

Comprehensive Historical Indexing: The Virginia Gazette Index

By STELLA DUFF NEIMAN AND LESTER J. CAPPON

Institute of Early American History and Culture

IF the compilers had been requested to discuss a year ago their processes of indexing in the preparation of the *Virginia Gazette Index*, published in the fall of 1950 by the Institute of Early American History and Culture, Williamsburg, Virginia, they would have met the idea with something less than enthusiasm. The actual publication of the two volumes, however, brings with it, besides the knowledge that our experience might be useful to someone else undertaking a lengthy indexing project, the courage to contemplate how it was done.

The work was designed to make the research labors of historians in early American history easier. Use of the newspaper as a primary source was developed comparatively late in the historiography of the American colonial period. This was undoubtedly because of the tiresome effort involved in searching pages of fine type for an isolated fact or two even though these facts might be of tremendous significance. An index to the *Virginia Gazettes* appealed to Colonial Williamsburg, Incorporated, its first sponsors, not only for this reason but because it would facilitate investigations into the location and nature of the buildings of Williamsburg which they were in the process of restoring. The Research Department of the organization therefore undertook a detailed index to the newspapers published in Williamsburg from 1736 to 1780.

After 1766 there were always two papers with the same title published by different printers. Three *Virginia Gazettes* competed for the public's attention in 1775 and part of 1776. The series has several gaps. No papers are now extant between 1747 and 1751, few issues by William Rind have survived for 1766 and 1767, and numerous single issues are missing throughout. Although the character of the news does not vary much in concurrent papers after 1765, its emphasis changes from the predominantly British reprints of William Parks, 1736-1746, to coverage of all colonies and British-American politics in the period after the Stamp Act of 1765.

The four-page issue of the average paper was crammed with information to be sorted out and entered by the indexer. The principles used in this process of selection were painstakingly developed by trial and error since there was no precedent in large-scale historical indexing. The system evolved could in essence provide a guide to any extensive project, and it is with this in mind that we offer a study of our methods.

Four principal factors should be taken into consideration when one is embarking on an index covering, in this case, approximately 6,200 large pages of material.¹ These are: finances, personnel, general policies of indexing, and specific procedures.

Financing of the undertaking is closely connected with the amount of time necessary to complete it. This, of course, is difficult to estimate and even a person fully acquainted in the field of the documents probably will not be able to guess at the duration with any degree of accuracy. A guess it will almost certainly be since there is a multitude of indeterminable factors involved in the process. This would be true even if an index of another colonial paper were under discussion because of variations among such publications, but the benefit of our experience would in some ways shorten the project. Probably the most satisfactory way of estimating the time required is to make a sample index of a portion of material and then to use the number of hours as a rough basis for computing the total time necessary. This test will also provide a basis for deciding how much money will be needed for the compilation. It will be a large amount, and there are very few ways of cutting it down once a decision has been reached as to the type of index to be made. It is important that this factor be clarified at the beginning and properly provided for. Often people outside the research field do not comprehend the complicated nature of such work. They may expect less experienced individuals, who offer a more rapid work schedule at less cost, to complete as worthy a product as those whose estimates seem high but in the long run are correct.

This point leads us to a discussion of the personnel problem. The number of documents involved plus the time limit will determine the number of indexers and typists employed. Three indexers, one of whom was in charge, and two typists proved most satisfactory during the major production period of the *Virginia Gazette Index*. On any large project this ought to be a workable number. It is essential to have one senior indexer who will be the arbiter in discus-

¹ There were about 1,700 issues, including numerous supplements, of the *Virginia Gazette* extant.

sions and maintain the greatest possible degree of uniformity throughout the work. Everything indexed ought to pass through his hands for checking since even with good instruction and the establishment of a body of rules,² it is impossible for three people to think exactly alike on such specific problems as are involved, for example, in the choice of main headings. If the indexers are reduced to two, no corresponding decrease can be effected among the typists since they must type, file, and check one another's work.

It is difficult to tell whether or not an individual will be a good indexer. Since few persons are already trained for such work, it is impossible to hire one of that "species" at a good employment agency. It is not so hard to delineate the qualifications, but whether a given person has them or not is another question. The indexer on a comprehensive technical project should have some background in the field germane to the subject. In our case, he did not need to know a great many facts about Virginia history or trades in the colonial period, but a general knowledge of the field of history and, preferably, of American history was essential. We found it impossible to use effectively a biology major, although she was an intelligent girl, simply because it took her too long to orient herself when she was faced with a discussion of the causes of the American Revolution. To find out whether an individual is familiar with the field is, of course, not the difficult part of this problem. A simple inquiry may settle that. We wish to stress, however, that a search until the right person has been found, will yield compensation in time and money, even at a higher salary, and in saving wear and tear on the senior indexer's disposition.

Along with assurance of the applicant's knowledge of subject matter must be combined an estimate of character and personality to find out whether he fulfills the remaining qualifications. No matter how much of an expert in early American history an indexer of the *Virginia Gazette* might have been, if he did not have an orderly mind and, above all, one which could see the relative importance of facts against the background of all history, we were better off without him. He must understand what he reads, not merely in itself but in relation to the established scheme of indexing. In addition, he must have the willingness and patience to concern himself with facts which, though of no special interest to him, some future investigator will be delighted to unearth. Lastly, the indexer must be able to phrase the entries in the briefest possible terms. Wordi-

² See p. 295 ff.

ness is not at home in an index. The importance of securing persons with these four qualifications cannot be overemphasized. Respectable salaries are a very good drawing card, and it is well worth while to pay them. The job cannot be done by a hack because it is "only an index." Any attempt at economy in wages, which results in sacrificing high standards, will only increase the cost of the project. If the senior indexer has to revise the entire work of an incompetent assistant, a double loss is involved.

Typists should be reasonably rapid, but they need not be speed demons. Since their duties will include filing, it is more important that they be intelligent enough to learn a complicated system of alphabetization than that they have a speed of 90 words per minute. With proper training and encouragement to take an interest in the project, typists can become very useful in catching clerical mistakes and errors in uniformity made by the indexers.

With the assembly of a good staff, the directors of the project must turn to their most difficult problem — the consideration of a general policy to guide the indexers in selecting material. The question, what to index, is the very essence of the task, and long and thoughtful deliberation must be given to it. Confronted with a mass of material — thousands, indeed millions, of facts — what does one choose to record for the research worker? To answer this question one must ask others. First, is the index for the use of some specific group? If so, then the problem is much simplified. The desires of this group and the means to satisfy them may be readily ascertained. When the *Virginia Gazette Index* was first undertaken, it was conceived as an aid to the research workers of Colonial Williamsburg, Incorporated, in finding information about the buildings, furnishings, and mode of life in eighteenth-century Williamsburg and Virginia. This made necessary very detailed indexing of the advertisements, even including architectural details like types of fences and kinds of paneling in houses. If one agrees to provide information for a particular group of specialists, they may dictate the details to be indexed. If they are genealogists who want merely a list of all personal names, the problem of selection is reduced to its simplest terms.

It is hoped, however, that if the material is worth indexing, the sponsors will feel it should be done as a broader *historical* work of reference. It would then provide assistance to people with many various types of research interests, not merely to one group alone whose needs may soon be satisfied. If this wider goal is set, two

questions more difficult to answer than those posed by a restricted index must be asked: what information is found in this material that is unique to it; and, is the remainder also of interest because it appears in this source? Intelligent answers to these questions depend upon familiarity with the contents of the documents. A few days spent reading and considering them against their background are necessary at this point. In the case of the *Virginia Gazettes*, it was noted that no colonial newspaper had ever been indexed, and therefore we could not depend on any work to cover part of ours. In addition, all of the papers published foreign and other colonial news to a greater degree than news of Virginia. For the latter, the advertisements had to be relied on; indeed, they furnished much local lore which would scarcely be preserved elsewhere.

When the information unique in the material to be indexed has been determined by this inspection, rules should be set up to analyze it completely for subject, personal name, and place. In the *Virginia Gazettes* an advertisement for a house for sale in Williamsburg was indexed under: Houses for sale; the name of the advertiser; the types of buildings on the lot; and the name of the estate, if any. All reports of elections to the Virginia House of Burgesses were indexed under the names of the persons elected and under the counties which they represented. Cross references from Elections to the names of places were used instead of entering each account under such a broad heading.³

Some items, on the other hand, will offer a problem in differentiation. These include facts for which no one would search in the source under consideration, either because they are not original in it, do not contribute to a better understanding of its background, or aid in interpreting it. Such details can be logically omitted. Announcements of awards by the British court, copied in the *Virginia Gazettes* from the London papers, were omitted in our plan. No historian would cite a colonial newspaper for a notice about the presentation of the Order of the Garter to Sir John Jones. Reports of speeches in parliament on a subject dealing with the colonies, however, were included under the name of the speaker and under the subject with which he dealt. Careful consideration of where to draw the line should be given to the problem as a whole and then numerous individual decisions will have to be made as the work progresses. It is impossible to anticipate all the types of information with which one must deal. Our general rule was to include

³ See explanation of cross references, p. 297 ff.

everything on Virginia and the other American colonies, both original and secondary information, whether from England or America, but foreign news items only in brief by subject, excluding foreign personal names unless they were well-known people. The extent of the colonists' acquaintance with Voltaire, for example, we thought worth noticing.

Having set up a frame of reference — that is, having determined in general what to include and exclude — regulations for those items to be indexed must be established. The ideal situation is to have on hand at this time all the indexers who are to engage in the project so that they may share in setting up patterns for their work. They will then remember the procedure more readily, because they have reasoned out the rules and not merely learned them by rote. Great care should be taken in arriving at the rules, both general and specific, for if they are altered during the course of the project to admit a new type of material or a new method of handling, what is to be done about the work already finished? If the rules are changed during the game, the score is suspect. A good index is based upon consistency. Without that characteristic, it cannot be reliable. When revision is imperative, *you must go back immediately over what you have done* so that all the earlier indexing is brought in line with the new policy.

A preliminary testing process, inaugurated after everyone understands and has agreed on the basic frame of reference, should determine the specific rules. In the case of the *Virginia Gazette* files, the test included ten issues each from 1736, 1751, 1766, 1776, and 1779. Each period offered a variety of problems but a survey of the whole enabled us to relate them to one another. We issue a word of warning against dividing each document into sections for indexing at different stages. Although this device may appear to be a timesaver, it really leads to many complications. Such a division makes it impossible to see at once all the factors in one document which must be fitted into the general scheme. In the *Gazettes*, for example, only advertisements were indexed during the early years of the project, thus creating many problems of inconsistency in relation to the indexing of the news, which would not have existed otherwise. As a result, considerable revision was necessary.

Each rule established should include a subject heading general enough to cover all items of a certain class. Subheadings must be used to give each reference specific value. They are essential in avoiding the pitfall of large blocks of unbroken references which

are anathema to the most patient scholar. Other specific main entries may be included as the needs arise. As an illustration, the rule for indexing material on crimes in the *Virginia Gazettes* reads as follows:

Example

- | | |
|-------------------------------|---|
| (1) Name of crime
location | Robberies,
in Williamsburg, Va.
If a trial held, use:
Robberies,
trials for,
in Williamsburg, Va.
If a conviction, use:
Robberies,
executions for,
in Williamsburg, Va.
and
Executions,
for robberies,
in Williamsburg, Va. |
| (2) Name of perpetrator | Jones, John,
guilty of robbery |
| (3) Item stolen and place | Silver,
stolen in Williamsburg, Va. |
| (4) Cross references | |
| (a) from place to crime | Williamsburg, Va., <i>See</i> Robberies in
Places were not used as main headings in the <i>Virginia Gazette Index</i> unless the item was directly connected with the character of the place or its government. Hence cross references from places to subjects were plentiful. |
| (b) from general to specific | Crimes, <i>See also</i> Robberies |

All cross references enable one to eliminate additional entries for the same item. The two types are illustrated above: (a) those from one specific heading to another. This occurs when the reader might expect to find material under a heading which, in practice, has not been used at all or only part of the time. He is, therefore, directed either to the substitute or alternative heading as the case may be; (b) those from a general subject to related and usually more specific subjects. From the user's point of view, they provide keys to certain facts pertaining to his general interest, while he may be unaware of their existence. Cross references may be used to al-

most any degree, though absurdities must be avoided. On the whole it is better to err on the side of too many than too few cross references of this type because of their usefulness.

Cross referencing is most satisfactorily done by years after all the entries have been checked and typed, rather than by individual items. One indexer can be assigned the job of rapidly picking up cross references on the completion of a year's span of the text. These can be typed directly on 3 x 5 slips of another color than that used for the index proper, and they may then be filed by the typists with the latter. This procedure saves confusion for the typist, who otherwise would have to type two kinds of slips at once; it also eliminates the possibility of cross referencing something which might be withdrawn in checking the index.

Each regulation for indexing, when decided upon, should be typed on a 3 x 5 card, in duplicate or triplicate, so that each indexer may have a set. The main subject of the item should appear as the heading. The card should include examples taken from the item for which the rule was originally set up. It is also wise to give the reference date of the example so that it may be checked in detail if necessary. An alphabetical subject file of these cards will enable the indexer rapidly to orient his current material with the policies of the project. When a new principle is developed, it should be added to the file.

Although the indexers should follow closely all the rules established for the project, they should be cautioned against trying to force items into an existing pattern when their main emphasis is elsewhere. What is *most* significant about a piece of information should determine its class. New rules or corollaries should be made when new problems arise. Accuracy of interpretation — another essential character of a good index — will thus be assured. While this point may not seem to need laboring, the tendency to classify in the most convenient pigeon-hole is, in fact, one of the errors most often committed by people learning to index.

When the questions, what and how to index, have been answered, there are several technicalities that require decisions. These include: reference symbols, abbreviations, spelling, capitalization, and punctuation. The criteria for good reference symbols should be brevity, clarity, and usefulness, and they form the equation, $a+b=c$. Long symbols increase the printing bill, confuse the consultant, and should not be used unless every part is vital to distinguishing one document from another. If the references themselves can be made

to serve in some way to identify the material, they will be all the more useful to the researcher. In the *Virginia Gazette Index*, the only possible reference was a lengthy one, but it did provide additional information for the reader by including the date of each item. The reader should be able to interpret the symbol without repeated consultation of the key.

Commonly accepted abbreviations should be used as much as possible. Where the line is to be drawn and what authority is to rule should be determined at the very beginning of the project, so that from the time the first typing is done the chosen abbreviations will be followed. Such procedure will eliminate one step from the final editing. Similarly, if spelling is made uniform as the index progresses, editing can proceed more rapidly. This will be necessary in cases involving old-style spellings and variants within the documents. A minimum of capitalization is desirable. If capitals are restricted to proper names, a better looking page of type will result. Numerous discussions of what is a proper name will undoubtedly arise. Should it be Long Island, N. Y., or Long island, N. Y.? Consistency may not be so easy to maintain as it might seem. Rules on punctuation should be determined preferably before any slips are typed.

The style of handling these technicalities should not be changed in the process of indexing unless you are prepared to go through the already existing slips and make the corrections. It is impossible to stress too strongly the importance of establishing policies at the beginning of the undertaking and of adhering to them.

Having outlined in detail what factors ought to be discussed thoroughly before the indexing actually begins, let us now turn to the actual work processes by which the task is accomplished. These, like the equipment needed to carry them out, are fundamentally simple. Typewriters and paper are the only essentials. Each indexer and each typist should have a typewriter. We found it helpful to have the paper perforated in strips of five 3 x 5 slips which could be easily torn apart. When used in strips of five to record the original indexing, these were called "tracings." The following illustration is taken from our "tracing" for Dixon and Hunter's paper of January 10, 1777, page 1, column 1 through page 2, column 2:

[Paine,
Thomas],
American crisis reprinted

D10Ja77:11-
P10Ja77:11-
P117Ja77:11⁴

⁴ If the same news item was printed by more than one editor, the references, con-

American crisis,
extract from reprinted

(1)

Sullivan's island, India,
naval engagements at

D10Ja77:22

(2)

Williamson,
Andrew, Col.,
treats with Cherokees

Cherokees,
treaty with S. C. negotiated

South Carolina,
treaty with Cherokees negotiated

(3)

Salt,
plentiful in S. C.

(4)

Gunpowder,
imported to Savannah, Ga., from W. I.

Imports
to Savannah, Ga., from W. I.

(5)

These sheets remain a permanent chronological record of what items from each document have been indexed, and are an invaluable aid on many unforeseen occasions. It would, indeed, be impossible to do any lengthy indexing job without keeping such a record. Only a book-length index can be done directly on slips which are immediately filed.

The perforated strips of paper can also be used by the typists to record the individual references, one 3 x 5 slip for each item. When a section of work has been completed, the slips will be torn at the perforations and filed alphabetically. Some 3 x 5 white cards for keeping files of rules will also be useful. Cross references are easily spotted by the printer and, during the process of indexing, by the staff, if they are in another color than the main entries. Slips in strips of five, perforated, would also be useful here. An item which makes for more comfortable working conditions, especially if the material being indexed is unwieldy in size, is a book-rest for each indexer's documents so that they may be read more easily upright.

ected by hyphens, were given for each paper. The same procedure was used for advertisements; but if an advertisement was duplicated in successive issues of a particular paper, only the first printing was indexed.

These can be made out of plywood. The typists should also have some sort of stand to hold the "tracings" from which they are to type.

We are now prepared to outline the step by step procedure of indexing. Each indexer should work on a fairly large segment of material (a year in the *Virginia Gazette* files) in order to gain a sense of continuity. When we had three indexers available, therefore, work was going ahead on three separate years' material at one time. The indexer reads an item, or paragraph as the case may be, the first time to ascertain the main subject and to relate it to the general principles which he is to follow. On the second reading the specific parts to be indexed are selected, the reference symbol typed at the top right-hand corner of the page, and the entries listed one after the other in the order in which they appear in the article. (See illustration on p. 299). All entries should be exactly as they are to be in the final copy unless some agreement is made with the typists with regard to capitalization, abbreviations, and the like. The main heading should appear on the top line, the first subheading on the second line with a two-space indent, the second subheading on the third line with another two-space indent, and so on for as many subheads as are used.

When each person has finished indexing his segment, the senior indexer should check the work of the other two for accuracy in reference dates, spelling, and other details as well as for understanding of the material, completeness in indexing, and uniformity in entries. In order to approximate the latter — and perfection is manifestly impossible in any long operation — the senior indexer must go over all the material before it is ready for typing. His own work should also be checked for accuracy by one of the indexers.

Then the typists receive the indexers' work which they may divide among themselves. Each entry is typed in the upper left-hand corner of a 3 x 5 slip with the reference date in the upper right-hand corner. Typing should be checked by exchange of work, and particular attention should be given to accuracy in the reference symbols. The slips are torn apart, scattered alphabetically, and filed by all the typists. Final filing in the master file should be done by one person only, well versed in the alphabet, to reduce the margin of error. Alphabetization is most important and should be done with the greatest care, with checking by the senior indexer until it is certain that the typist understands the process.

When all the documents have been indexed and the slips filed,

they may be edited and prepared for shipment to the printer. The latter will usually agree to set type from the slips providing they are properly numbered in their final order. In many cases several cards with the same main heading will have accumulated, some with the same main heading and first subheading, and some with the entire entry identical. In these cases, the editors (who will be the indexers — the typists' services are no longer required) will cross out whatever is identical with the first slip of a related group. Numbers can be used to identify what position of indentation the first remaining subhead should take in the printed column. Slips on which the whole entry is the same should be stapled together and the printer instructed to use only the reference symbol on all but the top slip. Examples from the *Virginia Gazette Index* follow:

slip 500	Houses, for rent, in Williamsburg, Va.	R26F68:21
" 501	Houses, for rent, In Yorktown, Va.	29S38:31
2		
" 502	Houses, for sale, in New Kent co., Va.	PD170c72:31
1		
" 503	Houses, for sale, in New Kent co., Va.	PD240c75:32

Note: Slips 502 and 503 are stapled together.

Punctuation can be omitted in the final editing if it has not been inserted during typing, provided an agreement is made with the printer that it is to be uniformly inserted by him. During the editing the alphabetization should be checked and a lookout kept for any cross references which may have been omitted. Inconsistencies in entries which have slipped by during the indexing proper should be eliminated now. With the completion of editing, the index will be ready for the printer. The *Virginia Gazette Index* was printed by photo-offset. In a long and complex text of this kind, the cost differential between this process and regular printing is not so great as in a job of ordinary length and composition.

The adoption of these principles and techniques of indexing, general and specific, should produce a good index. As the proof of the pudding is in the eating, the success or failure of an index is declared by the user. An index is a good one if it can answer the ques-

tions put to it by the majority of investigators. To do so it must have certain characteristics which the principles and their application, discussed above, will foster. These are: consistency of entries, accurate interpretation of material, accuracy of references, brevity of entries, and sufficient breakdown in the entries so that large numbers of block references will not appear. Consistency of entries means, in brief, that for a given type of information the same entry should always be used. If a research worker is interested in pursuing a study of crimes, for example,⁵ he should expect to find every item referring to a criminal act under its name, not sometimes under the name of the crime, sometimes under the perpetrator's name, and occasionally under the place where it was committed. Without this assurance no investigator can depend on having in hand all references to the items in which he is interested. His only alternative is to read the documents themselves, thereby defeating the purpose of the index. To maintain consistency among two or three workers over a period of three or four years' indexing is the most difficult challenge facing the senior indexer, but one which must be accepted and met.

Accurate interpretation of the material is of obvious importance. It is for this reason the indexer should have some familiarity with the field in which he is working so that he may read with intelligence and understanding the item he is to classify. Overlooking its significance or placing it under an incorrect main heading hides it from the user as effectively as ignoring it completely. Errors in reference symbols will as readily confound the reader. Although some are unavoidable in a large undertaking, extreme care should be taken in typing and checking to reduce inaccuracies to a minimum. Brevity of entries reduces the somewhat cumbersome quality of a long index and makes for more rapid consultation. It is one of the qualifications which can be met without much difficulty. Patience and attention will also assure a sufficient breakdown of entries. Nothing is more exasperating to the eager user than to look up John Doe's name and find 50, or even 15, references with no subheadings to indicate to him whether he is interested in all of them, some of them, or none of them. Limitations of space should never be pleaded as justification for block references. The additional words are well worth the extra expense. To serve the specific needs of the user is what indexing is all about.

It is the writers' hope that those who undertake an extensive in-

⁵ See above, p. 297.

dex, whatever its nature, may be able to find guideposts in this discussion of policies used in the *Virginia Gazette Index* which will lessen the complexity of their task. Even with these suggestions and warnings, their assignment will often be discouraging in the hundreds of little problems which will arise to obscure the main issues. But perhaps this outline of procedure will so lighten their burdens that they will be in fine fettle to celebrate the completion of their project.

The Nation's Standard Paper For Permanent Records

BYRON WESTON CO. LINEN RECORD

BYRON WESTON CO. LINEN RECORD has been tested and approved by generations of public and business record keepers. Made wholly of new white cotton and linen clippings, it possesses to the highest degree the strength, ruling, writing, handling and erasing qualities essential for endur-

ing satisfaction in bound or loose-leaf records, documents and similar usages. For other records, select the right paper for any purpose from the complete Weston line of rag content ledger, bond and index papers. Write for complete information.

BYRON WESTON COMPANY • DALTON • MASSACHUSETTS

Makers of Papers for Public and Business Records — Since 1863
