CHAPTER 6

Testing Techniques

The choice of the physical method to be used for searching was important, but difficult to make. Since the work was entirely concerned with index languages, it was essential that the physical form of the index should in no way impede the investigation by introducing any controls or restrictions of its own. Although it was not possible to forecast exactly the many different tests that would be made, it was clear that for each question there would be the necessity of obtaining several hundred sets of performance figures.

It was decided that a small test should be made soon after the project had commenced; this was to be done partly to check the indexing procedures but also to validate the proposed design of the tests and to provide experience that would assist in deciding on the physical form of the index. For this pilot test, 116 documents had been indexed, and fourteen questions were available for searching, for which there were 26 known relevant documents. It was planned to investigate five sets of recall devices and four sets of precision devices, based on the single-term, natural language indexing. These variables alone appeared likely to result in some 80 searches for every question, and when other variables were added in the main test, the potential number of different searches could run into several hundreds. It was unlikely that every combination of the various devices would be required, but the method used had to be flexible enough to provide for all possible variations of searches, since it would only be after some searches had been made that it would be known which were unnecessary.

Co-ordination was certainly the basic precision device, and some form of post-coordinate index was clearly required. For the pilot test, the decision was taken to prepare a peek-a-boo type index. This was done in a conventional way, but a complication arose due to the fact that, at this stage of the work, six different indexing weights were being used, and, to investigate the effect of these, it was necessary to have, for every term, six cards each of which represented a different weighting.

The first search for a given question was carried out on the natural language terms . Subsequent searches were made bringing in the various recall devices and precision devices; the nature of these searches is considered in more detail later in this chapter. The results of this test were interesting in themselves, but the main objective had been to obtain information on the techniques being used. In this respect, the test showed that the general test theory was reasonable and that the indexing was satisfactory for the objectives of the test. Quite definitely, however, it showed that a peek-a-boo index would be quite unsatisfactory for the main test.

This was because much of the testing involved use of increasingly large numbers of terms in the search as the recall devices were tested, with the continual need for co-ordination of all the different combinations. For example, if a question had five terms searched on initially, and each of the five terms had one synonym, two word forms and four quasi-synonyms, then in coordination of all five terms using all the recall devices, 32,768 different combinations are possible. After this, it would be necessary to search for any four of the five sets of terms, then any three and so on. It is true that by use of the lowest posted terms first, the number of coordinations to be done can be reduced considerably, but the use of natural language

for the file, together with weights, resulted in serious difficulties. Another problem that loomed large was that of recording the aggregate of the different documents retrieved out of all the possible coordinations at a given ordination level, since many documents would be retrieved several times. One possible solution to these problems was to prepare a new peek-a-boo index for each of the recall devices; that is to say, there would be one index for natural language terms, a second index with the synonyms controlled, a third index with word endings confounded, etc. However the manual re-punching of new indexes would have been a big task, and at that time no equipment could be found to aggregate a set of postings from a number of different cards all on to one card. Other considerations mitigating against a peek-a-boo index were the task of withdrawing and refiling large numbers of cards during a search, and the difficulty of performing more than one search at one time.

As a result other conventional index forms were considered but offered no satisfactory solution. At this point in the project, several people working on associative retrieval expressed interest in the possibility of using the indexing being performed on our collection for their own testing of statistical associative techniques, clumping, etc. With the agreement of the National Science Foundation, arrangements were made to make the indexing available in machine readable form, on magnetic tape. The format used for this is given appendix 6.1, and details of supplementary tests being made are given in Chapter 7. With the indexing available on magnetic tape, the use of this for computer searching for the testing was then considered.

A number of discussions were held with various groups, and we received cost estimates for programming and searching which varied by a factor of ten An effort was made to discover whether any suitable computer programme already existed, which could be used to do the required searches. Discussions were held during a visit of one of the project staff to the U.S.A., but no suitable programme was discovered to do the minimum of what was required. This led to a reconsideration of preparing programmes in this country, but not only were the cost estimates high in relation to the present project, but also the time factor was becoming critical. Particularly discouraging was to learn that the searches which we had requested would result in seven million lines of print out; for these reasons and our own lack of experience in the field, the idea of using computers was abandoned.

The flirtation with computers had not been entirely wasted, for by this time we had a clear idea of exactly what was needed, and this helped in producing a method which met the main requirements. At the time when the solution was first proposed, no similar method was known to exist, for it is quite unconventional and it is difficult to visualise any application in real life circumstances. However, it was later discovered that a somewhat similar suggestion had been made by Dr. John O'Connor, known as the 'Scan-column index' (ref. 31), although no actual example of its use in practice is known. It had the advantages of flexibility to meet changing circumstances, so that it would give results for the many different types of search, and also of permitting quite complex analyses to be done clerically.

The first stage in the preparation of the index was a complete posting of each single term used in indexing on to a set of cards. These cards also contained information regarding the weights assigned to each term. The indexing decisions regarding Document 2076 are shown on the master indexing sheet in Fig. 6.1. From this sheet, the single terms and their weights were posted on to cards, with a separate card for each term. Thus 'Insulated 10', together with the document code number (2076) would be posted on one card, 'Two-dimensional 10' on another card together with a code number and so on to every index term. These cards were then sorted into alphabetical order and sub-sorted into document number within each term.

PARTITIONS & THEMES		INTERFIX & CONCEPTS	TERMS & WEIGHTS		INTERFIX & CONCEPTS	TERMS & WEIGHTS	
A	a d e f g	a	Insulated two-dimensional surface	Insulated 10	i	Laminar boundary layer displacement thickness	Distribution 8
B	b d e f g			Two-dimensional 10			Displacement 8
C	c d e f g	b	Insulated flat plate	Surface 10			Thickness 8
D	a d e f h effect of i g	c	Insulated curved two-dimensional surface circular arc profile	Flat 8	j	Laminar boundary layer momentum thickness	Momentum 8
E	a d e f h effect of j g			Plate 8			Skin 8
F	a d e f h effect of i g compared with l			Curved 8	k	Skin friction	Friction 8
G	a d e f h effect of j g compared with l	d	Hypersonic flow	Circular 8			Wind 7
H	a d e f k g compared with l	e	Compressible boundary layer	Arc 8	l	Wind tunnel tests at Mach 6.86	Tunnel 7
				Profile 8			Test 7
		f	Linear velocity profile	Hypersonic 9			Mach 7
				Flow 9			6.86 7
		g	Approximate calculation	Compressible 9			
				Laminar 9			
		h	Pressure distribution	Boundary 9			
				Layer 9			
				Linear 9			
				Velocity 9			
				Profile 9			
				Approximate 9			
				Calculation 9			
				Pressure 8			

FIGURE 6.1   MASTER INDEXING SHEET FOR DOCUMENT 2076

The 361 questions which it was proposed to use for searching produced a total of 723 different terms, and these became known as 'starting terms'. As such they were terms used in the questions without being subjected to any controls, and were equivalent to the natural language index terms. For each starting term a set of sheets was provided, these sheets bearing the document numbers 1001-2400. As an example, consider the starting term 'Flow'. The pack of cards which had been posted with this term was taken, and the information transferred from the cards to the set of sheets. The code 1 was used to denote that it was the actual search term (i.e. Flow) that was being posted and Figure 6.2, which is an extract from the set of sheets dealing with 'Flow', shows that a large number of documents were indexed by this term. In particular it can be seen that document 1933 was indexed by Flow at a weight of 9, as were documents 1939, 1940 and 1941. Document 1942 was also indexed by Flow, but on this occasion the weighting is 8. After all the indexing by Flow had been entered, additional entries were made for terms related to Flow. The authority sheet for this is shown in Fig. 6.3, from which it can be seen that Flux and Stream are considered as synonyms. The packs of cards posted for these terms would be taken, and entered on the sheets for Flow. Referring to Fig. 6.2, it will be seen that, for example, document 1978 is marked A6. This indicates that Flux, (which is coded A in Fig. 6.3) was indexed in this document at a weight of 6, while document 1974 is one of several that was coded by Stream(B) The variant word ending, Flowing, (coded E) was used in document 1968; of the quasi-synonyms shown in Fig. 6.3, Motion (K) and Moving (M) are examples which both appear in document 1978. It will be noted that multiple posting can occur on one document number; 1978 has, in addition to Motion and Moving, also been posted with Flow and Flux. The reason for doing this will be explained later.

The completion of this meant that there now existed a record of every time the starting term Flow or any of its synonyms, word endings and quasi-synonyms had been used as index terms. Since the codes for these were always kept constant (A-D for synonyms, E-J for word endings and K-Z for quasi-synonyms), the staff always know to which group any particular entry belonged.

The posting had been done on foolscap sheets and these were now cut into narrow strips, $\frac{3}{4}$ in. wide, each strip being serially numbered so as to maintain the document sequence order. These sets of strips were then filed in two specially constructed 'beehive' cabinets (Fig. 6.4).

In effect, a separate index was now compiled for each question by the preparation of a set of search sheets. The production of these in relation to a particular question was controlled by the question starting term card, an example for question 181 being shown in Fig. 6.5. This listed the starting terms for the question and the order of the terms on the search sheets, this order being of importance in relation to some of the searching options. To prepare the search sheets, the sets of strips for each of the starting terms were obtained and assembled one page at a time by being clipped to a set of 23 prepared boards. These boards showed the document numbers at the extreme sides, and the strips were arranged in correct alignment with the numbers. When all 23 boards had been thus prepared, a xerox copy was made of each board; the result is shown in Fig. 6.6, which illustrates one of the 23 sheets for question 181 in relation to documents 1931-1992.

(13)	(14)	(15)	(16)	(17)	(18)
1745	1807 K6	1869 1.7	1931	1993 1.9 B9	2055
1746 B6	1808 1.9	1870	1932	1994 1.9 B9	2056
1747	1809 1.9	1871	1933 1.9	1995 1.9 B10	2057
1748 1.9	1810	1872 1.9	1934	1996 1.9	2058
1749 1.9	1811	1873 1.9 M9	1935	1997 B.10	2059
1750	1812	1874	1936	1998 1.9 B5	2060
1751 1.9	1813 K10	1875	1937	1999 1.9	2061 B9
1752 1.8	1814 1.6	1876	1938	2000 1.9 K8	2062
1753 1.9	1815 1.9	1877	1939 1.9	2001	2063
1754 1.8	1816 1.9	1878	1940 1.9	2002 1.9	2064 1.8
1755	1817 1.9	1879	1941 1.9 K9	2003 1.9	2065
1756	1818 1.9	1880 1.7	1942 1.8	2004	2066
1757 1.9 L8	1819	1881	1943 1.9	2005	2067
1758	1820 1.9	1882	1944	2006 1.10	2068
1759	1821	1883	1945 1.9 K9	2007 1.9	2069
1760	1822	1884	1946 1.9 B9	2008	2070
1761	1823	1885	1947	2009 1.9	2071
1762	1824	1886	1948 1.9	2010 1.9	2072 B8 M8
1763	1825 1.9	1887	1949 1.10	2011	2073
1764 K9	1826	1888	1950	2012	2074 1.9
1765	1827	1889	1951	2013	2075 1.9
1766	1828	1890	1952	2014 1.7	2076 1.9
1767	1829	1891	1953	2015	2077
1768	1830	1892 1.7 K9	1954	2016	2078 1.10
1769	1831	1893 1.5	1955	2017	2079
1770 1.9	1832	1894 B.9	1956	2018	2080 1.9 B8
1771 1.9	1833	1895 1.9	1957	2019	2081 1.9 K7
1772 1.9	1834	1896	1958 1.7	2020	2082 1.9 B8
1773 1.9	1835	1897	1959 1.9	2021	2083 1.9
1774 1.9	1836	1898	1960 1.9	2022	2084 1.9 B6
1775 1.9 B8	1837	1899 1.7 M9	1961	2023	2085
1776 1.7	1838	1900 1.9	1962 1.7 B9	2024	2086
1777	1839	1901 1.9	1963 1.8	2025	2087
1778	1840	1902 1.9 K9	1964 1.8	2026	2088
1779 1.7	1841	1903 1.9	1965 1.9	2027	2089
1780	1842	1904	1966 1.9	2028	2090
1781 1.7	1843	1905	1967 1.9	2029	2091
1782	1844	1906 1.7	1968	2030	2092
1783	1845	1907 1.7	1969 1.9 B9	2031	2093 1.9 B9
1784 1.9	1846	1908 1.6	1970 1.9 B9	2032	2094
1785 1.9	1847	1909	1971 1.9	2033	2095
1786	1848	1910	1972 1.9 0.9	2034	2096 A8
1787 K8 M9	1849	1911	1973 B9 1.9	2035	2097
1788 K9 M8	1850	1912 1.9	1974 B.9	2036	2098 A6
1789 B.9	1851	1913	1975 1.9	2037	2099 1.6 B8
1790 1.9	1852	1914 1.7	1976	2038	2100 1.6 B8
1791 1.9 K5	1853	1915 1.9	1977 1.9 K9	2039	2101
1792	1854	1916 1.9	1978 1.7 A6 K7 M9	2040 1.9	2102
1793 1.8	1855 1.9	1917 1.9	1979 1.9	2041	2103
1794 1.9 L7	1856 1.9	1918 1.9	1980 1.7	2042	2104 1.6
1795 1.9	1857	1919 1.9	1981 1.9	2043	2105 1.9 B8
1796 1.6	1858 1.9	1920 1.9	1982	2044	2106 1.9
1797	1859	1921 1.7	1983	2045	2107 1.9
1798 1.9	1860	1922 1.9	1984 1.9	2046	2108 1.9
1799 1.7 K7	1861 1.9	1923 1.9	1985 1.9 B9	2047 K8	2109 B7
1800 1.8	1862	1924 1.9	1986 1.9	2048	2110 1.9 K10
1801 L8	1863	1925 1.9 B5	1987 1.9 B9	2049	2111
1802 1.9	1864	1926	1988 1.9 B8	2050	2112 1.9 K6
1803 1.9	1865	1927 1.9	1989 1.9	2051	2113
1804 1.9	1866	1928	1990 1.9	2052	2114
1805	1867	1929	1991 1.9 B5	2053	2115
1806	1868	1930	1992 1.9	2054	2116
(13)	(14)	(15)	(16)	(17)	(18)

FIGURE 6.2   POSTING SHEET FOR 'FLOW' IN RELATION TO DOCUMENTS 1745-2116

Starting term	1	FLOW
Synonyms	A	FLUX
		STREAM
Word endings	E	FLOWING
Quasi-synonyms		MOTION
	L	MOVEMENT
		MOVING
	N	FLOWING

FIGURE 6.3   STARTING TERM AUTHORITY SHEET
SHOWING TERMS RELATED TO 'FLOW'
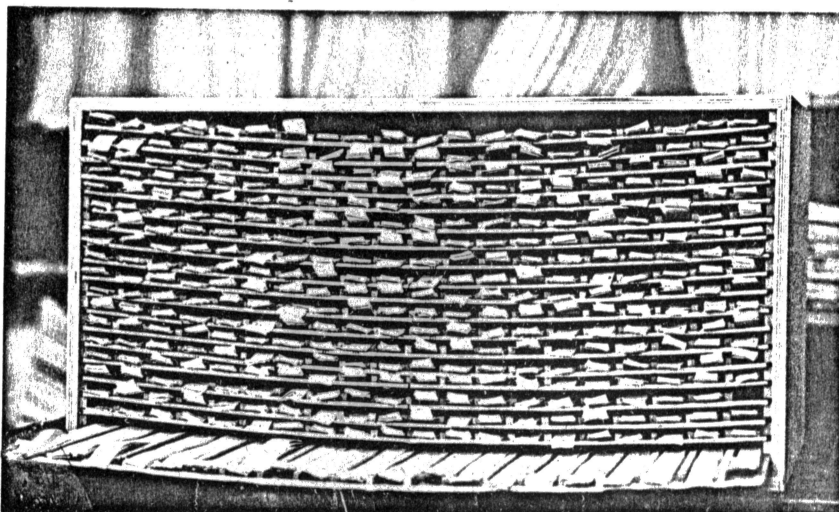
FIGURE 6.4   THE 'BEEHIVE' FILING CABINET

a	CHANNEL
b	VISCOUS
c	COMPRESSIBLE
d	FLOW
e	STRAIGHT
f	DETERMINATION
g	NATURE

FIGURE 6.5   SEARCH STARTING TERMS FOR
QUESTION 181

	Nature	Flow	Compressible	Viscous	Channel	Straight	Determination
1931							
1932			H.10				L.7
1933		1-9					
1934			H.10				1.7
1935			H.6				
1936							
1937							
1938			H.10				
1939		1.9	1.9		1.10		
1940		1.9					L.9
1941		1.9, K.9	1.10 N.5	1.10 E.8	1.10		
1942		1.8	M.8				
1943		1.9	1.9				
1944							
1945	K.5	1.9, K.9					
1946		1.9, B.9	N.9				
1947	L.10		N.10				L.6
1948		1.9	N.9				
1949	K.10	1.10					
1950						1.9	L.9
1951							L.8
1952							L.9
1953							
1954							
1955							
1956			H.10				
1957			J.6				
1958		1.7	H.8				
1959		1.9					
1960		1.9					
1961			1.10				
1962		1.7, B.9	M.7				
1963		1.8			1.8		
1964		1.8					1.9
1965		1.9					
1966	K.7	1.9	F.10		1.10		
1967		1.9	1.9	1.10			
1968	L.7	E.9					
1969		1.9, B.9	M.9				
1970		1.9, B.9	K.9 M.9				
1971		1.9	M.9 N.9				
1972		1.9, B.9	N.9				
1973		1.9, B.9	M.9				
1974		B.9	M.9				
1975	L.6	1.9					1.6
1976			N.9				
1977		1.9, K.9				1.9	
1978		1.7 A.6 K.7 M.9	N.6				
1979		1.9	M.9				
1980		1.7					
1981	K.9	1.9					
1982	K.7		N.10				
1983			N.10				
1984		1.9	1.10 J.10				
1985	L.9	1.9, B.9	1.10, J.10				
1986		1.9	J.10				
1987		1.9, B.9	J.10, K.9 M.9				
1988		1.9, B.8					
1989		1.9					
1990		1.9	1.9 J.10				
1991		1.9 B.5	L.10, M.5				
1992		1.9	M.9	1.8			
(16)							

FIGURE 6.6   SEARCH SHEET FOR QUESTION 181 IN RELATION
TO DOCUMENTS 1931-1992

It can be seen that for the search term Flow, the appropriate information which was first posted on the sheet shown in Fig. 6.1 for documents 1931-1992 has now been included in the second column of Fig. 6.6. The information relating to the other starting terms would have come from similar strips. As an example, the search sheet reveals that in document 1966 Nature did not appear, but the quasi-synonym Property (coded K) was indexed at a weight of 7. Flow was indexed at a weight of 9. Compressible did not appear, but it was present in the variant word form Compressibility (F) with a weight of 10, while Channel was indexed at a weight of 10. The remaining three starting terms did not appear in any way in this document.

When the search sheets had been printed, the 'boards' were dismantled, the strips sorted into order and redistributed into the beehive ready for further use with another search question. The boards finally used were of rigid hardboard, together with 'bulldog' type clips; earlier trials with cardboard sheets and perspex covers had failed because the strips moved out of position too easily. The time taken to mount a question on to the boards varied with the number of starting terms, but usually took between thirty and sixty minutes. The xeroxing and checking took ten to fifteen minutes, and redistribution of the strips a further ten to fifteen minutes. A minority of questions had more than eleven starting terms, and therefore needed two sets of sheets. It was usually possible to pick two questions with quite different sets of starting terms, so that both questions could be prepared at the same time. A system of double checking the search sheets was used to correct any errors which occurred; these were usually due to misfiling of individual strips in the re-distribution stage. While this method might seem cumbersome, it appears to have been justified by results, since it gave the flexibility that was required, and although expensive in man-hours was relatively cheap compared to what would have been the cost for any form of machine searches.

The end result of this exercise was that we had 361 sets of search sheets, 23 sheets in each set, posted with all the occurrences of the terms to be used in searching each question; there were, in fact, 361 question-indexes, and it was now possible to carry out the first series of searches. These were performed on single terms, and investigated three variables.

1. The recall devices of synonyms, word endings and quasi-synonyms, tested in six aggregations (known as 'index languages').

2. The precision device of simple coordination without any linking in the indexing, where the search rules allowed any combination of terms to be accepted, and every level of matching to be recorded.

3. The three levels of indexing exhaustivity, indicated by the weights (5-6, 7-8 and 9-10).

The six index languages investigated in the first series of tests were as follows:

Index Language	
1	Natural language terms (code 1)
2	Natural language terms + synonyms (codes 1 and A-D)
3	Natural language terms + word forms (codes 1 and E-J)
4	Natural language terms + synonyms + word forms (codes 1, A-D and E-J)
5	Natural language terms + synonyms + quasi-synonyms (codes 1, A-D and K-Z)
6	Natural language terms + synonyms + word forms + quasi-synonyms (codes 1, A-D, E-J and K-Z)

These six index languages appeared to cover all reasonable permutations, since it was not logical, for instance, to contemplate the use of quasi-synonyms without the use of synonyms.

The searches were carried out by clerical labour, and the results were recorded on a score sheet as shown in Fig. 6.7. The actual operation of carrying out a search became known as 'putting the ruler down the sheets', since the use of a straight edge to successfully uncover the postings for each document was found to be the best method. The searches were made on the sets of search sheets (as in Fig. 6.6), where each vertical column deals with one of the question starting terms, and shows not only the occurrence of the starting term itself, but also the related terms as described earlier. Often an examination of the postings for a certain question needed some care in working out, since in one operation the search results would be recorded for the six different index languages and for the three weights. However, after a relatively short learning period, the clerical staff had no serious difficulties. The time required to search a single question varied greatly; with this particular set of six index languages, it might be anything from ten minutes to one hour, being dependent on the number of starting terms, the frequencies of postings for each starting term, and the number of terms related to the starting terms.

The score sheets list the document numbers on the left hand side, and across the sheet space is given for recording the coordination level (i.e. the number of search terms that match with the document terms) of each document for each of the six index languages at each of the three levels of exhaustivity. The way this is done may be seen by examining a search sheet (Fig. 6.6) for question 181 'Has any work been done on determination of the nature of compressible viscous flow in a straight channel', in relation particularly to documents 1963, 1966 and 1978.

The search sheet shows that document 1963 has two of the search terms present, and a look at the codes shows that they are coded 1, the natural language terms, which are included in all six languages. Both terms have a weight of 8, and therefore do not come out at the lowest exhaustivity (weights 9 or 10), but do at the medium and high levels. The score sheet (Fig. 6.7) records this, the coordination score of 2 being put in every language at the medium and high levels of exhaustivity. Document 1966 has four of the search terms present; two natural language, (1) one word ending (F) and one quasi-synonym (K). So taking the highest level of exhaustivity (5-10), every index language will have a coordination score of at least 2; Index languages 3 and 4 will score 3, (1, 1 and F); Index language 5 will also score 3, (1, 1 and K), but Index language 6 scores the maximum, 4, (1, 1, F and K) since it accepts both word ending variants and quasi-synonyms. Considering now the various levels of exhaustivity, index languages 1 to 4 have all their terms weighted 9 or 10, and so keep the same coordination score at medium and low exhaustivity, but index languages 5 and 6 have the quasi-synonym weighted 7, so at low exhaustivity the coordination score drops to 2 and 3 respectively.

As a final example, for document 1978, one of the two search terms (Flow) is shown to be present in natural language at a weight of 7, as a synonym (A-6) and also as two quasi-synonyms (K-7 and M-9). All these, of course, only count as a coordinate score of one since they are all separate alternatives to one of the search terms, but the last quasi-synonym (M-9) is important because it is the only term at low exhaustivity. The coordination scores for this document in table 6.3 are 1 for index languages 1 to 4, and 2 for languages 5 and 6, with exhaustivity reducing these scores as shown.

Since the search rules at this stage allowed any combination of terms to be accepted, it was never necessary to note which search terms occurred. Some combinations accepted were obviously nonsense, e.g. document 1982 retrieved by the starting terms Nature and Compressible is not meaningful, and is even worse when the quasi-synonyms

Question 181

Index Languages	1 1			2 1, A-D			3 1, E-J			4 1, A-D, E-J			5 1, A-D, K-Z			6 1, A-Z		
Documents	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10
1956	-	-	-	-	-	-	1	1	1	1	1	1	-	-	-	1	1	1
1957	-	-	-	-	-	-	1	-	-	1	-	-	-	-	-	1	-	-
1958	-	-	-	-	-	-	2	2	-	2	2	-	-	-	-	2	2	-
1959	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1960	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1961	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1962	1	1	-	1	1	1	1	1	-	1	1	1	2	2	1	2	2	1
1963	2	2	-	2	2	-	2	2	-	2	2	-	2	2	-	2	2	-
1964	1	1	-	1	1	-	1	1	-	1	1	-	1	1	-	1	1	-
1965	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1966	2	2	2	2	2	2	3	3	3	3	3	3	3	3	2	4	4	3
1967	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1968	-	-	-	-	-	-	-	-	-	-	-	-	1	1	-	1	1	-
1969	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1970	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1971	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1972	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1973	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1974	-	-	-	1	1	1	-	-	-	1	1	1	2	2	2	2	2	2
1975	2	1	1	2	1	1	2	1	1	2	1	1	3	1	1	3	1	1
1976	-	-	-	-	-	-	-	-	-	-	-	-	1	1	1	1	1	1
1977	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1978	1	1	-	1	1	-	1	1	-	1	1	-	2	1	1	2	1	1
1979	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1980	1	1	-	1	1	-	1	1	-	1	1	-	1	1	-	1	1	-
1981	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1982	-	-	-	-	-	-	-	-	-	-	-	-	2	2	1	2	2	1
1983	-	-	-	-	-	-	-	-	-	-	-	-	1	1	1	1	1	1
1984	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1985	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
1986	1	1	1	1	1	1	2	2	2	2	2	2	1	1	1	2	2	2
1987	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
1988	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1989	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1990	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
1991	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2
1992	2	2	2	2	2	2	2	2	1	2	2	1	1	3	2	3	3	2

FIGURE 6.7 SCORE SHEET FOR QUESTION 181 IN RELATION
TO DOCUMENTS 1956-1992

on which it was retrieved are decoded as Property and Hypersonic. Intellect
was put in on later searches, to eliminate such unwanted combinations.

Contrary to the example shown in Fig. 6.7, in practice the score sheets for a
question rarely recorded documents with only one search term present, since this
would usually have involved recording the large majority of the documents in the
collection. The decision as to what coordination score to begin recording documents
varied for each question, depending partly on the number of starting terms in the
question. The objective was to examine an average of about 100 documents from the
collection (involving two or three score sheets), and this decision was fairly easily made
by looking at the density of postings on the search sheets. In some cases, when postings
were very heavy, a proportion of the collection only was examined (e.g. if half the
collection, the odd or even numbered documents only, etc.), and the results scaled up.
This was done to reduce the large clerical effort involved in searching so many questions
this way (involving looking at nearly 400,000 'documents' on the search sheets in this
first series of tests alone), but was only done when the results were statistically valid.
An exception to this was that the relevant documents were always fully recorded.

To obtain the final results for a question, the documents which had been assessed
as relevant were recorded on a separate score sheet, and deleted from those first
produced. The base document for the question being tested was deleted altogether
at this stage. Then the actual numbers of relevant and non-relevant documents
were totalled up, a separate total being obtained for each index language, at all
coordination levels and at each exhaustivity level. The final record is seen on
a Results Sheet, (Fig. 6.8). Here, for question 181, it is noted that the Search
rule is type A which, as stated previously, allowed any combination of terms to be
accepted; the question has 7 starting terms. The search sheets were examined
for all documents having a coordination score of 3 or more, and there are two
relevant documents sought in this question. Three tables of figures are given, for
the three levels of exhaustivity, each table recording the coordination score and
language variables. For example, using the highly exhaustive indexing (weights
5-10), a three term coordination score using language 3 retrieves both of the relevant
documents, and 60 non-relevant documents. At the next level of exhaustivity
(weights 7-10), the non-relevant documents drop to 45; at the lowest level of
exhaustivity, the non-relevant documents drop to 10. In this case the recall is
maintained throughout, but with index language 6, for instance, at a coordination
score of 4, the effect of moving from high exhaustivity to low exhaustivity is to lose
the one relevant document retrieved. It will be noticed that no non-relevant figures
are given for coordination scores 1+ and 2+, although the relevant documents are
shown here. In general, an attempt was made to cut down the clerical effort by
ignoring the count of non-relevant documents when the precision ratio was less than
3%, although, as will be recounted in the next volume, some sampling was done at
these low precision levels. The figures obtained from this particular question are
then ready to be totalled with those from other questions to provide results for a set
of questions. This, and the various methods for arriving at these totals, will be
considered in the next volume.

There were many additional tests, in which were investigated the effect of such
matters as the single term hierarchies, the set of concept languages, again incor-
porating the various recall devices such as alphabetical and hierarchical grouping,
and also the various searches with controlled terms. These other tests meant, of
course, that the preparation of the question-indexes had to be commenced from the
beginning. For instance, the single-term hierarchies resulted in a group of terms

QUESTION 181

SEARCH RULE A   RELEVANT DOCUMENTS 2

STARTING TERMS 7   LOWEST COORDINATION 3   DOCUMENTS TO ALL
FOR NON-RELEVANT   BE EXAMINED.

INDEX LANGUAGES 1-6

HIGH EXHAUSTIVITY (WEIGHTS 5-10)

COORDINATION LEVELS

INDEX LANGUAGE	1+		2+		3+		4+		5+	
	R	N-R	R	N-R	R	N-R	R	N-R	R	N-R
1	2	2	2		1	25	-	-	-	-
2	2		2		1	35	-	-	-	-
3	2		2		2	60	-	2	-	-
4	2		2		2	60	-	6	-	-
5	2		2		2	235	-	32	-	2
6	2		2		2	265	1	42	-	4

MEDIUM EXHAUSTIVITY (WEIGHTS 7-10)

INDEX LANGUAGE	1+		2+		3+		4+		5+	
1	2		2		1	20	-	-	-	-
2	2		2		1	30	-	-	-	-
3	2		2		2	45	-	2	-	-
4	2		2		2	45	-	4	-	-
5	2		2		2	175	-	20	-	-
6	2		2		2	195	1	26	-	-

LOW EXHAUSTIVITY (WEIGHTS 9-10)

INDEX LANGUAGE	1+		2+		3+		4+	
1	2		2		1	5	-	-
2	2		2		1	10	-	-
3	2		2		2	10	-	-
4	2		2		2	10	-	-
5	2		2		1	90	-	10
6	2		2		2	100	-	10

FIGURE 6.8   RESULTS SHEET FOR QUESTION 181 FOR
INDEX LANGUAGES 1 to 6

associated with a starting-term that was different from the group of synonym, word endings and quasi-synonym described earlier. There were some minor modifications in preparing the indexes, but in general the basic procedure described above was used for this further testing.

There was the additional necessity of investigating, on the single terms, the precision devices of interfixing and partitioning, which, as described earlier, are the two stages of links which were recognised, interfixing being concerned with single terms within a concept, while partitioning deals with concepts within a theme. This operation was done by examining the original indexing sheets for the relevant and non-relevant documents that had been retrieved as a result of the searches described above.

To illustrate the procedure adopted, Fig. 6.9 shows the processing of one of the relevant documents (2076) to question 51. This question has eleven starting terms; these are set out at the top of the table, with the double dividing lines indicating the concepts into which the question terms are divided, namely Displacement-Thickness; Plate-Flat; Flow-Compressible; Boundary-Layer-Laminar; Formula-Approximate. These concepts are the pairs and triplets of terms which must be interfixed within concepts. In testing partitioning, all the terms in the search are demanded to occur in one theme of the indexing. Each asterisked term in Fig. 6.9 is the basic term in its concept, and the search rules in operation at this stage of the test demanded that no subsidiary term (i.e. non-asterisked term) would be accepted unless the basic term was present. Thus in the index terms contained in document 2076 listed in the second row, the last term Approximate is not accepted, since Formula is not present. This row shows all the index terms in document 2076 that match with the terms requested in the search prescription, with the weights in brackets, this information resulting from decoding the entries on the search sheet. The index sheet of document 2076 (fig. 6.1) is examined next, the index terms in row 2 are located in the indexing, and the code letters assigned to the concepts in the indexing are recorded in the third row. The first two terms, Displacement and Thickness, both occur in Concept i, and therefore are interfixed; the fourth and fifth terms, Flow and Compressible, occur respectively in concepts d and e, so no interfixing is present. However, an alternative quasi-synonym acceptable in place of Compressible is Hypersonic; this occurs in concept d and thus interfixes with Flow. The fourth row shows the themes from the indexing that contain the greatest number of search terms; theme 02 does not include 'Displacement Thickness'; while theme 04 has this concept, it does not include 'Plate Flat', so both themes give the same results, since both eliminate one concept of two terms. From this data the results can be calculated for interfixing, for partitioning and for partitioning with interfixing, in all of the six index languages and at the three levels of exhaustivity. The results for this single document in regard to these devices are shown on the score sheet (fig. 6.10). This procedure was carried out on all the relevant documents in the questions tested, and also several of the non-relevant documents were examined. The totals of relevant and non-relevant documents for a question are again recorded on a results sheet as before, and from this can be seen the effect on recall and precision of these powerful precision devices.

The testing of the simple concepts involved more index languages than the single terms, since 16 aggregates of recall devices were tested. In this case the code letters used in the columns were each allotted to a single device, rather than a group of letters to a device. (e.g. B was synonyms, C was species, so that even if there were five synonyms or five species, they were all coded with B or C). This was done not only because of the large number of separate results wanted, but because the search

QUESTION 51   DOCUMENT 2076 (Relevance 2)

	*DISPLACEMENT	THICKNESS	*PLATE	FLAT	*FLOW	COMPRESSIBLE	BOUNDARY	*LAYER	LAMINAR	*FORMULA	APPROXIMATE
1. Starting terms	*DISPLACEMENT	THICKNESS	*PLATE	FLAT	*FLOW	COMPRESSIBLE	BOUNDARY	*LAYER	LAMINAR	*FORMULA	APPROXIMATE
2. Search terms in indexing of Doc. 2076, with weights	DISPLACEMENT (8)	THICKNESS (8)	PLATE (8)	FLAT (8)	FLOW (9)	COMPRESSIBLE (9) HYPERSONIC (N9)	BOUNDARY (9)	LAYER (9) FLOW (K9)	LAMINAR (9)		APPROXIMATE (9) but not accepted as FORMULA is not present
3. Concept codes assigned in indexing, showing interfixing	i	i	b	b	d	e d(N9)	e i j	e i j	e i j	-	-
4. Themes assigned in indexing, showing partitioning	Theme B - Theme D i	Theme B bdefg Theme D adefhig i	b -	b -	d d	e d d	e e	e e	e e	- -	- -

FIGURE 6.5   PROCESSING OF DOCUMENT 2076 IN RELATION TO QUESTION 51 FOR ANALYSIS OF INTERFIXING AND PARTITIONING

*   These terms must be present for other terms in the concept to be accepted.

QUESTION 51. DOCUMENT 2076

INDEX LANGUAGE Weights	1			2			3			4			5			6		
	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10	5-10	7-10	9-10
NO LINKS	9	9	5	9	9	5	9	9	5	9	9	5	9	9	5	9	9	5
INTERFIXING	8	8	4	8	8	4	8	8	4	8	8	4	9	9	5	9	9	5
PARTITIONING	7	7	5	7	7	5	7	7	5	7	7	5	7	7	5	7	7	5
INTERFIXING & PARTITIONING	6	6	4	6	6	4	6	6	4	6	6	4	7	7	5	7	7	5

FIGURE 6.10 SCORE SHEET FOR LINKS WITH
DOCUMENT 2076 FOR QUESTION 51.

STARTING TERMS

a. Compressible Flow

b. Viscous Flow

c. Channels

d. Straightness

SEARCH REQUIREMENTS

Only the following combinations of terms will be accepted.

2-term coordination   ac, bc. cd

3-term coordination   abc, acd, bcd

4-term coordination   abcd

FIGURE 6.11 INSTRUCTION SHEET FOR SEARCH WITH
CONTROLLED TERM VOCABULARY

prescriptions contained more related terms than the single term searches did, and would have required more divisions than the 26 in a single letter code. Another answer to the posting problem was not to post any related terms on a document when the natural language term or synonym term (both included in every aggregate of devices) was already there. This could be done provided that a related term did not improve the weights. For example, in document 1978 in fig. 6.6, Flow appears as such as 1-7. Because of this, A6 and K7 are really redundant, but on the other hand the posting of Moving (M) at a weight of 9 is required since this improves the performance in regards to weighting. This superfluous posting was done deliberately on the single terms to enable decoding of all search terms for the interfixing test, but no such requirement existed in the concept searches, and such posting was left off.

As stated, the first series of tests had been done using the minimum of intellect in the search programmes, with the result that many documents were retrieved on nonsensical combinations of terms. At later stages in the test, increasing intelligence was put into the search programmes; this is another way of saying that the requirements were more stringent. This was done in various ways, and each time the attempt was made to identify the particular intellectual decision which had been taken. One example of this is given in Fig. 6.11, where the search was being carried out on the Controlled Term Vocabulary. There are four starting terms, Compressible flow, Viscous flow, Channels and Straightness. Instead of any combination of these being accepted at the various levels of coordination, the search instructions specifically state, for instance, that Compressible flow and Viscous flow are not acceptable on their own. In fact, the definite requirement is that Channels must always be present.

This chapter has only considered the general techniques which were used in carrying out the tests. Quite inapplicable as far as can be seen to any operational situation, they gave, albeit with a large amount of clerical effort, all the flexibility that was required. One point which should be made clear concerns the prior knowledge regarding which documents were relevant to which question. This knowledge was not available to the indexers at the time of indexing, so therefore there is no question of the indexing being slanted towards a particular question. In theory it could have been available to Mills at the time when he was preparing the groups of related terms and the various hierarchies. In fact, Mills was doing this work in London while the indexes were being prepared and the searches were being carried out 50 miles away at Cranfield. Even if he had had access to this data and had attempted to use it in preparing these lists, we do not believe it would have made any significant difference to the results. With regard to the searching, the description given in this chapter of the methods used should make it obvious that its comprehensive nature precluded any possibility of influencing the results.