build should of course be as good specimens of their kind as could possibly be achieved. (If we were talking about athletes, "state-of-the-art" would mean the Olympic champion, not the typical Sunday athlete.) That was the rather lofty goal. The degree to which it was achieved—indeed, whether it was achieved at least partially—cannot be asserted with any confidence. A man's reach—a person's reach—a project's reach—should exceed his (his or her) (its) grasp.

The foregoing discussion of goals in developing the TALENT battery is motivated by the fact that the term "state-of-the-art" is misunderstood by many people, who confuse "state-of-the-art" with "mediocre," rather than understanding it to mean the most advanced state-of-the-art; and whether achieved or not, that was what we were aiming at!
C. **Content and organization of the battery**

Decisions on content and organization of the battery were based in part on the test materials assembled by Bob Craig during his year of planning, but mostly on the recommendations of the test panel and on the judgment of the staff members especially knowledgable in the area. Chief among these, of course, was the Principal Investigator, John Flanagan, whose expertise had been thoroughly demonstrated by, among other things, his former roles as Associate Director of the Cooperative Test Service and as Colonel Flanagan, Director of the Army Air Force Aviation Psychology Program during World War II.

It was established early in the planning of the test battery that it would include, among other things, a long, multi-purpose Information Test, achievement tests in math and English, a reading comprehension test, and tests of various other well-established aptitudes and abilities.

Permission was received from John B. Carroll, the principal author of the Psi-Lambda Language Aptitude Test (now called the Modern Language Aptitude Test), to develop parallel forms to the three paper-and-pencil subtests of the Psi-Lambda. (The other two of the five subtests could not be handled within the Project TALENT framework because they required use of a tape recording.) The decision to use three tests of the Psi-Lambda type was made on the recommendation of Marion Shaycoft, who had observed use of that instrument in an entirely different context and had been much impressed by the content validity of the test, its demonstrated effectiveness, and the novelty and ingenuity of the approaches.

Speaking of ingenuity, John Flanagan's FACT Battery (Flanagan Aptitude Classification Tests) included a promising Ingenuity test. It was decided to handle this test on the same basis as the three Psi-Lambda subtests—in other words to construct a parallel form for inclusion in the TALENT battery.

In connection with the parallel forms of these tests from other batteries, certain name changes were decided upon. The FACT Ingenuity

There were at least three departures from the decision to use only tests of well-established types rather than trying out experimental instruments. Henry Chauncey, the president of Educational Testing Service, who was a member of our Advisory Panel on Testing Problems, recommended that we include three tests that ETS had been using on an experimental basis and considered promising. One was called "Social Judgments" and required the examinee to answer a series of two-choice questions where the task was to indicate by choosing between two equally favorable (or equally unfavorable) adjectives the type of person he liked better (or disliked less). The score, which represented speed of making one kind of decision, was simply the number of items answered in a very limited time. (Of course, the nature of the scoring procedure is not revealed to the examinee.) The second test, which was called "Sentence Completion", called for studying a series of short sentences of a rather bland or neutral nature for a specified period to memorize them verbatim, and then filling in blanks with the missing words. The third test was called "Following Directions". We decided to try these tests since they sounded as if they might be useful and since they required relatively little time and therefore could be squeezed into the battery without much trouble. Our version of the "Social Judgments" test is called "Preferences;" our version of the "Sentence Completion" test is called "Memory for Sentences;" we had to make major revisions in the item format to change the memory items from write-in to multiple choice, but we believe the new format that we designed retained the essential features of the ETS test; i.e., it measured deliberate memorization rather than incidental memory, underlearned rather than overlearned material, and, most important of all, recall rather than recognition (despite the multiple-choice format)!
The "Following Directions" test which, like the Sentence Completion test, required a new item format to fit our requirements that every test item be machine-scorable, did not survive the item analysis tryout. (The other two tests did survive, and are in the final battery.)

As a matter of fact, in addition to the three "experimental" tests developed and tried out at Henry Chauncey's urging there was at least one additional test of a type that might, by some definitions, have been considered "experimental" rather than "state-of-the-art." The author of this History, however, who was also the author of the test in question, contends that the test in question, which happens to be the Arithmetic Reasoning Test (Part I of the 3-part Mathematics Test), is, if not a state-of-the-art test, demonstrably better than such a test. To be a little more specific, the Arithmetic Reasoning Test written for the TALENT Battery is virtually computation-free. Most conventional arithmetic reasoning tests at that time required (and most of the ones now being developed still require) that the examinee not only know the correct method of solving "word problems," but demonstrate that knowledge by actually doing the computation necessary to find the answer. Thus a student good at arithmetic reasoning but poor at computation might get an unduly low score which would (unfairly) cast doubt on his or her reasoning ability. The TALENT Battery Arithmetic Reasoning Test clearly separates reasoning from computation, and requires only the former. Thus if it was not precisely a state-of-the-art test, perhaps it advanced the state of the art (at least a little bit).

Prior to developing the tryout forms of the tests a very complete rationale was written for each test. The rationale explained what was to be measured and how it was to be measured, provided detailed specifications for the test items, indicated the distribution of the levels of item difficulty to be sought, and presented sample items. The concept

* The PCG Quantitative Reasoning Test is an exception. See Chapter 23.
of the rationale, developed largely by John Flanagan, should, in AIR's collective thinking, be an integral part of any test development activity.

D. The battery-building crew

Building the experimental form of the battery was an in-house activity, carried out largely by project staff, rather than a set of tasks to be farmed out to consultants, subcontractors, and other external personnel or agencies.

John Flanagan played a central role--the central role--in planning the outlines of the battery's content and organization. He worked closely with the staff members responsible for the various parts.

1. The test builders

Marion Shaycoft was the editor-in-chief of the aptitude and achievement tests. In this capacity, and in close conjunction with John Flanagan, she played a major role in establishing the outlines of the entire battery and making it fit within the physical and temporal constraints (test booklets and answer sheets representing the former kind of constraint and the two-day time limit for administering the battery the latter). She planned the details of all the tests, wrote the rationales for them, supervised the item writing, and edited the results. (John Flanagan, of course, did the final review of all rationales and all tests.) As things turned out, because the test development staff was quite small Marion Shaycoft also wrote many of the tests almost in toto and wrote large parts of most of the others. She was fortunate, however, that though members of the project staff available for this phase of the work were not numerous they did include at least one staff member who was both experienced and very competent in test development. That was Jo Edwards. She was the associate editor and a mainstay of the test development staff. Among other things, she is wholly responsible for the
item format and all the items in the main section of the Punctuation Test. It was decided to farm out those few segments of the test-writing that we felt could best be handled on that basis. The item-writing handled that way included a few of the areas in the Information Test—for instance the items on auto mechanics, an area about which neither Jo Edwards nor Marion Shaycoft nor any other staff member on the premises felt at all knowledgeable—and the mechanical reasoning items—which were drafted by Mary Willis, a highly experienced and expert test writer who at that time was not on the AIR staff but was later to join it. (Mary had been a test editor many years before, at the Cooperative Test Service (CTS)—one of the predecessors of the Educational Testing Service—when John Flanagan had been the Associate Director. It was at about the same time and in the same organization (CTS) that Marion Shaycoft had been initiated into the fine art of test writing under John Flanagan's aegis. Geraldine (Gerry) Spaulding, who was to be one of the reviewers of our tests, and a consultant on our questionnaires, was a test editor in the same organization, CTS, at the same time.

John Flanagan assumed personal responsibility for writing the Creativity Test, since it was to be modeled on the Ingenuity Test in his FACT Battery. He also wrote quite a few of the sports information items, as did John Dailey, since neither Marion Shaycoft nor Jo Edwards felt confident of her ability to deal with the technical aspects of such activities as football. John Dailey took personal responsibility for writing the Information Test items in the areas of Hunting and Fishing. (It is perhaps worth mentioning at this point that the Hunting Information score turned out to have a larger proportion of unique reliable variance than any other test in the battery. And for those readers who may be a little fuzzy as to whether unique variance is a good thing or a bad thing, perhaps we should clarify matters by stating flatly that unique variance is a very good thing indeed, provided it is also reliable variance, as it was in this instance.)
Probably because she had been instrumental in getting Jack Carroll's permission to prepare tests parallel to three of his Psi-Lambda subtests, Marion Shaycoft wound up writing the three tests. In granting permission Jack had imposed the entirely reasonable condition that he be allowed to review the tryout forms of the test and that his approval had to be secured before use. This proved to be an advantageous arrangement, since he readily gave his approval of the three tests, thus giving us the assurance we wanted that our tests were indeed parallel to his. This was particularly reassuring in the case of the Word Functions in Sentences Test, since its nature was such that it needed to have somewhat vague directions. For this test, whether one could comprehend the instructions was a legitimate component of the test score. (This is not true on most tests in the cognitive domain. On most such tests the instructions are supposed to be crystal-clear, so that all examinees can be assumed to understand them, thus permitting any reduction in score to be attributed to inability to cope with the items themselves, not with the directions.)

It was indicated at the beginning of this section that the test construction was largely an in-house activity. One partial exception to this should be mentioned. Because of our extremely tight schedule and because the art work on items requiring diagrams (visualization items, abstract reasoning items, object inspection items, mechanical reasoning items) is very time-consuming, it occurred to us that it would be worthwhile to explore whether the U.S. Civil Service Commission had available any good items of these types which they could spare for us. Accordingly Jo Edwards and I conferred with Civil Service Commission personnel (Albert Maslow, Harold McAdoo), who proved most cooperative. After looking into the situation they reported that they had quite a few visualization items of the "paper-folding type" (in Project TALENT's terminology Visualization in Three Dimensions
items) which had been item analyzed and were known to work well, which they could let us have. (They found they did not have suitable items for any of the other tests.) We accepted their offer and included a substantial number of their "paper folding" items in our experimental forms. Accordingly some of the items in the final form of the TALENT Visualization in Three Dimensions Test are in there through the courtesy of the Civil Service Commission, while others were developed by project staff.

2. Questionnaires for students

Dave Orr was assigned responsibility for developing and editing the SAI (the "Student Activities Inventory"--this was a code name for "personality inventory").

Izzy Goldberg had a corresponding role in regard to the other two student inventories--the SIB ("Student Information Blank," which provided background information) and the Interest Inventory.

3. School questionnaires

Dave Orr was responsible for developing and editing the General School Characteristics Questionnaire. Izzy Goldberg had the same role with respect to the other two school questionnaires--the Guidance Program Questionnaire and the Counselor's Questionnaire. One other person played a major role, on a day-to-day basis, in the development of these school-level questionnaires. That person was Dr. David F. Votaw, Sr. Dr. Votaw, who had been John Dailey's professor many years before, at the college from which he received his bachelor's degree, Southwest Texas State Teacher's College, had been invited to spend several months in Washington during the 1958-1959 period as a consultant to the fledgling project. In this capacity Dr. Votaw concentrated his efforts on the school questionnaires and made many important decisions in regard to them.
E. Tryout and item analysis

The experimental tryout was carried out in the spring of 1959, in 11 high schools varying widely in size, kind of community, and section of the country. All students in grades 9-11 were tested (although each student took only a part of the battery). Altogether, almost 6,000 students were involved; each form of each test was tried out on about 1,000 to 1,600 students.

Partly on the basis of the item analysis but chiefly because limitations on testing time required it, four tests were eliminated from the battery. The rest were revised and submitted to members of the Advisory Panel on Test Problems (see Figure 2-3 for their identities), as well as to other external reviewers; specifically, Paul Diederich, Gerry Spaulding, Mary Willis, Machlin Thomas, Max Engelhart, Paul Dressel. Their reviews were almost uniformly very complimentary. It will be recalled, too, that Jack Carroll reviewed the three tests based on his Psi-Lambda Battery. Thus all our tests received a very thorough review by nationally recognized experts, and passed with flying colors!

F. The end of the process

After the final review, final editing, and final revision of the final forms, we were ready for the penultimate steps—putting all the tests and other materials together and developing essential auxiliary materials such as a manual of instructions for test administration, and in general forming a cohesive battery out of a collection of separate components. The battery that resulted was organized in five separate test booklets: Booklets A, B, C1-x (an "expendable" booklet, i.e., one on which the students were permitted to write), C1, and C2. There were also five answer sheets: A, B1, B2, C, and Record form Z (the form containing basic identifying information such as name and address).* Answer sheets corresponded to the test booklets which had the same

* Record Form Z is shown in Figures 5-1a (front) and 5-1b (back). We hoped to get a Record Form Z for everybody in the class—even absentees—in order to complete our probability sample. Therefore the instructions in Figure 5-2 were incorporated in the Local Coordinators' Manual.
**Talent Testing Program**

**Record Form Z**

(Side 1)

---

**PLEASE PRINT:**
- **LAST NAME:**
- **FIRST NAME:**
- **MIDDLE NAME:**

**SCHOOL:**
- **CITY:**

**YOUR TESTING NUMBER:**

**TODAY'S DATE:**
- **Mo:**
- **Day:**
- **Year:**

**MAKE NO STRAY MARKS BELOW THIS LINE**

**DATE OF BIRTH:**
- **Mo:**
- **Day:**
- **Year:**

**YOUR NAME:**
- **Last name:**
- **First name:**
- **Middle name:**

**DATE OF BIRTH**
- **AGE AT LAST BIRTHDAY**
- **GRADE YOU ARE NOW IN**
- **LAST GRADE COMPLETED**

**Grade 1**
- **1st Half**
- **2nd Half**

**Grade 2**
- **1st Half**
- **2nd Half**

**Grade 3**
- **1st Half**
- **2nd Half**

**Grade 4**
- **1st Half**
- **2nd Half**

**Grade 5**
- **1st Half**
- **2nd Half**

**Grade 6**
- **1st Half**
- **2nd Half**

**College Freshman**
- **1st Half**
- **2nd Half**

---

**MARK WHETHER YOU ARE A BOY OR GIRL**
- **Boy**
- **Girl**

---

**MAKE NO STRAY MARKS BELOW THIS LINE**

**I. If you are a student in the school in which you are taking these tests, mark "Boy" or "Girl" here:**
- **Boy**
- **Girl**

**II. If you are not a student in the school in which you are taking these tests, mark "Boy" or "Girl" here:**
- **Boy**
- **Girl**

---

Copyright 1960 by University of chicago.
PRINT YOUR HOME ADDRESS BELOW:

ADDRESS ____________________________________________

CITY ____________________________________________ POSTAL ZONE __________

STATE __________________________________________

Now copy in the boxes below; one letter or digit per box. Leave a blank space between words. For instance: **LOS ANGELES**

--- YOUR HOME ADDRESS ---

<table>
<thead>
<tr>
<th>HOUSE NUMBER</th>
<th>STREET</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>POSTAL ZONE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

STATE
Blacken one space

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

CITY

STATE
Blacken one space

- Alabama
- Alaska
- Arizona
- Arkansas
- Calif.
- Colorado
- Conn.
- Delaware
- D.C.
- Florida
- Georgia
- Hawaii
- Idaho
- Illinois
- Indiana
- Iowa
- Kansas
- Kentucky
- La.
- Maine
- Maryland
- Mass.
- Mich.
- Minn.
- Miss.
- Missouri
- Montana
- Nebraska
- Nevada
- N.H.
- N.J.
- N.M.
- N.Y.
- N.C.
- N.D.
- Ohio
- Okla.
- Ore.
- Pa.
- Rhode Is.
- S.C.
- S.D.
- Tenn.
- Texas
- Utah
- Vermont
- Virginia
- Wash.
- W.Va.
- Wis.
- Wyo.
- Other

PLEASE...keep this record form clean.

Do not fold it or bend the corners.

MAKE NO STRAY MARKS BELOW THIS LINE

--- RECORD FORM Z (SIDE 2) ---
Figure 5-2. EXTRACT FROM LOCAL COORDINATORS' MANUAL

SPECIAL TREATMENT OF STUDENTS WHO DO NOT TAKE THE TESTS

1. Teachers should be instructed to fill out a Record Form Z for each student who is not tested, and to assign (but not hand out) an identification card to each such student, so that he will be identifiable by a six-digit "testing number."

2. On Record Form Z, Side 1, just to the left of the heading "TALENT TESTING PROGRAM" is a column of 39 answer space circles with the injunction "DO NOT MARK IN THESE SPACES."

This column is to be used by the teacher or Local Coordinator (or in some cases by the Regional Coordinator) to indicate the reason the student was not tested. One circle should be blackened in this column, using the following code:

Codes A - E: Absence
A. Excused absence: illness
B. Excused absence: other
C. Absence: habitual truant
D. Absence: reason unknown; has been to school at least once in the past month
E. Absence: reason unknown; has not been to school at all in the past month

Codes J - Q: Excused from taking tests (not necessarily absent)*:
J. Foreign student: non-English-speaking
K. Completely illiterate (though probably not mentally retarded)
L. Completely illiterate (presumably because of severe mental retardation)
M. Blind
N. Nearly blind
O. Paralysis, palsy, amputation, or some other severe handicap that prevents writing
P. Some other severe physical handicap that prevents taking the tests
Q. Temporary physical disability (e.g. broken arm)

Codes S - W: These codes apply only to the "special (15-year-old) sample," and primarily to those cases in it who are in college or not enrolled in any school. In general they do not apply to 15-year-olds in elementary school, although they may in some cases. These codes (S - W) should not be used if any of the codes J - Q is applicable.

S. Refused to take tests
T. Agreed to take tests but then failed to appear at scheduled time
U. In reform school, institution for delinquents, jail, etc. Permission for testing was withheld by the custodial institution.
V. Parent or guardian refused to permit the 15-year-old to be tested
W. Some other reason

3. In general, codes A - Q will be used by the teacher or Local Coordinator, to indicate why students enrolled in the schools were not tested. Codes J - W will be used by the Regional Coordinator to indicate why "special sample" cases were not tested.

*Before assigning any of these codes, the teacher or Local Coordinator should be thoroughly familiar with the contents of Paragraph IV on page 3 of the Teachers' Guide. In this paragraph principles to follow in determining whether to excuse a student from the tests are discussed.
letter designations—so it is clear that even though there were five of each there wasn't a one-to-one correspondence. In fact the whole thing was like a giant jig-saw puzzle; getting it all to fit together in a technically sound physical arrangement (e.g., with the test booklets organized in such a way that test items were not visible when test instructions were being read) was the very complex penultimate step in producing the battery. Since the volume involved was so large that one extra page in a test booklet could increase costs tens of thousands of dollars, we also had to be economical of space. The ultimate step was the printing—an enormous task in itself. The answer sheets were printed by MRC, of course, since only MRC was set up to produce answer sheets that would fit their machines within the very small tolerances allowed. Everything else was printed by the lowest bidder equipped to do a competent job—the Cuneo Press, in Philadelphia. It was done on such a tight time frame that it was necessary for someone (me) to stay on site in Philadelphia a couple of weeks to answer questions, look at galleys as soon as they were ready, and in general function as an interface* between TALENT headquarters in Washington and the massive printing operation. Eventually it got done—and we wound up with what we were told amounted to 18 freight carloads of printed paper! (Arrangements had been made for the materials to be shipped direct from the press to the schools all over the country, thus bypassing our headquarters. There was no way 18 freight carloads of paper could have been squeezed into our very tiny offices in Washington.)

G. What we ended up with

Readers who are interested in getting some idea of what sorts of instruments constituted the final battery and what sorts of measures they yielded are invited to look at Appendix A. Appendix A-1 shows, in tabular form, details concerning the final battery. Appendix A-2 consists of thumbnail descriptions of the principal tests.

* But in 1960 I didn't know I was an "interface"! The word interface used in that sense had not yet entered the general vocabulary.
   - Author
Chapter 6. PRE-"CRITERION-REFERENCED" CRITERION-REFERENCED TESTS
(The TALENT "Domain Tests")

A. The concept

The term "criterion-referenced tests" was coined by Robert Glaser in 1963 to contrast a certain kind of test score interpretation with a "norm-referenced" interpretation.

The basic Project TALENT battery, the construction of which was discussed in the previous chapter, consisted, of course, of conventional norm-referenced tests. None of the tests was, technically speaking, a criterion-referenced test (although a criterion-referenced interpretation possibly could be imposed on some of them, on an a posteriori basis). A criterion-referenced test is a test in which the scores have some sort of absolute meaning, rather than being meaningless except insofar as they can be related to the performance of some "norms group." Criterion-referenced scores (scores with absolute meaning) are numbers to which an interpretation such as one of the following can be attached:

A score of 96 means that the examinee can correctly compute 96 percent of all products of a 2-digit number by a 1-digit number.

A score of 60 means that the examinee can read the lead article in the Washington Post with "60 percent comprehension."

A passing score on the road test means that the driver can satisfactorily perform all the maneuvers one needs in one's repertoire in order to be a safe and competent driver.

A score of 120 on this vocabulary test means that one has an active vocabulary of 60,000 words.

A passing score on this end-of-unit examination means that the student has mastered the content of the instructional unit well enough to be able to cope with the next instructional unit, in which he will have to utilize the information he is supposed to have already acquired.
The above examples of criterion-referenced interpretations of test scores contain no particularly novel concepts—no concepts of sorts that were unheard of before the term "criterion-referenced" was coined in 1963. Clearly the notion of criterion-referenced tests has been around a lot longer than the name for them.

As a matter of fact, after construction of the regular Project TALENT Battery was well under way, John Flanagan proposed that in addition we develop criterion-referenced tests in several areas and equate them to the corresponding tests of the regular TALENT Battery, so that criterion-referenced performance (on the new tests) could be given norm-referenced interpretation (via the regular TALENT tests). Of course in presenting this idea John did not use the term "criterion-referenced". Remember, this was 1960, and the term did not exist yet. He called the proposed new tests "domain tests".

B. The tests

It was quickly agreed that the areas in which we should develop "domain tests" were Reading Comprehension, Vocabulary, and Spelling. There were to be two reading tests—one to measure comprehension of works of fiction and the other to measure comprehension of magazines. The vocabulary test was to be a measure of size of vocabulary and the spelling test a measure of the proportion of words one could spell correctly, in some suitable carefully defined population of words.

John Flanagan assumed personal responsibility for developing detailed plans for the two reading tests and getting them written, and Marion Shaycoft thus found herself with similar responsibilities in relation to the other two domain tests (Vocabulary and Spelling).

John arranged with Mary Willis (the same Mary Willis who, in the previous chapter, was identified as the author of our Mechanical Reasoning items and who also happened to be an experienced and expert writer of reading comprehension tests) that she would write the two domain Reading Comprehension Tests. For Test A-1 (Authors), John and Mary jointly
picked out 10 eminent writers of fiction, ranging from authors whose writings were very easy to understand to authors generally considered difficult. Specifically, the authors were Louisa May Alcott, Robert Louis Stevenson, Rudyard Kipling, Willa Cather, Sinclair Lewis, Jules Verne, Jane Austen, Joseph Conrad, Feodor Dostoevski, and Thomas Mann. (I leave it to the reader to decide whether that list is supposed to be arranged from easy to hard or from hard to easy.) Similarly for Reading Comprehension Test A-2 John and Mary picked 10 magazines, ranging from movie magazines to prestigious slick-paper publications. (Actually 11 magazines were picked, but two of them, which were both movie magazines, were treated as a single publication.)

For each of the 10 authors two books were selected, and one passage per book was chosen for the test. Similarly two passages were chosen from each of the 10 magazines. Five multiple-choice test items were written per passage, in accordance with a specific pre-defined pattern; for each passage there was to be one item of each of the following types:

1. An item testing general understanding of the passage—comprehension of its general idea or intent;
2. An item testing comprehension of one of the simplest points in the passage;
3. An item testing comprehension of a point of average difficulty (relative to the difficulty of the passage);
4. An item testing comprehension of the most difficult point in the passage;
5. An item testing appreciation of the passage, or an application of the ideas in it, or an inference drawn from it.

As for the Vocabulary Domain Test, which contained about 300 items, it was designed to provide an estimate of the number of word-meanings (not words) in the examinee's "comprehension vocabulary" (the meanings he understands, as opposed to his "active vocabulary," the meanings he uses in his speech or writing).

The 150-item Spelling Domain Test provided an estimate of the number of words, among the most frequently used 5,000, that the examinee could spell correctly.
All these tests required very detailed rationales, with precise rules for sampling (of passages in the case of the Reading Tests, word meanings in an unabridged dictionary in the case of the Vocabulary test and words in the Teacher's Word Book of 30,000 Words*, in the case of the Spelling test), and with very specific rules for item construction. Results and further details on the tests' construction are reported in the following publications:

For the two reading tests:


For the spelling test:


For the vocabulary test:

Chapter 7. EXECUTION OF THE PLANS

This chapter tells what was actually done—beyond putting the finishing touches on the Project TALENT Battery (the development of which is described in Chapter 5)—in the fall of 1959 and the spring of 1960 to convert Project TALENT from a concept to a reality.

A. Selecting the sample

Late in the fall of 1959 we were ready to select the sample. The sampling procedure had been decided upon (see Chapter 3, Section A). The National Center for Educational Statistics (NCES), at that time a branch of the Office of Education, had made available to us preliminary decks of cards containing the basic information we would need in our stratification procedures (and a lot of other information) from their 1958–59 National Survey of Public Secondary Day Schools. This was eventually to become a directory, providing basic information about all public secondary schools in the United States (both junior high schools and senior high schools). There were, in very round numbers, some 30,000 schools in this population, and NCES had six IBM cards chock-full of information about each school. We needed three of the six cards, so NCES obligingly made available to us over 90,000 IBM cards. If the same thing was being done today, all that information would probable be on just one 11½-inch reel of tape—but in 1959 cards were the basic unit, and one never bypassed them by going directly from data to tape. Anyhow perhaps the 90,000 cards were good practice for the time when about five million cards containing raw data from the administration of the TALENT battery (see Section E of this chapter) would be deposited in the lap of our programmer at the University of Pittsburgh Computation and Data Processing Center. But that is another (and longer) story.

To get back to the sampling, our staff programmer, Gary Lotto, wrote a program to select the sample on the IBM-650 (a terribly primitive computer by contemporary standards, but regarded in 1959 as practically the last word in computer technology).
He ran the 90,000 cards and lo and behold, his program produced a printout listing about 800 schools—the initial 800 public school primary sampling units. These were of course all senior high schools—i.e., schools that went at least as far as grade 11*—since that was how we had defined the "primary sampling units". The corresponding junior high schools (the "feeder schools") were to be identified later, after the cooperation of the senior high schools had been secured.

To select the sample of parochial schools, and of other nonpublic schools, we secured lists of these two populations of schools. The National Catholic Education Association provided the list of parochial schools and NCES gave us a list of the other nonpublic schools.

Since neither the parochial schools nor the other nonpublic schools were on IBM cards, selection of those parts of the sample was done clerically rather than by computer. The procedure, however, was identical, except, of course, that by prior decision the only stratification variable used for these two categories of school was geographical location. Randomization was effected through the use of tables of random numbers. In selecting these nonpublic schools, in effect humans were imitating a computer (rather than vice versa).

As for the population of public secondary schools, it turned out that even the 30,000-odd schools represented by the 90,000 cards provided by NCES did not constitute the complete population. Since NCES's 1958-59 survey was still in progress, questionnaires completed by schools were dribbling in to them. These late arrivals were of course not yet on IBM cards, but they could not be ignored. Therefore whenever NCES had accumulated a substantial number of these additional schools they notified us by phone. A project staff member would then go down there** and, again "imitating a computer", would select a sample. The very last dribble consisted of three "late late schools", for

* Grade 11 was used in the definition, instead of grade 12, to allow schools from 11-year school systems, if any, to be included. (These 11-year systems, which had been prevalent in the South, were rapidly being phased out at the time, and none was selected.)

** Fortunately, Project headquarters were in Washington at the time, so it was easy to get to NCES.
which questionnaires turned up at the very last minute—as a matter of fact not long before the period set aside for conducting the testing was almost over. Using a table of random numbers we selected one of the three, and fortunately, despite the short notice it agreed to participate. For students in this single late addition to the sample a weight of 3.0 (the reciprocal of the effective sampling ratio) had to be assigned, rather than the more usual 13.0, or 20.0, or 50.0 (or a still higher number, to compensate for nonparticipating schools).*

Users of the TALENT tapes encountering for the first time the record of a student or school with a weight as low as 3.0 almost invariably are convinced it is an error in the tape. It isn't.

A supplementary source of public schools was the Internal Revenue Service. The IRS had a list which included some schools for which NCES did not have questionnaires. Accordingly, we sampled from the IRS list too.**

But in spite of the physical problem of juggling 90,000 cards, and the nuisance of coping manually with late additions and with non-public schools, that was really the easiest part. Now we came to the hard part—or at least what we anticipated would be the hard part: getting the cooperation of the selected schools (and their feeder junior high schools). The next section of this chapter tells, among other things, about the pleasant surprise that was in store for us.

B. Getting the cooperation of the selected schools

As things turned out, getting cooperation was a lot easier than we had anticipated (and probably a lot easier than it would be today). This was in part because of the careful groundwork that had been laid in the years from 1957 to 1960 (see Chapter 2) and partly because of the dedication and competence of our Regional Coordinators.

* See Chapter 15 for a discussion of weights.

** Further details about the sampling procedure can be found in Chapter 3 of "Project TALENT: Designing the Study", by Flanagan, Dailey, Shaycoft, Gorham, Orr, and Goldberg (1960); or in Chapter 3 of "Design for a Study of American Youth" by the same authors (1962).
We had appointed about 90 Regional Coordinators—mostly members of Psychology or Educational Research faculties of colleges and universities all over the country. (The Regional Coordinators are listed in Chapter 8, Section C.) It was these Regional Coordinators who contacted the school superintendents and principals, told them about Project TALENT, and made them eager to participate. The activities of the Regional Coordinators were coordinated in project headquarters by Bill Gorham, who was in charge of getting the selected schools to cooperate and getting the tests administered.

First the good news!

Although the task of getting the schools' cooperation was not without problems, the overall success rate was an astonishing 93 percent. And in 19 states, including all six of the New England states, it was actually 100 percent!

As a matter of fact we got more than we bargained for in the way of cooperation. About one in 13 of New York City's high schools and about one in 20 of Chicago's had been selected. But headquarters of the New York City school system included a Bureau of Educational Research, headed by the very knowledgeable J. Wayne Wrightstone. Wrightstone and his staff pointed out that it would be very difficult, perhaps impossible, to disrupt an entire high school as large as those in New York City, for two whole days. Fortunately Wrightstone recognized the value of Project TALENT and wanted to participate. He immediately recognized what we had known right along—that if one in 13 of New York's schools was a good sample (as indeed it was) one in 13 of the students would be a superb sample. The reader will recall, from the discussion in Chapter 3, Section A, that we had decided to make the school, rather than the student, the primary sampling unit, not because we felt that that would have any theoretical advantages (which of course it wouldn't) but for purely practical reasons.
Thus we were absolutely delighted when Dr. Wrightstone proposed testing every 13th student in every secondary school (both junior and senior highs). And we were even more pleased when, somewhere along the line, the plan changed from every 13th student to every 12th student.

Authorities in the Chicago school system decided to follow suit, on a somewhat more restricted basis. Instead of the two schools that had originally been selected from the Second City's 38, the Chicago school authorities proposed that we select 10 times as many (i.e., 20 schools). They would then administer the TALENT Battery to every tenth student in those schools. Again, we were delighted, and again we acceded rapidly to this suggestion.

(It should be noted that via appropriate modifications in weights the additional schools tested in New York and Chicago could be treated as additional primary sampling units.)

Despite the better-than-we-asked-for cooperation from the two largest cities, little* Knoxville, Tennessee and the surrounding county put them to shame, by agreeing to test every student, in a five-grade range (i.e., grades 8-12 rather than the usual 9-12 range) in every school--public, and parochial, and private--and not just in the city of Knoxville but in the entire county. This coup was achieved primarily through the energy, initiative, and foresight of our Regional Coordinator in the Knoxville area, Louise Cureton. It was she who realized what a marvelous resource this would be for future research, and it was she who took the lead (after clearing the idea with Project headquarters) in selling the Knoxville and Knox County school authorities on the idea. Thus for that area we had the entire grade 9-12 population (plus grade 8) rather than just the two public schools initially selected.

* Little, at least, in comparison with those two metropolitan giants, New York and Chicago.
In addition to the more-than-we-expected instances of cooperation described above, there were other cases as well where school authorities asked whether a school that had not been selected for the sample might still be included. In many cases this occurred where there were only two schools in a community and if only one of them had been selected (as was likely to be the case) the school authorities felt it might be awkward to test just that one and not both. This was particularly likely to be the case in southern communities, where there might be one white and one black school (in 1960 there was still quite a bit of *de jure* segregation in the South despite the fact that it had been declared unconstitutional in 1954). Whenever a school system specified, as a condition of their permitting the selected school to participate, that we would have to allow the other school to participate too, we agreed. But since in most cases including the extra schools in the sample could have destroyed the sample's probability-sample characteristics, which would have been disastrous, we called the extra schools "non-probability schools". This meant that data from those schools would be left out of any data analyses that called for a probability sample. The non-probability data could, however, legitimately be included in any analyses that did not require the probability sample feature.

Actually most of the non-probability schools might have legitimately been included in the probability sample if a procedure developed by the author of this "History" had been applied. But that procedure, which is described in Appendix D1, wasn't even developed until about 15 years later.

Other non-probability cases occurred in some schools in the probability sample, where the school principal wanted to test eighth-graders, as well as the usual four grades, 9-12. When this request was made, we again acceded, and merely designated the eighth-graders "non-probability cases".
In addition to the "self-invited" nonprobability cases, the reader will recall that we had decided that in one-tenth of the school districts corresponding to probability-sample public schools we would test all 15-year-olds not in the grade 9-12 range. Many of these turned out to be in grade 8; some were in grade 7; and a few were in even lower grades. Some, of course, were not in school at all, but these were particularly hard to locate, since essentially they were truants, all states at that time requiring school attendance until at least age 16. In any event these 15-year-olds not in the grade 9-12 range clearly were not in the grade 9-12 probability sample—although, equally clearly, they were in the probability sample of 15-year-olds. In schools in which there were a very large number of 15-year-olds to be tested in grade 8, school authorities quite often opted to test the entire grade. That gave us still more eighth-graders in neither the grade 9-12 probability sample nor the probability sample of 15-year-olds. Besides the cases not in grades 9-12 tested because they belonged in our age 15 sample, there was at least one other group not in the main probability sample (grades 9-12) whom we invited to participate. This group consisted of selected schools trying one of two experimental programs in mathematics. One of the programs had been developed by Dr. Max Beberman, of the University of Illinois, and the other by Dr. E. G. Begle, who was then at Yale and is now at Stanford. The idea at the time was that the performance of students in the new program could be compared with that of appropriately matched students taking conventional mathematics courses.

And now for the bad news!

The good news, as related above, is that 93 percent of the invited schools agreed to participate, and we were bombarded with requests to permit additional schools or groups to participate. The bad news is that the success was not uniform. We did run into trouble (or at least a profound lack of enthusiasm) in a couple of

* No, this has not been done.
places—notably in New Jersey and California (particularly Los Angeles). The acceptance ratios among the sampled public schools in those areas were 63.6% (7 out of 11 schools) for New Jersey, 50.0% (2 out of 4) for Los Angeles, and 58.3% (14 out of 24) for the rest of California. There had been some clearly expressed opposition to Project TALENT in the state department of education in New Jersey because they viewed the project as Federal intervention. Frederick M. Raubinger, in a speech to the American Association of School Administrators early in 1959 said that Project TALENT would be "the camel's nose under the tent". Raubinger was New Jersey's Commissioner of Education at the time, and he probably was at least indirectly responsible for the comparatively poor cooperation rate in New Jersey. After all, it is not surprising if school personnel hesitate to take actions of which the state's Commissioner of Education has publicly indicated he disapproves. But the "bad news" wasn't really so bad when one considers that the opposition to Project TALENT substantially affected cooperation in only two states, and that even in those two states well over half the invited schools participated in the Project.

Identifying and contacting the corresponding junior high schools (feeder schools)

For those schools in our sample that regularly received as tenth-graders students from local junior high schools, it was necessary to identify the appropriate junior high schools, contact them, and secure their agreement to have their ninth-grade students participate in Project TALENT. This was carried out almost entirely by the Regional Coordinators, operating on the basis of instructions from Bill Gorham. (Largely because of the time pressure, most of these instructions were transmitted by telephone.) All of the invited junior high schools, as far as our records show, agreed to participate. (The fact that the senior high schools to which their students would go had already agreed to participate, and that this meant, in most
cases, that the school superintendent was backing the enterprise, must have helped considerably in achieving this 100% record. Nevertheless it is a record to which we can point with pride.)

However it must be admitted that if our grade 9-12 probability sample has one weak spot it is in the ninth-grade sample, and its source is the identification of the feeder junior high schools. In the first place the instructions for defining feeder schools, having been transmitted orally, almost certainly were not interpreted uniformly by all Regional Coordinators. In the second place the dilemma of how to cope with the partial overlap between junior and senior high schools had not been faced up to in formulating the instructions to Regional Coordinators. (Some senior high schools in the sample received their tenth-grade students from several junior high schools, and some junior high schools feeding students to the selected senior high schools also sent students to senior high schools not in the sample.) Furthermore, as far as I know, a technically immaculate solution to this sampling problem had not been developed at the time, so that even if the dilemma had been faced up to in developing instructions for the Regional Coordinators it might not have been resolved in a completely satisfactory manner. A technically sound solution to the problem now exists. The author of this "History" developed it some 15 years later in connection with work on an entirely different problem, not related to Project TALENT in any way. She dubbed the solution "The 'Closed School-Cluster' Procedure for Selecting a Probability Sample of Schools". It is described in Appendix D2. (This is an instance of what is referred to in the Preface as the "HIBK" phenomenon.)

It should be clearly understood, however, that despite the fact that the messy (and in 1959 unsolved) problem created by the frequent lack of one-to-one correspondence between senior high schools and feeder junior high schools resulted in a less-than-perfect ninth-grade sample, it is nevertheless a good sample—in fact a very good sample.
But if a second Project TALENT--what we sometimes refer to facetiously as "Son-of-TALENT"--or anything resembling it is ever undertaken, it is hoped sincerely (at least by the author of this History) that the "closed school-cluster" procedure will be used, so that the sample for grade 9 will be every bit as good as those for grades 10, 11, and 12--and nobody has ever suggested that the samples for those three grades are anything less than excellent.

As has already been implied, at the time that the Regional Coordinators determined which junior high schools were feeder schools, and should therefore be invited to participate, no data were formally collected in written form concerning how many tenth-graders each senior high school in the sample received from each junior high school and how many students each of the potential "feeder schools" thus identified sent to senior high schools not in the sample. These data were needed by Marion Shaycoft, who had completed her responsibility by then of getting the test battery developed, the test booklets organized, and the five answer sheets planned, and therefore was free to turn her attention to the problem of school weights. Thus at that point data were urgently required so that there would be some reasonable basis for determining as well as could be done under the admittedly not perfect circumstances the weights to be assigned to each feeder school so that the result would be as close as could be achieved to a flawless probability sample. (Bear in mind that the closed school-cluster method didn't yet exist, so there was no way the sample or the weights could be that good, but we were unaware at that time that such a solution was waiting to be found. We recognized only that we had to do the best we could with what we had.) Under Bill Gorham's direction, therefore, an effort was made by the Regional Coordinators in areas where junior high schools existed to obtain the needed data retroactively. When these data were examined the uncomfortable suspicion was confirmed that there was one segment of the grade 9 population for whom the probability of inclusion in the sample as drawn was zero, and thus not in accordance with one of
the main tenets of probability sampling, namely that every member of the population sampled must have known probability, greater than zero, of inclusion in the sample. The junior high schools with zero probability of inclusion that affected the quality of the ninth-grade sample were some of those in districts where the senior high school was selected for the sample. The junior high schools in question were the ones that did not meet whatever definition of "feeder school" had been locally operative in terms of the Regional Coordinator's operating principles.

This defect had to be corrected if at all possible. The way to do it was to identify the affected schools, select a random sample of them, and persuade them, at that late date, to participate. (By the time this particular crisis came to light it was late in the spring of 1960, and testing was almost over.) John Dailey at that point assumed personal responsibility for executing the selection of this small supplementary sample. He visited NCES, identified the junior high schools (almost 150) in the affected school districts, and then, using a table of random numbers, selected one in 20, or a total of seven. Subsequent checking against our records indicated that two of the seven were already in our sample. The remaining five constituted a technically sound probability sample of the segment of the population for which we were trying to get representation. The five junior high schools were contacted immediately, but because of the lateness of the date only two could participate. (These two schools, therefore, turned out to have extremely high weights. See Chapter 15.)

[Author's Note

As the author of this "History" I have used a fairly substantial amount of space and a fairly substantial number of words to relate the details of this particular irregularity. However, since I would not like to see its importance blown up out of all proportion it seems desirable to explain that I have gone into all this detail principally for two reasons: first, because I think future researchers should learn from our errors, and second, because to the best of my knowledge this particular irregularity has never really been documented, except in Project files.]
Figure 7-1 is the list of participating schools. Figure 7-2 is a map showing how they are distributed.

Almost all of the schools that declined to participate did so reluctantly, and only because of scheduling problems or other administrative difficulties. Therefore we decided to attempt to collect school data (i.e., response to the three school questionnaires) from them even though student data would not be secured. This effort was very successful. Almost every school that had declined to participate in the testing was entirely willing to complete the school questionnaires.

C. Administering the tests

"There is a school in the mountains of Eastern Kentucky which can be reached only by a gravel road bounded on either side by steep hills heavily wooded with pine. Occasionally the woods open sufficiently to reveal a cabin with an outbuilding, then the trees close in again, the steep slopes silently hostile to attempts to use the land for a livelihood. At a sharp fork in the road, one turns to the right and the road dribbles out and is no more. A wide shallow brook outlines the foot of the hill directly ahead. Across the brook are a school and accompanying dormitories. The dormitories are necessary, for the isolated and rugged country makes daily trips to and from school impossible. This is the Red Bird Settlement School.

"At the corner of West 205th Street and Golden Avenue in the Bronx, New York City, students gather from all over the city to study calculus and advanced courses in science. A high proportion of them will eventually earn doctoral degrees and assume key positions in science. This is the Bronx High School of Science.

"Off a residential street in Washington, D.C., a driveway sweeps into a broad arc up a hill in front of the main entrance of a modern parochial secondary school. The structure is of contemporary brick-and-glass design. The school appears to convey an impression of a quiet and efficient atmosphere. Students move from room to room with a bearing which seems to reflect their pride in the physical plant itself. This is the Archbishop Carroll High School."

--Design for a Study of American Youth, pp. 43-44
Figure 7-1. LIST OF SCHOOLS IN SAMPLE AS OF 22 JUNE 1960

ALABAMA
ALPINE—Winterboro HS
BURKE—Auburn HS
DOKE HS
ERNGRAM—Woodlawn HS
CHILDERSBURG—Childersburg HS
FIVE POINTS—Five Points High School
GADSDEN—Emma Sansom HS
GREENVILLE—Greenville Tng. HS
HEMEN—Heflin Tng. S
HUNTSVILLE—Oakwood College Acad.
JACK—Zion Chapel HS
JASPER—Walker Co. Tng. S
LAFAYETTE—Chamber Co. Tng. S
LANEET—Lanett HS
LEROY—LeRoy HS
MOBILE—Dunbar JHS
PISGAH—Pisgah HS
PLANTERSVILLE—Dallas Co. HS
RED LEVEL—Red Level HS
ROCKFORD—Coosa Co. HS
SATSUMA—Satsuma HS
TUSKEGEE—Tuskegee Inst. HS
VINCENT—Vincent HS
WOOL腰—Woodland HS

ALASKA
SITKA—Sitka HS

ARIZONA
DOUGLAS—Douglas HS
Douglas JHS
ST. MICHAEL—St. Michael’s HS
THATCHER—Thatcher HS

ARKANSAS
BATESVILLE—Batesville HS
DUMAS—Reed HS
HECTOR—Hector HS
HELENA—Lake View HS
LUKOR—Lure HS
MAGNOLIA—Magnolia HS
MARMADUKE—Marmaduke HS
MELBOURNE—Melborne HS
MONTICELLO—Childress HS
MT. HOLLY—New Hope HS
MT. VERNON—Mt. Vernon HS
MOUNTAIN VIEW—Mountain View HS
PINE BUFF—Coalmont HS
WOODROW Wilson JHS
POTTSVILLE—Pottsville HS
SMACKOVER—Smackover HS
TEXARKANA—Washington HS
WILSON—Wilson Trade HS

CALIFORNIA
ANAHEM—Marywood HS
ARLINGTON—Divine Word Seminary
BURBANK—Burbank HS
CARSON—T. Washington HS
CARLISLE—Army and Navy Academy
CARME2—Carmel HS
CENTRAL VALLEY—Central Valley HS
CHULA VISTA—Castle Rock JHS
CHULA VISTA—JHS
CROCKETT—John Swett Unified HS
DURHAM—Durham HS
FAIRFIELD—Armiio Joint Union HS
GORDON—Junipero Serra HS
LE GRAND—Le Grand Union HS
LOS ANGELES—Our Lady of Loretto HS
PHOENIX—Phoenix Union HS
WILMINGTON JHS
WOODLAND HS
MEATHUR—Folli River HS
MERCED—Our Lady of Mercy S
MODESTO—Modesto City Academy
MORGAN HILL—Live Oak Union HS
MOUNTAIN VIEW—St. Joseph’s College

NEEDLES—Needles Union HS
PORTOLA—Portola J-SHS
REDLANDS—Cape HS
RED HILLS HS
REDLANDS HS
SACRAMENTO—St. Francis HS
SAN FRANCISCO—Mercy HS
SAN MATEO—Junipero Serra HS
SANTA CRUZ—Holy Cross HS
TERRIF—White HS
YUCAipa—Yucaipa J-SHS

COLORADO
BETHUNE—Bethune HS
CROOK—Logan Co. Branch HS
DENVER—Cathedral HS
HORACE Mann JHS
LAKE JHS
NORTH SHS
SKINNER JHS
DOLORES—Dolores HS
JOES—Libby HS
LAFAYETTE—Lafayette HS
MORRISON—Bear Creek HS
NUCLS—Nuda HS
STEAMBOAT SPRINGS—Steamboat HS

CONNECTICUT
COLLINSVILLE— Canton HS
HARTFORD—Northeast HS
KENT—Kent School
LITCHFIELD—Litchfield HS
MERIDEN—Oville H. Pratt HS
NEW LONDON—St. Bernard’s HS
POMFRET—Pomfret School
PUTNAM—Putnam Catholic Academy
SOUTHBURY—Southbury HS

DELWARE
BRIDGELAND—Bridgeville HS
DELWARE CITY—Delaware City HS

DISTRICT OF COLUMBIA
WASHINGTON—Archbishop Carroll HS
ST. Anthony HS

FLORIDA
CENTURY—Century HS
FELSMIRE—Fellsmere HS
FT. LAUDERDALE—Central Catholic HS
GAINESVILLE—Buchholz JHS
GAINESVILLE HS
WESTMINSTER HS
GAINESVILLE—Greenlee HS
JACKSONVILLE—Isatiah Blocker JHS
JAMES W. Johnson JHS
NEW STANTON HS
LAKE BUTLER—Union Co. HS
MARIETTA—T. Washington HS
MULBERRY—Mulberry JHS
OCALA—Ole Miss HS
PALMETTO—Palmetto HS
PENSACOLA—A-V. Azusa HS
TRENTON HS
BRENTWOOD HS
BROOKSVILLE HS
ESCAPADE HS
GLENWOOD HS
Holla JHS

SANFORD—Seminole HS
WARRINGTON—Warrington JHS
YULEE—Yulee HS

GEORGIA
ATLANTA—Sylvan Hills HS
ATLANTA—Aquinas HS
BAINBRIDGE—Bainbridge HS
CAIRO—Washington Consolidated HS
CLAXTON—Evans Co. HS
DUBLIN—Dublin HS
FOREST PARK—Wm. A. Fountain HS
FORSYTH—Hubbard HS
FORT VALLEY—H. A. Hunt HS
GAINESVILLE—South Hall HS
HAWKINSVILLE—J. L. Busmay Tng. Ins.
MITTER—Conner Co. Tng. School
NEWnan—Newnan-Coweta Central HS
OCILIA—Irwin Co. HS
PERRY—Perry Co. Tng. School
RICHMOND HILL—Geo. Red JHS
ROSSVILLE—Rossville HS
ROSSVILLE—JHS
SMYRNA—Cobb HS

SPRINGFIELD—Effingham Co. HS
VANNA—Vanna JHS
ZEBULON—Pike Co. HS

HAWAII
HONOLULU—University HS
WAIAHALU—Ha-Neiahu High & Int. Sch.
WAILUKU—Moai—Baldwin HS

IDAHO
BOISE—Boise Junior Academy
GOODING—Gooding Academy
KENDRICK—Kendrick HS
WALLACE—Our Lady of Lourdes Academy

ILLINOIS
ARINGTON—Abington SHS
ALTON—Western Military Academy
ARCOLA—Arcola HS
AURORA—Marmion Military Academy
BELVIDERE—Belvidere HS
CARRIER MILLS—Carrier Mills HS
CHAPPAQUA—Chapppa School
EDISON JHS
FRANKLIN JHS
CHICAGO—Amundsen HS
(S缩号 Foornotes2) Carver HS
Crate HS
Dandle HS
Englewood HS
Farragut HS
Fenger HS
Francis W. Parker School
Harrison HS
Hirsch HS
Hyde Park HS
Kelly HS
Kelvin Park HS
St. Benedict HS
St. Patrick HS
South Shore HS
Steinmetz HS
Sullivan HS
Toft HS
Tuley HS
Walls HS
Wells Phillips HS
ELLISWOOD—Ellsworth HS
FISHER—Fishers HS
FAIRFIELD—East Leyden HS
FREEPORT—Aquin Central Catholic HS
GORGAM—Gorham HS
JOHNSTON—Johnston City HS
LENON—St. Vincent DePaul Seminary
LEROY—LeRoy HS
MACOMB—Edison HS
MACONB—Maconb HS
MARIASSA—Mariass HS
MACHEN—Michigan HS
MERIDIAN—Walther Lutheran HS
METROPOLIS—Metropolis HS
MONMOUTH—Monmouth HS
MT. OLIVE—Mount Olive HS
MT. CARMEL—Mt. Carmel HS
MT. MORRIS—Mt. Morris HS
MORING—West Leyden HS
OREGON—Oregon HS
PAYSON—Paysontown HS
ROSE—Rose HS
RED BUD—Fredmiu HS
RIVERO—Rivero HS
SESSON—Sydney-Baime Tng. HS
SHANNON—Shannon Co. Unit
THOMPSONVILLE—Thompsonville HS
VIOLA—Viola HS
WALNUT—Walnut HS
WITT—Witt HS
WOODHULL—Woodall HS
WOOD RIVER—East Alton-Wood River HS
ZIEGLER—Ziegler HS

INDIANA
ARLEDS—Dallas Twp. HS
BAINBRIDGE—Bainbridge HS

BATTLE GROUND—Battle Ground HS
DUNKIRK—Dunkirks HS
EAST GARY—St. Francis HS
THAS. A. Edison JHS
EVANSVILLE—Central HS
GREEN—Green HS
GREENSBORO—Sandkreck HS
HANNA—Hanna HS
HARRINGTON—Harford City HS
INDIANAPOLIS—Arsenal Tech HS
AVON HS
LAWRENCEBURG—Lawrenceburg HS

LOGANSPORT—Loganport HS
MADISON—Central HS
MADISON HS
MARENGO—Marengo HS
MARION—Marion HS
MARKS—Union Center HS
MAXWELL—Hancock Central HS
MEXICO—Mexico HS
MICHIE—Mitchell HS
NEW AUGUSTA—Pike Twp. HS
NEW CASTLE—New Castle HS
NEW GOSHEN—Fayette Twp. HS
PLAINFIELD—Espin HS
SELSA—Center HS

IOWA
ADEL—Adel HS
ANTHON—Anthon-Oto Ind. HS
CEDAR RAPIDS—Regis HS
CHARLES CITY—Charles City HS
CHARLOTTE—Charlotte Ind. HS
CLARENCE—Clarence HS
DANBURY—Danbury HS
DENISON—Denison HS
DOLIVER—Doliver HS
DYERSBURG—Evansville HS
ELOON—Elon HS
GOLDFIELD—Goldfield HS
GRISWOLD—Griswold HS
GUDBROUN—Begginsley HS
HACKENSACK—St. Mary HS
LAGOTA—Lagota HS
LAMONT—Lamont HS
LANKERSHIM—S. HS
LE CLAIR—Le Claire Ind. School
LE MARS—Gehlen HS
LIBERTY CENTER—Southwest Warren HS
MOMA—Muncie HS
MOUNTAIN—Mount Carmel HS
MT. VERNON—Mt. Vernon HS
PAULINA—Paulina HS
PROVINTION—Redland HS
REEDSBURG—Boone Valley HS
RICEVILLE—Riceville HS
TIPTON—Tipton HS

KANSAS
ARKANSAS CITY—Arkansas City HS
BELOIT—St. John HS
BREWSTER—Brewster HS
COLLIER—Cullier HS
CLAY CENTER—Clay Co., HS
ELK CITY—Elk City HS
EUPEKA—Eupeka HS
GREAT BEND—Great Bend HS
Harrison HS
ROosevelt HS
GREENLEAF—Greenleaf HS
HORSESHOE—Horseshoe HS
OLALA—Olala HS
WILCOX—Westley HS
WYANDOTTE—Wyandotte HS
ZAGREB—Zagreb HS

KENTUCKY
ALM—Almo HS
AUBURN—Auburn HS
AUDRIE—Audovie HS
BEVERLY—Red Bird Settlement HS
BLACKETT—Letcher HS

CHICAGO—by special arrangement with the Chicago Board of Education, 1 out of every 10 students in 25 of the 35 secondary schools in the city, were randomly selected and tested. These 25 schools were selected at random.
NEW YORK CITY—By special arrangement with the Board of Education, 1 out of every 12 students in every secondary school in the city was selected at random and tested.
Figure 7-1 (cont.)

**EASTOVER—Webber HS**
**GREENE—Greeneville HS**
**GREENWOOD—Greenwood HS**
**GREENVILLE—Greenville HS**
**LAWRENCEVILLE—Lawrenceville HS**
**MADISON—Madison HS**
**MARION—Marion HS**
**MAYSVILLE—Maysville HS**
**MAYWOOD—Maywood HS**
**TAMASSEE—Tamassee DAR HS**
**WELFORD—Florence Chapel HS**
**WHITMIRE—Whitmire HS**

**SOUTH DAKOTA**
**BERESFORD—Beresford HS**
**BRANDT—Brandt HS**
**DOLAND—Doland HS**
**MARTIN—Martin Co. HS**
**NEW HOLLAND—Dakota Christian HS**
**RAPID CITY—Rapid City HS**
**REDFIELD—Redfield HS**
**SCOTLAND—Scotland HS**
**SELBY—Selby HS**
**VIBORG—Viborg HS**
**WITTEN—Witten HS**

**TENNESSEE**
**BEECH BLUFF—Beech Bluff HS**
**BULLS GAP—Bulls Gap HS**
**CHATTANOOGA—The McCallie School**
**DANDRIDGE—Maury HS**
**DUNLAP—Sequatchie Co. HS**
**FOUNTAIN CITY—Fountain Halls HS**
**GOODLETTSVILLE—Goodlettsville HS**
**HARRIMAN—Harriman HS**
**KNOX COUNTY—Concord HS**
**LADY VIOLA—Lady Viola HS**
**MADISONVILLE—Madisonville HS**
**MAYSVILLE—Maysville HS**
**NEEDSVILLE—Needsville HS**
**OVERTON—Overton HS**
**PEARSALL—Pearsall HS**
**PORT ARTHUR—Port Arthur HS**
**QUEEN CITY—Queen City HS**
**RAYMONDVILLE—Raymonville HS**
**REDWOOD—Redwood HS**
**SAN ANTONIO—San Antonio HS**
**SEAGOVILLE—Seagoville HS**
**S. HOUSTON—S. Houston HS**
**SPANGDALL—Spangdall HS**
**SPOKANE—Spokane HS**
**WHITMIRE—Whitmire HS**

**UTAH**
**CIRKLINER—Cirkliner HS**
**HILLSDALE—Hillsdale HS**
**MONTICELLO—Monticello HS**
**MONROE—Monroe HS**
**NEEDSVILLE—Needsville HS**
**OVERTON—Overtown HS**
**PEARSALL—Pearsall HS**
**PORT ARTHUR—Port Arthur HS**
**QUEEN CITY—Queen City HS**
**RAYMONDVILLE—Raymonville HS**
**REDWOOD—Redwood HS**
**SAN ANTONIO—San Antonio HS**
**SEAGOVILLE—Seagoville HS**
**S. HOUSTON—S. Houston HS**
**SPANGDALL—Spangdall HS**
**SPOKANE—Spokane HS**
**WHITMIRE—Whitmire HS**

**WISCONSIN**
**BEAVER DAM—Beaver Dam HS**
**CLEAR LAKE—Clear Lake HS**
**DURAND—Durand HS**
**GREEN BAY—Green Bay HS**
**KENOSHA—Kenosha HS**
**KENDED—Kenedy HS**
**LADYSMITH—Ladysmith HS**
**LODI—Lodi HS**
**MADISON—Madison East HS**
**MARKESIN—Markesan HS**
**MILWAUKEE—Milwaukee HS**
**MINERAL POINT—Mineral Point HS**
**OAKFIELD—Oakfield HS**
**ONTARIO—Ontario HS**
**PRAIRIE DU SAC—Prairie Du Sac HS**
**PRESCOTT—Prescott HS**
**SALEM—Salem HS**
**SOLON SPRINGS—Solon Springs HS**
**TREMBLEAU—Trembleau HS**
**WEST ALLIS—West Allis HS**

**WYOMING**
**GREEN RIVER—Green River HS**
**SUNRISE—Sunrise HS**

---

**KNOX COUNTY—For a special regional study, all students in grades 8-12 were tested in all schools in Knox County and Knoxville.**
Figure 7-2. THE NATIONAL SAMPLE OF SCHOOLS
All three of the schools mentioned in the above quotation were in Project TALENT. They are mentioned only to give some notion of the tremendous diversity among the schools included. This diversity necessitated considerable flexibility in arrangements for test administration. In each participating school a Local Coordinator was appointed (usually by the principal in consultation with the Regional Coordinator). This Local Coordinator, working under the direction of the Regional Coordinator, trained the teachers who would actually administer the tests. He (or she) was also responsible for receiving and inventoring the test materials, distributing them to the teachers, collecting them after the testing sessions were completed, supervising the necessary checking operations, and shipping the answer sheets to Measurement Research Center (MRC), in Iowa City, where they would be processed.

The test battery required two full days, but there was considerable flexibility in how the requisite four half-days might be arranged. For instance in some schools the tests were administered during four consecutive mornings.

Each student in the sample was given—and asked to keep—an identification card (the stub of the IBM card on which his six-digit identification number was printed) identifying him as a member of the Project TALENT sample. The TALENT identification card bore, in addition to the six-digit number and the student’s name (which he filled in himself) the mailing address of Project TALENT, and a request that sample members keep us informed of address changes and name changes. The students seemed to like these ID cards. (Several years later, a sample member wrote us a most appreciative letter; he had lost his wallet in a movie theater and it was returned to him because it contained something identifying him—his Project TALENT identification card!)

There had been some concern that perhaps after the first day of testing a lot of the students would decide to be "sick" the second
day, and not show up. We needn't have worried. Overall attendance was a little better than average during the days of the tests, and in many schools it was noticeably better the second day than the first!

The Local Coordinators were requested to arrange make-up sessions for absentees if it was at all feasible to do so. Although in some schools this proved not to be possible, many schools did manage to have make-up sessions so that those who had been absent on one or both days could be tested.

D. E. F. Lindquist and Measurement Research Center

In the early 1950's, before Project TALENT was begun, the most advanced general-purpose scoring machine available was the IBM-805. This machine, which operated electrically (not electronically), transmitted current electrographically (i.e., across the graphite pencil marks made with a very soft lead pencil). It could score only one side of the answer sheet at a time. A second insertion was required for the second side. And both insertions were manual operations. There was no automatic feeding—even though IBM had perfected automatic feeding for IBM cards decades earlier. The machine could handle up to 150 five-choice items per answer-sheet side and up to three test scores or part scores and a total in a single insertion. Scoring was of the analogue type, not digital; the score was regis-
tered by a needle. The operator had to determine which integer the needle was closest to, and write it (with his or her right hand*) on a projecting edge of the answer sheet. An experienced operator could score a few hundred answer sheets in an hour. These scoring machines have not yet gone the way of the dinosaurs. Some still exist and are still operating—but in the early 1950's they were the only thing generally available, and the world of educational research was full of them.

Imagine, then, the excitement with which E. F. Lindquist's announcement, at the ETS Invitational Conference on Testing Problems in 1953, was greeted. He announced that he was building a scoring machine that would automatically feed and score 6,000 answer sheets an hour, would score both sides at once, and would yield as many as 14 scores, either raw or corrected for chance, on a single insertion. The 8½" x 11" answer sheet would hold roughly 1,000 five-choice items. The machine would operate via optical scanning (an electronic process) rather than via electrographic conduction. The output, too, would be vastly superior to anything then available.

When Lindquist made his announcement in 1953, he made it perfectly clear that the scoring machine was just in the planning stage, and that while the components all existed, and worked individually, they had not yet been put together and tested in combination.** He implied therefore that changes in design were not unlikely and that it would not be surprising if the scoring machine, when built, turned out to be somewhat different in operation from what was planned. It did turn out to be different, in some ways that were extremely advantageous and in other ways that proved somewhat inconvenient. The major improvement, as things turned out, was that miraculously, the direct output was neither answer sheets with the scores recorded on them nor printouts, but ready-punched IBM cards with the examinee's name, other identifying information, and test scores already on them!

---

* Individuals lucky enough to be left-handed were spared the "privilege" of operating this horse-and-buggy scoring machine.
** Lindquist made it clear that the design of this new scoring machine would be based, in large part, on developmental work done previously by Phillip J. Rulon.
Thus not only did it take only a tiny fraction of the time required by the old 805 machines, but it was far more accurate and, perhaps best of all, it eliminated the time-consuming, costly, error-producing process of key-punching.

Another improvement, of considerably less import but advantageous nonetheless, was that instead of having space for under 1,000 five-choice items, the answer sheet could hold (and the Scoring Machine could cope with) nearly 1,200.

As for the changes in the opposite (i.e., disadvantageous) direction, their impact was weaker. It turned out that only 13 scores, not 14, could be obtained on a single insertion. Somewhat more inconveniently, it turned out that for each test for which scores corrected for chance were desired, two separate scores—the number of items answered correctly and the number attempted—would have to be obtained, thus being likely to double the number of insertions required, and thus double the cost, if more than six tests were involved for which scores corrected for chance were desired. The scoring machine at that time was in a constant state of flux, continually being modified in some way by MRC's engineers, to add a new feature or improve it some other way. We were already drawing up final specifications for MRC processing of answer sheets A and C on the assumption that the earlier machine specs, according to which a set of scores corrected for chance would require no more insertions than uncorrected scores, applied, when we found out about the change.

We had wanted most of the scores "corrected for chance", alias "corrected for guessing". Note to reader: Yes, we were fully aware that "corrected for chance" and "corrected for guessing" are to some degree overstatements. Of course the correction isn't perfect. But we felt strongly that a simple correction (by the conventional formula) even though admittedly not perfect, was far better than the alternative (no correction). Unfortunately, however, the change in Scoring Machine specifications meant that the "simple correction" would not be so simple to achieve. Doubling the number of answer sheet insertions would double the costs. And there was no way that the project budget could absorb an increase of that magnitude. It
was clear that we would have to settle for "number right" scores except for the four very highly speeded tests* of various kinds of computational and perceptual speed and accuracy, for which it would be possible to get counts of both the number right and the number attempted. That would permit us to get formula scores which were deliberately set to overcorrect for wrong answers—a legitimate technique (indeed, a necessary technique) for that sort of test with the sort of predictive use for which those tests were intended.

Settling for having formula scores for only four tests was a compromise, but it was a compromise dictated by reality. And it wasn't a bad compromise. Number-right scores have a very high correlation with "corrected" scores. But even a very high correlation allows an occasional "outlier" case and a resultant injustice or serious mis-evaluation. We are aware, of course, that many otherwise admirable people think "number right" is the only way to go. However this "History" probably isn't the proper forum to debate the matter,** so we shall go on to something else, after contenting ourselves with pointing out that this compromise did not mean that we were abandoning for all time the idea of "corrected" scoring. (See Chapter 12.)

Before dropping (at least temporarily) the matter of Scoring Machine limitations we should mention one further problem—one which was to complicate matters considerably for us (and as a matter of fact still is causing complications. See Chapter 13.) That was the problem of distinguishing between legitimate zero scores (failure to answer even one item correctly) and blanks or "missing data" (i.e., failure to take the test at all). The Scoring Machine couldn't make this distinction. Missing data, like legitimate zero scores, were represented on the IBM card output by zero punches.

When the new scoring machine had been built and had been in use for a couple of years. Lindquist undertook to build another machine,

* These four tests are:
  1. Arithmetic Computation
  2. Table Reading
  3. Clerical Checking
  4. Object Inspection

even more advanced, whose card output would provide item responses, rather than test scores. This machine was called the "document reader" (as opposed to the "scoring machine"). Each of the 80 card columns of the IBM card output could register any one of 39 response codes: the 26 letters of the alphabet, the 10 digits, the two extra "overpunch" positions, or a blank. Each of the 80 39-character positions on the answer sheet could be reorganized to contain as many as six five- or six-choice items. This meant that even after allowing answer spaces for name and other identifying information it would still be possible to get well over 400 items on an answer sheet, some with as many as six choices. And in the case of items containing no more than six options each, it would be possible to double-punch the output cards so that each card column would contain responses to two items. (Items with more than six choices would also be possible, and equally easy to process, but each such item would cut down on the total number of items that would fit on the answer sheet; and since these items would preclude double-punching they would also reduce the number that would fit on an output card.)

When John Flanagan was developing the outlines of the plan for Project TALENT, early in 1958, he visited Dr. Lindquist at MRC to find out about the new scoring machine—how effective it was, how feasible it would be to handle Project TALENT's scoring on it, and in general to explore the possibilities. On that visit he learned about the Document Reader, which was then just in the blueprint stage; construction had not yet begun. It was immediately apparent from the description that the Document Reader would be a key element in Project TALENT. This incipient machine would have one major advantage over the Scoring Machine; it would provide greater flexibility since items could be combined and recombined in different ways, and different scoring formulas could be experimented with, if scoring were postponed and done later by computer, rather than being handled on the Scoring Machine.
As a matter of fact some of the most crucial features of Project TALENT—specifically the nearly 400 items of background information that constituted the Student Information Blank—would have been entirely out of the question if the Document Reader had not been promised.

In 1959 Lin was able to assure John Planagan that the Document Reader would be built and operating in time to handle those aspects of the Project TALENT data processing for which it was appropriate. The Scoring Machine, too, would be available. Because of its newness and its innovative character, problems—many problems—had been encountered in using the scoring machine. Lin told John that over a fairly long period of time there had been so much trouble with the scoring machine, and so much "down time", that it had only been usable about four hours a month. But—and this is a very important "but"—with just those four hours a month of usable time, it was possible to process the entire Iowa Every-Pupil Testing Program and keep up with all the many other testing programs using their services. That bit of substantive information, together with Lin's assurances that both machines would be available, functioning, and ready to process the projected TALENT data, provided a basis for firmer planning. Ideally, it would have been best to have all the answer sheets processed on the Document Reader. But as the plans developed and grew firmer, it became obvious that that would not be feasible. One important reason was that Lin could not promise that the Document Reader would be ready and working before the fall of 1960. But it was going to be necessary to get reports of results to participating schools as soon as possible. The only machine that would be available at a time that would make this possible was the Scoring Machine. Therefore, as a compromise, we decided to put the tests on which we would need scores right away on two answer sheets (Answer Sheets "A" and "C") that would go through the Scoring Machine. The remaining data—mostly responses to questionnaires (specifically the Student Information Blank and the Interest Inventory) but also responses to about 36 percent of the Information test items—would go
on two answer sheets (Bl and B2) that would be processed on the Document Reader. A fifth document, "Record Form Z", which would also be processed on the Document Reader, was designed to contain identifying information—name, address, sex, grade, age, and date of birth—as well as the six-digit student identification number used to link all the answer sheets for the same student. (Enough identifying information would also appear on Answer Sheets A and C, however, to permit students' scores on selected tests to be reported to the schools.)

[Author's Note (again):

A good deal of space has been devoted to a discussion of problems and difficulties that arose in connection with use of MRC's Scoring Machine and Document Reader, and I do not want to leave the wrong impression. Even though by today's standards both machines were primitive, in that they lacked a direct document-to-tape capability, and required, instead, an intermediate step involving the production of IBM cards, nevertheless in 1960 these two machines were in the vanguard of advanced technology. With them, processing the TALENT data was nightmarishly complex, but without them the project would have been impossible. (It couldn't have been done just one year earlier, because the Document Reader would not have existed.)]

E. Implementation

1. Before MRC (The "TALENT Laundry")

Answer sheets were shipped from each school direct to MRC by the Local Coordinator. Before they could be passed through the Document Reader or Scoring Machine they had to be inspected and "cleaned" (by human beings, not machines). The "cleaning" involved such activities as erasing stray marks on answer sheets, reconstructing answer sheets that had somehow been damaged in transit, and transferring responses to intact blank sheets if necessary; also making certain that the necessary identification information (school code, etc.) was on each answer sheets before it was separated from the identifying materials with which it was packed. The arrangements called for all this to be done under AIR
auspices, rather than by MRC. Therefore Marion Shaycoft spent a month in Iowa City in the spring of 1960, hiring and training a large crew of part-time clerks (mostly housewives who wanted to earn extra money) to do this processing. She had enough faith in the TALENT tests that she used an appropriate subset of them in selecting the personnel. This worked very well. In due course she appointed as the supervisor of this operation one of the clerks she had hired who had proven especially competent and who was able to work full-time. That was Marjorie Major. The operation she was to supervise came to be known informally as the "TALENT Laundry" because of the answer-sheet clean-up activities involved. It was a store-front operation, arrangements having been made to use for the purpose a vacant store a couple of blocks away from the University of Iowa building in which MRC's machines were located. The TALENT Laundry, under Marjorie Major's capable supervision, was to operate for the better part of a year.

In addition to operations of the answer-sheet-inspection-and-cleanup type that MRC required as a preliminary to the machine passes, one very important activity that the "Laundry" staff performed was to riffle through each pack of A and C answer sheets and pull out the ones for which the last two digits of the six-digit student identification number were 00, 25, 50, or 75. This provided a 4% sample. These answer sheets (the 4% sample) were to be the first ones processed on the scoring machine. Arrangements to give them priority would provide data on which interim percentile norms could be based, for use in reporting results to the schools, on 37 variables.

The remaining answer sheets of each kind were arranged in 11 groups—the 10 school subsamples into which the probability sample was divided (as described in Chapter 2, Section A) and the nonprobability cases—since the plan was that after the processing of the 4% sample A and C answer sheets, subsequent scoring machine
processing would proceed in subsample order. Document Reader processing would have to await completion of the Document Reader—i.e., getting the remaining bugs out of it.

2. **At MRC**

   Each of the answer sheets to be processed on the Scoring Machine (Answer Sheets A and C) required two passes to yield the required 52 scores. Answer Sheets B1 and B2 each required three passes through the Document Reader, with double-punched cards as output. Record Form Z only required a single pass.

   Thus we wound up with 11 IBM cards per student, or about five million cards! MRC scored the 4% sample as soon as it was provided, and also produced distributions in printout form, so that percentile conversion tables could be computed by grade. After Project staff had prepared the conversion tables, MRC scored the rest of the answer sheets, and prepared printout reports to send to the schools showing each student's raw score on 37 variables, the corresponding interim national percentile for the grade population, and the school mean for each grade.

3. **After MRC**

   When the five million cards were put on tape on the University of Pittsburgh computer (where, by this time, the IBM-650 had been replaced by the IBM-7070, which was far more advanced and had far greater capacity), Glenn Roudabush (the second programmer to have joined the Project staff) found that the easiest procedure was to make two tapes for each double-punched card (with the left half of the card, plus identifying information, on one tape, and the right half, plus the same identifying information, on the other half). Thus the five original documents (answer sheets), which had been inflated to 11 IBM cards, resulted eventually in 17 tape records.
Making and merging these 17 tape records for almost half a million students was an exceedingly complex and lengthy job (it lasted years, as a matter of fact), and it was greatly complicated by the fact that most of the project staff was in Washington, while the computer and programmer were 250 miles away, in Pittsburgh. Further complicating the operation was the fact that it had to take place in a milieu over which Project TALENT had no control—the University of Pittsburgh's Cathedral of Learning, which was where the Computation and Data Processing Center was located. Glenn Roudabush, on whom the burden of the computer work to plan and make the tape master file of 1960 data devolved, found he had to work nights, for months on end, because that was the only way he could get adequate access to the computer.

Meanwhile all kinds of logistical problems were interfering with speedy completion of the task. One problem was that a single operation couldn't be carried through on all cases at once (which would have been maximally efficient) because it was necessary to give top priority to completion of one of the 10 probability sample subsamples ("subsample 0") so that we would have at least one subset of data in fairly complete form, on which preliminary or exploratory work could be done. Another problem resulted from the lack of convenient space in which to store the IBM cards and keep processed and unprocessed ones separate. For instance on one memorable occasion after Glenn thought he had finished a particular sequence of computer runs, several boxes of unprocessed IBM cards from MRC were discovered in a broom closet! (He had to start the sequence all over again.)

To find out what happened next in this long-running soap opera, see Chapter 13!
Chapter 8. WITH A CAST OF THOUSANDS!

This chapter is intended to function as a guide to the cast of characters—the project staff and some of the others who played a role in the activities and events with which this History is concerned. The chapter title is perhaps only a slight exaggeration, even though somewhat under a thousand names are mentioned.

A. The staff

Figure 8-1 provides in tabular form an overview of the professional, technical, and supervisory staff. The titles shown in this figure are intended primarily as functional indicators. They are not necessarily identical in all cases with the actual officially assigned titles, since the official titles used have tended to change from time to time. For instance officially Dick Williams was the "Director of Computer Operations" while Gary Fulscher was "Manager of Data Processing". But it was essentially the same position, so both are shown under the "Director of Computer Operations" heading.

Figure 8-2 is a time-and-position framework, which gives some idea of what the composition of the staff was in any given period, at least with respect to the positions listed.
Figure 8-1. Project Professional, Technical, and Supervisory Personnel: 1957-76

<table>
<thead>
<tr>
<th>PRE-PROJECT PRELIMINARY PLANNING</th>
<th>INSTRUMENT DEVELOPMENT</th>
<th>EXPERIMENTAL TRYOUT</th>
<th>SAMPLE SELECTION</th>
<th>TEST ADMINISTRATION</th>
<th>SENIOR RESEARCH STAFF</th>
<th>DATA PROCESSING AND ANALYSIS</th>
<th>COMPUTER PROGRAMMER: DATA PROCESSING</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aptitude and Achievement Test Editor</td>
<td>Data Analysis Director: Marion F. Shaycoft</td>
<td></td>
<td></td>
<td></td>
<td>School Characteristics Studies: Isadore Goldberg</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dorothy S. Edwarde Mary B. Willis</td>
<td>Student Questionnaire Development: David B. Orr</td>
<td>Student Activity Inventory Editor</td>
<td>Isadore Goldberg</td>
<td>Interest Inventory and Student Information Blank Editor</td>
<td>Manpower and Guidance Studies: Isadore Goldberg</td>
<td>Gary Lotto</td>
</tr>
<tr>
<td></td>
<td></td>
<td>School Questionnaire Development: David Votaw, Sr. Consultant</td>
<td>David B. Orr</td>
<td>General School Characteristics Questionnaire Editor</td>
<td></td>
<td>Supervisor for Services: Clinton A. Neyman, Jr.</td>
<td>Glenn E. Boudabush</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Isadore Goldberg</td>
<td></td>
<td>Guidance Program Questionnaire Editor</td>
<td></td>
<td></td>
<td>George R. Burket</td>
</tr>
</tbody>
</table>

* On sabbatical from Hunter College

Continued on next page
### Figure 8-1 (continued)

**Principal Investigator:** John C. Flanagan

*Subsequent phases: Five-year and eleven-year follow-ups, and planning for subsequent stages*

<table>
<thead>
<tr>
<th><strong>Project Directors</strong></th>
<th><strong>Associate Directors</strong></th>
<th><strong>Data Bank Directors</strong></th>
<th><strong>Other Senior Researchers</strong></th>
<th><strong>Senior Editorial Services</strong></th>
<th><strong>Miscellaneous Administrative Services</strong></th>
<th><strong>Computer Programming</strong>*</th>
</tr>
</thead>
<tbody>
<tr>
<td>William W. Cooley</td>
<td>Marion F. Shostak†</td>
<td>Richard W. Holdenman</td>
<td>Louise W. Cureton</td>
<td>Mary B. Willis</td>
<td>Kenneth W. Buck</td>
<td>Data Analysis</td>
</tr>
<tr>
<td>Abraham Carp</td>
<td>James M. Richards, Jr.</td>
<td>Lyle F. Schoenfeldt</td>
<td>Paul P. Nak</td>
<td></td>
<td></td>
<td>Director:</td>
</tr>
<tr>
<td>John C. Flanagan</td>
<td>Donald H. McLaughlin†</td>
<td>John G. Clancy</td>
<td>Charles E. Hall</td>
<td></td>
<td></td>
<td>Jon van Wormer</td>
</tr>
<tr>
<td>David V. Tiedeman</td>
<td>Wendy M. Yon†</td>
<td>Donald H. McLaughlin‡</td>
<td>Paul K. Lohme</td>
<td></td>
<td></td>
<td>Marilyn S. Wingersky</td>
</tr>
<tr>
<td>William V. Clemans†</td>
<td>Sandra R. Wilson</td>
<td>G. Kasten Tallmadge</td>
<td>Alan E. Bayer</td>
<td></td>
<td></td>
<td><strong>General Supervision of</strong></td>
</tr>
<tr>
<td>1974-76**</td>
<td>1974-75</td>
<td>1972-73</td>
<td>1967-68</td>
<td></td>
<td></td>
<td><strong>Computer Operations:</strong></td>
</tr>
<tr>
<td></td>
<td>Lauretta L. Wise†</td>
<td>Donald F. Horst</td>
<td>Michael Wargo</td>
<td></td>
<td></td>
<td><strong>James M. Richards, Jr.,</strong></td>
</tr>
<tr>
<td></td>
<td>1975-76**</td>
<td>1972-73</td>
<td>1971-72</td>
<td></td>
<td></td>
<td>1969-70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Robert J. Rossi</td>
<td></td>
<td></td>
<td></td>
<td><strong>Director of</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1974-76</td>
<td></td>
<td></td>
<td></td>
<td><strong>Computer Operations:</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kevin J. Gilmer†</td>
<td></td>
<td></td>
<td></td>
<td><strong>Richard A. Williams, 1967-70##</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1975-76</td>
<td></td>
<td></td>
<td></td>
<td><strong>John C. Clardy, 1970-73</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>Gary V. Fulscher, 1973-74</strong></td>
</tr>
<tr>
<td>ASSOCIATE DATA BANK</td>
<td>Mark M. Greeno</td>
<td>G. Kasten Tallmadge</td>
<td></td>
<td></td>
<td></td>
<td><strong>Special One-Man Task</strong></td>
</tr>
<tr>
<td>DIRECTORS</td>
<td>1967-68</td>
<td>1972-73</td>
<td></td>
<td></td>
<td></td>
<td><strong>Forces (Seriation)</strong></td>
</tr>
<tr>
<td></td>
<td>Marsha D. Brown§</td>
<td>Donald F. Horst</td>
<td></td>
<td></td>
<td></td>
<td><strong>for Documentation:</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Robert J. Rossi</td>
<td></td>
<td></td>
<td></td>
<td><strong>Calvin E. Wright, 1972</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1974-76</td>
<td></td>
<td></td>
<td></td>
<td><strong>Donald F. Horst, 1972-73</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kevin J. Gilmer†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1975-76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Bank Research</strong></td>
<td><strong>Pittsburgh:</strong></td>
<td><strong>Pittsburgh:</strong></td>
<td><strong>Pittsburgh:</strong></td>
<td><strong>Pittsburgh:</strong></td>
<td><strong>Pittsburgh:</strong></td>
<td><strong>Junior Programmers,</strong></td>
</tr>
<tr>
<td>ASSISTANTS</td>
<td>Mary K. Garee</td>
<td>Janet C. Combs</td>
<td>Susan J. Becker</td>
<td>Carl Euelme</td>
<td>Michael J. Quirk</td>
<td><strong>Assistant Programmers,</strong></td>
</tr>
<tr>
<td></td>
<td>Palo Alto:</td>
<td>Carol J. Balcant</td>
<td>Ronza J. Sallade</td>
<td>Arelin M. Lichtenstein</td>
<td>Michael E. Barenfeld</td>
<td><strong>Programmers,</strong></td>
</tr>
<tr>
<td></td>
<td>Susan K. Cenog</td>
<td>Judith D. Miller</td>
<td>J. Fred Briner</td>
<td>Seon H. Cho</td>
<td>Patrick McDoull</td>
<td><strong>Programmers,</strong></td>
</tr>
<tr>
<td></td>
<td>Ardys J. Bloomquist</td>
<td>Deborah Tien</td>
<td>Suzanne R. Lewis</td>
<td>Francis Wimerley</td>
<td>Gary V. Fulscher</td>
<td><strong>Programmers,</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Robert H. McKillen</td>
<td></td>
<td>Jay A. Woods</td>
<td><strong>and Trainees</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Wendy B. Bartlett</td>
<td></td>
</tr>
</tbody>
</table>

*During this three-year period, John C. Flanagan was Acting Director as well as Principal Investigator.*

*† Director of Measurement Research 1960-65*

*‡ On project research staff earlier, in role other than Associate Director.*

*§ On project research staff earlier, in other role.*

*† On project research staff earlier, in nonsupervisory capacity.*

** Continuing**

*** The computer programming staff is shown up to 1974. In 1974, Project TALENT ceased to have its own programming staff and instead used ATR's Palo Alto computing staff.*
### Figure 8-2. The Changing Cast: 1957 - 1976

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Investigator</td>
<td>JCF</td>
<td>JCF</td>
<td>JCF</td>
<td>JCF</td>
<td>JCF</td>
<td>JCF</td>
<td>JCF</td>
<td>JCF</td>
</tr>
<tr>
<td>Project Director</td>
<td>RCC</td>
<td>JTD</td>
<td>JTD</td>
<td>WAC</td>
<td>WAC</td>
<td>WAC</td>
<td>AC</td>
<td>DVT</td>
</tr>
<tr>
<td>Associate Director</td>
<td>MFS</td>
<td>MFS</td>
<td>MFS</td>
<td>MFS</td>
<td>MFS</td>
<td>MFS</td>
<td>MFS</td>
<td>JNR</td>
</tr>
<tr>
<td><strong>Director of Measurement</strong></td>
<td><strong>Research</strong></td>
<td><strong>MFS</strong></td>
<td><strong>MFS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Editor</td>
<td>MFS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associate Test Editor</td>
<td>DSE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Bank Director</td>
<td>RWH</td>
<td>LFS</td>
<td>LFS</td>
<td>LFS</td>
<td>LFS</td>
<td>JGC</td>
<td>JGC</td>
<td>JGC</td>
</tr>
<tr>
<td>Associate Data Bank Director</td>
<td>NMG</td>
<td>NDB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer Operations Director</td>
<td>GL</td>
<td>JvW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For Data Analysis</td>
<td>GL</td>
<td>BGW</td>
<td>RAW</td>
<td>RAW</td>
<td>RAW</td>
<td>JGC</td>
<td>JGC</td>
<td>GVF</td>
</tr>
<tr>
<td>For Data Processing</td>
<td>GER</td>
<td>JvW</td>
<td>MSW</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tape improvement and documentation task forces</td>
<td>RTJ</td>
<td>DFR</td>
<td><strong>MFS</strong></td>
<td><strong>MFS</strong></td>
<td><strong>LLW</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Initially
** Later
*** Not a Project TALENT staff member
# For Data Bank
## For data tape improvement and documentation
### Other
#### Initial planning

**KEY**

- **NDB** Marsha D. Brown
- **GRB** George R. Burkat
- **AC** Abraham (Al) Carp
- **JGC** John G. Clauudy
- **WVC** William V. Clemans
- **WAC** William W. Cooley
- **RCC** Robert C. Craig
- **JTD** John T. Dailey
- **DSE** Dorothy S. (Jo) Edwards
- **JCF** John C. Flanagan
- **GVF** Gary V. Fulscher
- **MG** Mark M. Greene
- **RWH** Richard W. Holdeman
- **DFH** Donald P. Horst
- **RTJ** Richard T. Johnson
- **GL** Gary Lotto
- **DHM** Donald H. McLaughlin
- **JMR** James M. Richards, Jr.
- **GER** Glenn E. Roudabush
- **LFS** Lyle F. Schoenfeldt
- **MFS** Marion F. Shaycoff
- **DVT** David V. Tiedeman
- **JvW** Jon vanWormer
- **RAW** Richard A. Williams
- **SRW** Sandra R. Wilson
- **BGW** Barry G. Wingersky
- **MSW** Marilyn S. Wingersky
- **LLW** Laurrell L. Wise
- **WMY** Wendy M. Yen
1. **The Principal Investigator**

John Flanagan, as the person who conceived and planned the project, and as the prime mover behind it, has naturally been the Principal Investigator throughout the nearly 20 years that have elapsed (as of the end of 1976) since the Project started. The continuity that he has provided has kept the Project on course all these years. There have been some small shifts in plans, some minor changes in details along the way, but the general features and the goals have remained substantially the same. Without the constant element provided by having the same Principal Investigator from the very beginning this would have been difficult, if not impossible, to achieve.

2. **Project Directors**

As has already been indicated, Robert Craig was the first Project Director. During his single year in that role his activities were largely exploratory—investigating what sorts of procedures might be feasible and what sorts of tests and questionnaires might be useful.

He resigned in 1958 to accept an appointment at Marquette University. (He is now at Michigan State.)
John Dailey, the next Project Director, had had 16 years of experience in the Federal government. During World War II he had been a psychologist in the Aviation Psychology Program headed by John Flanagan. After the war, most of the psychologists left; he stayed at Lackland Air Force Base until 1951, directing the Personnel Research Unit. After that he had seven years in Washington, with the Bureau of Naval Personnel (1951-58), where he was technical director of the Research Division. He gave up Civil Service tenure to accept the appointment as Project Director of the project, still unfunded beyond the planning stage, and with no promises from any funding agency. John recognized that he was taking a big gamble, but the prospective project fitted in exactly with his professional interests, and directing it was, he says, what he most wanted to do. He did not regret his decision.

Around 1947, when he was at Lackland, he had been campaigning for a large-scale testing of a national sample. He had recognized the desirability of having data based on a general population, not just on those who happened to be in the Air Force. The project that he was leaving the Bureau of Naval Personnel to direct was exactly in line with that idea. Furthermore he had been one of the participants in the initial planning conference for the project, in February 1957. He had been excited by the idea unveiled by John Flanagan at that conference. These factors combined to make him enthusiastic about the project and the opportunity to direct it. And during his six years as director of Project TALENT that enthusiasm never wavered.

When the project moved to Pittsburgh, however, John decided his ties in the Washington area were too strong for him to be interested in making the move. Consequently he resigned from AIR (and the University of Pittsburgh). He had been TALENT's Project Director for a much longer period than anyone else before or since.
John accepted an appointment at George Washington University as Research Professor of Education, and Director of the Education Research Program. (In that capacity he carried out a number of studies that were based on Project TALENT data, and utilized the services of the Data Bank.) After four years at GW he became a psychologist for the Federal Aviation Administration, where he is now the Chief of the Analytic Branch of the Behavioral Sciences Division of the Office of Aviation Medicine. He has done some interesting work there, including developing a behavioral profile of potential skyjackers (or as the FAA prefers to call them, hijackers) which was instrumental in reducing the air piracy threat, at airports. (In 1972 he received the Flight Safety Foundation's top safety award for this work.) He is now working on the somewhat more general problem of protection against terrorist activity, and is also developing new types of tests for air traffic controllers.

Perhaps this would be a good place to mention that one of John Dailey's major contributions to Project TALENT was the concept of "follow-back studies" that he worked out. (This concept is discussed briefly in Chapter 21 and also in Chapter 13. It has not been used as extensively or as effectively as it should have been—but that is the fault of neither the concept nor its originator.)

Before going on to the next Project Director, we must say something about the important role that Fred Davis played. During the academic year 1963-64, Fred was on sabbatical from Hunter College. During that same year, members of the TALENT staff were scheduled to write a very large research report (it turned out to be over 700 8½" x 11" single-spaced pages) entitled "The American High School Student". Fred, who was a first-rate psychometrician, test-writer, and educational researcher, was also an expert editor. As a matter of fact, he had been the editor-in-chief of the entire
19-volume series of Army Air Forces Aviation Psychology Program Research Reports produced after the end of World War II. Fred was also a long-time member of AIR's Board of Directors. The coincidence of Fred's sabbatical coming at the same time as the enormous report-writing task that lay ahead of the project staff was a happy one for Project TALENT. John Flanagan persuaded Fred to spend his sabbatical year in Washington, as editor-in-chief of the forthcoming report and as director of an Institute for Educational Research (IRE), which didn't yet exist but would be created as one of the American Institutes for Research's several institutes. Project TALENT, along with some related activities, would be put under the umbrella of the Institute for Educational Research.

Thus during the last year in Washington (1963-64), Project TALENT had a tripartite management structure: Principal Investigator, Institute Director, Project Director. The arrangement worked very well; however after the project moved away from Washington the IRE Institute Director and the Project TALENT Project Director were invariably the same person.

We hoped that Fred would decide to stay with Project TALENT when his sabbatical ended, instead of going back to Hunter. As things turned out, however, he decided not to go back to Hunter, but also to withstand the heavy persuasion efforts to get him to stay at AIR. Instead, he accepted an appointment to the faculty of the University of Pennsylvania. After he left AIR, Fred maintained his interest in Project TALENT, and was a consultant to it on many occasions. He died in the spring of 1975, and we still feel the loss.

After John Dailey, the next Project Director was William W. Cooley, who had been teaching at Harvard. He came to AIR in 1964.
Bill's main professional interest was multivariate analysis. He and Paul Lohnes had written a book on multivariate analysis in which not only the statistical techniques but also computer programs they had prepared for implementing those techniques were presented. (More about those programs later; see Chapter 14.) For the three years that Bill Cooley was Project Director, most of the research done on TALENT data utilized multivariate methodology. When Project TALENT was about to move to California, Bill Cooley, like his predecessor, decided he would rather stay where he was. He preferred Pittsburgh to Palo Alto*. Therefore he resigned from ATR in 1967, to devote full time to his professorship at the University of Pittsburgh. (He had been on the faculty there since 1964, but directing Project TALENT had been his main occupational activity.) He is now Co-Director of the University of Pittsburgh's Learning Research and Development Center.

When the Project moved to Palo Alto, in the summer of 1967, Abraham (Al) Carp became the new Project Director. Like John Dailey's, Al's professional experience was heavily Air Force. He had been at AFPRC since 1951 (the year John Dailey left) in varying capacities, becoming in 1957 the technical director of the Personnel Laboratory, Aeronautical Systems Division.

Al was Project Director for only a year. In the fall of 1968 he became the Director of Research in ETS's Berkeley Office.

There was a three-year hiatus (1968-71) before his successor was appointed. During the interim, John Flanagan functioned not only as Principal Investigator but also as Acting Project Director.

In the summer of 1971 David Tiedeman resigned from a professorship at Harvard, to become Project TALENT's Project Director.

* Californians will not understand!
One of Dave’s principal activities while he was at AIR was to work on the preparation of the Career Data Book, a large volume designed to present Project TALENT’s 5-year follow-up data for career groups in a manner that would facilitate its use in career guidance. Dave devoted a lot of effort to the Career Data Book although it was far from finished with he left AIR. Dave also devoted considerable attention to getting a "Systems Manual" developed for use by IRE and specifically for use in Project TALENT operations. (The Systems Manual was written almost wholly by Calvin Wright, who had already been on the AIR staff for about 5 years when he joined Project TALENT for a very brief period.) Dave stayed at AIR for two years, leaving in the summer of 1973 to accept an appointment as professor at the University of Northern Illinois.

Several months later, William V. Clemans accepted an appointment as the next Project Director. Bill, a vice-president of Science Research Associates (a Chicago publisher of tests and educational materials), was no stranger to AIR. In the early 1950's he had worked for a short time on an AIR project that took him to Korea for several months. And in the late 1950's and early 1960's, when AIR developed some examinations for the National Board of Medical Examiners, Bill Clemans, who was then the Director of the Testing Service for that organization, monitored the project.

Like most of his predecessors Bill has a strong background in psychometrics, having received his doctorate at the University of Washington, where he was one of Paul Horst's students.

Unlike some of his predecessors, such as John Dailey and Bill Cooley, both of whom were heavily involved in doing research on Project TALENT data (as, too, the Principal Investigator has been)
Bill Clemans has regarded the job of Project Director as primarily administrative in nature. (This difference is mentioned only to point up the not too surprising fact that different project directors may have very different operating styles.)

Bill is the current Project Director.

3. **Associate Directors**

Marion Shaycoft, who became an Associate Director of Project TALENT (and of IRE) in 1965, had joined the project staff in 1958. She was Editor-in-Chief during the test development (i.e., until sometime in 1960) and then she became the project's Director of Measurement Research. Thus when she became Associate Director she was thoroughly familiar with the salient aspects of the project.

In 1969 it became advisable to appoint a second Associate Director, to take over one of the functions that the Project Director had previously performed—namely general supervision of the computer operations. This responsibility devolved upon James M. (Mac) Richards, a new member of the research staff. The Director of Computer Operations reported to him. Mac performed this Associate Director function for only about a year however.

In 1970 other responsibilities—specifically direction of a new project that had nothing to do with Project TALENT—required him to cut back sharply on the amount of time he could devote to Project TALENT. Mac left AIR in 1971, to accept a position at the University of Missouri at Kansas City. Again the Director of Computer Operations reported directly to the Project Director. Thus the project was down to one associate director again. But in 1974 there was a sharp inflation in the number. During the period from 1974 through 1976 there were usually three associate directors, and occasionally four.
Donald McLaughlin, who had been teaching at UC Berkeley for several years, was one of the new Associate Directors; Wendy Yen, a new Ph.D. from Berkeley, who had been on the TALENT staff about half a year, was another. She left after about half a year as Associate Director, to join the staff of CTB/McGraw-Hill (test publishers, in Monterey, California). Don McLaughlin, fortunately, is still associated with the project (primarily as Director of the Data Bank). Sandra Wilson, who had been on the AIR staff for several years but had not been associated with Project TALENT, transferred to the Project as another Associate Director, around the middle of 1974. (Project TALENT was the loser when Sandy left the project staff about a year later, to play a major role in AIR's development of new tests for medical school admission.) About that time Laurrell L. Wise was added to the lengthening list of associate directors. As discussed later in this chapter, in the spring of 1976 Marion Shaycoft almost completely discontinued active participation in the project's ongoing activities, to concentrate on other activities, including among other things the writing of a monograph on criterion-referenced tests (Shaycoft, 1976) and the writing of this History. In regard to the latter, possibly standing back from the thick of things improves one's perspective as the observer and historian of an activity; at least we hope that that is the most potent effect (although some impairment of the view is probably an inevitable byproduct of the distance).

4. Data Bank Directors

A few projects of the "Data Bank type"—projects in which computer output was provided at cost for some outside researcher (i.e., not an AIR staff member) who wished to use TALENT data—were done while Project TALENT was still based in Washington. These projects were handled primarily by George Burkett, one of our computer programmers. But it was not until the move back to Pittsburgh in 1964 that the Data Bank as such
was formally organized*. Its first director was Dick Holdeman, who was then a new Ph.D. from Purdue. He stayed about a year, then resigned to accept a job as an industrial psychologist. His successor was another new Ph.D. from Purdue, Lyle Schoenfeldt. Lyle stayed until about the end of 1968—which means that he was one of the staff members who made the move to Palo Alto in 1967 (see Chapter 9). He left to accept a position on the faculty of the University of Georgia. (He is now at Rensselaer Polytechnic Institute.) When Lyle directed the Data Bank, Mark Greene was Associate Data Bank Director for a year (1967-68) at the start of the Palo Alto phase, and Marsha Brown, who had started at the same time as Mark, was Associate Director of the Data Bank the following year (1968-69). Mark left in the fall of 1968, to join the staff of Northwest Regional Laboratory. Marsha left the following year, to go back to school for her doctorate. She was at the University of Wisconsin for a year, but got her doctorate at Harvard, where she worked with Christopher Jencks.) Lyle Schoenfeldt's successor as Director of the Project TALENT Data Bank was John Claudy, who was a new Ph.D. from the University of Tennessee. (John had been one of the students of Ted Cureton, who had been associated with Project TALENT as a consultant and had done some special studies involving TALENT data, in Knoxville.) John Claudy is still at AIR, but he transferred to other activities and discontinued his functions as Data Bank Director in 1975, when Don McLaughlin took over in that capacity. Don is the current Data Bank director.

5. Other recent senior staff members

Mary Willis (the same Mary Willis who was mentioned in Chapter 5 as having drafted the Mechanical Reasoning items, and in Chapter 6 as having written the two "domain tests" of reading) was a regular full-time member of the Project TALENT staff for three years (1971-74) after which she transferred to another AIR project again. During the three years Mary was on the Project TALENT staff she had editorial responsibilities of a varied nature. She was the editor of the annual edition of the Project TALENT

* See Chapter 22 for a discussion of the Data Bank.
News (a four-page "newspaper" sent to all project participants), and she wrote most of it. She also edited a Project TALENT Bulletin, and was heavily involved in the writing of the Career Data Book. Mary is a highly skilled writer. (She was a coauthor of a book entitled "Secret Archives of the Vatican" which, when it came out, was a Book-of-the-Month Club alternate selection.)

Kenneth Buck, who has a doctorate from Claremont Graduate School, and whose major fields were political science and education, was on the staff for about a year, starting in 1975. His role was as Senior Management Specialist; his duties were largely administrative.

6. Personnel in charge of computer operations

When the project was located in Washington except for the computer in Pittsburgh, we had three programmers—Gary Lotto, Glenn Roudabush, and George Burket. Glenn was primarily responsible for the data-processing tasks involved in preparing master tapes containing the raw data, while Gary and George were more involved in other activities, such as data analysis. (The distinction between "data processing" and "data analysis" may be a slight one, but it is convenient nonetheless. It was particularly convenient in Pittsburgh, because the University had two entirely different computers—the IBM-7070, which was used primarily for data processing, and the IBM-7090, which was used for data analysis.)

Gary, the first of the three to join the staff, was also the first to leave. (He left in 1963, to accept a position with IBM.) George Burket left in 1963 to accept a position in the test division of Science Research Associates, where he joined the staff of Bill Clemans, the present Project Director. In 1964, about the time that the project moved from Washington to Pittsburgh, Glenn Roudabush, too, resigned. He joined the staff of California Test Bureau (now
CTB/McGraw-Hill) where he remains. (George Burket has been there, too, since about 1967.)

Thus, except for a couple of programmer trainees there was no carryover in computer staff when the project moved to Pittsburgh.

Jon van Wormer was the first programmer hired; Bary Wingersky was hired shortly thereafter. Their responsibilities were split down the middle, with Jon being responsible for the data processing and Bary for data analysis. Bary, who is an extremely talented programmer, and the staff working under his direction, wrote all the programs that the research staff needed, including some very complex sequences of multivariate analysis programs. The programs were fast, accurate, and efficient; they worked beautifully. Ready-made packages (such as BMD) were never used while he was on the staff. In effect he developed his own package designed especially to meet the needs of the TALENT staff.

When Jon van Wormer left in 1966, Marilyn Wingersky (Bary's wife and a skilled programmer in her own right) was promoted to fill the vacancy. (Marilyn had been on TALENT's programming staff for about a year, so she was thoroughly familiar with the situation.)

When the project moved to Palo Alto in 1967, Bary and Marilyn Wingersky decided they preferred to live in the East, so they both accepted job offers from Educational Testing Service, in Princeton.

Meanwhile Dick Williams, one of the two staff members who had moved with the project from Washington to Pittsburgh*, had been learning programming. In Washington he had been an administrative assistant and then a research assistant; in Pittsburgh he had started out as a computer trainee; by the time his three years in

* The author of this History was the other one.
Pittsburgh were completed he was a full-fledged programmer. Thus when the move to Palo Alto left the project without a programming supervisor, Al Carp promoted Dick to that position. It was a logical decision; Dick was clearly more familiar with the TALENT tapes and TALENT procedures than any newcomer could have been. When Dick left in 1970 to accept a job as programmer for a pharmaceutical company John Claudy was his successor as Director of Computer Operations. John was already Director of the Data Bank, and he continued in that role, too. He continued to wear the "Director of Computer Operations" hat until 1973, and the Data Bank hat until 1975. In 1973 Gary Fulscher, a senior programmer on the staff, who was working for his doctorate at Stanford, was promoted to Director of Computer Operations. In 1974, by an administrative decision, the entire programming staff was made independent of Project TALENT, and that is why this particular section of the chapter ends at this point!

7. **Computer trouble-shooter task forces**

Now we come to the point where we should say something about our series of one-man task forces, in Palo Alto, for the solution of our tape documentation and other computer-related problems. As the reader will learn in Chapter 13, tape format documentation, records, and allied tasks associated with the management of computer operations were, it seemed, never-ending problems. Therefore at various times one-man task forces were appointed to make what progress they could towards amelioration of the trouble. Rick Johnson, a psychologist who was knowledgeable about computers and had been on the AIR staff for several years (though not on the TALENT staff) was the first appointee to this difficult assignment. And he made real progress on it, by producing the best tape inventory we had ever had (some might say the only one we had ever had). This task required hard work, carefulness, and accuracy—
all of which Rick provided. He also began work on a Standards Manual. When he left Project TALENT in 1972, Calvin Wright stepped into the breach; his assignment was to complete the Standards Manual that Rick had started. This was to be useful in maintaining appropriate standards of accuracy and thoroughness in programming, checking, and documenting, for all computer-related operations. As was touched upon earlier, Cal converted the embryo "Standards Manual" into a massive "Systems Manual", which was not limited in its scope to computer-related operations. The third one-man task force for TALENT's computer operation was Don Horst, who was assigned this role when he returned in 1972 from an AIR assignment in the Far East (mostly in Singapore). Don, unfortunately, did not have much time to devote to operation of the computer operation, since his other major assignment from the Project Director, which was to implement parts of the Systems Manual (mostly the non-computer-related parts) throughout Project TALENT, proved to be inordinately time-consuming—through no fault of Don's, needless to say, since his actions were fully prescribed by the Systems Manual. Don left Project TALENT, and AIR, in 1973.

In 1975 we received a generous grant from NIE explicitly for improving the tapes and their documentation, and carrying out certain related tasks such as the preparation of an improved handbook for Data Bank users. The grant was a portion of a considerably larger grant, and it was earmarked for this "tape clean-up" activity.

Marion Shaycoft, who had written the proposal for this activity and had campaigned for it (or for something like it) for years, was the Director. Her staff consisted of Lauress Wise and Wendy Bartlett. Wendy, who had been Gary Fulscher's successor as the Director of Computer Operations (after this ceased to be a TALENT function), had recently resigned to go back to school, but
she was available part-time as a consultant. Laurie Wise is a skilled programmer (as well as being one of Project TALENT's proliferating associate project directors).

Readers who have a morbid curiosity about some of the computer-related problems that have been touched on in this section are invited to read Chapter 13.

8. Other research staff members

a. Washington phase

The roles of Bill Gorham, Jo Edwards, Izzy Goldberg, and Dave Orr, who were among the earliest additions to the TALENT staff after John Dailey became Project Director have already been mentioned—Bill Gorham in Chapter 7 and the other three in Chapter 5. Of the four, Jo Edwards is the only one who is still at AIR.

Bill Gorham left AIR to accept a job with the U.S. Civil Service Commission, where he is now Director of the Personnel Research and Development Center. Izzy Goldberg is Professor of Community Education at Federal City College, in Washington. Dave Orr, after a number of other activities, is now at NCES, where the population data from which the TALENT sample was drawn came from.

Two other members of the professional staff should be mentioned. Clinton A. (Ted) Neyman, a graduate of the
U.S. Naval Academy, had retired as a captain from his first career, and chose educational research as his second. Actually, though, when he joined Project TALENT in 1960 as the Supervisor for Services (a largely administrative role) the field of psychometric research was by no means new to him. Captain Neyman, before his retirement from the Navy, had been assigned for many years to the Bureau of Naval Personnel, where he directed a section responsible for psychometric research. And under Navy aegis, he had been sent to Stanford University, where he had earned an M.A. in 1949 in educational research. Ted stayed at AIR until 1964, gradually adding research activities to his administrative assignments. He was a valuable staff member. When Project TALENT moved to Pittsburgh, Ted remained in Washington; he moved with John Dailey from AIR to George Washington University, where he was on the staff of the Educational Research Program. Simultaneously he worked towards his doctorate at GW. He received his Ed. D. there in 1971, and went on to direct the Educational Research Program at GW after John Dailey left. It is sad to have to report that after this rather impressive mid-life career-change level of achievement, he died, much too young, a few years later.

Another staff member from the project's Washington period was Dan Sillers, a counseling psychologist. Dan doesn't show up in Table 8-1, because he was associated with the project so briefly. The only reason he is mentioned here is that to the best of my knowledge he is the only former staff member who became a college president. After leaving Project TALENT, he joined the faculty of Jamestown College in North Dakota, and was appointed President a couple of years later.
b. Second Pittsburgh phase

Some members of the research staff when the project was in Pittsburgh have not yet been mentioned and should be. Paul Lohnes, who was Bill Cooley's coauthor on their book on multivariate analysis (Cooley and Lohnes, 1962) joined the project in 1965. Like Bill, Paul had a teaching appointment at Pitt. Like Bill, Paul was (and probably still is) a multivariate-analysis enthusiast. Like Bill, when the time came to move to Palo Alto, Paul decided he didn't want to leave the East (although he was willing to move away from Pittsburgh). He decided to rejoin the faculty of the State University of New York at Buffalo, where he had been immediately before joining the TALENT staff. His reason for preferring New York State to California wasn't just that he considered the Atlantic a better ocean than the Pacific. Probably no small consideration was the fact that at SUNY he would have more opportunity to teach than at Pitt--and he gives every indication of really enjoying teaching.

Charles (Chuck) Hall was also on the project staff from 1965 to 1967. After he left AIR he was with ETS in Princeton for a while. He is now with Abt Associates, to the best of my present knowledge.

Paul Mok was one of Project TALENT's more memorable staff members. Before joining the Project TALENT staff, Paul had been "chief of party" in a one-person "party" on an AIR project in Liberia, where he and his family spent a couple of years. Although he was only on Project TALENT about four months (late fall of 1964 to early spring of 1965) he left his mark on the project. A prolific writer, Paul had written several popular books about education (including "A View from Within" (1962) and "Pushbutton Parents and the Schools" (1964)) before joining the AIR
staff. When he left Project TALENT, it was to take a position in New York with a firm of consulting psychologists and to write a novel. If Dan Sillers is Project TALENT's only alumnus who became a college president, Paul Mok is probably its only published novelist.

c. Palo Alto phase

For about a year, Project TALENT had a resident sociologist on the staff: Alan E. Bayer. (He was already familiar with the project because he had been a Data Bank client previously.) While on the project staff, he used TALENT data for some interesting research on marriage and some research on college dropouts. When he left in 1968, it was to join the faculty of Florida State University.

Mike Wargo, who had been on the AIR staff for a couple of years and had done an excellent (though controversial) reanalysis and synthesis of compensatory education evaluation data (Wargo, Tallmadge, Michaels, Lipe, and Morris, 1972), joined the TALENT staff in 1971. His chief assignment was to develop a tentative plan for a study of the careers of scientists, since NSF had provided a grant for that purpose. When Mike left in 1972 to take a position with the Office of Education, Kast Tallmadge, who had been at AIR since 1966 and had done lots of interesting and worthwhile studies involving, among other things, evaluation of compensatory education, the development of training procedures, exploration of aptitude-treatment interactions, and instructional technology, joined the TALENT staff and took over on what Mike Wargo had been doing. (Kast left, in 1973, to establish and direct a Western office for RMC Research Corporation. Don Horst left at about the same time, to join him in that venture. Marion
Shaycoft inherited the NSF study from Kast, completed the planning, and completed the study.\(^\ast\)

Two other fairly recent additions to the Project TALENT staff are Bob Rossi and Kevin Gilmartin, both of whom received their doctorates in 1974. Bob's, from Stanford, is in philosophy of education (making him our first genuine certified philosopher!) while Kevin's, from Carnegie-Mellon, is in cognitive psychology—which, at Carnegie-Mellon, includes a heavy dose of computer modeling, so that Kevin may be the only one at AIR who knows much about that topic.

d. Washington, Pittsburgh, and Palo Alto

Louise Curenton has been on the staff since 1963, and even before that she was associated with Project TALENT as a Regional Coordinator (in Knoxville). Thus it is impossible to classify her as belonging in the Washington, Pittsburgh, or Palo Alto phase; she belongs in all three. Throughout all three phases, however, Louise has remained in Knoxville, where she directs our small but important "Knoxville detachment". Louise has carried out several studies based on the Knoxville data, supplemented by special follow-up data which she collected herself. Among the studies she has done are the following:

1) A study of the early identification of behavior problems

2) A study of welfare dependence

3) A study of the effects of having alcoholic parents

We hope she will continue doing these interesting and worthwhile studies, and that she will continue to take advantage of that valuable research resource, the TALENT data on the Knoxville and Knox County population, that she has done so much to make a valuable research resource.

\(^\ast\) Shaycoft (1975)
B. Early advisory panels

The counsel provided by our various advisory panels was crucial to the success of the project.

In Chapter 2 (in Figure 2-1) we listed the members of the four initial advisory panels—Guidance and Counseling Panel, Educational Research Panel, Testing Problems Panel, and Manpower and Sociology Panel—and (in Figure 2-3) the members of the Advisory Council. In Chapter 3 we listed our panel of advisors on sampling. Needless to say, we are very grateful to everyone in these groups; they helped us get the project started.

C. Regional Coordinators

As is explained in Chapter 7 (Section B) in making arrangements with the schools selected for our sample, and in getting the tests administered, we were, to a very great extent, dependent on our Regional Coordinators. The initial group is listed in Figure 8-3. (Over the years, we have continued to count on Regional Coordinators' assistance, in conducting our follow-ups. The list has changed gradually, and our gratitude extends to the newer Regional Coordinators as well as to the group listed.

D. OE and NIE personnel

Until the end of 1971, the chief financial support of Project TALENT came from the Office of Education; after that, the National Institute of Education was established, and from then on NIE has provided the support for the project. Obviously there had to be many people in both organizations who provided help, backing, and moral support, when help, backing, and moral support were what we needed. The project might not ever have gotten started if it hadn't been for the support of the Commissioner of Education, Dr. Lawrence Derthick, in 1958, and the Assistant Commissioner, Dr. Roy Hall, who was also
### Figure 8-3. REGIONAL COORDINATORS IN 1960

<table>
<thead>
<tr>
<th>State</th>
<th>Coordinator Name</th>
<th>Institution</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>Paul A. Givens</td>
<td>Birmingham-Southern College</td>
<td></td>
</tr>
<tr>
<td>Arizona</td>
<td>Richard E. Schutz</td>
<td>Arizona State University</td>
<td>Tempe</td>
</tr>
<tr>
<td>Arkansas</td>
<td>Carter Short</td>
<td>University of Arkansas, Fayetteville</td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>Alex D. Aloia</td>
<td>Loyola University of Los Angeles</td>
<td>Los Angeles</td>
</tr>
<tr>
<td>Connecticut</td>
<td>Joseph Raymond Gerberich</td>
<td>University of Connecticut, Storrs</td>
<td></td>
</tr>
<tr>
<td>Delaware</td>
<td>Arthur R. Delong</td>
<td>University of Delaware, Newark</td>
<td></td>
</tr>
<tr>
<td>Florida</td>
<td>Edward Caldwell</td>
<td>Board of Public Instruction, Manatee</td>
<td></td>
</tr>
<tr>
<td>Georgia</td>
<td>Cameron Fincher</td>
<td>Georgia State College for Business Administration</td>
<td>Atlanta</td>
</tr>
<tr>
<td>Idaho</td>
<td>Elwyn Dallaur</td>
<td>State Guidance Supervisor, Department of Education</td>
<td>Boise</td>
</tr>
<tr>
<td>Illinois</td>
<td>N. L. Gage</td>
<td>University of Illinois, Champaign</td>
<td></td>
</tr>
<tr>
<td>Indiana</td>
<td>N. A. Paum</td>
<td>University of Indiana, Bloomington</td>
<td></td>
</tr>
<tr>
<td>Iowa</td>
<td>Arthur Mctman</td>
<td>University of Iowa, Iowa City</td>
<td>Iowa City</td>
</tr>
<tr>
<td>Kansas</td>
<td>Kenneth E. Anderson</td>
<td>University of Kansas, Lawrence</td>
<td></td>
</tr>
<tr>
<td>Kentucky</td>
<td>Ernest McDaniel</td>
<td>University of Kentucky, Lexington</td>
<td></td>
</tr>
<tr>
<td>Louisiana</td>
<td>Robert N. Vidulich</td>
<td>Louisiana State University, Baton Rouge</td>
<td></td>
</tr>
<tr>
<td>Maine</td>
<td>David R. Fink Jr.</td>
<td>University of Maine, Orono</td>
<td></td>
</tr>
<tr>
<td>Maryland</td>
<td>Robert C. Lloyd</td>
<td>Baltimore Public Schools</td>
<td></td>
</tr>
<tr>
<td>Massachusetts</td>
<td>Seth Aksenian</td>
<td>Springfield College</td>
<td></td>
</tr>
<tr>
<td>Michigan</td>
<td>Claude L. Nemzek</td>
<td>University of Detroit</td>
<td>Detroit</td>
</tr>
<tr>
<td>Minnesota</td>
<td>Ralph F. Berdie</td>
<td>Director, Student Counseling Bureau</td>
<td>University of Minnesota</td>
</tr>
<tr>
<td>Mississippi</td>
<td>Roscoe A. Boyer</td>
<td>University of Mississippi, University</td>
<td></td>
</tr>
<tr>
<td>Missouri</td>
<td>Joseph L. French</td>
<td>University of Missouri, Columbia</td>
<td></td>
</tr>
<tr>
<td>Montana</td>
<td>William A. Garrison</td>
<td>Eastern Montana College of Education, Billings</td>
<td></td>
</tr>
<tr>
<td>Nebraska</td>
<td>Warren Baller</td>
<td>University of Nebraska, Lincoln</td>
<td></td>
</tr>
<tr>
<td>New Hampshire</td>
<td>Paul Mcintire</td>
<td>University of New Hampshire, Durham</td>
<td></td>
</tr>
<tr>
<td>New Jersey</td>
<td>Donald W. Campbell</td>
<td>Board of Education, Newark</td>
<td></td>
</tr>
<tr>
<td>New Mexico</td>
<td>Virginia Keehan</td>
<td>Department of Education, Santa Fe</td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>Warren W. Coxe</td>
<td>University of Buffalo</td>
<td></td>
</tr>
<tr>
<td>North Carolina</td>
<td>Roy N. Anderson</td>
<td>North Carolina State College, Raleigh</td>
<td></td>
</tr>
<tr>
<td>North Dakota</td>
<td>Ralph H. Kolstoe</td>
<td>University of North Dakota, Grand Forks</td>
<td></td>
</tr>
<tr>
<td>Ohio</td>
<td>Howard B. Lyman</td>
<td>University of Cincinnati</td>
<td></td>
</tr>
<tr>
<td>Oklahoma</td>
<td>W. R. Brown</td>
<td>University of Oklahoma, Norman</td>
<td></td>
</tr>
<tr>
<td>Oregon</td>
<td>J. Spencer Carlson</td>
<td>University of Oregon, Eugene</td>
<td></td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>L. Kathryn Dice</td>
<td>Department of Public Instruction, Harrisburg</td>
<td></td>
</tr>
<tr>
<td>Rhode Island</td>
<td>Frances L. Dunn</td>
<td>Brown University, Providence</td>
<td></td>
</tr>
<tr>
<td>South Carolina</td>
<td>R. L. Kalmbach</td>
<td>Columbia Public Schools, Columbia</td>
<td></td>
</tr>
<tr>
<td>South Dakota</td>
<td>V. Gregory Rosemont</td>
<td>Huron College, Huron</td>
<td></td>
</tr>
<tr>
<td>Tennessee</td>
<td>George E. Copple</td>
<td>Vanderbilt University, Nashville</td>
<td></td>
</tr>
<tr>
<td>Texas</td>
<td>Robert P. Anderson</td>
<td>Texas Technological College, Lubbock</td>
<td></td>
</tr>
<tr>
<td>Utah</td>
<td>Hyrum M. Smith</td>
<td>Department of Public Instruction, Salt Lake City</td>
<td></td>
</tr>
<tr>
<td>Virginia</td>
<td>Edward L. Beard</td>
<td>University of Virginia, Charlottesville</td>
<td></td>
</tr>
<tr>
<td>Washington</td>
<td>William C. Budd</td>
<td>Western Washington College of Education</td>
<td></td>
</tr>
<tr>
<td>West Virginia</td>
<td>Walter Jaroske</td>
<td>University of West Virginia, Morgantown</td>
<td></td>
</tr>
<tr>
<td>Wisconsin</td>
<td>Elder A. Bond</td>
<td>Milwaukee Public Schools</td>
<td></td>
</tr>
<tr>
<td>Wyoming</td>
<td>R. Duane Andrews</td>
<td>Department of Education, Cheyenne</td>
<td></td>
</tr>
</tbody>
</table>
the Director of OE's Cooperative Research Program. After Roy Hall left, his successors as Director of the Cooperative Research Program continued to provide support and help. Dr. David Clark and Dr. Francis Ianni are among those in this category.

At most stages, however, the person with whom we had closest contact was the project monitor. Over the years there have been a great many monitors, some who filled this role for several years and some for just a few months. As a group they were interested and helpful—and their help was indispensable. In mentioning specific monitors we must start with Alice Scates (of the Office of Education)—not only because she was our first monitor, but also because she was our most frequent monitor, having filled that role at least three separate times, with other people functioning as monitors in the interims; and not only has Alice been our only "recidivist monitor" but she has continued to maintain a lively interest in the project even at other times.

Having named the first monitor, let us now mention the last one who held that position during Project TALENT's Office of Education period—or better yet, let us name the last three: Judith D. Weinstein, Susan Klein, and Laurence G. Goebel. And lest anyone get the idea from the fact that three of the four monitors named have female names that some sort of sex discrimination (against men) was operating, let me assure the readers that all of the numerous monitors between Alice Scates and Judy Weinstein were male—as were all the monitors in the periods between Alice's several hitches in that role.

The list of monitors in TALENT's NIE period is shorter, perhaps because the period itself is shorter (only about 5 years, as compared with OE's 14). When NIE was started, Larry Goebel, TALENT's last monitor under OE, transferred to NIE, and continued as monitor for quite a while. This was important because it helped provide con-
tinuity. The next monitor was Garry McDaniels, who was the director of NIE's Teaching Division, in the Basic Skills Group. He continued in a general overseeing role, after he appointed a member of his staff, Jane David, as the next monitor. When Jane left NIE, Andrew C. Porter, who had recently become Acting Director of the Basic Skills Group, inherited responsibility for Project TALENT: he appointed Richard Harbeck as monitor. All of the NIE people named--Larry, Garry, Jane, Andy, Dick--have been interested and helpful--and we needed their interest and their help. Dick Harbeck continues as monitor.

The part of the study represented by NSF's supplementary funds had its own monitor, Dr. Charles H. Dickens of NSF. He and a member of his staff, Naomi Sulkin, have had a continuing interest in Project TALENT.

E. TALENT post-doctoral fellows

Chapter 9. THE GEOGRAPHY OF THE PROJECT

If by "the geography of Project TALENT" were meant the location of the members of the sample, the correct description would be that they are all over the country. Strangely enough, the correct description is almost the same if by "the geography of Project TALENT" is meant the location of Project headquarters. The physical location of project headquarters has changed almost as many times in the years since the project started as has the identity of the Project Director.

In 1957, when the project obtained its first funding (a grant for use in planning) John Flanagan was living in Pittsburgh, where AIR's headquarters were at the time. He appointed Robert Craig (a member of the AIR staff at the time) as Project Director. The project was then a very small one; the planning activity was virtually a one-man operation. In the spring of 1958, when Bob Craig resigned to join the faculty of Marquette University, John Dailey, who had been Technical Director of the Research Division of the Bureau of Naval Personnel, resigned from that position to become Project Director of what was later to become "Project TALENT". At that point, since Bob Craig's departure left no project staff in Pittsburgh except the Principal Investigator, who habitually spent part of every week in Washington anyhow, and since the new Project Director was in Washington, project headquarters transferred to that city. Within a few months the project staff enlarged considerably from its one-man size, and it has never again shrunk that small (although its size has fluctuated wildly from time to time). Four members of AIR's research staff transferred from other activities to the new project. They were Dorothy S. (Jo) Edwards, William A. Gorham, David B. Orr, and the author of this History.

The geography got complicated about 1959, when Project TALENT started to need a computer. Though the project staff (by then, a fairly sizable group) was in Washington, it became clear that the computer we were going to have to use would be in Pittsburgh (at
the University), and for a very good reason. Because the Cooperative Research Grant was to the University of Pittsburgh rather than to AIR, the Project would have free use of the University's computer. Free computer time is a nice thing to have, but as we shall see (in Chapter 13) the distance of 250 miles between the location where the Project Director and most of the project staff were and the University's Cathedral of Learning, where the computer was ensconced, caused problems. In 1964 these particular problems were resolved by moving project headquarters back to Pittsburgh (where they had started out). Only two staff members (Marion Shaycoft and Dick Williams) made the move to Pittsburgh. The rest all left the project (either for other activities at AIR or to transfer elsewhere). For the next two years the Principal Investigator, the Project staff, and the computer were all based in the same city--Pittsburgh.

In the summer of 1966 the Principal Investigator moved across the country to Palo Alto while the rest of the project remained in Pittsburgh for another year. New OE regulations that permitted a nonprofit organization such as AIR to receive direct support for Project TALENT without the necessity of using the University as an intermediary led to a formal severing of the Project's relations with the University, the transfer of the contract to AIR (effective the beginning of 1967), and a half-year later, another move of Project headquarters--this time to Palo Alto, where the Principal Investigator already was. Since the severing of the Project's ties to the University meant that we no longer had access to virtually unlimited computer time on the University's computers, the computer operation as well as the rest of the project could make the move to Palo Alto in the summer of 1967. The project (including Principal Investigator, staff, and computer) has been there ever since.

The geography of Project TALENT--in other words who and what were where when--is summarized in Table 9-1.
<table>
<thead>
<tr>
<th>Computer and Computer Staff</th>
<th>Principal Investigator</th>
<th>Other Project Staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1957-58 ---</td>
<td>Pittsburgh, commuting to Washington weekly</td>
<td>Pittsburgh</td>
</tr>
<tr>
<td>1958-59 ---</td>
<td>Pittsburgh, commuting to Washington weekly</td>
<td>Washington</td>
</tr>
<tr>
<td>1959-64 Pittsburgh</td>
<td>Pittsburgh, commuting to Washington weekly</td>
<td>Washington</td>
</tr>
<tr>
<td>1964-66 Pittsburgh</td>
<td>Pittsburgh</td>
<td>Pittsburgh</td>
</tr>
<tr>
<td>1966-67 Pittsburgh</td>
<td>Palo Alto</td>
<td>Pittsburgh</td>
</tr>
<tr>
<td>1967-76 Palo Alto</td>
<td>Palo Alto</td>
<td>Palo Alto</td>
</tr>
</tbody>
</table>

TABLE 9-1. The geography of Project TALENT
(Who and what were where when)
Chapter 10. EARLY REPORTS

In this chapter we discuss, very briefly, the research reports which were produced by the TALENT staff up to 1966. Chapter 20 ("Later Studies") does the same thing for subsequent reports. Admittedly the division point between "early" and "later" is to a great extent arbitrary.

However the reports discussed in the present chapters are ones which, it is hoped, provide a basic understanding of the project, the data, their scope, and their potential. (Only major reports are discussed here, not short articles, or papers presented at professional meetings.) In neither Chapter 10 nor Chapter 20 are the findings of the studies discussed, since that is beyond the scope of this History.

A. Basic reports

Five reports fall in this category. Two of the five are introductory reports, presenting an overview of the plans for the project, and describing the preliminary phases. The first of these two is "Project TALENT: Designing the Study" (Flanagan, Dailey, Shaycoft, Gorham, Orr, and Goldberg, 1960, 448 pages). The second is a considerably abridged and re-edited version that has been published by Houghton Mifflin, as a hard-cover book, under the title "Design for a Study of American Youth" (Flanagan, Dailey, Shaycoft, Gorham, Orr, and Goldberg, 1962, 248 pages).

The other three of the "basic" reports deal respectively with school characteristics, high school student characteristics, and the results of the study of the age 15 sample. These reports are, respectively:

Studies of the American High School
(Flanagan, Dailey, Shaycoft, Orr, and Goldberg, 1962, 375 pages)

The American High School Student
(Flanagan, Davis, Dailey, Shaycoft, Orr, Goldberg, and Neyman, 1964, 738 pages)
Studies of a Complete Age Group—Age 15
(Shaycoft, Dailey, Orr, Neyman, and Sherman, 1963, 288 pages)

B. Equating studies

The Air Force and the Navy both gave us contracts to equate their tests to the TALENT battery, and perform related analyses. The resultant reports are, respectively:

For the Air Force:

Calibration of Air Force Selection Tests to Project TALENT Norms
(Dailey, Shaycoft, and Orr, 1962)

Calibration of academic aptitude composite to Project TALENT results
(Shaycoft, 1963)

For the Navy:

Comparisons of Navy Recruits with Male High School Students on the Basis of Project TALENT Data
(Shaycoft, Neyman, and Dailey, 1962)

C. Subsequent studies applying the results of the equating

Following the equating of the Air Force tests to the TALENT Battery, the Air Force commissioned us to apply the results of the equating in various ways. Some of the resultant reports are:

Comparison of Air Force Recruits with Segments of the High-School-Age Population
(Shaycoft, 1963)

Normative Distributions of AQE Indexes for High-School-Age Boys
(Tupes and Shaycoft, 1964)
Chapter 11. RETURNING TO THE SCENE

The original plans for Project TALENT had included nothing about retesting. But during the school year 1962-63, the year that our original 9th-grade cohort was due to graduate, we decided that it would be highly desirable to retest some of them; it would be our last chance while they were still in school. And we felt that much could be discovered about the factors that affect the amount of learning and other changes that occur in high school.

A. Designing the retest study

Having decided to retest, the only questions remaining were who, what, where, when, how much, and how.

1. Who

Obviously we were limited to members of the 9th-grade cohort, since they were about the only members of the original sample who were still in school. But how many? We decided to aim at testing in roughly one-tenth of the original schools in our probability sample and to test all 12th-graders in those schools.

2. Where

It was decided that not all categories of schools should be included. Nonpublic schools would be omitted, and because there were indications that it would be very difficult to get permission to retest in the five largest cities it was decided not to include those cities (except for some vocational high schools).
Those exceptions eliminated four of the 19 categories in the taxonomy of high schools developed by John Dailey and first reported on in "Studies of the American High School" (Flanagan, Dailey, Shaycoft, Orr, and Goldberg, 1962); but 15 categories remained. It was decided to select the sample of schools by stratifying into those 15 categories, to insure that all of them were represented. Accordingly, a sample of 144 schools was selected. This was a little more than we wanted, but we anticipated that some of them would decline. Some of them did, but not as many as we had expected—so we found ourselves with 108 schools to test. This was a somewhat larger group than we had anticipated, but we were pleased that the discrepancy was in that direction, not the other.

3. What

The question we had to answer was: What tests and questionnaires shall we retest with? We decided that we wanted to use almost the entire original battery. The only instruments that would be omitted were the Memory for Sentences test, the Preferences test, and the Student Activities Inventory (the personality inventory). We also decided to revise and shorten the Student Information Blank.

4. When

The best time for the retesting would be the spring, so that the data would be comparable, in regard to time of year collected, with the 1960 data.

5. How much

In contacting the schools selected for the sample, we had not asked for two days of testing time (as was done for the 1960 testing) but for just one day, since we felt schools would be readier to participate on that basis. This meant that each student could take only about half the tests in the battery.
6. How

We didn't want to just divide the battery in half and give half the students one half and the other half the other half. That would have meant we couldn't get correlations between the tests in one half and those in the other half. Because we felt it was necessary that every test be taken in conjunction with every other, we divided the battery in four parts (the four original half-days) and paired every one of the four resulting half-day half-batteries with every other, to get six full-day batteries (Retest Batteries A, B, C, D, E, and F). For good measure (no pun intended), we put the Abstract Reasoning Test in all six batteries, to provide a common metric. Each of the batteries was administered in about one-sixth of the schools (except that a seventh battery—Battery V—was used in 10 vocational high schools, added later).

B. Processing the answer sheets

Things had progressed rapidly at Measurement Research Center since their processing of the answer sheets from the 1960 testing. By 1963 they had a document-to-tape optical scanner that could handle the kind of answer sheets that in 1960 had required the Scoring Machine; the new optical scanner would put every item response directly on tape. From the tapes containing item responses, we could thus obtain whatever scores we wanted, using our own (or rather, Pitt's) computer to calculate them.

This meant that in addition to studying change over the 3-year period, we would also be able to obtain split-half reliability coefficients where appropriate and we could also, at last, get empirical data on the degree to which the tests were speeded for 12th-grade students.

C. Data analyses and final report

The data were analyzed in many different ways, and from many different angles. Several methodological innovations were introduced for analyzing the data. The final report, "The High School
Years: Growth in Cognitive Skills" (Shaycoft, 1967, 376 pages), discusses both the methodology and the substantive results in detail.

D. Test reliability data

One important point about the report mentioned above is that at present it is the best source of test reliability data. Columns 5 and 6 in Table 4-7 (Shaycoft, 1967) contain split-half reliability coefficients for the grade 12 norms group. Researchers using TALENT variables and needing reliability coefficients are urged to use the values found in that table; but it is suggested that they apply the conventional correction-for-range formula to obtain reliability estimates applicable to the set of cases on which the corresponding research data (test scores) are based.

It might be mentioned, incidentally, that though some of the tests are very short (since they were designed for research use, not for decisions on individual students) and may therefore have only moderate reliabilities, the reliabilities are high for most of the tests, and on a per-item basis they are, on the whole, excellent, even for the shortest tests.
Chapter 12. THE 4% SAMPLE

When we were arranging with MRC for the processing of the answer sheets to be used in the 1963 retesting discussed in the previous chapter, we found out about the advances in MRC's equipment (see Chapter 11, Section B). These advances meant that the same answer sheets used in 1960 and processed on the Scoring Machine, in other words answer sheets A and C, could now be processed on an optical scanner which would produce output containing individual item responses, rather than merely 13 scores. Furthermore the output would be tape, not cards; and thus only one pass of an answer sheet would be necessary to get all the item responses recorded.

We decided to take advantage of this improved equipment by excavating answer sheets A and C for our original 4% sample--i.e., cases for which the six-digit student identification number ended in 00, 25, 50, or 75--from the catacombs where they were stored, and shipping them to MRC, where we could have them processed very inexpensively on the new equipment, by arranging for them to be "piggy-backed" on the corresponding sets of answer sheets from the 1963 retest. (It was not necessary to have any 4%-sample answer sheets other than A and C reprocessed, since, it will be recalled, the other answer sheets from the 1960 testing (answer sheets B1 and B2, and Record Form Z) had been in a different format and had originally been processed on the Document Reader then existing, which had produced card output containing individual item responses. (The Document Reader had been unable to handle the format used for answer sheets A and C, which had been designed for the Scoring Machine.)

Thus for 4 percent of the cases we would be able to develop a tape containing every item response, as well as the standard set of scores, augmented by whatever additional scores we might wish to obtain either for experimental purposes (including methodological research) or because we considered them useful for substantive research.

To produce this magnificent data tape containing the complete set of data would require that the new tape containing the answer sheet A
and C data be merged with the item responses and other data from answer sheets B1, B2, and Z already in the tape master file. This would not be an easy job. At that point our tape files were in a rather primitive stage of development—not even in perfect numerical order—so that merging could not be handled efficiently. Furthermore not everybody's record had as yet been incorporated in the tape master file. The same difficulty in merging would apply of course to the retest data tapes.

It was decided, therefore, that the sensible thing to do would be to lay aside the new tapes received from MRC (i.e., the 4%-sample tapes for answer sheets A and C and all the 1963 retest data) until the master file tapes were in better shape, before proceeding with efforts to merge and edit. Unfortunately, by the time the master file was getting into fairly usable shape the project had moved to Pittsburgh and no longer had the same Project Director it had had in Washington. None of the subsequent Project Directors agreed that a complete and fully edited 4% sample data tape would be a valuable asset to the project—and so they never assigned this task high priority; they never assigned it even medium priority; it was always shunted to the bottom of the list. And there it sat—for years! Occasionally some member of the programming staff would be assigned responsibility, at a low level of priority, for some sort of superficial editing and formatting, but nothing much ever came of it.

The author of this History continued to campaign (incessantly, some might say) for getting the 4%-sample tape file in first-rate shape. Thus when she had an opportunity, about December 1973, to write a proposal to be submitted to NIE (at their invitation) to revamp the TALENT tapes and their documentation (see Chapter 13, section D2) she made certain that fixing up the 4%-sample tape and documentation was an integral part of the work proposed. It was funded in the summer of 1975. For further details see Chapter 13.

Before dropping the subject of the 4%-tape even temporarily we should point out that not the least of its advantages would be that it would yield eminently sound reliability estimates for any variables for which internal consistency estimates (e.g., split-half reliability, KR-20 reliability)
would be appropriate. For many of the TALENT variables (including most of the test score variables) there was no way of determining reliability on the basis of the regular 1960 data. The 1963 retest data would to a certain extent share these advantages with the 4% sample data, in regard to the determination of reliability, but the 4% data would be preferable for the following reasons:

1. It would be available for all four grades, not just for grade 12.

2. It would be for a probability sample of the grade.

3. It would be available for all variables in the TALENT battery (including some that hadn't yet been computed, and even some that hadn't yet been defined) rather than being limited by the composition of the six retest batteries.
Chapter 13. TAPES: TRAUMA, TROUBLES, TRAVAIL, TRIALS, TRIBULATIONS, AND T.E.A.R.S.

A. TALENT's tapes: Ancient history

1. Iowa City (1960)

When the original answer sheets had been processed at MRC in 1960 the tapes didn't yet exist, but about 5 million IBM cards (11 per student) did. One of the limitations of the optical scanning equipment that was used was that unless we had arranged to get counts of "number of items attempted" on every test* rather than merely counts of the number answered correctly, there was no way that it would be possible, on the basis of the output cards alone, to distinguish between zero scores, denoting that the test was taken but no item was answered correctly, and missing data (no item attempted, suggesting, though not indicating with certainty, that the student had been absent from the testing session). Both missing data and genuine zero scores were represented on the IBM cards by zero punches.

2. Coping with geography

It will be recalled (see Chapter 7 section E3, and Chapter 9) that when the answer sheets were being processed at MRC we (most of the project staff) were in Washington, while our computer staff and the computer we were using were both in Pittsburgh--about 250 miles away. This sort of splitting of the operation is really a bad idea, but it was a logistic necessity. We had to use the University of Pittsburgh computer because we were getting virtually unlimited free time on it. However I would like to offer this

*For budgetary reasons doing this would have been out of the question, since it would have doubled our scoring costs.
advice to future researchers: If you have any choice at all in the matter, the computer staff should be at the same site as the other staff members—and the computer should also be nearby—preferably within a few miles and certainly no more than 15 or 20 miles away. (Not everything in a project of the magnitude of Project TALENT can be handled reasonably on a terminal.)

3. **Zeros and blanks**

The fact that the output of the MRC Scoring Machine would not distinguish between zeros and blanks had been foreseen. We (in Washington) were aware of the Scoring Machine's limitation in this respect, and had intended it to be taken care of in the tape editing process, by making the assumption, in the case of answer sheets that had been processed on MRC's Scoring Machine, rather than the Document Reader, that if a student's scores were zero for all the tests administered in a half-day testing session, the student had been absent that half-day, and his "zeros" should be converted to whatever code was being used to indicate "blank" or "missing data". But presumably because of the 250 miles separating that intention (in the heads of some people in Washington) from the programming operation (by personnel in Pittsburgh) the distinction between zeros and blanks didn't get edited into the tape. In the years that followed, sporadic efforts were made to get the necessary corrections made, and this was largely accomplished in the case of students who lacked an entire answer sheet which might (or might not) indicate absence for a full day, but not for students who missed only a half day and were not missing an answer sheet.

While this problem applies to a very small number of cases (certainly less than 1/2%) it has been a continuing concern. Recent efforts at reducing or eliminating such problems are discussed later in this chapter.
4. Building a tape master file

Zeros and blanks (and the difference between them) didn't constitute the only problem. Far from it! From some of the answer sheets two cards were supposed to have been made, and from others, three. But when the time came to merge the tape records for cards which had been made from the same answer sheet, and which therefore should have matched perfectly, there were lots of unmatched records. And when, because of double-punched columns, two tape records were made from the same card (see Chapter 7 sections E2 and E3) there should have been perfect matching between the two. Inexplicably, there wasn't. (The word "inexplicably" is not intended to imply that there was no possible explanation. As a matter of fact several explanations were possible; but because of inadequate control procedures, there was no way of knowing which was the correct one.) Of the 17 sets of tape records no 2 matched perfectly. And again, partly because of the lack of good communication between the Washington staff and the programming staff in Pittsburgh, bad procedures were being used and bad results were being obtained, in the efforts to merge the 17 tape records for each student to create a "master file" containing the students' complete records. The master file had to be created first for "school subsample 0" as indicated in Chapter 3 section A4, so that we would have a 10% sample on which to do preliminary analyses. The remaining 10 school subsamples (9 probability subsamples and one non-probability) could of course have been processed all together--each step being done for all ten at once. However, this would have greatly increased the costs of coping with hardware errors, a not infrequent problem in those days. Instead, it was decided to continue on the one-subsample-at-a-time routine--although this meant that each of the very numerous steps was carried out 10 times instead of just once. And then it turned out that the 4% sample, which had been given special treatment initially so that preliminary norms
could be obtained for use in reports to the schools, had been omitted from subsequent processing; this meant that more backtracking was necessary to prepare master file records for these cases. Small wonder that the master file production seemed endless! It was endless (figuratively speaking). It went on for years (literally).

Some of the cases that should have merged didn't, because of either bad card records or bad tape records produced from good card records, or bad sorting. (The latter could wreck the merging not only of a missorted case but of every case subsequent to it in the same file—a fact which was ignored for far too long, resulting in far too many failures to merge and far too many unnecessary steps.)

By sometime in 1963 all the merging operations had been completed at least once, resulting in merged master file records for 292,050 cases in the grade 9-12 probability sample, plus some other cases, together with unmerged remnants for 100,000 cases or more. The merging efforts had involved, primarily, matching on student identification number and beginning of last name.

Efforts to merge what data there were for 100,000 cases represented in the remnants was a more complicated operation, since many of the failures to match were due to errors that had been made in position-coding a digit of the student identification code or a letter of the name, on the answer sheet, or to failure of the optical scanner to pick up one or two characters of the identification data. Thus visual inspection of printouts of the unmatched cases, inferences as to what was wrong, hand-matching, and keypunching the correction cards were all integral parts of the ever-so-laborious, ever-so-time-consuming process of converting cases in the remnant filed to cases in the master file. (In retrospect, a major part of this could have been obviated through different procedures in the earlier phases—but that is water over the dam.)
Efforts to merge the 100,000 or more cases represented in one or more of the various remnant files were continued, and were still going on (with steady progress being made) after the project moved from Washington to Pittsburgh in 1964; and it was still going on after the project moved from Pittsburgh to Palo Alto, in 1967! But before that happened, a part of the problem was eliminated by the Project Director, who ordered that certain tapes which contained unmerged or partly merged "remnants" (in some cases as many as about 20,000 cases*) be scratched. This was an unfortunate "solution," as many of these unmerged or partly merged remnants contained solid sets of data in their own right—for instance all of the Information Test scores, of which there are about 40. But getting rid of these remnants certainly made the tape situation seem neater. The order to scratch remnants did not apply to all of them. Some were retained, for future efforts at merging.

After the move to Palo Alto, Dick Williams, who had become the Director of Computer Operations, worked doggedly on the "remnants of the remnants" and made excellent progress. He was able to clean up most of the remaining remnants, to the point where what was left unmatched could be regarded as negligible. Dick deserves the credit for getting 83,148 probability sample cases (plus some other cases) added to the master file, thus bringing the probability sample total up to 375,198 and the grand total for grades 9-12 up to 404,218. Even by today's standards this is a very considerable accomplishment!

The point of this discussion is that even though data processing technology has improved vastly since the original answer sheets were scanned, researchers are frequently unversed in proper data control procedures. When a project is the size of TALENT, poor procedures can cause very large headaches that are difficult if not impossible to correct.

*If memory serves
5. **Design of the tape master file**

Some readers may wonder why this section doesn't precede the previous one, entitled "Building a tape master file". Preferably, designing a tape should precede building it, not follow it, but it doesn't always happen that way. The data in the master file were not arranged on a logical basis that would make the format easy to understand and the tape easy to use—the format was not logical. Rather, it had just "happened". And that has been another continuing source of problems over the years. Around 1965 one of the programmers was apparently given the assignment of revising the format and, of course, the tapes—but the changes made were rather superficial and hardly improved matters at all. However, that was the format that existed until recently. As we shall see, a major effort to overhaul the tapes and improve their documentation was initiated in 1975 and completed by mid-1977.

6. **Other instances of data loss**

There were other instances over the years in which decisions were made that resulted in some data being lost or destroyed. While the loss of any data is unfortunate, the need to have a manageable data system and the presence of cost and other constraints on what could be done must be recognized. These instances are briefly described below to illustrate the kinds of problems encountered.

**Data about non-examinees.** It will be recalled from Chapter 5 that teachers were requested to fill out Record Form Z for absentees and other students who weren't tested, and to position-code the reason for their not having been tested. Regional Coordinators were requested to do the same thing for 15-year-olds
not in grade 9, 10, 11, or 12. Unfortunately, these data no longer exist.

**Redundancy reduction.** A lot of redundancy had been built into the set of data collected from the students in 1960. This was done deliberately, as insurance against undetected error in our records of important bits of data (e.g., sex, grade, and age, all of which were recorded in at least two different ways, sometimes on different answer sheets). The idea was that if the two position-codes for, say, sex, agreed we could have confidence in the response. If they disagreed one of them would obviously be wrong but we wouldn't know which. The problem might be resolved, however, by inspection of a printout showing first names, although this is a very expensive procedure. An alternative would have been to have a code to indicate that both answers had been given, and then to wait for the follow-up to resolve the dilemma. What happened, instead, was that an arbitrary decision was made as to which of the two sources of the student's sex would be used (e.g., Record Form Z), and the other source (e.g., Answer Sheet A) was ignored. The same thing happened with the record of what grade the student was in.

**Conversion to the IBM 360.** In 1967, when we were getting ready to move to Palo Alto, the data processing staff was busy converting data tapes so that they would be usable on the 360 we would be using in Palo Alto, rather than the 7070 at Pitt. Since the workload of the data processing staff was heavy, the Project Director decided to lighten it by having them "scratch" (wipe out) certain data tapes rather than convert them. The principal tape file chosen for this form of oblivion was the one containing the data for 15-year-olds not in school. The justification that was offered for this action was that the report on 15-year-olds had already been written, so nobody would have any use for the data. It is fortunate that this irreplaceable data was a small file, not a large one.

In addition the file of 15-year-olds in school was scratched, on the grounds that it could be reconstructed from the 15-year-olds' duplicate records on the appropriate tapes organized by grade and sex. On similar grounds the Project Director ordered that on the
tape containing the 1963 retest data and the 1960 data for the same cases, only the 1963 data be converted. His explanation was that in Palo Alto the retest data could again be matched with corresponding 1960 data. As it turned out, after the editing and correcting of the "age" variable in connection with the tape editing and reformatting (see Section D below), the 15-year-old file needed to be reconstituted anyway.

**Nonrespondent data.** In the early follow-up surveys, information on the special sample nonrespondents (e.g., the reason for nonresponse) was not added to the data tapes and was subsequently scratched. Beginning with the 5-year follow-up of the 11th grade cohort, information on special sample nonrespondents has been included in the final data tapes. The information has proved essential to the new procedure for adjusting weights to correct for nonresponse bias (see Chapter 15).

**Mailer tape**

Some follow-up data has gotten lost because it was obtained in connection with efforts to update the mailer file. In this category is some of the information that would cause deletion of a sample member from the mailer tape. When sample members were flagged for deletion the reason (e.g., reported death--surely a relevant item of follow-up information) was not transferred to the data tapes. Also it is not clear to what extent records have been retained, on tape, of old names and addresses after their replacement on the mailer tape by newer ones. The outdated information needs to be preserved for use in follow-back studies (see Section C below); but not all of it has been.

B. **Interim Alphabetic Directory**

Despite the title of this chapter we certainly do not want to give the impression that everything that was done in collection with data processing was done badly. Many things were done well. The alphabetic directory was one of them, even though it does not meet current needs well. The alphabetic directory was prepared about 1963. It is a print-out containing names of members of the TALENT sample in alphabetic order, together with other identifying information, and one of its intended
functions is for use in follow-back studies, which are discussed in paragraph C below.

When the alphabetic directory was prepared in 1963 or thereabouts, the programmer quite properly designated it an "interim directory."
He was aware that the master data files from which he prepared it were at that stage far from complete. It will be recalled that in late 1963 the master file, which now contains our 375,000 cases, contained only about 292,000 cases. The alphabetic directory created then may have contained even fewer probability-sample entries,* certainly it contained no more. But it served its immediate purpose—which was to function as a stopgap.

Unfortunately that "interim" directory is still our only directory, 13 years and probably 90,000 cases later. It is an enormous thing—about 34 side-inches of printout bound in 23 volumes, with inelegant but sturdy and effective covers made out of printout-sized (i.e., 11" x 15") rectangles of the kind of brown cardboard used in the manufacture of ordinary cartons; each volume is held together at the spine by a "home-made" binding of gauze and glue.

Although the alphabetic directory was fine when it was created it now suffers from the following deficiencies:

a. It lacks some unknown number of entries, but probably at least 90,000; possibly more.

b. For the purposes for which it is used, it would be more effective if each entry contained more information about the individual.

c. After 13 years of very hard use its physical condition is deteriorating badly. Parts of it have almost disintegrated.

* It contains "nonprobability cases," as well.
A complete update was planned several years ago to include not only a new alphabetic printout containing more information and formatted for good readability, but also a directory tape containing the same information, and also a printout in numerical order, which we also sometimes need. Unfortunately, overall project cost constraints have precluded updating the directory and preparing a directory tape.

C. The concept of follow-back studies

The follow-back technique is designed for studies in which the interest is in members of the TALENT sample who subsequently became members of some specific group or category—for instance people who went to prison; or graduated from medical school; or became patients in a mental hospital; or played professional football; or won some kind of award; or were elected to state legislatures; or were members of any other delimitable group that might be of interest. The follow-back procedure is simple. The researcher provides a list of the members of the group he is interested in, who happen to be in the right age range to have been in the TALENT sample. He also provides auxiliary identifying information if possible—for instance date of birth, high school attended, maiden name if female, father's first name, etc. If the age range is right, presumably about 4 percent of the group might be in the TALENT sample. A search is therefore made of the TALENT tapes to locate the tape records of anyone on the list provided who also was in TALENT. (The tape search is usually supplemented by a search of the alphabetic directory of TALENT sample members). Data from the TALENT records of the cases identified as members of the group that is of interest in the particular study are then analyzed.

This follow-back technique is a potentially powerful one. But primarily because of the inadequacies of the alphabetic directory it has not worked as well as in theory it should. In some studies only about 1½ percent of the cases in the right age range have been found, instead of the expected 4 percent.
D. Tape Editing and Reformatting System (T.E.A.R.S)

The Tape Editing And Reformatting System was a major effort designed
to resolve many of the problems with the data files referred to above.
The production and documentation of the reformatted files was a signifi-
cant accomplishment and greatly enhanced the usability of the data. The
activity started in 1975 but the idea for it, or to be a bit more precise,
the idea for a major part of it, goes back at least 10 years to 1964 or
1965, when the author of this History, having moved to Pittsburgh and
thus having gotten a little closer to the computer operation, perceived
an urgent need for reorganizing the tapes (see Section A5 above) and
improving the documentation, and began campaigning for it. Subsequent
developments are described below in four phases: (1) the NIE Panel;
(2) the proposal; (3) September 1975 – April 1976; (4) April 1976 – May 1977.

1. The NIE Panel

Since most of the project directors were more concerned with
research issues than with the need for improving the tapes and their
documentation, nothing happened until late in 1973, when NIE brought
in a panel of outside consultants, one of whose functions was to
review Project TALENT's proposal for continued funding (see Chapter 17,
Section C).

Most of the panel members had some prior familiarity with Project
TALENT, and at least one had been a Data Bank client (see Chapter 22).
The panel made a recommendation along the exact lines that Marion
Shaycoft had been urging; they recommended a thoroughgoing reorganization
and documentation of the data tapes. The recommendation seemed to
imply, also, although it did not so state explicitly, an updated
Data Bank handbook was needed.

2. The proposal

As a result of the Panel meeting, the project monitor requested
the author of this History to draft a proposal for tape improvement
and documentation. She immediately (and enthusiastically) complied
with this request. The proposal was apparently tabled for about a
year, but in the spring of 1975 we were asked to submit a formal proposal. The proposal Marion Shaycoft prepared in response to this request organized the work in separate modules with built-in priorities, as shown in Figure 13-1.

A request was received from NIE to cut the scope of the proposed work for the first year, because of budgetary limitations. The modular arrangement made this easy to do. In the revised proposal the following modules were retained:

- #1a. 1960 master file
- #1d. 4% sample
- #1b. Follow-up tapes
- #1c. Retest tape
- #5. Data Bank Handbook

Reorganization, expansion, and documentation of these tapes

Figure 13-1 contains additional details.

The staff had to be cut a bit below the Panel's recommendation of one full-time director and two full-time programmers.

For administrative reasons the proposal was incorporated in a larger proposal that also covered other aspects of Project TALENT. It was funded, with Marion Shaycoft as Director of this part (i.e., the tape improvement and documentation, and the preparation of the new Data Bank handbook), and with Laurie Wise and Wendy Bartlett as her colleagues on this activity. It was at this time that Laurie Wise developed the acronym for the task, T.E.A.R.S.

3. September 1975 - April 1976

Work began on the project September 1975.* Marion Shaycoft drew up plans for formatting which would, in her opinion, meet the necessary standards for good documentation. The principle she followed was that in documenting computer-related information for the use of researchers who are not necessarily programmers, a necessary (but not sufficient) condition for good documentation is that the answer be affirmative to all three of the following questions:

* The Data Bank handbook aspect of the project is discussed in Chapter 22, Section F.
Figure 13-1. Priority chart* for each additional step**

*No step should be undertaken unless every step to which it is connected by a line and which is above it in the chart is also undertaken.

**The step numbers shown in the above chart have the following meanings:

1. Modification and documentation of student data tapes
   a. Master file for 1960
   b. Follow-up tapes
      1) 1-year follow-up tapes (containing 1-year follow-up data and 1960 data)
      2) 5-year follow-up tapes (containing 5-year follow-up data, 1-year follow-up data if available, and 1960 data)
      3) 11-year follow-up tapes (containing 11-year follow-up data, 5-year and 1-year follow-up data if available, and 1960 data)
   c. Retest tape
      This tape should contain both the 1960 master file data for the students when they were in grade 9 and the grade 12 data obtained in 1963.
   d. 4% sample tape

2. Directories and directory tapes
   a. Alphabetic
   b. Other

3. Self-weighted samples for 1960 master file

4. Norms, item response distributions, and reliability coefficients
   a. Compute those that haven't been computed
      1) Percentile norms, with means and standard deviations, for all test variables, a priori composites, and inventory scales
      2) Reliability coefficients
      3) Item response distributions for inventory and questionnaire items
   b. Prepare for publication

5. Revision, expansion, and updating of Data Bank Handbook

6. Modification and documentation of the school data tapes
a. Is the documentation correct?

b. Can it be read easily and understood fully by someone who is adept at reading English but illiterate in computerese?

c. Does it give him or her all the information he or she needs?

The answers to the following questions also have to be affirmative if the documentation is to serve the needs of the programmer:

a. Can the documentation be used easily by a competent programmer who has little or no knowledge of research?

b. Can he use the documentation without having to ask a lot of questions of the person who prepared it?

c. Can he proceed with his work without having to ask the person or persons who supplied the tape that is documented questions about it?

To meet these standards the plan called for documentation in three parts.

Part A was to be as described in Figure 13-2.

Part B was to give the information about what kind of information is on the tape, how it is represented on the tape, and where on the tape it is. More specifically, it was intended to provide full details about what information is in each byte position; exactly what the codes are (both on the source documents and on the tapes; how omitted items (in case of item response data) are handled; what representation is used for missing data; how many decimal places there are; what the possible range of the scale is for any variable; and any references and other explanatory notes that seem necessary or useful.

Several forms, all of them variants of the same general form, were drawn up for use in Part B. The idea was that the particular form used would depend on the general nature of the material to be formatted on that sheet—e.g., a different form would be desirable for questionnaire item responses from the one designed for indicating scores on test variables.
Part A will describe the general characteristics of the tape, including, among other things, such features as:

1. Physical characteristics, e.g.,
   a) seven-track or nine-track
   b) number of bytes per inch (bpi)

2. Physical record length

3. Block size

4. Kind of notation used (e.g., EBCDIC, packed binary, etc.)

5. Cases (or logical records) included in the file — described fully both in terms of inclusion and (insofar as necessary for clarity) exclusion.

For instance the cases included in one of the 13 to 16 tape files to be prepared under the present grant that consist of 1960 master file records might be described as: "All Grade 11 males in the Grade 9-12 probability sample." Another of the 13 to 16 tape files — the one briefly described as the "nonprobability" file — would be more fully described as "all cases included in the 1960 TALENT testing who are not included in any of the following files:

a) the grade 9-12 probability sample file

b) the age 15 probability sample file

c) the Knoxville-and-Knox County saturation sample file

d) the grade 8 semi-probability sample file."

[The term "grade 8 semi-probability sample" will be fully explained in the documentation for that file.]

6. Order in which the cases (or logical records) are arranged. For instance, the arrangement of cases on a tape file might be described as "by student identification number, low to high."

Another arrangement might be described as follows:

<table>
<thead>
<tr>
<th>Level</th>
<th>Characteristic on which arrangement is based</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAJOR CLASSIFICATION</td>
<td>Grade</td>
<td>Low to high</td>
</tr>
<tr>
<td>INTERMEDIATE CLASSIFICATION</td>
<td>Sex</td>
<td>1, 2, Blank</td>
</tr>
<tr>
<td>i.e., classification within major classification</td>
<td>Student Name</td>
<td>Alphabetic</td>
</tr>
<tr>
<td>MINOR CLASSIFICATION</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 13-3a shows a Part B form intended primarily for test scores; figure 13-3b shows a Part B form for other kinds of information including questionnaire items.

Part C of the tape documentation was to consist of whatever appendices are needed to round out and complete the documentation. Appropriate parts of the appendices were to be referenced at appropriate points in Parts A and B. It was anticipated that in both Part A and Part B it would be necessary to add a considerable amount of cross-reference information (to appendices, etc.) for more information about variables, coding systems, specific codes, etc.

The principle was established that instead of spending 8 or 9 minutes planning the tape preparation and 8 or 9 years executing it, as seemed to have happened in the original preparation of the master file tapes,* we would spend at least three-quarters of the available time planning and the remaining quarter executing the plan. This meant among other things, that all the formats would be designed and documentation completed before the new tapes were made.

In December 1975 a 100-page packet of materials, consisting of fairly detailed plans for the tape revision and documentation and a sample format, were sent to three former users of the Data Bank services, whom we had invited to serve as consultants. The three consultants were:

John Hause (University of Minnesota)
Lloyd Humphreys** (University of Illinois)
Christopher Jencks (Harvard)

The reviewers reactions were very favorable.

Things proceeded along those lines until April 1976, when the

* See paragraph A5 above and consider the fact that cases were still being added to the master file about 1969.

** Lloyd Humphreys was unable to review the materials when we sent them, but each of the other two consultants arranged to have someone knowledgeable in addition to himself look at them so that in effect there were four active reviewers.
Director of Project TALENT made some administrative shifts in personnel. To give Marion Shaycoft more time for the other major responsibilities she then had*, and still permit completion of T.E.A.R.S. (the tape improvement and documentation activity and a new Data Bank Handbook) by 30 September 1976, Laurie Wise was made full director for the T.E.A.R.S. effort.


Edited and reformatted tapes for the 12th grade probability cases containing 1960 data and data from all 3 follow-ups were completed in October 1976 along with a draft of the documentation for these tapes. The tapes and their documentation were put to immediate use in a study directed by J.J. Card of the consequences of adolescent childbearing.

Work on the data for remaining cohorts plodded along. As new problems were discovered, investigated and resolved with the testing of each new edit program, it became clear that the required programming time had been significantly underestimated. Consequently, it was decided to focus on completing the tapes of greatest current interest--the 1960 and follow-up survey data on the entire probability sample.

During the course of the project Laurie made several changes in the scope of the work reflecting changes in priority based in part on the results of initial tests. Specific changes in project activities were as follows:

**ADDED TASKS:**

**Test record forms.** A record form was filled out for each of the roughly 13,000 groups tested in 1960 containing information on the test timings, the dates tested, and any group or individual irregularities. Nothing much has ever been done with these forms. Marion Shaycoft had planned to use them to identify cases where testing irregularities might have invalidated some scores. Under her direction a preliminary plan for this was developed. Laurie Wise implemented the plans, and in some ways expanded on them. In order to correct for the irregularities noted and to add test date and group information to the individual records, data from

---

* Writing of this History and completion of a monograph on criterion-referenced tests (Shaycoft, 1976).
these forms were put onto keytape, edited and corrected, and merged with the individual student data. The information from the test record forms provides an important part of the efforts to distinguish zero test scores from missing data.

**Consistent follow-up formats.** The follow-up data were stored on tape in questionnaire item order which varied widely from one grade to the next. The codes used for each variable also varied from one grade to the next in many cases. This meant that when the same analyses were run on different grade cohorts, vast changes in the computer programs were required because of the different formats. In addition, just finding a variable on the tape formats was difficult if you didn’t already know the questionnaire item number.

It was decided that the follow-up data should be rearranged into a common format for all 4 grade cohorts. The common format developed was based on a topical ordering of the follow-up items (within follow-up) so that an item taxonomy was generated as an additional bonus.

**Putting the formats online.** It was decided that, insofar as possible, the new tape documentation should be stored on magnetic tape in a form that could be interpreted by computer programs as well as by people. The formats were streamlined somewhat for this purpose, as shown in Figure 13-4. Laurie indicates that the following advantages have resulted from this decision. The edit/reformat programs used the online formats in defining the input and output structures, saving considerable programmer time and reducing the possibilities for errors. Laurie also says that "The final tape format could be created from the work tape formats with a few simple text editor commands." Work tape formats for Data Bank projects have similarly been created using the text editor, saving the typing or extensive cutting and pasting that was necessary heretofore. The online text was used in generating the tables for Volume 2 of *The American Citizen*, saving enormous time and headaches.
<table>
<thead>
<tr>
<th>E</th>
<th>QUESTIONNAIRE</th>
<th>VARIABLE</th>
<th>TEXT OF QUESTION</th>
<th>NO.</th>
<th>POSSIBLE RANGE</th>
<th>MISSING VALUE</th>
<th>UNITING VALUE</th>
<th>DEC.</th>
<th>FL.</th>
<th>LOW</th>
<th>HIGH</th>
<th>VALUE</th>
<th>CODE</th>
<th>CODE</th>
<th>CODE</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>ITEM NUMBERS AND HELP FOR OPTIONS</td>
<td>ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>FOR EACH GRADE</td>
<td>CODE</td>
<td>TEXT OF OPTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positions</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>TAPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**X. XII EDUCATION**

| 1707-1708 | 1 | 3 | 3 | J | 4 | E442C | I decided to transfer to another kind of program in junior college. | 0 | 1 | b | b | 2221 |
| 1709-1709 | 1 | 4 | 4 | J | 5 | E442D | I decided to transfer to another kind of school (not a junior college). | 0 | 1 | b | b | 2222 |
| 1709-1709 | 1 | 1 | 4 | J | 3 | E442E | I decided that I couldn't get a job in the field even if I completed the program. | 0 | 1 | b | b | 2223 |
| 1709-1709 | 1 | 1 | 1 | J | 6 | E442F | I took a job. | 0 | 1 | b | b | 2224 |
| 1709-1709 | 1 | 1 | 1 | J | 7 | E442G | I couldn't afford to stay in school. | 0 | 1 | b | b | 2225 |
| 1709-1709 | 1 | 1 | 1 | J | 0 | E442H | The training was not satisfactory. | 0 | 1 | b | b | 2226 |
| 1709-1709 | 1 | 9 | 9 | J | 9 | E442I | Other | 0 | 1 | b | b | 2227 |
| 1709-1709 | 1 | 4 | 25 | 4 | 0 | E442 | If the junior college program was intended to prepare you for a specific occupation, what occupation? | 0000 | 0999 | b | b | 2228 |
| 1709-1709 | 1 | 4 | 25 | 4 | 0 | E442 | Did you get a job in the field for which the junior college trained you? (Mark all that apply.) | 0000 | 0999 | b | b | 2229 |
| 1709-1709 | 1 | 4 | 25 | 4 | 0 | E442 | Coding for question E442: | | | | | |
| 1709-1709 | 1 | 4 | 25 | 4 | 0 | E442 | Marked | | | | | |
| 1709-1709 | 1 | 4 | 25 | 4 | 0 | E442 | Not marked | | | | | |
| 1709-1709 | 1 | 1 | 1 | a | 8 | E413A | No, I never looked for a job in that field. | 0 | 1 | b | b | 2230 |
| 1709-1709 | 1 | 1 | 1 | b | 8 | E413B | No, I couldn't find a job in that field and am now in a different field. | 0 | 1 | b | b | 2231 |
| 1000-1000 | 1 | c | c | c | 8 | E413C | No, but I am still trying to. | 0 | 1 | b | b | 2232 |
| 1000-1000 | 1 | d | d | d | 8 | E413D | Yes, and I am still in that field. | 0 | 1 | b | b | 2233 |
| 1000-1000 | 1 | c | c | c | 8 | E413E | Yes, and I have advanced to a higher level. | 0 | 1 | b | b | 2234 |
| 1000-1000 | 1 | f | f | f | 8 | E413F | Yes, but later I changed to a different field. | 0 | 1 | b | b | 2235 |
DELETED TASKS:

Reformatting the 4% sample. The 4% sample, including a separate scanning of Answer Sheets A and C, afforded an excellent opportunity for assessing the reliability of the scanning equipment. Tests run on these records indicated close agreement between the two scanings, except on the highly speeded tests. The data from both scanings were compared to the original answer sheets to determine the source of any discrepancies, but errors were found in roughly equal number for both scanings. The errors resulted chiefly from the machine reading an answer where none of the options were marked. Efforts to use one scanning to flag errors in the other were therefore dropped.

The original plans called for a large number of new variables to be created from the item data on the 4% file. These plans were dropped for lack of funds (although the scoring programs were stored away for future use).

The retest tape. The 1963 retest data are unquestionably valuable for a thorough assessment of changes occurring during high school. Because the cases on the file do not constitute a true probability sample of high school students and because reformatting these data depended on the reformatting of the 1960 data, work on this file was scheduled after the main 1960 and follow-up files, and we never got to it. More will be done with these data as time and money can be scraped together to support such work.

The school data tapes. Work on the school data file was not included in the original grant for tape cleanup, but was included in the grant awarded a year later. An abbreviated school file was created for merging with the student data and considerable work was done in checking and correcting the weights assigned to each school and the test number ranges for each school. These data were merged with the Test Record Form data for each school. Further work was suspended in order to complete the main tape documentation and revisions to the Data Bank Handbook.
With the completion of the recent grant for data tape reformatting and documentation, the TALENT data base is in far better shape than at any time previously. This is not to say that further improvements could not be made. Several editing tasks could not be carried out within the scope of the recent grant but are, nonetheless, important.
Chapter 14. COMPUTERS AND COMPUTER PROGRAMS

Everybody who uses computers has problems with them. Project TALENT is no exception. However, there have been bright spots too. This chapter is about both.

A. Problems with programs

We discovered early that most standard program packages were not well suited to use with large data sets, such as TALENT. When Bill Cooley joined the TALENT staff as Project Director in 1964, he brought with him the set of multivariate analysis programs developed in connection with his textbook (Cooley and Lohnes, 1962). It didn't take Bill long to decide that the programs would not work well on TALENT data. It was clear to him that the best approach would be for Project TALENT to develop its own data analysis programs. Therefore he added an expert programmer, statistician, and mathematician (all combined in one person) to the staff in 1965. That was Bary Wingersky. Bary wrote a very fine set of multivariate analysis programs and other data analysis programs as well. At that time we were using the Pitt facilities, which included two IBM computers--a 7070, which we used for data processing, and a 7090, which we used for data analysis. Thus Bary's programs were designed for use on the 7090.

When we were preparing to move to Palo Alto, where we would be using the IBM 360/50 that AIR was planning to secure for the joint use of Project PLAN and Project TALENT, it was necessary to convert Bary's programs since the 360, it turned out, was not compatible with the 7090. Some of the programs got converted, but unfortunately a great many didn't. Bary left AIR at about the time we were switching from Pittsburgh to Palo Alto, and from a 7070 to a 360/50. Consequently in Palo Alto we lost the use of a great many of the programs that Bary and his staff had developed. Eventually Dick Williams found time to convert some of them, but most of them never were.

Thus in recent years (particularly since about 1970, when Dick Williams left) we seem to have become primarily dependent on program packages again.
Of course the packages have improved since the early years, but there are still many problems associated with their use.

B. Keeping up with technology

Over the years we have used many different computers. As Section A above may suggest some of the transitions (like the switch from the 7090 to the 360) have been traumatic events—even though in most of the transitions the newer computer (e.g., the 360 as a replacement for the 7090) was bigger, better, faster, and more versatile than whatever it replaced. Some of the transitions, on the other hand have been so uneventful that the change was hardly noticed.

As has already been indicated the computer that Pitt had when Project TALENT first needed access was a 650. That was soon replaced by 7070, which was designed primarily for data processing (as opposed to data analysis) and was excellent for the purpose. After a while, Pitt added a 7090, which unlike its stable-mate, the 7070, was intended primarily for data analysis, not data processing. Since the move to Palo Alto we have had the use of a series of 360's and 370's, and we have used the same computer for data processing and data analysis.

AIR no longer has its own computer, having found that rented time on a computer service agency's computer works much better in the long run. Currently we are using an IBM 370/165 and a 370/158.
Chapter 15. STRUGGLING WITH WEIGHTS
(Not a Chapter About Weight-Lifting)

Because not all schools in the sample were selected on the basis of the same sampling ratio, and therefore not all students in the sample had had an equal a priori chance of selection, it is necessary to weight sample members differentially if one wishes to eliminate systematic error, or at least reduce it substantially. However, in so doing one must realize that there is a trade-off. Every time differential weights are used one is increasing the random errors of the type usually called sampling errors. Minor variations in weights will have only a negligible effect on random error, but extreme variations can increase it seriously. Table 15-1, which classifies errors in a two-by-two table, provides a framework that may be helpful. (In this chapter we are concerned exclusively with the two top cells of the table—i.e., with sampling error, not measurement error.)

A. Sets of weights that have been developed

1. Weights for 1960 data (where no follow-up data are being used)

At least six separate sets of weights (Weights A, B, C, D, E, and F) have been developed for use in conjunction with the 1960 data. Weights B, C, and F are intended primarily for use in computing statistics in which the school is the unit of interest, while Weights A, D, and E apply to students.

Weights A and E are almost identical, differing only for students in the two junior high schools referred to in chapter 7 (on page 7-11 immediately above the "Author's Note"). Weight A for students in these schools is the same as the weight for the senior high school selected from the same school district. Weight E is 20 times Weight A for students in these two schools. Weight E gives as close to an unbiased estimate as it is possible to achieve with these data. But the weights for the two junior high schools are enormous, and for most purposes tend to inflate the systematic
TABLE 15-1. Taxonomy and etiology of statistical error*

<table>
<thead>
<tr>
<th>SOURCE OF ERROR</th>
<th>KIND OF ERROR</th>
<th>RANDOM ERROR</th>
<th>SYSTEMATIC ERROR (BIAS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using a sample instead of the entire population</td>
<td>Sampling error**</td>
<td>1. Not having a large sample</td>
<td>a. Selecting the sample in such a way that not all population members have an a priori probability of selection greater than zero</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Selecting the sample in such a way that differential weights must be used. (The more extreme the differences in weights the greater the sampling error.)</td>
<td>b. Not using weights that are the reciprocal of the a priori probability of selection***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>c. Certain sampling techniques (e.g. cluster sampling) may cause systematic error even though the differential weights are mathematically correct****</td>
</tr>
<tr>
<td>Using imperfect measuring instruments</td>
<td>Measurement error</td>
<td>Using instruments that lack perfect reliability</td>
<td>Using instruments that lack perfect validity*****</td>
</tr>
</tbody>
</table>

* It is recognized that this table is to some degree an oversimplification.

** The term "sampling error" is usually reserved for random errors due to sampling. In this table, however, it is convenient to use the term, for the nonce, to include both random and systematic error.

*** In the case of multistage sampling, the probability of selection is the product of the various sampling ratios.

**** However with large samples systematic error from this source is likely to be negligible.

***** Bias in an instrument may be regarded as a special kind of invalidity.
errors far too much to be compensated for by the slight reduction in random error. Therefore Weight E is seldom used. Weight A is almost always preferable even though its failure to weight those two junior high schools precisely results in a little bit of systematic error. Random error with Weight A should be much less than with Weight E.

Weight D is the student weight used to get estimates based on the probability sample of 15-year-olds.

In regard to the weights for use in determining statistics about the schools, Weight B is identical with Weight A except for New York City and Chicago, in which, to compensate for the inclusion of extra schools, only a sample of the students in each school were tested. For schools in these two cities Weight B is much smaller than Weight A. (In New York City, where every secondary school was tested, Weight B equals 1.)

Weight B is 0 for schools that declined to participate in the testing. Weight C is the weight to be used when the statistics are to be based on all schools selected for the sample—not just those that agreed to participate.

Weight F equals Weight B except for the two schools for which Weight E exceeds Weight A; for those two schools Weight F is 20 times Weight B. As in the case of Weights A and E, Weight F will yield essentially unbiased statistics about schools, but with much larger random errors than Weight B.

When used in conjunction with follow-up weights, Weight A is usually referred to as "1960 Weight A".*

2. Weights for follow-up data

The follow-up weights differ somewhat from the 1960 weights because not everyone responds to the follow-up questionnaires.

* In mathematical formulas both the 1960 weights and follow-up weights are represented by \( W \) with a subscript representing time—"5" for 5-year follow-up, "11" for 11-year follow-up, etc.; and another subscript representing type of follow-up weight—e.g., "A" for follow-up weight A, etc. In this context the time code for 1960 is "0". Thus in contexts where five-year follow-up weight A is designated \( W_{5A} \), 1960 Weight A would be designated \( W_0 \).
At least four separate sets of follow-up weights have been developed. They are Weights A, B₁, B₂, and C.

Follow-up Weight A is the same as 1960 Weight A, except for nonrespondents to the mailed questionnaire. For mailed-questionnaire nonrespondents not selected for the special sample follow-up, Weight A is 0. For special-sample members, the 1960 Weight A is multiplied by the reciprocal of the sampling ratio used in selecting the special sample. This applies whether or not the person cooperated in the special-sample survey.

Weight B₁ is the same as Weight A except for members of the special sample. Special-sample members who cooperated have their weights increased by a factor sufficient to compensate for those special-sample members who didn't. Weights for these latter members are reduced to 0.

Weight B₂, like Weight B₁, is 0 for special-sample members who don't cooperate. But to compensate for them, instead of just raising the weights of the other special-sample members, the weights of all respondents, whether to the mailed questionnaire or to the special-sample survey, are increased proportionately. Thus for cooperating members of the special sample Weight B₂ is less than Weight B₁ while the opposite is true of mailed-questionnaire respondents.

Follow-up Weight C* was developed in an effort to improve on follow-up Weights B₁ and B₂, by correcting in a more sophisticated way for noncooperating members of the special sample. Weight C has proved very effective. For most variables it gives far more accurate estimates than any of the other sets of follow-up weights we have tried. Like Weight B₂, Weight C distributes the weight for special-sample noncooperators among the individuals for whom questionnaire responses are available, whether via the mailed questionnaire or the special sample. Weight C is described in Appendix E2.

* Developed by the author of this History.
B. Whether and how to weight

With the multiplicity of weights described above, how does one know which one to use for a particular purpose? In fact, how does one even know whether to use any? Should one be using unweighted data instead?

There aren't any good rules of thumb for this last question. As has already been implied one has to consider the purpose of the research, and the trade-off between systematic and random error.

In general, however, weighting is essential when one is interested in normative or distributional data; in other words, when one wishes to be able to make statements of the following sort:

The average spelling score for 12th grade students who later majored in English in college was ________.

The 84th percentile for 9th-grade girls on the Scientific Attitude Scale in 1960 was ________.

Eleven years after high school the average salary of members of the 10th-grade cohort who had become teachers was ________.

Generally speaking, for data analyses of this type, 1960 Weight A would be the preferred weight if the entire 1960 sample for the segment of interest (e.g., 10th-grade girls) or a representative subsample of it is being used. If the analysis is limited to cases with follow-up data, follow-up Weight C is usually the best one to use (at least in the opinion of this author). Note that 1960 Weight A is to be preferred to 1960 Weight E even though Weight A does not yield quite as "pure" results (i.e., results free from systematic error). That minor disadvantage is more than cancelled by the reduction in random sample errors.

But if one is interested in being able to make statements of a relational or predictive rather than normative type, it is the opinion
of this author* that on the whole one is better off not using weights at all. Weighting the cases differentially will not have much effect on correlation coefficients except to make them less stable—-in other words, to increase the random error. And if a statement of the following type,

"Students with high math scores but low scores on Mechanical Information and Mechanical-Technical Interest are more likely to be successful and happy in (Occupation A) than in (Occupation B)."

were to fall to pieces just because it was based on a sample that wasn't perfectly representative of the national population, it wouldn't be worth much to begin with.

C. Self-weighting: A solution to the trade-off dilemma

In Section B above, there was some discussion of the trade-off between not weighting—-in other words, using unit weights—-and using differential weights. Unit weighting has the advantage of reducing sampling errors but it may introduce significant systematic errors. The solution to the problem was initiated by John Flanagan when he proposed that we develop "self-weighted subsamples" of the TALENT sample. These would be subsamples selected in such a way that each case in the TALENT sample would have a probability of selection proportional to its 1960 Weight A. (For a self-weighted sample of the cases with follow-up data, presumably follow-up Weight C would be used rather than Weight A.) The advantage of such a subsample would be that the optimum weights would be unit weights. In other words, the subsample would be "self-weighted."

Marion Shaycroft developed an algorithm for selecting a subsample that would meet this requirement and that except for the differential-probabilities feature would be a simple random sample. Don McLaughlin

* I feel obliged to point out that not everyone agrees. Some very knowledgeable people have a contrary opinion on this point.

—Author
proposed an alternative algorithm that would yield a systematic (rather than simple random) subsample having the self-weighting feature. It was unanimously agreed (by Flanagan, McLaughlin, and Shaycoft) that Don's approach would have major advantages. Consequently it has been used frequently since then. It was, incidentally, the procedure John Flanagan would later use in selecting a 1000-case sample from the probability sample of 15-year-olds to interview 15 years later for a quality-of-life study of 30-year-olds. (See Chapter 21, Section E)

The same approach was utilized in a slightly modified form (i.e., with stratification) to select the special sample of nonrespondents to the 11-year follow-up of the 9th-grade cohort.
Chapter 16. OTHER STATISTICAL PROBLEMS ENCOUNTERED AND SOLVED

The title of this chapter is not meant to imply that we have solved every statistical problem that we have encountered. What it means, instead, is that we are judiciously limiting this chapter to the problems we have solved. In accordance with the HIBK syndrome (see Preface) the solution to a few of the problems, unfortunately, came a bit too late to use in the present project. Ah, but in Son-of-TALENT....

Since we were discussing sampling in the last chapter, we shall start with that topic.

A. Sampling

1. How to handle schools that are in the project on a specially invited or volunteer basis

As discussed in Chapter 7 Section B, quite a few schools participated in the project that hadn't been in the initially selected sample. Some were invited by us (for instance, all the schools in Knoxville that hadn't originally been selected) and some invited themselves. We treated the students in all these schools as not being members of the probability sample. This was technically sound but was it really necessary to lose all the good data from these schools when doing analyses requiring a probability sample? The answer is that a method was developed by the author of this History, a couple of years ago, in connection with an activity unrelated to Project TALENT, whereby it would have been possible to treat any of these "non-probability-sample schools" that had been in the original universe (i.e., all secondary schools in the United States) as members of the probability sample, through appropriate adjusting of the weights. The method is described in Appendix D1.
2. The "feeder junior high school" problem

As discussed in Chapter 7 Section B (page 7-9), a special procedure has now been developed to cope with the problem of selecting the "feeder" junior high schools for the sample. The procedure, which was developed far too recently to have been available for use in selecting the TALENT sample in 1959, is called the "closed school-cluster procedure." It is described in Appendix D2.

3. Coping with complexly structured samples in computing standard errors

Although this problem is concerned primarily with computation, it is put under the sampling heading because it has implications for sampling. (It involves picking two wholly parallel subsamples.) Like the procedures discussed in paragraphs 1 and 2 above, this one has not been tried yet. It is described in Appendix D3.

4. Is the Special Sample of nonrespondents necessary? (An empirical validation of follow-up Weight C)

Collecting data from the special sample of nonrespondents is extremely costly. Therefore there has been serious consideration at one time or another of whether we really needed it; there have also been at least two empirical studies.

The first was an APA paper presented by Marion Shaycroft in 1964 (reprinted as Appendix E1), called "When is Bias Better?" This study was an effort to determine whether by stratifying on 1960 data and weighting the questionnaire respondents in each stratum to also represent the nonrespondents it would be possible to dispense with the special sample. The answer was a qualified "no."

* It was not an unqualified "no," because the study was somewhat restricted by the fact that the range of computer programs available to us was very limited and did not permit stratification on more than one variable.
The second empirical study turned out to be a validation of follow-up Weight C (described briefly in Chapter 15 Section A2). This was a study by Don McLaughlin and Wendy Yen, in which statistics obtained by using Weight C were compared with corresponding statistics computed using respondents only, but adjusting their weights. Weight C turned out to be superior to any of the other procedures used. The study is reported in "Project TALENT's Special Sample: Is It Necessary?" (McLaughlin, Fulscher, and Yen, 1974).

5. Can the selection of the Special Sample be improved?

As indicated in Section 4 above, empirical evidence suggested that use of a special sample should be continued. However a substantial improvement has been introduced recently by using a modified self-weighting scheme with stratification on variables of particular interest and judicious oversampling of certain cells.

B. Reducing sampling error—both random and systematic

As has been implied at several points, use of a special sample of nonrespondents has tended to increase systematic error, though it reduces random error. Efforts have been more or less continuous to find ways to moderate the increase in random error and thus decrease in total error.

The development of Weight C (see Chapter 15, section A2) was one such effort. Use of a self-weighted special sample (as in the 11-year follow-up of the grade 9 cohort) was another.

A third was initiated in 1975, when Laurie Wise and Don McLaughlin, in an in-house memo, proposed a formula for getting a weighted mean of data from mailed-questionnaire respondents and special-sample respondents, with the weights for the two components being different from those which, like follow-up Weight C, are intended to get as close to an unbiased estimate as possible. The goal of the McLaughlin-Wise procedure was to minimize total error. But their formula for determining the weights
for the two components was a preliminary one, requiring many assumptions, some of which seemed dubious.

Starting with their idea, the author of this History developed a set of formulas and procedures, applicable under various conditions and not (we hope) requiring untenable assumptions. The outcome was a paper entitled "Minimizing Total Error: A Trade-Off Problem" (Shaycoft, Wise, & McLaughlin, 1975), which is reproduced as Appendix E3.

C. Reliability coefficients and related statistics

1. Reliability coefficients

Parallel form reliability coefficients were not feasible for the TALENT tests because there were no parallel forms. Retest correlations are unsuitable for use as reliability coefficients for cognitive variables. That meant we were limited to some sort of internal consistency reliability coefficient—either split-half or one of the KR coefficients. It is our judgment that for most of the tests in the TALENT battery, split-half coefficients with a priori splits are preferable to KR-20 or KR-21. In the first place K-R coefficients would tend to systematically overestimate the reliability of certain tests such as Reading Comprehension for which split-half coefficients with an appropriate item split would not. Furthermore, formulas such as KR-20 are based on the assumption that all the items are fungible units. That assumption may be valid for certain scales in the battery, such as Visualization in Two Dimensions (for which the reliability cannot be determined anyhow, because it is a somewhat speeded test and therefore a parallel form or separate administrations would be needed) but it certainly isn't true for the information scales and achievement tests, for which, it will be recalled from Chapter 5, very detailed rationales were prepared in advance, which spelled out the proposed distribution of item content. Therefore the method of choice for determining the reliability of the TALENT tests is split-half, using a priori splits where appropriate, with Angoff formula 16 (Angoff, 1953) as the correction formula. This has limited our data sources
to the 4% sample (see Chapter 12) which provides suitable data for all variables except, of course, those that are substantially speeded and those few for which there is some other factor precluding reliability determination (see Shaycoft, 1967, Chapter 4, pp. 4-12 through 4-15) or to the 1963 retest data (see Chapter 11) for the variables for which grade 12 data are available.

2. Coefficients of internal precision

Around 1968, Lyle Schoenfeldt was working on a study of twins in the TALENT sample. One day he and I were discussing some strange results he was getting, which neither he nor I understood. When he tried to apply a conventional correction for attenuation to the correlation between twins' scores he got suspiciously high correlations; some of them were over 1. Such results are possible only when one is using as the divisor an underestimate of the geometric mean of the two reliability coefficients. But we had no reason to think that our reliability coefficients were systematic underestimates. The explanation suddenly occurred to me, and a lot of other things fell into place. The explanation lay in the fact that when the correlation is between two measures obtained with the identical test the appropriate coefficient to use in correction for attenuation is not the reliability coefficient, because using the reliability coefficient means that a correction is being applied for, among other things, the random effects of having a different sample of items. This is an overcorrection when one has the same items in the two variables correlated. In a Eureka-type of reaction, it occurred to me that the same explanation would account for some inexplicably high corrected-for-attenuation correlations, many of them far in excess of 1, that I had obtained in the retest study (Shaycoft, 1967, Table 5-4, columns 3 and 4), and which I had found baffling at the time. The correlations were between the grade 9 and grade 12 scores of the same student on the same test.

The solution to the problem of these excessively high "corrected" correlations—not only the ones Lyle had obtained in his research on twins and those I had obtained in research involving retesting with the same test, but probably others as well—was to derive a formula
for a coefficient that would be like a reliability coefficient except that it would not include the concept of sampling errors in selecting the particular item sample as a source of error. Hence, the coefficient of internal precision! (See Appendix C.) These coefficients have been computed for a limited number of tests in the TALENT battery, but eventually they should be computed once and for all, for all of the tests for which they are suitable. This would make them available to any researcher using the TALENT data, since the same correction-for-range formula applies as for reliability coefficients, thus making it possible to adapt the "once and for all" coefficients to the particular sample of students being used.

D. Inferring score distributions for an unselected age group

In 1963 we had a contract with the Air Force, which called for inferring score distributions for the total population of 18-year-olds in the United States. The solution, which is presented in Appendix F, would presumably be generalizable to other age groups (such as age 17).

E. Problems in methodology related to career guidance

Appendix G consists of papers by the author of this History on two problems of a methodological nature, to which approaches that are believed to be largely new were developed in the course of Project TALENT.

The first is a new approach to the use of a battery such as TALENT in career guidance; it is called propinquity analysis. See Appendix G1; also for further information, see Chapter 7 and Appendix N of Five Years After High School (Flanagan, Shaycoft, Richards, and Claudy, 1971).

The second is a suggestion on the grouping of career categories. (See Appendix G2). Attempts at hierarchical analysis have led to the conclusion that it isn't a particularly useful approach for the TALENT data. (Flanagan, Shaycoft, Richards, and Claudy, 1971, Chapter 5.)
Chapter 17. MANAGEMENT PROBLEMS

Whatever problems are routine in almost any research management situation are greatly magnified in a longitudinal study. And if it is a large-scale longitudinal study, the problems multiply.

This chapter is concerned not with management problems in general but with those problems that apply particularly to the management of a large-scale longitudinal study.

A. Record-keeping problems

1. Identification numbers and other codes

   In a longitudinal study it is completely impractical to rely on students' names or school names or almost anything else that is spelled out in words, if one wishes to maintain continuity— for instance to be able to match new follow-up data with the corresponding set of earlier data. Some sort of identification code is necessary—either a serial number or some kind of segmented number where each segment has a meaning, or perhaps a code that is a combination of numbers and letters. (For computer use all-number codes are usually somewhat preferable to number-and-letter combinations.)

2. Some comments about redundancy

   Whatever the code system is, one has to bear in mind the fact that everyone makes errors sometimes and some people make errors very frequently, so that matching records on the basis of identification number is not error-proof. Someone may have made an error in recording a number or in keypunching or keytaping the number—or if the number is position-coded on a machine-scannable form, the scanner may fail to pick it up completely. The same caveats apply not just to identification numbers but to other important categorizing or identifying information.
Therefore it is important to build enough redundancy into the system to make it possible to cope with error—both human error and machine error.

In Project TALENT lots of redundancy was built into the system. In addition to the six-digit identification number, each answer sheet contained either the student's full name, or his last name and his initials, or, at the very least, the first several letters of his last name. The student's sex was position-coded on two of the five answer sheets, as was his grade. In addition each set of answer sheets was preceded by a "Group Identification Sheet" filled out by the teacher or other examiner, and containing in position-coded form the grade as well as other useful information. The answer sheet form containing the student's full name, address, and other basic information contains in position-coded form not only his date of birth (month, day, year) but also his age in years at his last birthday. This particular bit of redundancy seemed essential even though we made it impossible for these spring 1960 examinees to record their year of birth as 1960. (We prevented that particular error by not providing a position for position-coding '60.)

But it seems that the more carefully one builds redundancy into the system, the more likely it is that some well-meaning person will come along and wipe it out before anyone gets a chance to take advantage of it (see Chapter 13). He will wipe it out because somewhere along the way he has been taught that redundancy is a bad thing.

Therefore our advice to future longitudinal researchers who want to profit from our experience is that by all means they should build plenty of redundancy into the most important parts of their record-keeping systems, but that they should also make certain that anyone who is in a position to wipe out the redundancy prematurely is informed that it is there, and is not there by accident, and they should also make certain that he understands what it is there for. (This underlines the importance of good two-way com-
3. Confidentiality

In releasing research results or individual data of any kind, it is important to maintain commitments on confidentiality and not infringe on sample members' rights of privacy. This requires vigilance. Routine deletion of names and dates of birth from released information is not sufficient. It is necessary to make certain that identity cannot be inferred through the use of mathematical logic. For instance in one system of classifying states into regions the Far West might consist of Washington, Oregon, and California, while in another system the Far West would be defined as Washington, Oregon, California, and Nevada. If the same data were released by region both ways, it would be easy to determine the values for Nevada by subtraction. Researchers and research administrators have to be alert to this sort of possibility, and keep it from happening.

4. Retrievability

In a longitudinal study the need sometimes arises to correct an error in an individual record or to update the record without waiting until there are a hundred other records to correct or update at the same time. Or the need arises to get a printout of the record for a single case. A data processing system that permits handling such situations efficiently and economically is worth developing.

B. Securing and maintaining cooperation

Project TALENT has developed certain procedures and techniques to help keep members of the sample interested and cooperative. These include such devices as an ID card identifying the person as a member of the TALENT sample (see Chapter 7 section C); and the Project TALENT News, a 4-page newspaper mailed to the entire sample once a year (see Chapter 18, section E1).
In order to be able to keep sample members interested and cooperative, one has to know how to contact them. In other words, up-to-date name-and-address files are essential. One thing this means is that about once a year (or at least once every year-and-a-half) something—a postcard, a form letter (preferably, though, one that doesn't look too much like a form letter), a newsletter or newspaper, a questionnaire, something—should be mailed to everyone in the sample, with "Address Correction Requested" on the envelope. That is an excellent way to help keep address files up-to-date. The reason it should be done once every year or year-and-a-half is that postal regulation require post offices to keep address changes on file for two years; after that they throw them away. It is worthwhile to assign someone on the staff responsibility for keeping informed about changes in postal regulations. There is nothing immutable about those regulations—and changes one is unaware of or changes one finds out about too late can raise the project's mailing costs by thousands of dollars, or can result in having to drop thousands of entries from the project's mailer tape because they have turned into dead-end addresses.

It is also worthwhile to do a little experimenting on what mailing "styles" are most likely to result in cooperation. For instance, if one is mailing a questionnaire first-class, does a postage stamp have any advantage over metered postage? Does a commemorative stamp have any advantage over an ordinary postage stamp? Some evidence was obtained years ago that the answer to both of these questions was affirmative. Whether it still is isn't entirely clear, since collective moods change with the times and so do prevalent mailing procedures. (Metering is used much more frequently for first-class mail than it used to be, so the sight of an envelope with metered postage stamped on it isn't as likely to stimulate the mental response "junk mail" or the reflex action of tossing the thing in the circular file.) As a matter of fact, the entire problem of when it pays to mail something first-class that could legally go by bulk rate has to be considered. As postal rates keep rising it becomes more and more important to send mail by the most economical way that will accomplish whatever you are trying to accomplish through that particular mailing. And it is probably not feasible to develop a "once-and-for-all" answer, since not only absolute costs but the relative costs of first-class and bulk-rate mailing keep changing.
Furthermore before sending something out by first-class mail that could go by bulk-rate, you should know just what it is that you are paying extra for. Are you buying the free forwarding? a favorable impression on the recipient? more careful handling by the post office? faster delivery? Whatever benefits you expect first-class mail to deliver, it is important to be aware of your expectations so that you can take steps to find out whether they were fulfilled, and perhaps learn from experience.

Closely related to the question of how the researcher mails things out is how he plans to get them back. Should he enclose a self-addressed "business reply mail" envelope with a first class mailing permit? Or should the return envelope have a commemorative postage stamp affixed to it? The mailing permit means there will be no charge for return postage in the case of nonrespondents—but for each respondent it will cost more than a postage stamp would have. And furthermore, there is evidence that having a stamp already attached to a return envelope increases the likelihood of response. (Of course the extent to which this is true must depend, to at least some extent, on the nature of the group on which one is doing the longitudinal study.)

The next chapter contains some further discussion of Project TALENT's mailing procedures for follow-up questionnaires. The reader is also referred to "Methodology of the Project TALENT 11-Year Follow-Up Study," (Rossi, Wise, Williams, and Carrel, 1976.)

C. Getting funding (the cliff-hanger serial)

In Chapter 2 (Sections B and D in particular) the problems of getting funding to get the Project started were discussed. In April 1959 (just a little more than two years after the idea first began to take shape) a $480,000 grant was received, to get the project going. But obviously that $480,000 was not going to put a permanent end to the funding problem. On a project that it was anticipated wouldn't end until, perhaps, about 1985 (2 years after the 20-year follow-up of the high school class of 1963, in other words the grade 9 cohort), it was clear that getting funding would be a recurrent problem—indeed, a nearly continuous problem. At various times the project's primary funding agency (OE initially, then NIE) has renewed the funding, sometimes with a five-year grant, sometimes with
a two or two-and-a-half year grant, sometimes with just a one-year grant. (What the length of the grant is probably depends more on what the fiscal situation, the prevailing administrative regulations, and the general mood are at the funding agency at grant-renewal time than it does on any specific formal steps that AIR can take (such as writing a 5-year proposal versus a one-year proposal).

Of course the best insurance on getting the funding continued is doing as good a job as possible on the study—but of course even that provides no sure guarantee against loss of funding.

One way or another, the project has had continuous funding since 1957—but on several occasions there have been delays of many months in our receiving the signed document, so that we have been forced to mark time, rather than going full speed ahead with our plans.

The most troublesome delay occurred when we were completing the 5-year follow-up and were seeking funding for the 10-year follow-up, which should have started in 1970. At that point it wasn't clear where our funding was going to come from. The feeling in certain quarters at OE was that Project TALENT should certainly be continued, but since the sample members would then be about 28 years old, and thus well out of the age range that was of most concern to the Office of Education, it wasn't clear that OE should continue to be the prime funding agency. What about the Department of Labor? How about the Department of Defense? Any why not National Science Foundation? (We duly explored all these leads.) The part of OE that had in recent years been providing the support for Project TALENT was the National Center for Educational Research and Development (NCERD). Most of the people we had dealings with at NCERD, including the monitors (Sue Klein and then Larry Goebel) understood the need for continuing the project and were strong supporters. But since money was very tight at NCERD it was suggested that even if OE were to provide the main support for the project perhaps the funding could be a joint venture of several parts of OE, rather than coming exclusively from NCERD. It was suggested that we explore this possibility with several other parts of OE—for instance the Bureau of Education for the Handicapped (BEH), the Bureau of Elementary
and Secondary Education (BESE), the Bureau of Higher Education (BHE),
the Bureau of Educational Personnel Development (BEPD), and the Bureau
of Adult Vocational and Technical Education (BAVTE).

We conferred with representatives (usually the Director) of all these
groups and submitted special "sub-proposals" showing how Project TALENT
data could fit in with their particular needs and interests. The persons
with whom we conferred were all quite interested, and they recognized how
Project TALENT data might be useful to their agencies; but unfortunately
most of them had no unbudgeted discretionary funds available. The monitor
(Larry Goebel) also explored with OE agencies other than NCERD the possi-
bility that they might provide supplementary support for Project TALENT.
Unfortunately his efforts were no more successful than ours. It soon
became apparent to us that if OE was going to provide support for the
10-year follow-up, the part of OE that was going to provide that support
would have to be NCERD. That became apparent to NCERD, too. But since
the fiscal situation was too tight to enable them to manage it with funds
available in FY '71 (the fiscal year running from July 1970 through June
1971) it was necessary to wait until the FY '72 funds became available.
And that is why we had an 11-year follow-up instead of a 10-year follow-
up! In the intervening year it was possible to keep the project in
existence, and thus maintain staff continuity, via time extensions on the
previous grant, from which unexpended funds still remained.

At just about the time that NCERD was arranging to fund the 11-year
follow-up, one of our efforts to get supplementary funding from another
agency finally proved successful; we received word that NSF was going to
supplement the OE grant (approximately $1,300,000 for the first 30 months)
with a very welcome additional $75,000. (NSF arranged, via an interagency
agreement, to have their grant administered by OE.)

The next major episode of uncertainty and delay (though not as long
a delay) came when the grant for the first 30 months of the 11-year follow-
up was about to expire. By then NIE was in existence and Project TALENT
had been transferred to it from OE. NIE had appointed a panel on longi-
tudinal research studies and methodology. One of the functions of this
panel was to serve as a Review Panel for Project TALENT's proposal for
continued funding. A meeting was held in Washington on 18 October 1973, at which TALENT staff members presented reports, and the committee made recommendations. The meeting was attended by:

Panel Members
Jerry Bachman
Robert Crain
Christopher Jencks
Ned Gramlick
Herbert Parnes
Michael Timpane

NIE staff
Garry McDaniels (chairman)
Larry Goebel (TALENT monitor)
Andrew Porter
Marshall Smith

TALENT staff
John Flanagan
Bill Clemans*
John Claudy
Don McLaughlin
Marion Shaycroft

The Panel made several recommendations. The most important one (at least from our viewpoint) was that funding be continued; however they recommended that it be earmarked primarily for data collection and processing, with only a comparatively small amount being allocated for data analysis. Most of the rest of their recommendations reflected concerns about the data tapes, with reference particularly to their organization, documentation, and availability—in other words many of the same concerns that a few of the staff members had already had and that are touched on in Chapter 13.

* Bill, at that time, had agreed to become TALENT's next Project Director, but he had not yet joined the staff.
Specifically they recommended that data tapes should be made available to researchers on request, and that AIR prepare a proposal for revamping the tapes and preparing thorough documentation. (They suggested that a full-time project director and two programmers could complete this work in a year--an entirely reasonable estimate, in the opinion of the author of this History.)

About a month after the October 18 meeting Marion Shaycoft was pleased to receive a request from the TALENT monitor, Larry Goebel, to draft a proposal for getting the tapes and their documentation in good shape. She did this immediately, incorporating in the proposal the recommendation that a new Data Bank handbook be prepared--along with recommendations for a comprehensive overhauling of the tapes and the tape-format documentation. (This proposal was finally funded about a year-and-a-half later. For details about what happened next, go back to Chapter 13.)

A grant was received for a year's work starting October 1976, to develop plans and instruments for a 17-year follow-up. That work is now in progress, and hopes are high that funds will be received for implementing the plans--in other words for actually carrying out the follow-up. Tune in next week.

Total funding received thus far for "Mainstream Project TALENT" (i.e., the basic funding to keep the project going) is a little over 7 million dollars. Other TALENT-related studies done by AIR (exclusive of "Data Bank studies") total a little over a million dollars, bringing the combined total to something over 8 million dollars. Table 17-1 shows the breakdown.

D. Problems related to staffing

In addition to the problems touched on in this section the problems discussed in Section E, which follows, also have some implications for personnel management.
TABLE 17-1. TALENT Funding* for first 20 years (1957-1976)**

<table>
<thead>
<tr>
<th>Funding agency</th>
<th>&quot;Mainstream TALENT&quot;</th>
<th>Other TALENT-related studies***</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>OE</td>
<td>$4,833,000</td>
<td>$30,000</td>
<td>$4,863,000</td>
</tr>
<tr>
<td>NIE</td>
<td>1,920,000</td>
<td>211,000</td>
<td>2,131,000</td>
</tr>
<tr>
<td>NIMH</td>
<td>116,000</td>
<td>148,000</td>
<td>264,000</td>
</tr>
<tr>
<td>NIAAA</td>
<td>-</td>
<td>33,000</td>
<td>33,000</td>
</tr>
<tr>
<td>NICHD</td>
<td>137,000</td>
<td>250,000</td>
<td>387,000</td>
</tr>
<tr>
<td>Welfare Office****</td>
<td>-</td>
<td>67,000</td>
<td>67,000</td>
</tr>
<tr>
<td>NSF</td>
<td>90,000</td>
<td>166,000</td>
<td>256,000</td>
</tr>
<tr>
<td>ONR</td>
<td>35,000</td>
<td>31,000</td>
<td>66,000</td>
</tr>
<tr>
<td>AFPRTC</td>
<td>-</td>
<td>94,000</td>
<td>94,000</td>
</tr>
<tr>
<td>ARI</td>
<td>-</td>
<td>96,000</td>
<td>96,000</td>
</tr>
<tr>
<td>Other (approx.)</td>
<td>-</td>
<td>100,000</td>
<td>100,000</td>
</tr>
<tr>
<td>TOTAL</td>
<td>$7,131,000</td>
<td>$1,226,000</td>
<td>$8,357,000</td>
</tr>
</tbody>
</table>

* To the nearest thousand dollars

** As of December 1976

*** Excluding "Data Bank studies". (See Chapter 22.)

**** Part of HEW
1. Turnover

Obviously the longer a study lasts the greater the turnover and the problems associated with turnover. Project TALENT has had a lot of turnover as can be seen from a perusal of Chapter 8—and particularly from a look at the number of columns in Figure 8-2. But the fact that the Principal Investigator has been a constant has helped to keep to a minimum the problems that having half a dozen Project Directors would otherwise have caused.

2. Personnel selection

In this section I do not intend to talk about the importance of careful selection of staff in general because although that is obviously an important matter it is no more important for a longitudinal study than for any other kind of research.

Instead I shall focus on the lowly coding clerk—because there has sometimes been a tendency to underestimate the value of their work and its long-lasting effects. A researcher who wishes to do a study comparing career plans one year after high school with the same groups' career plans 11 years after high school is dependent for the one-year follow-up career plans on the work done by coders 12 to 15 years ago; and for the 11-year follow-up codes, on the work of coders 2 to 5 years ago. The two sets of coders are obviously different sets of people (that is the nature of the job), but it is desirable to be able to have confidence that they followed the same general principles—and that they did their job well. The coding is not a strictly routine job; doing it right requires intelligence and judgment, as well as accuracy. Moreover as the nature of the work load shifts from month to month, at some stages calling for questionnaire check-in, then coding, and then, months later, perhaps telephone interviewing of nonrespondents, obviously members of the temporary "coding staff" have to have the flexibility to cope with these shifts in assignment.
Therefore the more attention is paid to ability in selecting the coders the better the results are. At times there has been a tendency among those who do the hiring, and among the immediate and not-so-immediate supervisors of those who do the hiring, to overlook this. There are very large individual differences in applicants in terms of the specific abilities required; and these differences cannot be determined by looking at an application blank or conducting an interview.

E. Management problems related to computer and computer-related operations

1. Management of keypunching operations

The author of this History takes a dim view of large-scale keypunching operations if there is any effective and economical way of bypassing them. But sometimes there isn't. For instance in the one-year and five-year follow-ups, we had the enormous task of keypunching all questionnaire responses. (As is explained in Chapter 18, Section B 1, by the time of the 11-year follow-up we were able to switch to an optically scannable questionnaire form, but this option wasn't available to us on any practical basis at the time of the one-year and five-year follow-ups.)

If one can't avoid undertaking a large keypunching task the next question concerns who should do it. Should the organization (e.g., AIR) have its own keypunches and its own keypunching staff, or should the task be farmed out, to an agency whose primary function is keypunching? We have done it both ways. When Project TALENT was in Washington we had our own keypunches and keypunch operators. (That was during the one-year follow-up of the grade 12, grade 11, and grade 10 cohorts.) In Pittsburgh, and subsequently in Palo Alto, on the other hand, we farmed the keypunching out (to a different agency almost every year).

Most service agencies seem unwilling to guarantee that the error rate will be smaller than 1%, in terms of cards containing errors. The rate is much lower in terms of percentage of keystrokes that are wrong, but this is irrelevant since a single keystroke error in certain positions can make the entire card (or tape record) worthless.
There is no doubt in the author's mind that doing the keypunching right on the premises is by far the better procedure. One is able to set, and maintain, much higher standards of accuracy when the operation is under one's control than most of the outside agencies even attempt to meet (regardless of their claims). The costs of the two approaches are roughly comparable, and the greater accuracy of the cards produced in an in-house operation results in savings in subsequent processing--to say nothing of the research advantages of having significantly more accurate data.

The purpose of this discussion has been to give future researchers who have to decide whether to farm out keypunching or do it themselves the benefit of Project TALENT's experience (or at least the author's interpretation of that experience).

2. Coordinating between the research staff and the computer staff

As anyone who has read Chapter 13 will probably realize, a great many (though by no means all) of Project TALENT's problems with the data tapes had their start when there were 250 miles separating the computer staff (in Pittsburgh) from the rest of the project staff (in Washington). In a study such as Project TALENT, good two-way communication between the research staff and the computer staff is of primary importance. It is also desirable that the programmers, and particularly the programming director, be recognized as central members of the project staff rather than just being on the fringe of it. (What the situation has been in Project TALENT in this respect has varied from time to time, and the results have varied accordingly.) And as has been mentioned in previous chapters the computer and computer staff should not be 250 miles away!
3. **Backup tapes**

Having a duplicate set of tapes to serve as backup in case anything happens to the primary set is absolutely crucial in a project such as TALENT. Without such backup major subsets of the TALENT data would have been accidentally, but permanently, destroyed years ago.

F. **Pitt leaves TALENT before TALENT leaves Pittsburgh!**

It will be recalled from Chapter 2 that because a Federal law limited OE grants under the Cooperative Research Program to universities and state education departments the Project TALENT grant went to the University of Pittsburgh, not to AIR. Although this arrangement had advantages for the project (e.g., free computer time at first, then computer time at extremely low rates) it also created problems (for the project staff, and probably for the University's administrative staff as well). Therefore we were all pleased when the law was changed in such a way that the Project TALENT grant could be transferred to AIR. The three concerned parties—OE, Pitt, and AIR—jointly agreed to the transfer, which was to be effective January 1, 1967. The details were spelled out in a document called a "Novation Agreement", which represented a tri-partite agreement (Pitt, AIR, and the Unites States of America). In return for turning over administration of the project to AIR, Pitt requested, and received from AIR, a commitment that for the duration of the grant then in effect, any reports produced by Project TALENT would indicate that they represented the joint efforts of the American Institutes for Research and the University of Pittsburgh. And that is why our report entitled "Five Years after High School," (Flanagan, Shaycoft, Richards, and Claudy, 1971) bears on its cover the words "American Institutes for Research and University of Pittsburgh."
Chapter 18. THE FOLLOW-UPS

The original plan called for follow-ups one, five, ten, and twenty years after the grade cohort was scheduled to graduate from high school. In this chapter we shall talk about the procedures that have been used to execute this plan (and about changes that have been made in the plan itself).

A. Mailing of the questionnaires

Basically our follow-up approach has been to use a mailed questionnaire. The procedure has involved four "mailing waves." The first wave goes out in the early fall, typically at the end of September or the beginning of October; then about a week later a reminder postcard is mailed; then, about 2 weeks after that, the second wave of questionnaires is mailed to those from whom responses have not been received from the first wave. Then about 7 weeks later the third wave is sent to those who still haven't responded; the fourth wave is sent after another 2 months or more, to the "hard-core nonrespondents". Questionnaires in the four waves are identical, except that they carry an inconspicuous code indicating which wave they are, and each wave is usually printed in a different color; also if we are using an optically scannable form (as we have been doing on everything subsequent to the five-year follow-up) each questionnaire carries a set of machine-readable "skunk marks"—a series of broad black bands separated by white space—to designate the mailing wave.
Each questionnaire was accompanied by a letter—with different letters for different wave.

Further details about the mailing are given in various TALENT reports (e.g., Rossi, Wise, Williams, & Carrel, 1976, Chapter 2; Flanagan, Shaycoft, Richards, & Claudy, 1971, Chapter 1; Flanagan and Cooley, 1966, Chapter 2).

B. Processing the returned questionnaires

1. Keypunching vs. optical scanning

Our early follow-ups (one-year and five-year) were not on optically scannable forms. This meant that instead of having the completed questionnaire forms run through an optical-scanning machine that would convert the responses (and position-coded identifying information) directly to magnetic tape, every questionnaire had to be keypunched. Some years three IBM cards were required per questionnaire, and some years four. The first card was always a check-in card, which was used to identify respondents to earlier waves, so that their names could be eliminated from the mailings of subsequent waves. Subsequent cards contained item responses.

Those who have read Chapter 7 Section D can readily (and correctly) infer that the author of this History takes a dim view of any large-scale keypunching operation if there is any reasonable way of bypassing it in getting from document to magnetic tape. Thus it is probably obvious to the reader that if in each of eight follow-ups (one-year and five-year, four grades each) we undertook the horrendous task of getting 100,000 or more cards keypunched and verified, there was, in our judgment, no good way of avoiding it. We were continually checking on what the new technological
developments in optical scanning were, in order to decide whether the time had at last come that we could switch from a questionnaire form that required keypunching to one from which output tape could be produced directly. It wasn't that good optical scanning machines didn't exist. They did exist, and they were getting better all the time. But during our earlier follow-ups the kinds of forms they could handle were not suitable for our needs. They were usually too restricted in size to allow room for all the questions we needed in the questionnaire, and too restricted in allowable format to give us the flexibility we needed.

By about the time we were getting ready to plan the 10-year follow-up, however, National Computer Systems (NCS), in Minneapolis, had developed optical scanning equipment, document printing procedures, and preliminary processing routines that would meet our needs. Using their services would enable us to get our data tapes prepared more economically and with substantially better accuracy than we could achieve via the keypunching route. Furthermore the NCS machines could handle questionnaires folded accordion-style, with as many as six 7" x 11" sheets (12 pages). This would actually give us more space for questions than the keypunch-style documents, which were typically 8-page 8½" x 11" booklets with the letter occupying the front page and the mailing address and return address the back page, would allow.

Taking everything into consideration, we decided to use NCS's services starting with the 10-year follow-up. As has already been noted, however, there never was a 10-year follow-up—so we didn't initiate use of an optically scanable questionnaire until the 11-year follow-up.
2. **Coding**

Regardless of whether the questionnaire was to be key-punched or optically scanned, some preliminary processing of the returned questionnaires, at AIR, was necessary. For certain questions (e.g., name of college attended, kind of job, long-range career plan) a multiple-choice item format on the questionnaire was not feasible; the respondent had to write the answers to such questions. This meant that the responses had to be converted to numerical codes, and in the case of optically scannable questionnaires the code then had to be recorded on the questionnaire form by blackening appropriate circles on a grid, to produce a "position code." This is done by one coding clerk, and checked independently by another; the coding supervisor, of course, is responsible for spot-checking as much as necessary, and making certain that appropriate standards of accuracy are met. Coding is a seasonal occupation, and it is performed by part-time clerks but it is by no means a strictly routine job. Doing it well requires speed and accuracy—but also intelligence and judgment. (Most of our coding clerks are college students who want part-time jobs; quite a few have eventually become members of AIR's permanent research staff, and are doing research studies themselves.)

3. **Name and address changes**

In addition to the coding of responses to open-end questions, an important part of the coding clerk's job is to prepare rosters for changes in a sample member's name or address or both. Correction cards are then keypunched from the rosters, so that the mailer tape can be corrected.
C. Mail response rates

On the very first follow-up survey we did—the one-year follow-up of the grade 12 cohort—almost two-thirds of the recipients completed the mailed questionnaire. Actually not everyone who should have been mailed the questionnaire was, since our mailer tape (the tape with names and addresses) wasn't in as good condition as it should have been, so that the response rate to the mailed questionnaire was only about 62 percent of the total 12th grade sample. That rate, though,—the "only 62%" rate—was really extraordinarily good, in view of the fact that response was purely voluntary. Never again have we even come close to a 62% response rate. At the time we achieved it, we thought it was a very good response rate; now we think it was sensational! (Of course, the fact that for this follow-up the least time had elapsed helped. Not only was the Project TALENT experience freshest in the sample members' minds, but our address records were freshest too.

For the one-year follow-ups of the other three grade cohorts, the response rates were in the 40's. The overall mailed questionnaire response rate for the one-year follow-up was about 50 percent.

For the five-year follow-up, grade 12 was still the one on which the response rate was best, but it was only about 38 percent. The overall response rate for the five-year follow-up mailed questionnaire was about 33 percent.

For the 11-year follow-up the downward slide continued—but fortunately at a decelerating rate; the overall respondent rate for the 11-year mailed questionnaire was down to 23 percent.
D. Special samples of nonrespondents

In order to maintain the probability-sample character of the data, it was necessary to get the nonrespondents represented somehow. This was accomplished by picking a probability sample of the nonrespondents and then contacting them directly, usually by phone, and interviewing them to get their answers to the questionnaire items. The weights for these cases had to be increased by an appropriate factor, to take account of the fact that each of these contacted nonrespondents to the mailed questionnaire represented not only themselves but also a large number of other non-respondents, who weren't in the special sample.

The manner of selection of the special sample was quite similar in most of the follow-ups although in a few of them something slightly different (and usually a little more complex) was done. The most prevalent procedure was to arrange the nonrespondents in student number order and then to pick a "systematic" sample of them--i.e., every umpt-th nonrespondent--to get a sample of the desired size (usually about 2500 to 3000 cases). Modifications of this procedure occurred in the case of the 1-year follow-up of the grade 9 cohort, the 5-year follow-up of the grade 11 cohort, and the 11-year follow-up of the grade 9 cohort. In all three of these follow-ups, the procedure was exactly like the general procedure outlined above, with the exception that the following modifications were used.

For the 1-year follow-up of grade 9

The nonrespondents were stratified on the basis of whether they had been included in the 1963 retest sample or not. For the ones that had been, a higher sampling ratio was used.

For the 5-year follow-up of grade 11

The nonrespondents were stratified on the basis of whether they had responded to the one-year follow-up questionnaire. For the ones that had, a higher sampling ratio was used.
For the 11-year follow-up of grade 9

The nonrespondents were stratified jointly on their "general academic aptitude composite" score and on the percentage of minority students who were enrolled in their high school in 1960. The sample was selected with selection ratios four times as high for high-minority-enrollment schools and twice as high for medium-minority-enrollment schools as for other schools, i.e., those with low minority enrollment and those that provided no information about minority enrollment. With those exceptions, the 9th-grade 11-year follow-up special sample was "self-weighted". (See Chapter 15, section C, for a brief explanation of "self-weighting".)

Obviously the special procedure that was used in the 5-year follow-up of grade 11 was designed to increase the proportion of the sample members on whom both one-year and five-year follow-up data would be available, so that results could be analyzed longitudinally, across follow-ups. Because the desirability of having follow-up data obtained at different times available on the same individuals is obvious, some readers may wonder why we didn't make more extensive use of the oversampling procedure—or use it to a more extreme degree. For instance why didn't we decide that to represent those 11-year nonrespondents to the mailed questionnaire who had also been nonrespondents to the 5-year mailed questionnaire, we would select as members of the 11-year special sample only those 5-year mailed-questionnaire nonrespondents who had been in the 5-year special sample? The logic seems impeccable; we would merely be taking advantage of the fact that the 5-year special sample already represented the 5-year nonrespondents not in the special sample. And the advantage is clear; our 11-year special sample would consist wholly (or almost wholly) of sample members on whom 5-year follow-up data were available. The answer to the question is that we did consider the possibility of selecting the 11-year special sample in the manner outlined above; we considered it very seriously. But we decided that before making a definite decision we should do one empirical check. We knew that at the time of selection the 5-year special sample was
representative of all the nonrespondents to the 5-year mailed questionnaire. But were they still representative after the special follow-up interview experience? It seemed likely that they would be; after all it didn't seem likely that one interview—perhaps a 15-minute conversation on the phone—would change behavior six years later. But being empiricists we decided to make certain. There was just one kind of behavior on which we had data (the early returns to the 11-year follow-up mailed questionnaires) that could be compared for 5-year special sample members and 5-year nonrespondents not in the special sample. That behavior was the act of completing and returning the 11-year mailed questionnaire. Accordingly we computed a chi-square coefficient on the following four-fold table, based on all nonrespondents to the 5-year follow-up mailed questionnaire:

<table>
<thead>
<tr>
<th>In 5-year special sample?</th>
<th>Respondent to 11-year mailed questionnaire?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

Much to our surprise (or, at least, much to my surprise) the chi-square coefficient was statistically significant. A larger proportion of the 5-year special sample members than nonmembers returned the 11-year questionnaires. Since that one type of behavior had apparently been changed by the experience of the special-sample interview, we had to assume that perhaps other kinds of behavior subsequent to the interview had also been changed. We decided, therefore, not to apply the otherwise attractive idea of limiting the 11-year special sample to cases with 5-year follow-up data (plus those few 5-year special-sample members on whom interviews had not been completed).

As for methods of contacting members of the special sample, there has been continual experimentation. Among the different methods tried,
either alone or in combination with others, in various follow-ups, have been the following:

1. Follow-up by the Regional Coordinator, by telephone interview.

2. Follow-up by the Regional Coordinator in a face-to-face interview. This approach was used only in the very early years of the project, and then only rarely. Whether the contact should be by phone or in person was left to the judgment of the Regional Coordinator, who generally would telephone except if the nonrespondent couldn't be reached by phone or if there was some other problem that couldn't be handled any other way than by a face-to-face interview.

3. Follow-up by Retail Credit*. Since this method was expensive it was generally reserved for nonrespondents on whom the Regional Coordinators—or sometimes the TALENT staff member who was doing the telephoning (as discussed in paragraph 4 below)—had to give up. But in the one-year follow-up of the 9th grade cohort, the Project Director decided to use Retail Credit's services for the entire special sample. It didn't work well (only two-thirds of the nonrespondents were interviewed) and was never tried again. (The decision to abandon this approach was good one, in the opinion of the author, for more reasons than just the low response rate. It seemed reasonable to suppose that our carefully selected Regional Coordinators, or other interviewers selected and trained by us, would be more skilled at the sort of interview required than would Retail Credit's employees, presumably trained for an entirely different type of telephone conversation.)

4. Phone calls from the Project TALENT office by junior staff members (research assistants, administrative assistants, coders), who were trained by senior staff members for this interviewing. This procedure of using staff members to make the calls was reserved, in the early follow-ups, for cases that the Regional Coordinators had given up on, or even sometimes, for cases that Retail Credit (see paragraph 3 above) had not been able to contact. In recent years, however, TALENT staff members have been used in preference to Regional Coordinators, except in areas where the Regional Coordinator has proven very effective in previous follow-ups.

* Retail Credit has recently changed its name to "Equifax".
5. Mailing of another questionnaire to a member of the special sample, after he or she has been contacted by phone and has agreed to fill it out promptly and return it. (This method is used only under special circumstances.)

For further information on methods of contacting members of the special sample see "Methodology of the 11-Year Follow-Up" (Rossi, Wise, Williams, and Carrel, 1976).

The completion rates for interviewing of the special sample, like those for the mailed questionnaire, have varied from year to year. The all-time high, 99.5%, was of course for the very first follow-up—the one-year follow-up of grade 12—when our address records were freshest and therefore mostly likely to be correct. Honors for the all-time low are shared by the five-year follow-up of grade 9 (under 63.0%) and the 11-year follow-up of grade 12 (just a bit over 63.0%).

E. Efforts to maintain and improve response rates

1. For mailed questionnaire

Completion rates depended, of course, not only on the state of the address file but also on the nature of the contact procedures used—as well as various factors out of our control. Such factors included the collective mood of the particular cohort of young people, and what proportion of them were unavailable because they happened to be in Viet Nam at the time of the follow-up.

Nevertheless there was quite a bit we could do to improve response rates. Some experimentation has been done to find out the optimum mailing schedule when there are several waves. For instance should there be two weeks between waves? Or a month? When should the reminder postcard go out? We have not used precisely the same mailing schedule in every follow-up of every cohort, but the schedule described in section A above represents our current thinking. It is the schedule that was used in the
ll-year follow-up of grade 10. (A somewhat different schedule was used for grade 9, because of the operation of Murphy's well-known Law*.)

We have learned a lot about effective use of the mails, since our first follow-up. In the more recent follow-ups we have tried to use a reasonably sophisticated mix of first-class mailing (for [hopefully] speed and dependability) and bulk rate for economy.

For the ll-year follow-up of the grade 10 cohort, the mailing schedule was as shown in Table 18-1.

The Project TALENT News (the four-page "newspaper" that we mail to all members of the TALENT sample annually) is intended to maintain continuing interest in the project, and to help promote good will towards it. It is also an effective device for maintaining and updating address files. (Our current mailing procedure is to use bulk rate for mailing the newspapers, together with "address correction requested".)

2. For special sample

Efforts to improve the completion rate for the special-sample interviews have two distinct aspects: (a) developing better ways of locating hard-to-find people and (b) developing ways of securing the cooperation of the special-sample member, once he or she has been located. Getting their cooperation probably depends at least as much on the attitude and manner of the interviewer as it does on anything else. As for the procedures that have been developed for locating sample members, they are described in considerable detail in "Methodology of the Project TALENT ll-Year Follow-Up Study" (Rossi, Wise, Williams, and Carrel, 1976, Appendices B through D).

* Whatever can go wrong will.
TABLE 18-1. Mailing schedule and procedure for 11-year follow-up of Grade 10 cohort, and TALENT News

<table>
<thead>
<tr>
<th>Item mailed</th>
<th>Wave</th>
<th>Mailing date</th>
<th>Interval in weeks</th>
<th>To whom sent</th>
<th>Method of mailing</th>
</tr>
</thead>
</table>
| TALENT News  | -    | 6/18/73      | 13½              | Entire TALENT sample | Bulk mail  
Address correction requested |
| Questionnaire| 1    | 9/20/73      | 1                | Entire 10th-grade cohort* | First class |
| Reminder postcard | - | 9/27/73 | 2 | Entire 10th-grade cohort | Bulk mail  
Address correction requested |
| Questionnaire| 2    | 10/11/73     | 7                | Entire 10th-grade cohort | Bulk mail  
Address correction requested |
| Questionnaire| 3    | 11/29/73     | 10               | Members of 10th-grade cohort who hadn't yet responded** | Bulk mail  
Address correction requested |
| Questionnaire| 4    | 2/7/74       |                  | Members of 10th-grade cohort who hadn't yet responded (except that those for whom two dead-end address unknown returns were received were eliminated from this mailing) | Bulk mail  
Address correction requested |

* Prior to mailing, the mailer tape was updated on the basis of name and/or address changes received as a result of the Project TALENT News.

** Prior to mailing, the mailer tape was updated on the basis of name and/or address changes received from prior mailings.
F. Content of special-sample questionnaire

Should the special sample questionnaire be the same as the mailed questionnaire or should it differ in some way? Four different levels of similarity may be recognized:

Level 1. The special-sample questionnaire uses an entirely different set of questions from that used in the regular questionnaire. (They are worded differently, and may have different options.)

Level 2. The special-sample questionnaire consists of a subset of the regular questionnaire's items, with minor changes in wording.

Level 3. The special-sample questionnaire consists of a subset of the regular questionnaire's items, identically worded.

Level 4. The special-sample questionnaire is identical with the regular questionnaire.

Gradually, we have progressed from Level 1 (the initial procedure) to Level 4 (the current procedure). Someone decided to use the Level 1 procedure in the one-year follow-up of grade 12. It worked out badly, for obvious reasons; if the special sample is supposed to provide information as to how the nonrespondents to the regular questionnaire would have responded to that questionnaire if they had responded, it is clear that the same questions have to be asked.

The idea behind Levels 2 and 3 (a subset of the regular questionnaire's questions) is that the amount of time that an interviewee can be expected to spend answering questions is limited. But the Level 4 approach allows us to have our cake and eat it too. It includes all the questions but they are asked in descending order of importance, and with the "sensitive" items of limited importance last, so that if the interviewee terminates the interview before answering all the questions, he has already answered all the questions we would have if the questionnaire had been abridged to start with. And if, on the
other hand, the interviewee is willing to answer all the questions, he is given the opportunity to do so.

G. Ten becomes eleven; twenty equals seventeen?

It will be recalled from Chapter 1 that because of delays in funding, the 10-year follow-up, it became the 11-year follow-up instead. It now appears that the long-planned 20-year follow-up is about to turn into a 17-year follow-up, at the suggestion of NIE personnel.
Chapter 19. TALENT POST-DOCTORAL FELLOWSHIPS

During the academic year 1966-67, Project TALENT had a grant from the Office of Education to select three post-doctoral Fellows and to operate a post-doctoral training program for them, in computer and multivariate applications to educational research.

The three winners, selected from many applicants, were Dr. William A. Love (a social psychologist), Dr. David Kapel (an education major whose doctorate was in curriculum development), and Dr. Douglas Stewart (a sociologist). They received their doctorates from, respectively, the University of Texas, Temple University, and the University of Pittsburgh.

The entire senior staff of the project* participated in setting up the training program and providing the training, but the person who played the leading role in this activity was Paul Lohnes. Paul is a natural teacher, with a great enthusiasm for teaching and a real gift for conveying enthusiasm about whatever he is teaching—whether it happens to be multiple discriminant analysis, or Project TALENT procedures, or canonical correlations, or the elements of computer programming—to whoever his students happen to be. Whether the post-doctoral program was a success probably depends in large part on whether the three Fellows think it was a year well spent. To the best of our knowledge they did. Looking back retrospectively they were uniformly enthusiastic, regretting only that the program didn't last longer.

Their principal responsibility during the year was to complete a research study on a suitable topic, and write a report. Dave Kapel did a study that resulted in a report entitled "Effects of Negro Density on Student Variables and the Post-High-School Adjustment of Male Negroes"; it was an interesting effort, though unfortunately the usefulness of the report was severely reduced by the fact that data for

* The senior staff included: Bill Cooley, Paul Lohnes, Lyle Schoenefeldt, Barry Wingersky, Charles Hall, and Marion Shaycoft
the study were available for so few cases. (A smaller percentage of blacks than whites responded to the follow-up questionnaires.)

Bill Love and Doug Stewart were so infected by Paul Lohnes' enthusiasm for multivariate analysis that they worked together to try to develop a method of determining what proportion of variance is accounted for by each canonical correlation, and to apply their method to the TALENT sociometric data. Their report is entitled "Interpreting Canonical Correlations: Theory and Practice".

After their year in Pittsburgh, at AIR, they scattered, of course. Doug Stewart traveled the least distance. He returned to the Sociology Department at his alma mater, the University of Pittsburgh. Dave Kapel joined the faculty of Glasboro State College, in New Jersey. (Since then he too has moved back to his alma mater, Temple University.) Bill Love accepted an appointment at Nova University, in Florida.
Chapter 20. LATER STUDIES

In this chapter we discuss studies by the TALENT staff that were done subsequent to those reported in Chapter 10. Though there is certainly no sharp distinction between the two groups of studies, possibly those discussed in the present chapter have less of a "necessary preliminaries" flavor and more of a "let's-get-on-with-it" flavor. But that distinction, if it exists at all, exists only tenuously. In any event, to really know what the various reports are like, the reader will have to look at the reports themselves. Therefore my advice is not to worry about the distinction between what went into Chapter 10 and what goes into Chapter 20, but instead, "let's get on with it"! As in Chapter 10, only major reports are discussed in this chapter. Furthermore it is limited to "mainstream Project TALENT studies". The term "mainstream Project TALENT" is used to designate that part of Project TALENT which was supported by OE or NIE.

A. The one-year follow-up studies

"The American High School Student"* (Flanagan, Davis, Dailey, Shaycoft, Orr, Goldberg, and Neyman, 1964) includes a presentation of preliminary results from the one-year follow-up of the grade 12 cohort. But until "Project TALENT One-Year Follow-Up Studies" (Flanagan and Cooley, 1966, 365 pages) was written there was nothing pulling together follow-up data from all four cohorts.

B. The MAP studies

Paul Lohines factor-analyzed 62 variables--grade, sex, and 60 score variables (mostly cognitive) in one analysis and 40 variables--grade, sex, and 38 noncognitive variables--in another. He came up with 13 factors in each analysis--grade, sex, and 11 "abilities factors" in the first, and grade, sex, and 11 "motives factors" in the other. He called these 22 factors the MAP factors (MAP = "Measuring Adolescent Personality"). MAP is a clever acronym; but as Truman Lee Kelley

* See Chapter 10
wrote many years ago, no two factor analysts agree on how factor analysis should be done—and therefore it is perhaps a foregone conclusion that the author of this History greatly prefers the factor analysis methodology she used in "The High School Years: Growth in Cognitive Skills" (see paragraph C below, or Chapter 11) and thinks it produced more meaningful results. (See also paragraph E below for yet another factor analysis of the TALENT battery.) Anyhow, two reports resulted from the MAP factor analysis—"Measuring Adolescent Personality" (Lohnes, 1966, 225 pages), and "Predicting Development of Young Adults" (Cooley and Lohnes, 1968, 233 pages). The first of these describes the factor analysis and the second applies the results to five-year follow-up data in an effort to relate factor scores to such things as career plans and career plan changes. The factor scores in many of the analyses are analyzed by multiple discriminant analysis, so that essentially the procedure used is a multivariate analysis of a multivariate analysis.

It might be mentioned in passing, to forestall a deluge of requests by potential Data Bank clients (see Chapter 22) who wish to obtain and use MAP scores in their own research, that this is impossible. Paul Lohnes unfortunately did not present in the report the matrix of coefficients that defines the factor scores; nor did he present the complete tables of factor loadings* which would have permitted reconstruction of the matrix of coefficients; nor the original correlation matrix he used, which would have permitted reconstruction of the whole thing; nor did he provide the TALENT staff with any of these unpublished items. And since he did his own programming and managed his own computer operations personally, the MAP factors are irretrievably lost.

In view of my earlier comment that I prefer the factor analysis that I did on the 1963 retest, perhaps it should be pointed out that factor scores from that factor analysis aren't available either, but

* Only loadings of at least .35 were published.
for an entirely different reason. They are not available because a matrix of coefficients to define factor scores was never computed. And the reason a matrix was not computed is that this author is opposed to such a use of factor scores. She thinks it just obscures results that would be clear and meaningful if presented in terms of the original variables. What value factor analysis has in substantive research lies in its use to gain a better understanding of the original variables—rather than in replacing those original variables with artificial abstractions called factor scores, that often don't even represent the extracted factors very well. Thus, in the opinion of the author, there is hardly ever any value in computing factor scores, even when the factor analysis from which they would be derived is sound and useful in helping to understand the underlying nature of the original variables.

C. The 1963 retest study

As described in Chapter 11, about 7500 students who had been tested in 1960 as 9th-graders were retested as 12th-graders in 1963. "The High School Years: Growth in Cognitive Skills" (Shaycoft, 1967, 376 pages) is a comprehensive report on the results and the methodology.

D. The five-year follow-up studies

The two major reports which came out of the five-year follow-up other than the already mentioned report by Cooley and Lohnes (1968) were "Five Years After High School" (Flanagan, Shaycoft, Richards, and Claudy, 1971, 492 pages plus a separate 359-page Appendix) and "The Career Data Book" (Flanagan, Tiedeman, Willis, and McLaughlin, 1973, 397 pages).

"Five Years After High School" not only presents substantive results from the follow-up but also describes some methodological
innovations (e.g., propinquity analysis, an idea developed by Shaycoft).

"The Career Data Book" is discussed in paragraph F below.

E. Ted Cureton's two TALENT studies

Edward E. (Ted) Cureton has never been on the Project TALENT staff, but he has been a consultant and has been sufficiently interested in the project that he arranged to do two important studies using the data. One was an equating study and the other was a factor analysis.

The equating study, which was carried out on Knoxville students who had already taken the TALENT battery, involved administering three other batteries to these students, so that tests in those batteries could be equated to the TALENT tests. The three batteries were the FACT Battery, the Differential Aptitude Tests, and the General Aptitude Test Battery. The methodology and results are reported in "Project TALENT Tests as Anchors for Equating Other Tests" (Cureton, E. E., 1973, 60 pages).

The factor analysis was based on the same data as the equating, plus one additional battery, the Essential High School Content Battery. The methodology and results are reported in "A Factor Analysis of Project TALENT Tests and Four Other Test Batteries" (Cureton, E. E., 1968, 79 pages).

F. The Career Data Book and the supplement to it

"The Career Data Book" (Planagan, Tiedeman, Willis, and McLaughlin, 1973, 397 pages) presents TALENT score profiles for groups of people in various career fields on the basis of the five-year follow-up. These are intended for use in career guidance.

The Career Data Book supplement, which is entitled "Using the TALENT Profiles in Counseling" (Rossi, Bartlett, Campbell, Wise, and
McLaughlin, 1975, 216 pages), does essentially the same thing, except that 11-year follow-up data (for the Grade 11 and Grade 12 cohorts) are used instead of 5-year data, the career groups are defined a little differently, and the profiles are presented in a slightly different format.

G. The 11-year follow-up studies

In addition to the Career Data Book supplement described above, the three chief substantive reports that have come out of the 11-year follow-up thus far are "The American Citizen: 11 Years After High School" (Wilson and Wise, 1975, 163 pages), "Post High School Education and Career Development" (Yen and McLaughlin, 1974, 90 pages), and "Careers in Science: A Project TALENT Study" (Shaycoft, 1975, 132 pages).

The Yen-McLaughlin report compares students with varying amounts and kinds of post-high-school education, with respect to such career-related variables as salary, job satisfaction, and employment rate. (Socioeconomic status and academic aptitude are used as control variables.)

"The American Citizen" presents response distributions for the 11-year follow-up questionnaire items, for the Grade 11 and Grade 12 cohorts.

The "Careers in Science" report, as its title suggests, is concerned with the use of Project TALENT data in choosing a career in science, and with problems of career guidance in high school, for potential scientists. Although funds for this study were provided by NSF rather than by OE or NIE, it is still, in a sense, a part of "mainstream Project TALENT", since the NSF funds were administered by OE (and then later by NIE) via an interagency agreement. (See Chapter 17.)

In the next chapter some further, quite varied, uses of Project TALENT data are discussed.
Chapter 21. SOME OTHER USES OF TALENT DATA

Over the years a great many studies involving Project TALENT data in one way or another have been carried out (and a great many that might have been carried out haven't been). Researchers from numerous disciplines—not only educational researchers but also psychologists, economists, sociologists, home economists, library science specialists—have used TALENT data in their studies.

This chapter is a potpourri of studies of various kinds which have very little in common except that they all involve Project TALENT. The chief purpose of this chapter is to give some idea of the breadth of applications. Again, as in Chapters 10 and 20 we shall make no effort to report the findings of the various studies.

In the sections that follow, the studies are classified into loose categories (since they don't fall into tight categories very well).

A. Equating studies

Three of the five equating studies have been briefly discussed in previous chapters. The following instruments have all been equated to appropriate TALENT Tests or composites in order to obtain norms approximately equivalent to those that would have been obtained had the test been administered to the TALENT sample.

1. For the Air Force:
   
   Armed Forces Qualification Test (AFQT): 1 variable
   Air Force Officer Qualification Test (AFOQT): 27 variables
   Airman Qualifying Exam (AQE): 4 variables

2. For the Navy:

   Navy Basic Test Battery: 5 variables
3. **Knoxville equating**

   Equating of tests in three other batteries (FACT, DAT, GATB) to TALENT tests. (This equating was done by Ted Cureton under a grant from OE.)

4. **Inter-American Tests**

   This battery by Herschel T. Manuel was equated to the TALENT tests.

5. **PCG equating**

   Tests of the Planning Career Goals (PCG) battery have been equated to the TALENT tests. For a discussion of the PCG battery, see Chapter 23.

B. **Knoxville follow-up studies**

   Louise Cureton has done three separate studies involving special data collection efforts for the Knoxville sample. The first was a study of the correlates of juvenile delinquency; the second was a study of welfare dependency; and the third, a study of alcoholic problems among adolescents in the TALENT sample, and their parents.

   Other researchers besides Louise Cureton have also done studies for which they found the Knoxville sample peculiarly useful. For instance John Claudy, David Gross, and Rebecca Strause (three AIR staff members) prepared a report entitled "A Study of Married Couples in Knox County, Tennessee" (1974).

C. **Other studies of marriage, the family, and parenthood**

   Claudy, Gross, and Strause also did another population study (this one not limited to the Knox County population) which resulted in a report entitled "Family Size, Birth Order, and Characteristics of Young Adults" (1974).
Josefina J. (JJ) Card* has completed a study (Card, 1977) of the consequences of adolescent childbearing for the young parent's future personal and professional life. She is now doing a related study, on long-term consequences for children of adolescent parents.

D. Changes in norms

John Flanagan has personally carried out two studies which involved going back to a sample of the TALENT schools 10 or more years later, giving one or more of the same tests, and comparing performance with that in 1960. The findings of both studies have had considerable bearing on the current controversy concerning test score decline.

The first was the "Progress in Education" study. This involved, among other things, going back to 134 of the TALENT schools in 1970, giving the Reading Comprehension Test again (along with some of the Student Information Blank items) to all 11th-grade students or a representative sample of them, and comparing the results with those achieved in 1960. The resulting report is entitled "Progress in Education: 1960-1970" (Flanagan and Jung, 1971). This study was funded in part by NIH through general research support funds.

In 1975, in connection with the equating of the PCC battery (see Chapter 23) to the TALENT battery it was necessary to administer both batteries to about 28 schools, and for the purpose some schools from the original TALENT sample were used. Findings relevant to the test-score-decline controversy are reported by Flanagan in "Changes in School Level of Achievement: Project TALENT Ten and Fifteen Year Retests" (Flanagan, 1976). This study was financed by AIR itself, rather than by a government grant or contract.

E. Quality-of-life studies

In recent years, John Flanagan has personally directed two quality-of-life studies that utilized TALENT cases and TALENT data (as well as new follow-up data). The first of these studies, which was done by John Flanagan and Sandy Wilson, was based on about 250 TALENT cases

* She is a member of the AIR staff.
who had Army experience; it was sponsored by the Army Research Institute (ARI). Results are reported in "Quality of Life as Perceived by 30-year-old Army Veterans" (Wilson, Flanagan, and Uhlaner, 1975). The second was a study of the quality of life of 30-year-olds; it was based on a 1,000-case subsample from TALENT's probability sample of 15-year-olds (500 men and 500 women), who were interviewed 15 years later. The principal purpose of the study was to help formulate educational goals. It was sponsored by NIE. The results are reported in "An Empirical Study to Aid in Formulating Educational Goals" (Flanagan and Russ-Eft, 1975).

F. Follow-back studies

Several of the Data Bank studies have been of the follow-back type. Unfortunately, however, these studies have been neither as numerous nor as successful as the potential of the method might lead one to expect. The reasons have been discussed in Chapter 13.

G. A guidance study for NSF

This study, unlike the Shaycoft study (1975) mentioned in section G of Chapter 20, was done directly for NSF rather than being administered by OE or NIE. But like the Shaycoft study, it is concerned with career guidance in secondary schools. The report is entitled "Development of Scientific Careers: The High School Years" (Gilmartin, McLaughlin, Wise, and Rossi, 1976, 216 pages). Even though this study and the Shaycoft study deal in large part with the same problem, their approaches and methodologies, and the specific questions they seek to answer, are entirely different.

H. Sex differences in achievement

J. J. Card (see section C above) is currently doing a study of sex differences in realization of achievement potential. It is sponsored by NIMH.
I. All-volunteer military forces

In a study done by John Claudy, Project TALENT data were analyzed to provide insight into problems associated with the establishment of an all-volunteer force.

J. Project PLAN

Project PLAN was a major enterprise conceived and directed by John Flanagan. PLAN is an acronym for "Program for Learning in Accordance with Needs". The purpose of the project was to develop, evaluate, and demonstrate a comprehensive system of individualized education for use in grades 1 through 12. One important goal of the program is to help the students develop and clarify their educational and career plans; educational programs are then individualized partly on the basis of these plans. This aspect of the program leans very heavily on TALENT data.

The development of the PLAN materials was financed primarily by Westinghouse Learning Corporation, but also to a substantial extent by AIR. The 14 participating school systems also made a substantial investment, in the form of contributed staff time.

K. Data Bank studies

The variety and scope of these studies is considerable. The reader is referred to Chapter 22.

L. Miscellany

Over the years, Project TALENT staff members have presented a total of over 100 TALENT-related papers at professional meetings, and have had about 60 articles published, all of which have contributed to the breadth of scope of the project.
M. Undone studies

In the first paragraph of this chapter there was a reference to studies "that might have been carried out [but] haven't been". In section F above, it was indicated that some potential follow-back studies might have fallen in this category.

Among other kinds of uses of TALENT data that are potentially promising but have hardly been tried are studies involving the development of special score scales, tailor-made for a particular investigation. Such studies could be done using the 4%-sample tape, which, it will be recalled from Chapter 12, contains item response data for all items. Unfortunately, however, the 4% tape has not yet been put in as good shape as it should be, nor in as good shape as would be necessary for many studies of the type indicated. For a discussion of the 4%-sample tape and related matters, see Chapter 13.

For many potential studies requiring item-response data, the 1963 retest tape would theoretically be a good substitute for the 4%-sample tape, if the researcher were willing to limit his study to grade 12 cases. Unfortunately, though, the retest tape is in no better condition right now than the 4% tape. (See Chapter 13 again!)

We still have hopes that it will be feasible to get both tapes in first-rate shape. It would be helpful not only to potential Data Bank clients but to Project TALENT staff as well. (Actually, it would be helpful to the entire educational research community to have these resources available in first-rate condition, including first-rate documentation.)
Chapter 22. THE DATA BANK

The term "Data Bank study" loosely means, in TALENT terminology, a study carried out by a researcher who is probably (though not necessarily*) not a member of the Project TALENT staff and who uses TALENT data (print-outs or tapes) especially prepared to his specifications, for use in his own research.

It is true that some project staff members, particularly in recent years, have been treating the term as if it had a much broader meaning, subsuming within itself all TALENT research, and all uses of any of the TALENT tapes. But the term is not used that way anywhere in this History.

A. History of the Data Bank

Although the data bank concept was embedded in the early plans for the project, the Data Bank operation did not get organized immediately. In the first place nothing could be done along those lines until there were tapes ready to use in data analysis.

However by the middle of 1962 things were in good enough shape that we could accept our first data bank contract. This was a contract to provide J. Alan Thomas, then at Stanford, with data for use in an analysis of school and community characteristics as predictors of test scores and other variables. Thomas provided the input data on community characteristics; Project TALENT provided all the other data. There were about four other data bank studies done during the Pittsburgh phase. One was for the Institute of Life Insurance. A second was for the National Merit Scholarship Corporation, where John Holland was doing a study of National Merit Scholarship finalists. The third was carried out for the Russell Sage Foundation, where David Goslin was doing a study of the impact of testing. Two large

* Even John Flanagan carried out a Data Bank study (in connection with Project PLAN)—proof positive that TALENT staff is not necessarily excluded from the ranks of Data Bank users!
reports resulted: a technical report (Brim, Goslin, Glass, and Goldberg, 1964) and a book reporting the substantive results (Goslin, 1963).

The fourth data bank study was done for NSF, and it was on academic degree projections. Actually this fourth study was not really a "data bank study" in terms of the definition, since it was not just a matter of obtaining computer output and presenting it to the sponsor. A member of the Project TALENT staff, specifically Izzy Goldberg, together with John Dailey, helped develop the research plan, analyzed the data, and wrote the report.

But though even with the NSF project excluded at least four studies of the data-bank type were completed during the Washington phase, the "Data Bank" (with a capital D and capital B) was not formally organized until the move to Pittsburgh. As mentioned in Chapter 8, Dick Holdeman became the Data Bank Coordinator when he joined AIR in the fall of 1967. Since his departure a year later there have been three Data Bank Directors: Lyle Schoenfeldt, John Claudy, and Don McLaughlin.

B. Extent of use

About 140 Data Bank studies have been completed, since the Data Bank began its formal existence, in the fall of 1964. Of these about 40% have dealt with education, 10% with career development in general, 5% specifically with science careers, 9% with family and intergenerational effects, 7% with methodology, and the remaining 29% with a wide variety of miscellaneous topics, including, for instance:

- Participation in athletics
- Gifted students
- Black male adolescents
- Teachers
- Airline stewardesses
- Cognitive style
- Mental patients
- Suicide
- Social mobility.
Many of the Data Bank clients are repeaters. For instance J. William Asher, who is at Purdue, and is possibly the heaviest Data Bank user, has been responsible for at least half a dozen studies based on TALENT data; the studies have been mostly concerned with driver education and auto accidents.

John Folger, Alan Bayer*, and Helen Astin, who were on the staff of the now defunct Commission on Human Resources and Advanced Education, also did many Data Bank studies, which dealt mostly with educational attainment beyond the high school level, and the factors to which it is related. As a matter of fact in terms of institutions rather than individuals the Commission on Human Resources was probably the heaviest user of the Data Bank—except for AIR itself and the University of Pittsburgh—having commissioned at least nine separate projects. (If it seems to the reader that the counts of projects are expressed in slightly vague terms—e.g. "at least' nine", etc.—that is because in the case of closely related projects it is not always clear where one leaves off and another one starts.) The three former Commission on Human Resources staff members—Folger, Astin, and Bayer—produced a book entitled "Human Resources and Higher Education" (1970), which was based in part on TALENT data. After Helen Astin left the Commission on Human Resources, she continued to use the Data Bank, too, in her new role as a member of the staff of the Bureau of Social Science Research.

Another multiple user is John Hause, of the University of Minnesota; he is among the economists who view TALENT data as a useful resource for their research, and his interest in the project has continued over a very long period.

Two former directors of Project TALENT—John Dailey and Bill Cooley—were users of the Data Bank’s services after departing from AIR.

* Al Bayer subsequently joined the AIR staff (see Chapter 8, section A8).
Christopher Jencks used TALENT data in his well-known study leading to his book entitled "Inequality" (Jencks et al., 1972). He also used the outcomes of a Data Bank study in an article of major importance entitled "Effects of High Schools on Their Students" (Jencks and Brown, 1975).

C. **Charging for Data Bank services**

Data Bank projects have been charged for at cost right since the beginning. As a matter of fact except under special circumstances there is no charge for preliminary consultations that take place with a researcher who decides not to go ahead with a Data Bank study. Thus, since the fixed costs of operating the Data Bank have to be covered, charges to the users have to be set high enough to cover the costs of possibly lengthy consultations that do not lead to Data Bank projects.

Before a project is actually started a cost estimate is prepared and a contract is signed. Most of the contracts call for charges on a cost-reimbursement basis, but (usually at the client's request) some contracts have been written on a fixed-price basis.

D. **Policy on use of TALENT Battery**

Some potential Data Bank clients wish to do research that requires administration of part (or all) of the TALENT battery to a new group. It will be recalled from Chapter 3, section C, that it had been agreed that a limited number of copies of the tests would be preserved for use in possible future research, but that the battery would not be used in competition with commercial tests. In line with this understanding, we have permitted administration of TALENT tests or inventories if the proposed use falls in one of the following categories:
1. The purpose is to equate some other test (commercial or otherwise) to the TALENT battery, so as to be able to take advantage of the TALENT norms or (indirectly) of the TALENT follow-up data.

or

2. The proposed research requires the use of one or more TALENT tests because no instrument available commercially would meet the needs of the research as effectively.

or

3. The proposed research fits in with our planned program for Project TALENT research, and will further progress towards those goals.

It will be observed that routine operational use does not qualify under any of the three categories. Operational use is not permitted and never has been.

When permission is given to a researcher to use TALENT tests, either of the following options may be arranged:

1. The researcher may borrow an appropriate number of test booklets and accessories from Project TALENT, paying only a nominal "rental fee" for them. The fee is set to cover our costs, including the costs of storing the booklets, shipping costs, etc.

2. The researcher may instead prefer to get permission to reprint the tests or other instruments he wishes to use. (In this case there is no rental fee.)

E. Policy on use of Data Bank services

Two major and continuing problems which have had to be dealt with in operating the Data Bank have been the necessity of maintaining the confidentiality of data, and the entirely separate problem of whether tapes containing raw data (after suitable purging of identifying information) should ever be released to Data Bank users so that they can do their own analyses or whether, instead, the Project TALENT staff (specifically the Data Bank staff) should be
responsible for getting the computer analyses done, and should pro-
vide the Data Bank client only with data analysis printouts and/or
data analysis summary output tapes.

These two problems are discussed in turn, below.

1. **Maintaining confidentiality**

   Perhaps the one immutable principle in operation of the Data
Bank has been that no data or tapes will be released in a form
that violates the confidentiality of either the 1960 data or any
follow-up data for any individual, or for any participating sec-
ondary school, or school system, or community, or state. Data
cannot be released for any geographical unit smaller than region
of the country, except where the geographical unit is neither
identified nor identifiable. Furthermore we have honored our
commitment not to release data that would permit a comparison of
the three categories of school control: public, parochial, and
other private.

   One of the things these commitments mean is that before re-
leasing tapes containing TALENT data for individual students, they
must be purged of identifying (or partially identifying) informa-
tion such as student's name, date of birth*, address, mother's
maiden name, father's first name, and school attended. (Our reg-
ular school codes have to be replaced by dummy codes.) Further-
more we have to be careful that analyses (either printouts or out-
put tape) are not in such a form that individual schools, commu-
nities, states, or other entities about which data should not be
released cannot be identified even by inference.

2. **Raw data tapes versus data analysis printouts**

   The initial policy of the Data Bank was that raw data tapes
would not be released, even after removal of identifying infor-
mation, except under very special circumstances. Data tapes

* Age in years is retained on the tape.
would be released only if in the opinion of Project TALENT's special Data Bank Policy Review Committee (John Flanagan, Marion Shaycoft, the Data Bank Director, and the Project Director) it wouldn't endanger our commitments on confidentiality and if furthermore it was advisable for at least one of the following reasons:

1. Project TALENT didn't have and couldn't develop, or acquire economically, a computer program suitable for the type of analysis the researcher wished done.

2. The computer schedule or the schedule of Data Bank personnel was so crowded that it wouldn't be feasible to complete the analysis within a reasonable time.

We wished to keep tape release to a minimum because we felt that providing some users with tapes instead of doing their analyses for them would raise the costs of those Data Bank clients who didn't have convenient access to a computer and therefore needed to be provided with printouts, not tapes. Furthermore we felt that we could run analyses of TALENT data more economically than other people could because we were more familiar with the data and understood it better.* Although those reasons for preferring other alternatives to the release of data tapes still have some validity, our policy has loosened up considerably on this issue since about 1974. The reason for this is that many of the Data Bank users are graduate students who, wishing to base a master's thesis or a doctoral dissertation on TALENT data, don't have funds at their disposal to pay for extensive Data Bank analyses, but do have enough funds to pay for a data tape, after which they can get free time on their university's computer. This was one of the reasons a committee appointed by NIE to review various longitudinal studies, including Project TALENT, in the fall of 1973 recommended letting researchers have easier access to TALENT data tapes. We have changed our policy in accordance with this recommendation, although we recognize that there is a trade-off

* See section F of this chapter. Also see Chapter 13 again!
involved. In liberalizing the policy on data tape release we are making costs lower for some (not all) of those clients who receive tapes (mostly just for the ones that have access to free computer time at their own universities), but, since fixed costs of operating the Data Bank have to be covered, the costs become higher for the other clients.

F. Handbooks, brochures, etc.

It was obvious when the Data Bank was formally initiated that some sort of brochure or handbook would be needed, outlining the kinds of data available, the services offered, and the administrative arrangements.

The first effort along these lines was a 28-page 8½" x 11" handbook entitled "The Project TALENT Data Bank" produced in the spring of 1965, when Dick Holdeman was in charge of the Data Bank. In 1968, when the Data Bank had its next director, Lyle Schoenfeldt, a small brochure was prepared that folded to fit in a standard business envelope. This was intended to supplement, not replace, the handbook prepared three years earlier.

A few years later, as use of the Data Bank grew, it became apparent that the handbook would have to "grow", too. A larger more comprehensive one was needed. Consequently a new handbook was prepared. John Claudy, the Data Bank Director at the time, prepared much of the first draft; the general outline and the content were a joint effort of the senior staff (Claudy and Shaycof t writing; Shaycof t, Tiedeman, and Flanagan reviewing). John Claudy was responsible for the format. This handbook, 8½" x 11" again, was three-and-a-half times as long as the first one, even if one excludes the pages on which the follow-up questionnaires are reproduced,* leaving 96 pages.

* The follow-up questionnaire pages have to be excluded in comparing successive Data Bank handbooks, because the number of follow-up questionnaires that have been produced and have to be reproduced keeps increasing, and because the number of pages they fill depends on whether the reproduction is page-for-page (which usually produces a clear easy-to-use reproduction) or two-to-a-page (which is a slightly more economical mode but considerably less readable).
It will be recalled from Chapter 17 that the longitudinal study review panel that met on October 18, 1973 to review Project TALENT recommended comprehensive data tape documentation so that persons not on the TALENT staff who acquired TALENT data tapes, and other Data Bank clients who wished to plan their data analysis efficiently and economically [i.e., without having to make a trip to Palo Alto to confer with Project staff and without having to make numerous lengthy and therefore expensive long-distance calls to the Project TALENT office] could do so.

As mentioned in Chapter 17, the October 18 meeting and other considerations resulted in a proposal submitted to NIE that called for, among other things, the reformatting and improved documentation of the data tapes. An associated activity was the production of an updated Data Bank Handbook. The grant that was awarded to do this work was incorporated in a larger grant which started in July 1975. Marion Shaycoft drew up plans for a Data Bank Handbook that would meet the needs not only of Data Bank clients and potential clients but also of the programmers for the project and researchers on the project staff.

The responsibility was shifted to Laurie Wise and a complete draft of the revised handbook was submitted to NIE on July 16, 1976 as scheduled. After 6 to 8 weeks of review the draft was accepted without a single negative comment. The final printing was delayed until March 1977 in order that some internal suggestions could be incorporated.
Chapter 23. BRIDGING THE GAP:
The PCG (Planning Career Goals) Battery

It will be recalled that very early in Project TALENT—in fact while it was still in the planning stage—a commitment was made that the Project TALENT battery would be reserved for research purposes only, and would not ever be used operationally. This commitment was made for very good reasons, but it created problems nevertheless. Our five-year and 11-year follow-ups had produced data that would make the TALENT battery, or a battery more or less parallel to it, extremely useful in career guidance of high school students. But the schools couldn't use it very well because there was no test battery available that they could use operationally. We became thoroughly convinced of the difficulty of using other tests, more or less parallel to TALENT tests, but unequated to the TALENT battery, when efforts were made to use the Career Data Book (see Chapter 20, section F) under those circumstances. The Career Data Book had been widely distributed (every one of the over 1300 schools that had participated in Project TALENT had received one, as well as many counseling specialists on university faculties), and we had been inundated with reports that it was difficult to use them without having TALENT test scores for the students (or scores on directly parallel tests that could be expressed as TALENT equivalents), so that the Career Data Book data would apply directly. Because of the difficulties encountered in using the Career Data Book, it was arranged that Don McLaughlin and Wendy Yen would run a series of workshops for counselors on "How to Use the Career Data Book". But even these workshops, though carefully planned and expertly run, didn't help much. Career Data Book users still felt they were handicapped by the fact that scores for their students were not available on the TALENT battery. But the TALENT battery was barred from operational use by the commitment made many years before; and no parallel battery existed. It was almost a "Catch-22" situation—but not quite! A classic "Catch-22" situation has no possible solution; but the dilemma described above could be, and was, resolved.
A. Resolution of the dilemma

It was clear that the only sensible solution was to develop a battery, parallel to the TALENT battery, that could be used in the schools. Early in 1974 it was therefore agreed that AIR would undertake the construction of such a battery.

Since a two-day battery would be impractical from the schools' point of view, the battery would be designed to be administered in a single day. This meant that not everything in the TALENT battery could be administered. But that did not strike us as a serious problem. After all, after more than a dozen years of doing research on and with the TALENT battery, we felt we knew (and still feel that we know) which components of the battery were the most important and most useful ones. Those were the components we would concentrate on, in developing an abridged battery, to be given in a single day.

The plan called for equating the new tests to the corresponding tests of the TALENT Battery, by an equipercentile procedure, so that TALENT norms and TALENT follow-up data could be translated into norms and follow-up data for the new battery.

It was agreed that the new battery would be developed by the only three members of the then-current staff in Palo Alto who had been involved in development of the original battery; they were John Flanagan, Marion Shaycoft, and Mary Willis.

CTB/McGraw-Hill, in Monterey, California, was interested in publishing and distributing the new battery. In due course, arrangements were made with them.

That left us with hardly any problems except to plan the battery, develop it, and name it!
B. Planning the battery

Let's start with the hardest problem: naming the battery! It must have been the hardest problem; it certainly took the longest to solve. From the time that it was first decided to construct the battery to the time when the name was finally settled on, which was at almost the last minute, very shortly before materials had to go to press in final form, over a year had elapsed, and we had gone through about four separate and distinct battery names—adopted them for a couple of weeks or a couple of months, and then rejected them. One thing was clear. The name of the battery had to reflect the fact that it was intended primarily for use in career guidance. The name we finally settled on wasn't an acronym for anything (all the previous names had been) but it did make the purpose of the battery clear. The name was:

Planning Career Goals (PCG)

Fortunately we didn't have to postpone work on the other problems (the "easy" ones) while we struggled with the problem of battery name. We proceeded immediately with planning the battery.

It was decided that the battery would be scored primarily in terms of the 12 career groups into which John Flanagan had classified careers in connection with his research on Project TALENT data, and around which certain aspects of Project PLAN had been built. These 12 career groups had been used extensively in TALENT research, and had been found to work well. (The 12 groups are shown in Table 23-1.)

Accordingly it was decided that the Information Test to be developed and the new Interest Inventory would each yield 12 scores—one for each of the 12 career groups. In addition to the 12 interest scales and the 12 information scales, 10 ability measures would also be developed.
| TABLE 23-1. |
| List of the Occupations Included in Each of the 12 Career Groups |

1 **ENGINEERING, PHYSICAL SCIENCES, MATHEMATICS, ARCHITECTURE**
- Aeronautical Engineer
- Architect
- Chemical Engineer
- Chemist
- Civil, Hydraulic Engineer
- Computer Engineer
- Computer Systems Analyst
- Electrical, Electronic Engineer
- Mechanical, Automotive Engineer
- Mining, Petroleum, Metallurgical Engineer, Engineering Manager
- Physical Scientist: Geologist, Meteorologist
- Physicist
- Statistician, Mathematician
- Teacher: College Science, Mathematics

2 **MEDICAL, BIOLOGICAL SCIENCES**
- Conservation Specialist: Wildlife, Forestry, Agriculture
- Dentist
- Dietitian
- Life Scientist: Biologist, Microbiologist, Biochemist
- Medical, Dental Technologist
- Medical Specialist: Veterinarian, Optometrist, Osteopathic Physician
- Pharmacist
- Physical Therapist, Occupational Therapist
- Physician
- Registered Nurse

3 **BUSINESS ADMINISTRATION**
- Accountant, Auditor, Comptroller
- Advertising Worker
- Banking, Finance Worker
- Business Executive, Manager
- Certified Public Accountant
- City Planner
- Commercial Pilot
- Computer Programmer
- Efficiency Expert
- Hospital Administrator
- Manufacturing Manager
- Market Analyst
- Officer: U.S. Air Force, Army, Coast Guard, Marines, Navy
- Personnel Administrator
- Public Administrator
- Public Relations Worker
- Retail Store Buyer
- Sales Manager
- Stockbroker
- Wholesale, Retail Trade Manager

4 **GENERAL TEACHING, SOCIAL SERVICE**
- Clergy, Other Religious Workers
- Counselor
- Educational Researcher, Evaluator
- School Administrator
- Social Worker
- Speech Teacher, Speech Therapist, Reading Specialist
- Teacher:
  - Art
  - Elementary School
  - Of the Handicapped
  - High School Commercial Education
  - High School English
  - High School Foreign Language
  - High School Home Economics
  - High School Mathematics
  - High School Physical Education
  - High School Science
  - High School Social Studies
  - High School Trade, Industrial Education
  - High School (Other Subjects)
  - Music
  - Preschool (Not Elsewhere Classified)
  - Vocational Counselor

5 **HUMANITIES; LAW; SOCIAL, BEHAVIORAL SCIENCES**
- College Administrator
- Editor
- Journalist
- Lawyer
- Librarian
- Psychologist
- Social Scientist: Physical Anthropologist, Economist, Geographer, Historian, Political Scientist, Sociologist
- Teacher:
  - College English, Foreign Language
  - College Social Science
  - College (Other Subjects)
  - Writer (not including Journalist)

6 **FINE ARTS, PERFORMING ARTS**
- Commercial Artist
- Interior Decorator
- Musical Artist
- Performing Artist

7 **TECHNICAL JOBS**
- Computer Operator
- Computer Services Repairer
- Dental Hygienist
- Drafter
- Electronics Technician
- Laboratory Technician
- Medical, Dental Technician
- Photographer
- Surveyor

8 **PROPRIETORS, SALES WORKERS**
- Automobile Sales Associate
- Business Supervisor
- Farm, Ranch Owner or Manager
- Independent Contractor
- Independent Proprietor
- Insurance Sales Agent
- Purchasing Agent
- Real Estate Sales Agent
- Route Delivery Sales Associate
- Salesclerk, Checker
- Sales Associate (Including Sales Engineer)

9 **MECHANICS, INDUSTRIAL TRades**
- Airline Mechanic
- Appliance Repairer
- Automobile Mechanic
- Clothing and Fashion Trades Worker
- Diesel, Other Specialized Mechanic
- Electrician
- Industrial Machine Repairer
- Machinist
- Office Small Machine Servicer
- Printing Trades Worker
- Telephone Installer, Repairer

10 **CONSTRUCTION TRADES**
- Bricklayer, Mason, Roofer, Plasterer
- Carpenter
- Explosives Handler, Construction Machinery Specialist
- Metal Trades Worker
- Miner, Oil Well Driller
- Plumber, Pipefitter
- Road Construction Worker
- Shift Leader

11 **SECRETARIAL-Clerical, Office Workers**
- Accounts, Records Worker
- Bank Clerk, Teller
- Bookkeeper
- Catheter
- Clerk
- Key punch Operator
- Legal Secretary
- Medical, Dental Secretary
- Office Machine Operator; Shipping, Stock, Mail Clerk
- Office Receptionist; Hotel, Reservation Clerk
- Secretary
- Secretary-Bookkeeper
- Stenographer, Court Reporter
- Telephone Operator
- Typist, Transcribing Machine Operator

12 **GENERAL LABOR; PUBLIC, COMMUNITY SERVICE**
- Barber
- Butcher, Meat Cutter
- Domestic Worker
- Enlisted Personnel: U.S. Air Force, Army, Coast Guard, Marines, and Navy
- Farm, Ranch Worker
- Firefighter
- General Laborer
- Hairdresser, Manicurist
- Mail Carrier
- Medical Aid Officer
- Practical Nurse
- Railroad Worker
- Service Worker: Waiter, Hospital Attendant, usher
- Taxi, Bus, Truck Driver
Table 23-2 lists the PCG instruments, shows the number of items in each, and indicates the corresponding TALENT instrument. It will be noted that for most of the scales, except those few that are already very long in the TALENT battery, the number of items is somewhat greater in the PCG than in the corresponding TALENT test. This was planned in order to give the PCG scales the additional reliability they would need to make them useful in individual guidance, not just in research.

In general the reliability of the PCG tests turned out to be about equivalent, on a per-item basis, to that of the TALENT tests.

C. Developing the PCG Battery and equating it to TALENT

The three writers of the PCG battery split the test development work among themselves as follows (arranged in ascending order of volume):

Mary Willis--primarily responsible for:
Reading Comprehension Test
Creativity Test
Mechanical Reasoning Test

Marion Shaycoft--primarily responsible for:
Vocabulary Test
English Test
Visualization Test
Abstract Reasoning Test
Quantitative Reasoning Test
Mathematics Test
Computation Test

John Flanagan--primarily responsible for:
Information measures
Interest Inventory
All manuals
Examiner's manual
Student's Guide (Planning Your Career)
Career Handbook
Counselor's Handbook
Technical Bulletin #1
Planning Profile
Planning Report
Arrangements for and execution of equating
Other arrangements
TABLE 23-2. PCG and Corresponding TALENT Instruments

<table>
<thead>
<tr>
<th>P.C.G. Scales</th>
<th>Corresponding TALENT Scales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of instrument</td>
<td>No. of items*</td>
</tr>
<tr>
<td><strong>Ability Measures</strong></td>
<td></td>
</tr>
<tr>
<td>1. Vocabulary</td>
<td>60</td>
</tr>
<tr>
<td>2. English</td>
<td>48</td>
</tr>
<tr>
<td>3. Reading Comprehension</td>
<td>40</td>
</tr>
<tr>
<td>4. Creativity</td>
<td>24</td>
</tr>
<tr>
<td>5. Mechanical Reasoning</td>
<td>25</td>
</tr>
<tr>
<td>6. Visualization</td>
<td>25</td>
</tr>
<tr>
<td>7. Abstract Reasoning</td>
<td>22</td>
</tr>
<tr>
<td>8. Quantitative Reasoning</td>
<td>22</td>
</tr>
<tr>
<td>10. Computation</td>
<td>76</td>
</tr>
<tr>
<td><strong>Information Measures</strong></td>
<td></td>
</tr>
<tr>
<td>12 Career Groups</td>
<td>(240)</td>
</tr>
<tr>
<td></td>
<td>20 ea</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Interest Inventory</strong></td>
<td></td>
</tr>
<tr>
<td>1. Occupations</td>
<td>(104)</td>
</tr>
<tr>
<td>2. Occupational Activities</td>
<td>(99)</td>
</tr>
<tr>
<td>3. Current Activities</td>
<td>(97)</td>
</tr>
<tr>
<td>12 Career Groups</td>
<td>25 ea</td>
</tr>
</tbody>
</table>

* Numbers in parentheses do not correspond to scales on which scores are obtained.
The tests were developed in draft form, tried out, item analyzed, revised and put in final form, and then equated to the corresponding TALENT tests. The equating took place in 1975; it was based on data obtained in 28 schools where both PCG and TALENT tests were administered. Of the 28 schools, 17 had been in the TALENT sample in 1960. It was from these 17 schools that the "15-year retest data" on which John Flanagan reported in his paper entitled "Changes in School Achievement: Project TALENT Ten and Fifteen Year Retests" (Flanagan, 1975) were based. (See Chapter 21, section D.)
POSTSCRIPT

As we explained in the Preface, the History needed a cut-off point, and so the conclusion of activities associated with the 11-year follow-up data collection and processing was selected to serve that function. But events that have occurred since then ought also to be mentioned, if only briefly in order to provide as up-to-date a picture as possible of the current status of Project TALENT.

In the first place, there is the 17-year follow-up. In April of 1976 an agreement was reached between TALENT staff and NIE staff to move up the plans for a 20-year follow-up of TALENT participants and plan instead for a 17-year follow-up. A grant was awarded AIR in November, 1976, to cover the necessary planning activities, including sample selection and instrument development, and associated data bank maintenance activities.

To review the plans a panel was assembled and a meeting was held in New York City in April, 1977. The advisory panel included:

Ronald Abeles
Social Psychologist
Social Science Research Council

David Featherman*
Professor of Sociology
University of Wisconsin

Paul Barton
Senior Consultant
National Manpower Institute

John Flanagan
Chairman of the Board
American Institutes for Research

Orville G. Brim, President
Foundation for Child Development

Hilda Kahne, Economist
Radcliffe Institute

Glen Elder
Professor of Sociology
University of North Carolina
On leave at Boystown Center for the Study of Youth Development

Herbert Parnes*
Professor of Economics
The Ohio State University

*Unable to attend.
Also attending the meeting in an advisory capacity were:

Marion Shaycoft, Principal Research Scientist
American Institutes for Research

Sandra Wilson, Senior Research Scientist
American Institutes for Research

In addition, the following three senior AIR staff members with expertise in areas of interest to the survey were consulted:

Richard Rowe, Vice President
American Institutes for Research

Paul Schwarz, President
American Institutes for Research

Harold Sheppard, Principal Research Scientist
American Institutes for Research

Following this review the plans were amended to incorporate the suggestions of the advisors. A formal request for funds to support the data collection and analysis for the survey, as described in the revised plans, was submitted in the summer of 1977. This request was still pending as of this writing.*

At least one other event of considerable significance to Project TALENT has occurred in 1977. The project's Principal Investigator has "retired." "Retired" is in quotation marks because John Flanagan continues to work essentially full-time and he continues to direct research projects for AIR, on problems in which he is particularly interested. And even though he has relinquished the role of Principal Investigator on Project TALENT, he continues to maintain a lively interest in the project and to act as a consultant to it. We hope Project TALENT will continue for many years to come, and that his interest in it will continue during all those years.

* September 1977
REFERENCES


Shaycoft, Marion F. Guide to the development, evaluation, and use of criterion-referenced tests. 1976. Note: This monograph, which was prepared under the sponsorship of the U.S. Office of Education (Office of Planning, Budgeting, and Evaluation), is not yet available in published form.


Tupes, E. C., & Shaycoft, Marion F. Normative distributions of AQE aptitude indexes for high-school-age boys. (Technical Report PRL-TDR-64-17, Personnel Research Laboratory, Aerospace Medical Division, AFSC, Task 771705, Lackland Air Force Base, Texas, 1964.)


### APPENDIX A1

**COMPOSITION OF THE PROJECT TALENT BATTERY**

<table>
<thead>
<tr>
<th>Instruments</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APTITUDE AND ACHIEVEMENT TESTS</td>
<td>A1-2</td>
</tr>
<tr>
<td>Information Test</td>
<td>A1-2</td>
</tr>
<tr>
<td>Other Aptitude and Achievement Tests</td>
<td>A1-3</td>
</tr>
<tr>
<td>MISCELLANEOUS</td>
<td>A1-3</td>
</tr>
<tr>
<td>INVENTORIES</td>
<td>A1-4</td>
</tr>
<tr>
<td>Student Activities Inventory</td>
<td>A1-4</td>
</tr>
<tr>
<td>Interest Inventory</td>
<td>A1-4</td>
</tr>
<tr>
<td>Student Information Blank</td>
<td>A1-4</td>
</tr>
</tbody>
</table>
### APTITUDE AND ACHIEVEMENT TESTS

#### Information Test

<table>
<thead>
<tr>
<th>Part I</th>
<th>Options per item</th>
<th>No. of items</th>
<th>No. of working time**</th>
<th>No. of scales***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(395)</td>
<td>(125)</td>
<td>(43)</td>
<td></td>
</tr>
<tr>
<td>Subscales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 101</td>
<td>Screening</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 102</td>
<td>Vocabulary</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 103</td>
<td>Literature</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 104</td>
<td>Music</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 105</td>
<td>Social studies</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 106</td>
<td>Mathematics</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 107</td>
<td>Physical science</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 108</td>
<td>Biological science</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 109</td>
<td>Scientific attitude</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 110</td>
<td>Aeronautics and space</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 111</td>
<td>Electricity and electronics</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 112</td>
<td>Mechanics</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a. Tools and construction***</td>
<td>(10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. Motors and mechanisms***</td>
<td>(9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 113</td>
<td>Farming</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 114</td>
<td>Home economics</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a. Cooking***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. Other***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 115</td>
<td>Sports</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 190</td>
<td>Part I Total</td>
<td>(252)</td>
<td></td>
<td>(1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part II</th>
<th>Options per item</th>
<th>No. of items</th>
<th>No. of working time**</th>
<th>No. of scales***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(143)</td>
<td>(35)</td>
<td>(25)</td>
<td></td>
</tr>
<tr>
<td>Subscales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 131</td>
<td>Art</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 132</td>
<td>Law</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 133</td>
<td>Health</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 134</td>
<td>Engineering</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 135</td>
<td>Architecture</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 136</td>
<td>Journalism</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 137</td>
<td>Foreign travel</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 138</td>
<td>Military</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 139</td>
<td>Accounting, business, sales</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 140</td>
<td>Practical knowledge</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 141</td>
<td>Clerical</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 142</td>
<td>Bible</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 143</td>
<td>Colors</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 144</td>
<td>Etiquette</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 145</td>
<td>Hunting</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 146</td>
<td>Fishing</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 147</td>
<td>Outdoor activities (other)</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 148</td>
<td>Photography</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 149</td>
<td>Games (sedentary)</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 150</td>
<td>Theater and ballet</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 151</td>
<td>21. Foods</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 152</td>
<td>22. Miscellaneous</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overlapping Subscales</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 162</td>
<td>Vocab. (overlapping other scales)</td>
<td>(9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K 163</td>
<td>Outdoor activities (Hunting + Fishing + other)</td>
<td>(19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R,K,L 192</td>
<td>Part II Total (inc. 10 misc. items)</td>
<td>(143)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parts I + II combined</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 172</td>
<td>Vocabulary scale (Variables 102+162)</td>
<td>(30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 100</td>
<td>Grand Total</td>
<td>(395)</td>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Variable &amp; Code</td>
<td>Options per item</td>
<td>No. of items</td>
<td>No. of minutes working time</td>
<td>No. of scales</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-----------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>Kind of digit score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>R</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>211</td>
<td>Memory for sentences Study</td>
<td>- (40 sentences)</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>5</td>
<td>16</td>
<td>10</td>
</tr>
<tr>
<td>212</td>
<td>Memory for words Study</td>
<td>- (24 words)</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Practice</td>
<td>- (24)</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>5</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td>220</td>
<td>Disguised words</td>
<td>5</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td><strong>English</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>231</td>
<td>1. Spelling</td>
<td>5</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>232</td>
<td>2. Capitalization</td>
<td>2</td>
<td>33</td>
<td>-</td>
</tr>
<tr>
<td>233</td>
<td>3. Punctuation a. Punctuation marks***</td>
<td>3-5</td>
<td>(16)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>b. Sentence structure***</td>
<td>3</td>
<td>(11)</td>
<td>1</td>
</tr>
<tr>
<td>234</td>
<td>Punctuation Total</td>
<td></td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>235</td>
<td>4. English usage</td>
<td>3-5</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>236</td>
<td>5. Effective expression</td>
<td>3-5</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>230</td>
<td>Total</td>
<td></td>
<td>(113)</td>
<td>1</td>
</tr>
<tr>
<td><strong>R</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>Word functions in sentences Directions</td>
<td>-</td>
<td>-</td>
<td>2½</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>5</td>
<td>24</td>
<td>15</td>
</tr>
<tr>
<td>250</td>
<td>Reading comprehension</td>
<td>5</td>
<td>48</td>
<td>30</td>
</tr>
<tr>
<td>260</td>
<td>Creativity</td>
<td>5</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>270</td>
<td>Mechanical reasoning</td>
<td>3-5</td>
<td>20</td>
<td>11</td>
</tr>
<tr>
<td>281</td>
<td>Visualization in 2 dimensions</td>
<td>5</td>
<td>24</td>
<td>4</td>
</tr>
<tr>
<td>282</td>
<td>Visualization in 3 dimensions</td>
<td>5</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>290</td>
<td>Abstract reasoning</td>
<td>5</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td><strong>Mathematics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>311</td>
<td>Part I. Arithmetic reasoning</td>
<td>4-5</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>312</td>
<td>Part II. Introductory h.s. math</td>
<td>5</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>320</td>
<td>Subtotal (Parts I + II)</td>
<td>(40)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>333</td>
<td>Part III. Advanced h.s. math</td>
<td>5</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>334</td>
<td>H.S. Subtotal (Parts II + III)</td>
<td>(38)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>340</td>
<td>Total (Parts I + II + III)</td>
<td>(54)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>R,A,F</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>410</td>
<td>Arithmetic computation****</td>
<td>5</td>
<td>72</td>
<td>9</td>
</tr>
<tr>
<td>420</td>
<td>Table reading****</td>
<td>5</td>
<td>72</td>
<td>3</td>
</tr>
<tr>
<td>430</td>
<td>Clerical checking****</td>
<td>2</td>
<td>74</td>
<td>3</td>
</tr>
<tr>
<td>440</td>
<td>Object inspection****</td>
<td>5</td>
<td>40</td>
<td>3</td>
</tr>
</tbody>
</table>

**MISCELLANEOUS**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Options per item</th>
<th>No. of items</th>
<th>No. of minutes working time</th>
<th>No. of scales</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Preferences test</td>
<td>2</td>
<td>166</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Themes</td>
<td>-</td>
<td>(2)</td>
<td>(10+)</td>
</tr>
<tr>
<td></td>
<td>- 910 My views about an ideal occupation</td>
<td>1</td>
<td>5+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- 920 What high school means to me</td>
<td>1</td>
<td>5+</td>
<td></td>
</tr>
<tr>
<td>Variable # Code</td>
<td>Kind of digital source</td>
<td>Options per item</td>
<td>No. of items</td>
<td>No. of working minutes</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>600</td>
<td>Student Activities Inventory</td>
<td>5 (150)</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regular scales</td>
<td></td>
<td></td>
<td>(108)</td>
</tr>
<tr>
<td>R</td>
<td>601</td>
<td>Sociability</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>602</td>
<td>Social sensitivity</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>603</td>
<td>Impulsiveness</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>604</td>
<td>Vigor</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>605</td>
<td>Calmness</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>606</td>
<td>Tidiness</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>607</td>
<td>Culture</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>608</td>
<td>Leadership</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>609</td>
<td>Self-confidence</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>610</td>
<td>Mature personality</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Experimental scales</td>
<td></td>
<td></td>
<td>(18)</td>
</tr>
<tr>
<td>R</td>
<td>611</td>
<td>Conventionality***</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>612</td>
<td>Theoreticality***</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>R</td>
<td>613</td>
<td>Group centeredness***</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>700</td>
<td>Interest Inventory*****</td>
<td>5</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Part I: Occupations</td>
<td></td>
<td></td>
<td>(122)</td>
</tr>
<tr>
<td></td>
<td>Part II: Activities</td>
<td></td>
<td></td>
<td>(83)</td>
</tr>
<tr>
<td></td>
<td>Scales: (&quot;core item&quot;)</td>
<td></td>
<td></td>
<td>(173)</td>
</tr>
<tr>
<td>R,P</td>
<td>701</td>
<td>Physical science, engineering, math</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>702</td>
<td>Biological science and medicine</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>703</td>
<td>Public service</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>704</td>
<td>Literary-linguistic</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>705</td>
<td>Social service</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>706</td>
<td>Artistic</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>707</td>
<td>Musical</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>708</td>
<td>Sports</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>709</td>
<td>Hunting and fishing</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>710</td>
<td>Business management</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>711</td>
<td>Sales</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>712</td>
<td>Computation</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>713</td>
<td>Office work</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>714</td>
<td>Mechanical-technical</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>715</td>
<td>Skilled trades</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>716</td>
<td>Farming</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>R,P</td>
<td>717</td>
<td>Labor</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td></td>
<td></td>
<td>(32)</td>
</tr>
<tr>
<td></td>
<td>&quot;non-core items&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>800</td>
<td>Student Information Blank*****</td>
<td>2-36</td>
<td>80</td>
</tr>
<tr>
<td>P*</td>
<td>801</td>
<td>Socioeconomic index</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>802</td>
<td>H.S. curriculum: acad. vs other</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>803</td>
<td>H.S. courses taken (acad. orient.)</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>820</td>
<td>H.S. grades</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>822</td>
<td>Guid. received in h.s.</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>823</td>
<td>Guid. received elsewhere</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>827</td>
<td>Study habits and attitudes</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>828</td>
<td>Self-perception of writing skills</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>829</td>
<td>Self-perception of reading skills</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>830</td>
<td>Amount of extracurricular reading</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>831</td>
<td>Var. of extracur. group activities (except sports)</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>832</td>
<td>Degree of participation in extracur. group activities (except sports)</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>P</td>
<td>833</td>
<td>Variety of hobbies (except sports)</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>834</td>
<td>Degree of act. in hobbies (exc. sports)</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>835</td>
<td>Participation in sports</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>836</td>
<td>Leadership roles</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>837</td>
<td>Social life</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>838</td>
<td>Work activities (chores &amp; jobs)</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

* The "kind of score" column shows the kinds of scores routinely available for Project TALENT cases in general. The meaning of these code letters is explained in Appendix B-2. Certain additional scores may be available for certain cases, such as the "4% sample" (Subsample A-04.0-25) described in The American High School Student (Flanagan et al., 1964, Chapter 2).

** Does not include the time used for giving directions except where otherwise indicated. (The exceptions occur where comprehension of directions is considered an integral part of the testing time allowance.)

*** In the regular TALENT testing (1960), scores on this variable are available for the "4% sample" (Subsample A-04.0-25) only.

**** The term "scale," as used in this table, means set of items. The numbers appearing in the "No. of Scales" column represent the number of scales, including composites, for which scores are available routinely for the Project TALENT cases in general.

***** For information about scoring procedures on these variables see Appendix E.
Appendix A2. THUMBNAIL DESCRIPTIONS OF THE TESTS

Variable Number

Information Test: Part I
This test contains 252 items which are divided into 15 subscales, as follows:

101 Screening. This scale contains questions which are extremely simple and are basic knowledge to all elementary school children. It was designed to help in identifying mentally retarded students, others who are functionally illiterate, and those who took the test with a flippant or apathetic attitude.

102 Vocabulary. This score gives some indication of the relative size of the student's general vocabulary.

103 Literature. The purpose of this score is to measure familiarity with the world of literature. Both prose and poetry are included. Some, but not all, of the literary works on which test items are based are required reading in many schools, and they are on recommended reading lists in other schools. However, the items are not limited to reading that is likely to have been done in school. The test's broad coverage makes it likely that a student who has acquired the habit of recreational reading will get a reasonably good score regardless of what specific books are required reading in his school.

104 Music. This score is intended to indicate amount of musical information, not musical talent. Those who enjoy going to concerts and opera, or listening to serious music on radio and phonograph, are as likely to get a good score as those having formal training in music.

105 Social Studies. This scale covers facts and concepts from the fields of history, economics, government and civics, geography, and current affairs.
Mathematics. These items are concerned with definitions, the vocabulary of mathematics, mathematical notation, other kinds of factual information, and the understanding of mathematical concepts. None of the items requires computation or reasoning or calls for solving a problem, since these abilities are covered by other tests in the battery (the Arithmetic Computation Test and the Mathematics Test).

Physical Sciences. This scale includes items about chemistry, physics, astronomy, and other physical sciences. Many of the items cover information that might readily be acquired in other ways than through formal instruction.

Biological Sciences. This scale includes items about botany, zoology, and microbiology. A few items about nature lore are included, though most of the items are concerned with more formal aspects of biological science.

Scientific Attitude Scale. These items provide a subscore which should be indicative of how the individual views the world—whether he views it as a place where there are logical cause-and-effect relationships, or whether he regards it as a place where consequences are illogical and arbitrary. In the latter category fall the modern-day equivalents of the primitive beliefs that have been called "sympathetic magic". Also in the category of illogical and arbitrary consequences are the premises of palmistry, astrology, and numerology.

Each item consists of a description of a phenomenon or an occurrence. Five explanations are presented, and the task is to select the "best" one. Only one of the five explanations is reasonable and logical. Of the other four, some involve common superstitions, magic, concepts incompatible with the scientific viewpoint, or belief in the occult. Selection of other distractors may primarily imply muddled thinking on the part of the examinee.
Aeronautics and Space. These items are on such topics as flying technique, navigation, jet planes, and space exploration. Much of the information that the student has in this area is likely to have been acquired out of school.

Electricity and Electronics. These items stress information that is acquirable through direct experience in the construction and maintenance of electrical and electronic equipment. Students who have worked on radios, hi-fi sets, or other electronic equipment, or on mechanisms with electric motors, should get good scores. Since the range of content covered is fairly broad, it seems likely that the average high school student will be able to answer many of the questions if he is at all interested in the area.

Mechanics. These items tap a wide range of information. Many of them are concerned with automobiles; others are concerned with other common machines and tools with which people who are interested in mechanical activities are likely to be familiar. The emphasis is on information that is likely to be acquired through direct experience with tools, engines, and motors. The scores should also be influenced by the amount of experience and training in mechanics the student has had.

Farming. These items are intended to give some representation to the information that children who grow up on farms or ranches are likely to acquire. The items might give some advantage to those who have an interest in rural life.

Home Economics. These items test information on cooking, sewing, caring for babies, cleaning, and other activities of a domestic nature. Students who have engaged in such activities, whether as a chore, or part-time job (e.g., babysitting), or just because they enjoy them, as well as students who have had formal instruction in home economics, should have an advantage on these items.
Sports. These items are intended to measure general familiarity with a wide range of sports and sports terms. Knowledge of specific sports figures is not required.

Information Test: Part I Total. The 15 scales described above are combined to give a Part I total score.

Information Test: Part II
This test consists of an additional 143 items which can be divided into 23 subscales mainly testing information acquired out of school. These subscales are numbered 131 through 152, and 162. Scale 162 (Vocabulary) partially overlaps some of the other scales.

Information Test: Parts I and II Combined
Total Vocabulary. Vocabulary scales from Part I and Part II are combined to yield an overall vocabulary score.

Information Test Total. The total score, which represents breadth of general information, is based on the 395 items in Part I and Part II. Total score on an information test may be regarded as one very good indicator of general ability to learn, something called "intelligence", at least in the case of students whose cultural and educational backgrounds have been normal. However, students from extremely deprived backgrounds, who have not had average opportunities to acquire a broad range of information, will probably have more general ability than is indicated by their score on this test.

Memory for Sentences.
This score indicates one kind of memory—the ability to memorize simple descriptive statements and recall a missing word when the rest of the sentence is provided sometime later. It is important to recognize that only one very specific kind of memory is tested, and to avoid assuming that the score indicates how well the student can do in all kinds of memory tests.
Variable Number

212 Memory for Words.
The purpose of this test is to measure another kind of rote memory—the ability to memorize foreign words corresponding to common English words. This ability is obviously directly relevant to the learning of a foreign language. It is also presumably related to the ability to learn many other kinds of material.

220 Disguised Words
This test is designed to measure the ability to form connections between letters and sounds. This is believed to be related to the ability to puzzle out from context and appearance the meaning of a word which is probably reminiscent of a familiar English word. This ability is probably one aspect of aptitude for learning shorthand or a foreign language.

English Test
The purpose of this test is to measure ability to express oneself adequately in English. A total score and five separate subscores are reported: spelling, capitalization, punctuation, English usage, and effective expression. The test is primarily concerned with written English, but presumably some generalization to spoken English on the basis of the English Usage and Effective Expression subscores is justifiable.

231 Spelling. This score is intended to indicate ability to spell—not size of vocabulary. Students who do not have especially large vocabularies but who are able to spell most of the words they have encountered should get good scores.

232 Capitalization. This score indicates degree of mastery of the rules of capitalization.

233 Punctuation. The purpose of this test is to measure knowledge of the appropriate use of all standard punctuation marks. Considerable emphasis is placed on whether the student has mastered the concept of what constitutes a sentence.
English Usage. This score measures knowledge of preferred usage.

Effective Expression. This score is intended to measure recognition of good prose expression--in other words, recognition of whether an idea has been expressed clearly, concisely, and smoothly.

English Total. This score is the sum of the five subscores described above. It indicates overall achievement in various aspects of English expression which can be measured by objective test items.

Word Functions in Sentences.
This test is intended to measure the student's sensitivity to grammatical structure. The fact that the terminology of grammar is not used at all in the test helps reduce the effects of formal training to a minimum. To score well, one must understand something about structure of a sentence and recognize the function of each word or phrase in the sentence. This ability is probably also related to the ability to learn the formal rules of English grammar. The test was made difficult deliberately, since there is reason to believe that a hard test would predict ability to learn foreign languages better than an easier one.

Reading Comprehension
The purpose of this test is to measure the ability to comprehend written materials. The ability measured is the ability to read with comprehension, rather than mere ability to mouth or recognize the printed word without understanding the fact, idea, or concept that the writer is attempting to convey. The test includes passages on a wide range of topics. The student reads the passage and then answers a number of questions about it, referring back to the passage as often as he likes. None of the items is answerable without reading the passage. The ability measured by this test is a good predictor of school success in an academic or liberal arts curriculum. In the unusual situation where a student scores low on the reading comprehension test, but obtains good scores on other types of intellectual tests, he may have a specific reading disability that can be corrected by special remedial training.
Creativity
The purpose of this test is to measure the ability to find ingenious solutions to a variety of practical problems. High scores on this test should be interpreted as indicating inventiveness or creative ingenuity.

Mechanical Reasoning
The purpose of this test is to measure the ability to visualize the effects of the operation of everyday physical forces (such as gravitation) and basic kinds of mechanisms (for instance gears, pulleys, wheels, springs, levers). A kind of reasoning which is related to mechanical aptitude is involved. Although all of the items can be answered without experience in woodworking or other crafts, or in working with motors, past training and experience must nevertheless be borne in mind in interpreting the results.

Abstract Reasoning
This is a non-verbal test to measure one kind of abstract reasoning ability—the ability to determine a logical relationship or progression among the elements of a complex pattern, and to apply this relationship to identify an element that belongs in a specified position in the pattern. Some youngsters who have had little schooling or who have had serious cultural or environmental handicaps which result in low scores in vocabulary and reading comprehension may make a high score on the abstract reasoning test, indicating greater potential for academic work than has been developed. Among typical students, who have had a normal environmental background, the abstract reasoning score should be considered just another element in the general domain of intellectual ability.

Visualization Tests
The general purpose of these two tests is to measure spatial visualization. The two tests are described briefly below.

Visualization in Two Dimensions. This test measures the ability to visualize how diagrams would look after being turned around on a flat surface, in contrast with the way they would look after being turned over.
Visualization in Three Dimensions. This test measures the ability to visualize how a figure would look after manipulation in three-dimensional space (more specifically, how a two-dimensional figure would look after it had been folded to make a three-dimensional figure).

Mathematics Test
The mathematics Test has three parts.

Mathematics Part I (Arithmetic Reasoning). This test is designed to measure the ability to do the kind of reasoning required to solve arithmetic problems. Computation, except at the very simplest level, is excluded from the test.

Mathematics Part II (Introductory High School Math). The purpose of this subtest is to measure achievement in all kinds of mathematics generally taught up to and including 9th grade, with the exception of the areas covered in the Arithmetic Computation Test and in Mathematics Part I (Arithmetic Reasoning). The primary emphasis of this test is on elementary algebra; other topics include fractions, decimals, percents, square roots, intuitive geometry, and elementary measurement formulas. While the topics covered are taught in Grade 9 or earlier in most schools, curricula differ considerably in regard to grade placement of various topics.

Mathematics Part III (Advanced High School Math). This subtest covers topics normally taught in Grades 10-12 in college-preparatory courses. The items are intended primarily to test understanding and application of basic concepts and methods, not rote memory. A wide range of subjects is included: plane geometry, solid geometry, algebra, trigonometry, elements of analytic geometry, and introductory calculus. It should be noted that some of these subjects are not offered in most high schools. However, students who have taken college preparatory mathematics beyond Grade 9 level, and have really understood it, should be able to score well.
Variable

Mathematics Total (I + II + III). This is the total mathematics score. It is a comprehensive measure of achievement in mathematics.

Tests of Clerical and Perceptual Speed and Accuracy

Arithmetic Computation. The purpose of this test is to measure speed and accuracy of computation. The test is limited to the four basic operations (addition, subtraction, multiplication, and division), and to whole numbers. The aptitude measured is an important one. However, a student does not necessarily need an outstandingly good score on this test in order to do well in mathematics in high school and college.

Table Reading. The purpose of this test is to measure speed and accuracy in a non-computational clerical task, involving obtaining information from tables. This kind of clerical aptitude is somewhat more complex than that measured by the Clerical Checking Test described below.

Clerical Checking. This test is designed to measure speed and accuracy of perception in a very simple clerical task. The test involves comparing pairs of names to determine whether they are identical.

Object Inspection. The purpose of this test is to measure speed and accuracy in perception of form. More specifically, the test is intended to measure the ability to spot differences in small objects quickly and accurately when comparing them visually.

Other

Preferences Test. One purpose of this test--frankly an experimental one--is to measure the speed with which a person can reach at least one kind of decision. The test, with a time limit of three minutes, contains 166 items, each consisting of a pair of adjectives describing a person. The student is required to select the adjective from each pair that he would prefer to have applicable to his friends. The two characteristics are intended to be about equal in social acceptability.
In considering the results derived from this test, bear in mind that the test was still in the experimental stage of development when given. We must remember, too, that only one limited kind of decision-making is being tested. This trait may, or may not, be important. It is hypothesized that the test locates the individual on a continuum that has "snap judgments" at one end and "indecision" at the other.
Appendix B. CHOOSING WHICH 1960 VARIABLES TO USE

A. Some general advice

In planning a study using TALENT data it is necessary to decide which of the vast number of 1960 variables to use, since it is seldom either feasible or desirable to incorporate all of them in a single study. In studies in which (whether for reasons of cost or for some theoretical reason) it is only feasible to incorporate a very limited number of variables, the choice of precisely which ones to include become crucial. And in making a decision of this sort we cannot emphasize too strongly that it is inadvisable to depend on instrument name or scale name. For instance the score on the "Preferences Test" is not intended to tell anything about the student's preferences; rather it tells us something about how fast he makes decisions (in a certain limited area). But we could not name it the "Speed-of-Decision-Making Test"; that would wreck it completely by revealing its purpose to the examinees. Similarly, the Student Activities Inventory does not tell, and was not intended to tell, anything about what activities the student engages in. "Student Activities Inventory" was selected as a cover name for a personality inventory.

Of course it is possible to read the thumbnail descriptions of the tests that appear in Appendix A2, and find out the facts indicated above, but thumbnail descriptions can go only so far. For many (though not all) purposes it is important that the researchers have a more exact knowledge of the character of the instrument--precisely what kinds of items make it up, what areas are covered, etc.

Researchers who need more information than the instrument names or scale names provide (and virtually all researchers using TALENT data fall in this category) are referred to the thumbnail descriptions in Appendix A2. Those who need more information than that appendix
provides (and many will) are referred to "Design for a Study of American Youth" (Flanagan, Dailey, Shaycoft, Gorham, Orr, and Goldberg, 1962). And those who require still more information about the tests, inventories, and questionnaires are urged to look at the instruments themselves. Specimen Sets of the TALENT Battery are available for this purpose.* For many purposes they should also familiarize themselves with the reliability of the tests. The best currently available source of this information (at least for those variables included in it) is "The High School Years: Growth in Cognitive Skills" (Shaycoft, 1967), table 4-7, columns 5 and 6, in which grade 12 reliabilities are shown. (The corrected-for-range procedure may be applied in order to obtain reliability estimates applicable to other grades or subgroups.)

In the above paragraphs we have urged researchers to recognize the importance of making a wise choice of variables to be included. In making such a choice, some researchers may find it useful to base their decisions in part on suggestions derived from past research results and extensive use of the TALENT variables. Providing such suggestions is the intended purpose of the paragraphs that follow.

B. Some more specific suggestions

1. Variables with broad-spectrum usefulness

   Among the most useful general-purpose or multiple-purpose variables are the following background and test variables.

   **Background variables**
   
   P*801 Socioeconomic index

---

* Specimen Sets may be ordered by writing to: American Institutes for Research
   P. O. Box 1113
   Palo Alto, CA 94302

   The current price is $3.00.
Test variables

R-190  Information Part I Total*
R-100  Information Test I Total*
R-172  Vocabulary Total
R-230  English Total
R-250  Reading Comprehension Total
R-260  Creativity
R-270  Mechanical Reasoning
R-280  Visualization in Three Dimensions
R-290  Abstract Reasoning
R-311  Arithmetic Reasoning
R-312  Introductory High School Math

and to a somewhat lesser degree:

F-410  Arithmetic Computation

A priori test score composites

C-002  General Academic Aptitude Composite
C'004  Math Composite
C'005  Technical Information Composite

2. Using the above list

The background variable singled out for listing above, Socio-
economic Index (P*801) has proven useful mainly as a control vari-
able or covariate (a variable where effect is to be partialed out,
through the use of partial or part correlations, or covariance
analysis).

The test variables listed in section B1 above are the ones
that measure abilities and kinds of information acquisition that
are useful in a wide variety of contexts. Researchers who for
one reason or another must limit their studies to a comparatively
small number of test variables might do well to make their selec-
tion primarily from those listed above, supplemented, perhaps, by
a small number of appropriate variables from the Information Test.

It should be noted that in the list of test variables above,
Arithmetic Computation (F-410) is given a somewhat restricted
endorsement. This is because its unknown (though probably fairly
high) reliability may limit its usefulness in certain contexts.

* One or the other of these two variables (R-190 and R-100) might
be used in a study, but it would usually be undesirable to use
both in the same study (except for the purpose of obtaining
empirical data as to which of the two would be preferable for
the particular purpose being investigated).
C. **Answers to specific questions about cognitive variables**

Sometimes researchers who wish to use a fairly small number of cognitive variables, perhaps five to ten, find themselves bewildered by the almost astronomical number available. On the basis of our familiarity with the variables—how they behave and what their properties are—we can offer some advice in situations of this sort. Here are some of the questions that might be asked, together with the answers that represent the best advice we can give.

**Question 1.** Should I use individual variables, such as

- R-250 Reading Comprehension
- R-290 Abstract Reasoning
- R-311 Arithmetic Reasoning
- R-312 Introductory High School Math
- R-230 English
- R-112 Mechanical Information
- R-270 Mechanical Reasoning

or should I use a *priori* composites, such as:

- C-001 I.Q. Composite
- C-004 Quantitative Aptitude Composite
- C-003 Verbal Composite
- C'-005 Technical Information Composite
- C-005 Technical Aptitude Composite
- C-006 Scientific Aptitude Composite
- C-002 General Academic Aptitude Composite

or should I use some of each?

**Answer:** If you are going to use five or ten variables, all of them should be selected from among the individual variables. To the extent you use composite variables either in conjunction with other composites, or in conjunction with individual variables, you would be artificially and unnecessarily introducing overlapping of variables into the battery with possibly misleading results. For instance the intercorrelations among the variables would be spuriously high. The only exception to this might be if you wanted to use several individual
variables and in conjunction with them one or two com-
posites that do not overlap either each other nor any
of the individual variables. An example of such a set
of variables would be:

R-250 Reading Comprehension
R-230 English
R-260 Creativity
R-290 Abstract Reasoning
R-109 Scientific Attitude
R-270 Mechanical Reasoning
C'004 Math Composite
C'005 Technical Information Composite

One important point in conjunction with all of the advice
offered above is that it is intended to be applicable to
substantive research only. Of course, if you want to do
some sort of methodological research involving an empirical
study of the effects of overlapping on the correlations
of composites,** you would want to incorporate several
overlapping composites in the set of variables in your
study.

** Question 2. If my research design permits only one cog-
nitive variable and the question the research is supposed
to answer is such that a variable representing "general
academic ability" or "overall ability" or something equally
nonspecific is needed, what one variable would you recom-
mend?

** Answer: Variable C-002 (General Academic Aptitude Composite)
is an excellent variable for the purpose. It is an overall
ability composite incorporating many different kinds and
aspects of ability, and it yields extremely reliable scores.

* Readers: Please do not copy this set of variables to use for
your research, unless you have some reason for thinking it is
appropriate for your particular study. It is presented here
merely as an example of a nonoverlapping set involving composites,
and not because it is believed to have any universal merit for
research studies in general.
Question 3. If I want some sort of general information score, which of the following three is best?

R-190  Information Part I Total
R-192  Information Part II Total
R-100  Information Total (which equals R-190 plus R-192)

Answer: For most purposes, R-100, being the most general, is best. Actually the other two composite information scores, R-190 and R-192, were computed largely as a matter of practical convenience, rather than because there was any particular rationale for them. In organizing the battery in terms of test booklets and answer sheets, it had been necessary to put part of the Information test on one of the answer sheets that was to be scored immediately so that reports of results could be sent to the participating schools. But since there wasn't room on the answer sheet for the entire information test, it had to be split in two, with the second portion put on one of the answer sheets scheduled for later processing. However because the first part was to contain the subscores that were to be reported to the schools, most of the directly school-related topics are in Part I of the Information Test, while Part II contains more of the items that are more likely to be learned outside of an academic program in high school than in such a program. What this means is that while R-100 (the Information Total) is probably the best all-purpose general information score, for a ready-made score that emphasizes (but is not based solely on) school-taught subjects, R-190 is useful; and R-192 likewise may provide a useful score on general information probably acquired elsewhere than in formal education.
Question 4. Which of the following three vocabulary scores is best to use?

- R-102  Vocabulary I
- R-162  Vocabulary II
- R-172  Vocabulary Total (equals R-102 plus R-162)

Answer: The answer to this question is almost, but not quite, parallel to the answer to Question 3 above. Since Vocabulary I is a subset of the Information Part I items and since Vocabulary II is a subset of the Information Part II items, the Vocabulary Total score (R-172) is the best overall vocabulary score. However, there is a special consideration that is applicable to R-162 (and consequently to R-172) that is not applicable to any other subscore routinely yielded by the Information Test; in Part II of the Information Test, none of the items of the vocabulary scale, R-162, is in that scale exclusively; instead, each such item is jointly in the Vocabulary Scale and in one other Part II subscale. Thus if the nature of the data analysis to be carried out is such that overlapping scales should not be included, and if any of the Information Part II scales that include items also in Vocabulary II are to be included, then clearly the only vocabulary scale that should be seriously considered is Vocabulary I (R-102).

Question 5. If I want to include one or more measures of English ability should I use the English Total score (R-230) or some or all of the following part scores,

- R-231  Spelling
- R-232  Capitalization
- R-233  Punctuation
- R-234  English Usage
- R-235  Effective Expression

or both?
Answer: It depends on what the purpose of your study is. For most purposes using just the English Total will be entirely adequate. But if you are concerned with the question of predicting ability in a specific occupation or group of occupations in which different aspects of English are specifically important in their own right, rather than just as contributing to overall ability in English, you will probably want to use some or all of the part scores.

Suppose, for instance, that the occupation you are interested in is secretary; or suppose it is book editor. For many secretarial jobs and for almost all editing jobs, ability to spell well is important—and spelling does not have a particularly high correlation with the other components of the English test.

Remember, however, that if you are planning to compute a correlation matrix and then use it in some sort of multivariate analysis (under which heading we include multiple correlation, factor analysis, components analysis, multiple discriminant analysis, canonical correlations, etc.) you will not want to include both the total score and all its components in the same multivariate analysis.

Question 6. There are a lot of different mathematics variables. Which should I use?

Answer: A wide variety of data analyses that have been carried out involving both Math Information (R-106) and Introductory High School Math (R-312) suggest that for reasons that are not crystal-clear, the two measures have essentially the same factorial composition even though
they were constructed to different specifications. Thus if the number of test variables you want to include is strictly limited, you might want to include one of these two variables, but almost certainly not both.

If you wish to limit your consideration to the kind of math that is involved in knowing how to solve "word problems" that can be solved without the use of algebra, use the Math I score (Arithmetic Reasoning). If you are concerned about the level of mathematical skills acquired by students (especially upperclassmen) who have taken or who might have taken math beyond grade 9, use the Math III score (Advanced High School Math).

Question 7. If I want to include some sort of a math composite or overall math score should I include the "Math Total" score (R-340), which is the raw score total of R-311, R-312, and R-333, or should I use "Math Composite" score (C'004) which is a weighted sum of the same three part scores that go into R-340, plus the "Math Information" score (R-106)?

Answer: For reasons explained in the answer to Question 6 above, it doesn't make a lot of difference whether you use R-340 or C'004. The latter may be slightly preferable because it is based on more items and is therefore a bit more reliable. But both have excellent reliability. It should be noted that in the opinion of the author, for most purposes the C'004 math composite is preferable (by a microscopic amount) to C-004 (the Quantitative Aptitude Composite), which has exactly the same components but with slightly different weights.
Question 8. There are two tests called memory tests:

R-211 Memory for Sentences
R-212 Memory for Words

If I want to use only one memory test, which one should I use?

Answer: Both of these are tests of memorization ability, and the question as to which one is preferable is in this instance easy and unambiguous. By all means use Memory for Words (R-212). The Memory for Sentences total was included in the battery for experimental purposes only, as indicated in Chapter 5 section C. Although the nature of the test is such that its reliability cannot be determined, there is good reason to suppose its reliability is very low. Perhaps for that reason, it has not been found to have predictive value for anything at all. The Memory for Words test, on the other hand, is parallel to a test that has proved its usefulness in Carroll's Psi-Lambda Battery, as a measure of aptitude for learning of foreign language, and has also proved to have predictive value in other contexts. Although this test, like Memory for Sentences, is one for which a reliability coefficient cannot be computed, the test's correlations with some other variables are sufficiently high to warrant the inference that it has good reliability.

Question 9. There are also two "visualization" tests:

R-281 Visualization in Two Dimensions
R-282 Visualization in Three Dimensions

Which is the better one if just one is to be used?

Answer: These two tests appear to measure somewhat different functions. Besides the difference in dimensionality, the two-dimensional test (R-281) is considerably more
highly speeded than the three-dimensional test (R-282). In general, the Visualization in Three Dimensions test is the one that has turned out to have better predictive value. Thus if you feel you are limited to one visualization test, R-282 is generally to be preferred to R-281. However, it should be borne in mind that it might be worthwhile to use both.

**Question 10.** I want to measure mechanical ability. There are two variables with the word "mechanical" in their names:

- R-112  Mechanical Information
- R-270  Mechanical Reasoning

Which should I use?

**Answer:** Use both if possible. These tests are both useful, and they differ substantially. Their names happen to be quite descriptive. R-112 measures information, while R-270 measures ability to understand and reason about mechanical relations, without requiring any substantial amount of specific technical information.

* D. Noncognitive variables

The question sometimes arises as to which set of noncognitive measures--the Interest Inventory scores or the scores on personality traits yielded by the SAI--is more useful, when one doesn't want to use both. Here the answer is unambiguously in favor of the Interest scores. Although we do not yet have reliability data on either of these instruments, and therefore our conclusions have to be a little more guarded than if we did, all the correlational evidence now available suggests that of the ten regular scales in the SAI only the impulsiveness scale has any substantial degree of uniqueness.
The other nine SAI scales all seem to be measuring one common factor (which is probably best represented by the Mature Personality scale, if one wishes to choose a single scale from the SAI). If the design of one's research study allows room for two SAI variables, the author of this History suggests that for most purposes the best ones to include might be Mature Personality and Impulsiveness.
GLOSSARY OF ACRONYMS AND ABBREVIATIONS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AERA</td>
<td>American Educational Research Association</td>
</tr>
<tr>
<td>AFOQT</td>
<td>Air Force Officer Qualification Test</td>
</tr>
<tr>
<td>APFTRC</td>
<td>Air Force Personnel and Training Research Center</td>
</tr>
<tr>
<td>AFQT</td>
<td>Armed Forces Qualification Test</td>
</tr>
<tr>
<td>AIR</td>
<td>American Institutes for Research</td>
</tr>
<tr>
<td>APA</td>
<td>American Psychological Association</td>
</tr>
<tr>
<td>AQE</td>
<td>Airman Qualifying Exam</td>
</tr>
<tr>
<td>ARI</td>
<td>Army Research Institute</td>
</tr>
<tr>
<td>BAVTE</td>
<td>Bureau of Adult, Vocational, and Technical Education, O.E.</td>
</tr>
<tr>
<td>BEH</td>
<td>Bureau of Education of the Handicapped, Office of Education</td>
</tr>
<tr>
<td>BEPD</td>
<td>Bureau of Educational Personnel Development, O.E.</td>
</tr>
<tr>
<td>BESE</td>
<td>Bureau of Elementary and Secondary Education, O.E.</td>
</tr>
<tr>
<td>BHE</td>
<td>Bureau of Higher Education, Office of Education</td>
</tr>
<tr>
<td>BuPers</td>
<td>Bureau of Naval Personnel</td>
</tr>
<tr>
<td>CTB</td>
<td>California Test Bureau</td>
</tr>
<tr>
<td>CTS</td>
<td>Cooperative Test Service</td>
</tr>
<tr>
<td>DAT</td>
<td>Differential Aptitude Tests</td>
</tr>
<tr>
<td>ETS</td>
<td>Educational Testing Service</td>
</tr>
<tr>
<td>FAA</td>
<td>Federal Aviation Administration</td>
</tr>
<tr>
<td>FACT</td>
<td>Flanagan Aptitude Classification Tests</td>
</tr>
<tr>
<td>FY 71</td>
<td>Fiscal Year 1971 (7/1/70 – 6/30/71)</td>
</tr>
<tr>
<td>FY 72</td>
<td>Fiscal Year 1972 (7/1/71 – 6/30/72)</td>
</tr>
<tr>
<td>GATB</td>
<td>General Aptitude Test Battery</td>
</tr>
<tr>
<td>GW</td>
<td>George Washington University</td>
</tr>
</tbody>
</table>
HEW  Department of Health, Education, and Welfare
HIBK  Had I But Known...! (see Preface)
IRE  Institute for Research in Education (one of AIR's institutes)
KR  Kuder-Richardson (a kind of reliability coefficient)
MAP  Measuring Adolescent Personality
MRC  Measurement Research Center
NCERD  National Center for Educational Research and Development
NCES  National Center for Educational Statistics
NCS  National Computer Systems
NIAAA  National Institute on Alcohol Abuse and Alcoholism
NICHD  National Institute of Child Health and Human Development
NIE  National Institute of Education
NIH  National Institutes of Health
NIMH  National Institute of Mental Health
NSF  National Science Foundation
OE  Office of Education
ONR  Office of Naval Research
PCG  Planning Career Goals
Pitt  University of Pittsburgh
PLAN  Program for Learning According to Needs
PRB  Personnel Research Branch
SAI  Student Activities Inventory
SES  Socioeconomic status index
SIB  Student Information Blank
UC  University of California
TALENT  Not an acronym!
T.E.A.R.S.  Test Editing and Reformatting System