Visual display of Quantitative Phenomena

John W. Tukey

Princeton University
Fine Hall
Washington Road
Princeton, NJ 08544-1000

U.S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

Important aspects of the many sensor problem have been explored. Shape theory has been consolidated and extended; new techniques have been studied. Simple resistant smoothing in two (or more) dimensions has been studied and applied. A graphical approach to exploratory analysis of variance, emphasizing single degrees of freedom has been developed and discussed. Techniques for showing distributions of large samples and large letters in strips have been developed that are more sensitive and more flexible than previous techniques. The importance of impact in designing graphical displays has been stressed; the likely evolution of graphical techniques has been described.
Best Available Copy
Visual Display of Quantitative Phenomena

by

John W. Tukey

FINAL REPORT
May 1, 1988 — May 31, 1991
DAAL03-88-K-0045

supported by the

U. S. Army Research Office (Durham)
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>1</td>
</tr>
<tr>
<td>- Books, chapters in books</td>
<td>2</td>
</tr>
<tr>
<td>- Papers</td>
<td>2</td>
</tr>
<tr>
<td>- Submitted for publication:</td>
<td>3</td>
</tr>
<tr>
<td>- Papers delivered and in preparation:</td>
<td>7</td>
</tr>
<tr>
<td>- Technical Reports: Department of Statistics</td>
<td>5</td>
</tr>
<tr>
<td>- Princeton University</td>
<td>5</td>
</tr>
<tr>
<td>- Technical Reports: Department of Civil Engineering</td>
<td>5</td>
</tr>
<tr>
<td>- Princeton University</td>
<td>5</td>
</tr>
<tr>
<td>4. Shape Theory (Goodall)</td>
<td>6</td>
</tr>
<tr>
<td>5. Smoothing</td>
<td>6</td>
</tr>
<tr>
<td>6. Data analysis and display in the context of many sensors</td>
<td>7</td>
</tr>
<tr>
<td>7. Graphical methods in ANOVA</td>
<td>7</td>
</tr>
<tr>
<td>8. Graphic display of linked aspects</td>
<td>7</td>
</tr>
<tr>
<td>9. Columns divided into bars and their competitors</td>
<td>7</td>
</tr>
<tr>
<td>10. Designing graphics for impact</td>
<td>8</td>
</tr>
<tr>
<td>11. Visual display in the years to come</td>
<td>8</td>
</tr>
<tr>
<td>12. Graphical display of distributions: distribution strips</td>
<td>8</td>
</tr>
<tr>
<td>13. &quot;Lo—tech&quot; dynamic displays</td>
<td>8</td>
</tr>
<tr>
<td>14. Graphical techniques in multiple comparisons</td>
<td>8</td>
</tr>
<tr>
<td>15. Graphical techniques in describing plant-breeding trials</td>
<td>9</td>
</tr>
<tr>
<td>16. Mapping and other cartographic techniques</td>
<td>9</td>
</tr>
</tbody>
</table>
Executive Summary

Work under DAAL03-88-K-0045 focussed mainly on the following topics:

- shape theory and techniques (5 items published, 2 to appear, 2 submitted, 4 delivered or in preparation);
- smoothing (4 items published, 1 submitted);
- graphical presentation of classes of data analytic results (1 item published, 1 submitted, 1 presented and to appear);
- classical graphical techniques (4 technical reports);
- distribution strips (1 item submitted, others in preparation);
- visual display in the years to come (1 item published);
- designing graphics for impact (1 item presented).

Special attention should be drawn to:

- The Technical Report on summarization and iconization in the context of many sensors.
- The work on consolidating and extending shape theory, especially as described in the RSS paper with discussion on procrustes methods, and the 4 papers "to appear" or "submitted".
- The work on smoothing in two or more dimensions expanding the scope of "head banging" as a simple non-linear smoother.
- The work on graphical methods in exploratory analysis of variance, which offer an important new approach focussed on single degrees of freedom, and later extended to the robust case (by Eugene Johnson, not supported by ARO).
- The work on distribution strips, which is producing highly effective displays for large samples and, more generally for large batches.
- The emphasis on impact as an essential element in the design of graphic displays.

John W. Tukey
Princeton 31 July 1991

Faculty

Colin R. Goodall 1988-91
John W. Tukey 1988-91

Research Assistant

E. Olszewski 1988-91


Books, chapters in books:


Papers:


NOTE: Letters used with years on John Tukey's papers correspond to bibliographies in all volumes of his collected papers.


Submitted for publication:


Papers delivered and in preparation:


**Technical Reports: Department of Statistics, Princeton University**

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Author and date</th>
</tr>
</thead>
<tbody>
<tr>
<td>295</td>
<td>Graphic comparisons of several linked aspects: alternatives and suggested principles</td>
<td>John W. Tukey October 1989 (in draft)</td>
</tr>
<tr>
<td>297</td>
<td>Columns divided into bars and their competitors</td>
<td>John W. Tukey July 1989 (in draft)</td>
</tr>
<tr>
<td>300</td>
<td>Combining CL (confidence limits) and PL (partial limits) about non-linear smoothers</td>
<td>John W. Tukey June 1990</td>
</tr>
<tr>
<td>301</td>
<td>Graphical displays of: Are the (x,y) pairs compatible with a linear dependence?</td>
<td>John W. Tukey June 1990</td>
</tr>
</tbody>
</table>

**Technical Reports: Department of Civil Engineering, Princeton University**

<table>
<thead>
<tr>
<th>Title</th>
<th>Author and date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data analysis and iconization in the context of many sensors.</td>
<td>Colin R. Goodall 1988</td>
</tr>
<tr>
<td>I. Iconization, summarization, and available schemata</td>
<td></td>
</tr>
</tbody>
</table>
4. Shape Theory (Goodall)

Among the topics treated were:

Eigenshape analysis of an iterative mapping scheme applied to shapes of triangle
(paper in *SIAM J. Appl. Math.* 51)

Growth-curves models for the shapes of triangles of landmark points (joint with K. Lange
and Moss, paper submitted to *Applied Statistics*).

Procrustes models for shapes, including an overview of shape theory and the connection
of shape theory with the QR decomposition (paper with discussion, *J. Roy. Statist. Soc. B*
53:).

Probability models for shapes and forms; shape density (joint work with K. Mardia,
papers (a) to appear in *Advances in Applied Probability*, (b) submitted to the *Annals of
Statistics*, (c) submitted to *J. Multivariate Analysis*.

Conformal mapping of tooth shapes (joint with S. H. Kim, L. Moss-Salentijn,
and M. Shinozuka, paper to appear in *J. Theoretical Biology*).

Exploring shape spaces with high-interaction graphics (paper at International Statistical
Insitute meeting, to appear in *Proceedings*).

Interface).

5. Smoothing

A survey of smoothing techniques was prepared and published (Goodall).

A paper on resistant smoothing of irregularly spaced data by "head banging" was
prepared and published (Goodall and Hansen).
A paper on head-banging techniques as applied to rainfall fields was reported and published (Goodall and Phelan).

A related paper on non-linear least squares and model assessment was prepared and published (Goodall and Phelan).

Smoothing and related topics were discussed at the inaugural S workshop in New Zealand (Goodall).

Investigation of smoothing county cancer rates continued. (Goodall, Kafadar, and Tukey).

6. Data analysis and display in the context of many sensors

A paper was written (Goodall and Tukey) and presented (at an IST/SDIO symposium, by Goodall). The new techniques were implemented in new S.

7. Graphical methods in ANOVA

A substantial paper on graphical and robust analysis of factorial data, making use of E. G. Johnson’s new insights, has been written and submitted. (Johnson and Tukey).

8. Graphic display of linked aspects

A substantial technical report (#295) is in near final draft. (Tukey)

9. Columns divided into bars and their competitors

A substantial technical report (#297) is in near final draft. (Tukey)
10. Designing graphics for impact

A presentation has been drafted and used.

11. Visual display in the years to come

An invited paper for the Sesquicentennial of the American Statistical Association has appeared in their proceedings and in *Statistical Science*. (Tukey)

12. Graphical display of distributions: distribution strips

A technical report, on scrawl-strips and letter or B-letter strips, has been revised and issued (#299). (Tukey and Veitch)

A first paper on strip displays with a rectangular reference has been completed and submitted. (Tukey and Tukey)

A second paper, on strip displays with other references, is in a late stage of drafting.

(Other papers are contemplated.)

13. "Lo--tech" dynamic displays

A paper has been completed and submitted

14. Graphical techniques in multiple comparisons

Work in this area forms part of two papers (The philosophy of multiple comparisons - - Where should multiple comparisons go next?) much of whose thrust is no explicitly graphical, but in which the importance of graphical output is stressed. (Tukey)
15. Graphical techniques in describing plant--breeding trials

Work with Kaye Hasford (University of Queensland, Australia) has made substantial progress.

16. Mapping and other cartographic techniques

Extended correspondence between Mark Monmonier (Syracuse) and Tukey is raising some new issues and helping to settle old ones.