CONTENT ANALYSIS FOR CLASSICISTS

by T. F. Carney
(University of Manitoba)

This study has two aims. One is to give colleagues in Classics some idea of the potentialities of the technique known as content analysis. The other is to indicate important works – classics in their own way – in this general area. The author gained what insights he has in these matters by a time-wasting trial and error process, involving much casting about for helpful books, but can now teach students the rudiments quite quickly. Consequently it seemed a good idea to present colleagues with a systematic exposé which would enable them, should they wish, to read their way into the subject quickly and in depth. The study commences with an illustration of what content analysis can do. It then goes on to discuss the mechanics of using it as a technique. It concludes by reviewing the advantages and disadvantages involved in using this technique.

Let's start, then, by showing how content analysis can make a new and more meaningful type of source investigation possible. In the course of our work we all of us, sooner or later, have to make allowance for the personalities of the writers or figures whom we study. Most often we do this by applying psychoanalytic theory, in some form or other. This type of approach, however, is methodologically unsatisfactory because of the amount of inferring, on the part of the would-be psychoanalyst, which takes

1. I am glad to acknowledge the interest and support of Professor C. P. T. Naude, with whom I have corresponded, in bringing the technique to the attention of readers in South Africa.

2. The recent controversy in JHS (88(1968)136–38) seems to indicate that a guide to the 'literature' on content analysis for Classicists is needed. In this and in adjacent fields I have tried to restrict references to works which have in actual experience proved meaningful (and readable!) to humanists. Few of these works are on classical themes – which is why it is so difficult for us to find out about them – but, hopefully, their relevance to our basic concerns as literary scholars or historians should become apparent on reading.

3. Perhaps the epitome of this type of approach among classical scholars is to be found in the work of N.O. Brown (e.g. Life against Death, Vintage, N.Y. 1959). Generally the only school of psychological thought on personality known to the layman is the psychoanalytic, so that, even if he rejects it, he still thinks in its terms of reference. For a review of the main schools of thought in psychology, see I.G. Sarason, Personality: An Objective Approach, Wiley 1966, 27–117.

4. This is not to deny the significance of the basic concepts identified by Freud (e.g. rationalization, projection, displacement, repression and reaction formation), but to distinguish between these and Freudian theory on personality, which is currently in scant regard outside of purely literary circles (though the post-Neo-Freudian Ego-psychologists – scholars such as E. H. Erikson – fare somewhat better).
place. For instance, if a son dislikes his father, he is assumed to have an Oedipus complex; but if he likes his Father, he is assumed to be suppressing (and therefore to have) this complex. Such approaches can no more be disproved than they can be proved, and so one study of this nature can be piled atop another without winning anyone over to conviction.

Contrast this approach with what is known as psycho-biography. The latter takes the following form. One takes, say, four in some way typical writers from a certain period and one examines everything they have to say on, for instance, five or six aspects of a major social or political issue. Then one takes another four comparable writers a generation later and does the same. One thus has, from the overlaps among the views of each group, a picture of the issue as it appeared to this type of person at each period. One also has, from the disparities between the two pictures, a measure of the change in viewpoint which has occurred.

This psycho-biographic approach is built upon a set of assumptions about language, perception and personality and a method of cross-comparing documents, however lengthy or complicated these may be. The method is content analysis — but as a technique it too is grounded in these same assumptions. One cannot understand one without knowledge of the other, so let us quickly pass these assumptions under review.

Supposing your car touches bumpers accidentally with someone else’s. It rapidly becomes evident how perversely one’s fellow man can misperceive reality. Psychological experiments show in fact that no two persons perceive the same event in quite the same way. Apparently what happens is that we selectively perceive the myriad stimuli with which our senses are bombarded, and we also make sense of what we have perceived by rearranging it according to our views on nature and causality etc. These views are a sort of image of reality which our childhood, school and adult working experiences have built up in us. No two men have ever had quite the same set of such experiences, so no two images are identical. But images have much in common. After all, all men go through similar maturing stages, have common physical needs etc. And all men in a specific society have gone to similar schools, have shared similar family backgrounds, cultural values and so on. Observably certain viewpoints and stereotypes are shared by the men of this society. These can be identified by looking for characteristics, in viewing this or that aspect of reality, which the group shares. What a man defines


152
as reality is reality for him. It is a fact, it has consequences etc., so one can establish the 'reality' of such and such a group at such and such a period. Secondly, language and the communications process are no longer seen as neutral tools automatically and perfectly transmitting the desired message. An abstract word, for instance, has a central core of meaning and an aggregation of nuances. The overall pattern of emphases in this amalgam varies over time. Such words alter in meaning in the flux of language in the flow of time. Also, as well as denoting things, such words have emotional overtones for their user. Thus such a word has that meaning which its user gives it by the way he employs it, not a fixed meaning which one can assume. Moreover, a message becomes garbled in the process of communication or in memory over time. Give someone from our industrialized society a folk-tale from another culture to read, then either ask him to relate it to another person (and that other to someone else again, and so on), or check his memory of it at weekly intervals thereafter. Typically, certain things happen. Striking details are exaggerated, prosaic ones forgotten; the sheer bulk of details is drastically curtailed down to a certain residual limit and the tale is assimilated to the background assumptions (re causation, for instance, or how people are motivated) of the communicator.

Thirdly, humans tend to see things in patterns, rather than in ones and twos. What a man means by 'democracy' for instance can only be understood by examining that concept in conjunction with a whole group of related concepts, practices and figures to which he may make reference. Personality is probed by looking for syndromes (related patternings) of...
characteristics. Degrees of emphasis can only be made out in relation to the totality of emphases. Moreover, until one is sure of everything that is in a man's picture of a particular aspect of reality, one cannot see the all-important things which are not there.

If one reads for impressions, assuming one knows what the writer's words mean for him and adopting 'common-sense' (or Freudian) views of personality, one is merely selecting out of the data that version of reality which represents a pattern of sensitivities idiosyncratically important to oneself. Consider Gibbon's version of the Fall of Rome, for instance, or Buchan's Augustus. In conducting a lengthy study impressionistically (that is, without conscious controls upon one's selectivity in perception) with time new aspects of the same issue come to seem important, and, throughout, certain facts, to which one is particularly sensitive, are given undue prominence. The focus of attention wanders, the data are partially considered. And so there can be two schools of thought about a given historical figure or incident, although the two scholars who produced them drew upon the same body of evidence. Imagine a spotlight on a danseuse in a can-can line-up. The beam of light doesn't waver or flicker, dim or brighten; the patch of light doesn't expand or contract; the focus remains unshifting on one piece of the action only right through the show, whatever its length. This is the kind of focus which content analysis gives. What it also does is extract all the evidence (against as well as for), on specified aspects of the specific issues on which it is thus uniformly focussed, from all of a defined body of data. Now we in Classics do not have overly many analytical techniques at our disposal. (If you think we have, try to name five.) So we cannot afford to ignore this one. Its implications, as has been seen, are revolutionary for our particular field of studies.

A word of warning is necessary at this point. Content analysis is a technique for methodically abstracting and correlating all the data relevant to a specific set of questions from a specified body of text(s). It does not supply the questions (only ways of posing them) or make inferences from the data which it gathers (it 'merely' gathers all the relevant data, to enable inferences to be made on a sound basis). Questions and inferences must come from the insight of the person using the tool. Content analysis is a clinical, rather
than a scientific, technique. One's skill in using it develops with practice; one has to know a lot about the background to the questions one wishes to ask and the sources one proposes to use. The more familiar one is with technique, subject matter and sources, the more easily one perceives new ways of posing questions and new aspects of the subject and source material. If one has no questions to ask, the technique won't suggest any. If one asks silly questions, mere busy work results. In this as elsewhere 'Discovery favours the prepared mind': only knowledge in depth of subject and sources enables one to reap maximum benefit from the welter of details generated via the technique.

Let's take a very basic problem and see how content analysis helps us to deal with it. Supposing one wants to teach one's students how to look for the picture of a politician (say Marius) in a series of writers (say Cicero, Valerius Maximus, Plutarch and John the Lydian). Just how, in detail, are they to go about doing this? They should not have each to start from scratch and invent his or her own way of going about it, because there is a considerable literature on the analytical problems involved. If one were to attempt the above task, of looking for several writers' picture of a politician, in ignorance of the basic modern works on that politician, one's work would be deservedly open to criticism for its bibliographical shortcomings. Likewise, to attempt the task in ignorance of a basic proven technique relevant to the analytical problems involved, leaves one equally open to criticism for not having done one's bibliographical homework. Furthermore, use of this technique enormously expands the range, versatility and subtlety of one's investigations by giving access to a consolidated body of findings, involving many ingenious ideas and procedures, on this analytical problem.

Probably the quickest way to show the sorts of things that content analysis can do in such an investigation is to illustrate with a diagram (no. 1). This

17. This type of question was chosen because experience here in Canada indicates that, although it interests students, teachers seem incapable of giving instructions as to how to go about dealing with it: see A. B. Hodgetts, What Culture? What Heritage? Ontario Institute for Studies in Education, 1968, pp. 20, 24–26 (esp.), 99, 102 and n. 6 on p. 111. Nothing in my experience suggests that this situation is peculiar to Canada.
shows how the first stage of analysis can be managed. The aim of this study is to enable readers to set out and manipulate such an organized analytical structure when complex sets of data have to be extracted and compared.

In looking at the ‘mechanics’, as it were, of the technique, one finds that it can be broken down into (i) its component pieces, (ii) how you put these together, and (iii) the problems which occur at the various stages of so doing. Let’s identify the ‘pieces’ first, then look at the other two matters both at the same time. Diagram no. 2 should help readers to visualize what is involved.

The ‘pieces’ comprise things you can count – which come embedded in passages which establish their meaning in context – and compartments into which you can sort them, together with some form of grid or matrix to sift the things counted into this or that pattern of siftings. The ‘things you can count’, or coding units, can be words, themes or interaction-units. Words are self-evident. A theme is a subject discussed in a certain specific way (care must be taken, when using themes as coding units, to define them very clearly): for instance, an ikon defending its city’s walls.15 An interaction-unit is an exchange between two or more parties. For instance, in a previous number of Acta Classica, I reviewed Terence’s use of the words sodes and quaeso.20 Who used the word and to whom? Was it a male to a male or to a female, or vice versa? Was it a person of superior status to an inferior or vice versa? The interaction unit in this case was the persons involved in cases where the words were used.

Now words, themes and such interaction units have to be seen in context before their full significance can be appraised. This ‘context’ is termed the context unit, for purposes of content analysis, and has also to be carefully defined. It may be a sentence, when a word is involved as coding unit, or a paragraph (when an interaction unit is involved, say) or a page, or even a chapter (for a theme).

The ‘compartments’ are the classification system under which different kinds of data about the coding-units are assembled. They are categories, categories of aspects of the topics or texts which are being investigated (see diagram 1). The grid or matrix is generally obtained by breaking up a theme into smaller component parts which can be more easily defined, and therefore identified, in analysis (read downwards in diagram 1). These components are each categorized in regard to how they are presented in the literature in a series of different aspects (read across).

As to ‘putting the pieces together’, the reliability of the study, that is the faithfulness to its prescriptions with which it has been carried out, depends on the care with which coding units, context units and categories are defined. One way of checking on reliability is to have a second analyst redo part of

the study, and see how far his findings agree with the first man's. 21 Another way is to use a computer and eliminate the human error factor. All this means that a feature of the technique is the care taken throughout to be specific about assumptions, definitions and operations.

It is, in particular, of the utmost importance to be specific in the matter of sampling, that is to say of establishing what parts of what texts are to be considered, and why such a choice is justified in terms of the study to be undertaken. Especially in cross-comparing two bodies of literature one has to ensure that bulk, emphases, complexity and so on are relatively equal. An obvious step, readers will be thinking. Well, just consider how frequently, in your own experience, this step has been explicitly taken in studies you have read (and it isn't until one is quite explicit in regard to sampling that the full range of problems involved becomes evident).

Problems arise with regard to the categories, which are all-important for the analysis. Basically, categories are of two main types, one of which is more objective than the other. The more objective type involves straightforward quantification of one sort or another (i.e. is the theme discussed at length or skimpily; does the discussion contain errors and inconsistencies or not?), or dating (e.g. in regard to the ikon-theme mentioned before, discussions can be categorized as occurring in the early, middle or late stages of the Iconoclast Controversy), or simply description (for instance, categorization by source or origin of statement). The less objective type of category involves qualitative assessments of the direction of bias or of the intensity of a statement.

It is rarely possible to avoid qualitative assessment entirely, and the competence of one's analysis largely depends upon the sophistication with which such assessments are conducted. Various ways have been evolved of reducing subjectivity. Possibly the most effective is the simplest (if one's source-material allows one to do it, that is): one redefines direction of bias so that it becomes a quantitative matter. An illustration will show what is meant. For instance, one might rephrase the question 'Is Plutarch biased against Marius?' as follows: 'In the various battles in which Marius was involved, to whom does Plutarch attribute responsibility for success?' One can then proceed so to organise one's categories as to allow emerging suggestions of bias to compound one another. Thus, one might further ask: What sources are cited in such cases (and where else and with what comments are they employed in Plutarch's works)? Are emotionally charged words used, and, if so, is the emotion directed in favour of or against Marius? Sulla's career contrasts with that of Marius: how does Plutarch treat it, in regard to the above points? 22

21. McClelland (op. cit. 74) describes a study in which this was done.
22. A good example of this kind of approach can be found in A.G. Woodhead, 'Thucydides' Portrait of Cleon', Mnemosyne 13 (1960) 289–318, where the treatment of Pericles and Cleon in regard to a number of items - gossip (favourable and unfavourable) reported, successes and failures recounted etc. - is contrasted.
Here are the findings of such an analysis, to indicate just how tell-tale they can be. Passages deriving from the contemporary memoirists, all Marius’ *animi* and partisan writings elsewhere little used by Plutarch, attribute success in Marius’ various battles always to someone other than Marius. Very often these passages employ terminology which is unusually emotional for Plutarch’s narrative passages. The Graecophile Sulla is better treated on all points in these regards.

Another way of handling this type of assessment is as follows. When someone has to make a value judgement three courses are open to him. He can be non-committal and opt out of judging; or he can adopt a clear-cut position, pro or con; or he can adopt a qualified position, pro or con. So a scale to assess value judgements must allow for at least these five positions along a dimension running from negative to positive (indications are that a seven point scale – very negative, negative, somewhat negative, neutral, etc. – distinguishes reactions best). Such scales make it more easily possible to assess degrees of bias and to ‘panel test’ one’s assessments by having them, in part or whole, redone independently by someone else.

Finally, to assess a word’s emotional charge for a writer, one can locate it in ‘semantic space’ (see diagram no. 3):23 a word can be placed along a dimension running from strong to weak and in relation to another dimension, ranging from passivity to activity; it can further be appraised as ‘good’ or ‘bad’ (words appraised as ‘good’ might be capitalized). One can count outwards from the central point in each dimension,24 thus sorting words into neutral on this dimension (i.e. between 0 and 10 on a dimension scaled to run from zero to thirty), moderate (i.e. 11 to 20) and extreme. Admittedly, such placement is subjective – to a degree; striking words can easily be located – but it is explicit and can be tested by a panel test. Also one can compare the word’s usage elsewhere in the writer, to see how frequently he uses it and in what kinds of context.

Other problems relating to categorization are as follows. Categories must be inclusive. A bulging ‘etc.’ or ‘rag-bag’ category for odds and ends is likely to mean that one has not thought out how to compartmentalise the full range of aspects relevant to the subject-matter and the analysis. And categories must be relevant to the subject matter. It can easily happen, for instance, that one’s frame of reference makes one oblivious to the significance

23. See J.B. Carrol, *Language and Thought*, Prentice-Hall, N.J. 1965, 102–104 for the *rationale* behind the diagram (the ‘dimensions’ used are, apparently not culture-bound; rather they are characteristic ways of appraising, found in culture after culture, across the ages).

24. This device was adopted when I found students assigning plusses and minuses to numbers assigned to highly charged words. It could thus happen that a writer who employed highly emotional verbiage, but indulged both in praise and censure, came out with a total score resembling that of a determined fence-sitting employer of neutral words (the plusses and minuses cancelling each other out). This zoning device sorts writers out into high, medium and low scorers without possibility of such confusion.
of things by which a person in classical antiquity set high store. Hence all content analyses should be commenced by (generally several) pilot-studies or pretests, in which one tries one's analytical framework out on a part of the text or data to see if it produces meaningful results. Usually some redefinition of one's questions or of coding units, context units and categories results.

The arrangement of the categories should be such as to proceed from the readily quantifiable to the more qualitative. The reason for this is that, as the material will have to be processed several times, the easiest processings should be done first. Then, as familiarity with the material builds up, the more difficult types of processing occur only at a point when one is no longer simultaneously struggling to interpret the passages of text involved. It is most surprising what can be in a text when one looks hard and methodically at it, so the repeated processings help one fully to assess the import of the various passages.

Next comes the business of counting, and in some way weighing the relative importance of (this is known as 'scaling') the various data extracted from the texts during analysis. At this point a basic assumption must be made explicit. It is this. Counting with a zero point and a numerical scale is not regarded, in this study, as the only way to count but as the best way, the end point on a dimension of counting operations which commences with the crude simple operation of seeing whether item X is there or not. An example of the practical usefulness of such a question in my own experience is: is the canard that Marius was a coward before assemblies though brave on the field of battle found only at Plutarch Marius 28, 2? (Answer: yes.)

From this type of 'counting' one progresses to asking how much of item X is there (e.g. does the so-called 'Marian Massacre' become more overdrawn the further one gets away from the event in time?). 'How much' involves scaling, which will be discussed shortly. Let's get to the next 'counting' stage, which, though still not employing a zero point etc., will readily be admitted to be a very rigorous process. It is as follows: it takes a graduated scale of frequencies (high, medium, low), and, by this criterion, ranks each of, say, one to four aspects of each of one or more issues in comparing two or more writers. Thus I compared John the Lydian, a middle-level bureaucrat, with Cassiodorus, a top-level bureaucrat, in regard to their treatment of the three crucial bureaucratic issue-areas (sportulae, suffragium and praescriptio fori), each issue-area being broken down into its component aspects. This is a simple form of multivariate analysis and it is how one generates one's matrix, the matrix which allows patternings to emerge and to be compared.


Let us see what this means by examining diagram no. 1. The aim is to present what each writer has to say about the politician in question in such a way that (a) each writer’s pattern of emphases becomes apparent (at a glance: complex presentation of a complex issue merely heightens the reader’s inability to make out the wood for the trees) and (b) so that the various patterns which result can be compared. The first objective is most easily attained by some form of matrix. This is achieved by breaking the politician’s life up into a number of component themes (reading down) and categorizing each theme in a number of different ways (reading across). Incidentally, multivariate analysis is the only way in which one can correlate two or more items which each contain a range of variation within themselves. Think of diagram no. 1 and notice how its framework handles the following problem—very literally a problem of ‘content analysis’. You are an archaeologist on a dig. The dig exposes five levels of habitation. There are a scatter of artifacts from different poleis and different epochs. You have to compare this distribution with those excavated on digs in five other regions.

Comparing the various patterns, to revert to aim (b) above, is also best done by some form of matrix. What is involved is conflating the detailed findings of each individual pattern in some simplified way so that all the writers involved can be compared in terms of a uniform set of criteria. Instead of categories (reading across), one now has a list of writers. One may retain all the themes (reading down) or only those to which at least one writer gives some specified degree of prominence. Such foreshortening sacrifices only the less significant findings (you will have noted separately items which are either not mentioned or only receive fleeting comment) to achieve greater intelligibility in the form of a less complicated layout. What goes into the ‘boxes’ of the matrix? Well, one can either invent some form of points-system for each of the categories of the original matrices and simply enter an aggregate points total or, better (inasmuch as this procedure shows how each entry is made up, a thing which a mere number cannot show), one enters a number of symbols. For example, one might use a letter (H - high) for frequency of mention, hatching for fullness of detail, a symbol (+, 0, —) for attitude, coloured (red) if the verbiage is striking, with an X to indicate the existence of errors. The number of entries in the ‘box’ thus indicates the amount and variety of attention given to the item by the writer.

---


28. ‘The simplest type of property by which an object can be characterised is a dichotomous attribute, such as voter/non-voter, white/non-white…. It is always possible to simplify a more complex property by reducing the number of classes which are distinguished’, Lazarsfeld and Rosenberg, op. cit. p. 41. So here one can simplify the data, so that one can record them easily, by only registering High frequency (i.e. splitting High, Medium and Low into a dichotomy: High/not high), fullness of detail, striking verbiage, existence of errors.
in question, and 'boxes' with many entries contrast strikingly with those containing few and so cause patterns (of frequent and infrequent annotation) to stand out at a glance at the matrix.

It is worth observing that most content analyses are improved by a pause, or maturing phase, between the detailed analysis and the final summing up of the major trends within the findings. One tends to become so preoccupied with details as to be blinded to the large scale trends which would be discernible beneath the details if only one could readjust one's perspective. Without such a pause, content analysis studies tend not to produce good summings up of the individual matrices. One cannot be dispassionate enough in discarding detailed findings which took such effort to discover, apparently, so the trends fail to emerge from beneath the welter of detail.

As promised, a note on scaling.29 There are four scales. The 'nominal' merely assigns an identifying term or number; it's a labelling device, which enables one to assess the mode (lowest common denominator of the labels shared by all items). The 'ordinal' scale indicates that item 1 has more of X than item 2, so is higher than item 1 on a scale for X. This scale enables you to select the median, or item in the middle, relative to the other items. The 'interval' scale ranks items and quantifies the distance between them (e.g. property qualifications of the various classes on a censor's list of voters). An absolute zero point gives a 'ratio' scale and allows of the full range of mathematical operations – which are possible on an interval scale if you arbitrarily assign a zero point.

It is particularly helpful when one has to synopsize to realise that, basically, the individual and group data which one is trying to conflate have three types of properties.30 There are analytical properties, based on data about each member. For instance different stages of Greek civilization have been demarcated according to the amount of ‘achievement-oriented’ themes appearing in the writers of each stage.31 There are structural properties, based on data about the relationships among members. For example the gentes maiores can be identified by comparing the numbers of consulships which these families attain with the numbers gained by other families competing with them. Finally there are global properties based on data attributable to a group but not to any one of its members. An illustration might be communities-having-their-own-bishoprics among the Christians of the late Roman Empire.

Even quite a superficial awareness of scaling procedures and individual versus group properties of data enormously expands one's ability and

29. The study by Budd & co. (chap. 5, esp. pp. 31–33) has an excellent survey of this whole matter of measurement and scaling.
30. The discussion on 'Systematization of Group Properties' in Lazarsfeld & Rosenberg (pp. 290–96) is helpful on this matter.
facility in marshalling details. A modicum of sophistication in this general matter of 'counting' makes a world of difference in one's overall analytical capabilities. These matters are to content (or any other quantifying form of) analysis as syntax is to Latin prose composition.

By now readers should have an idea of the technique and its significance. It should perhaps be pointed out that the degree of objectivity of one's analysis varies with the source material and the type of analysis undertaken. Imagine two intersecting lines. One represents a range of 'literature' varying in ease of dissectability from a policeman's report of a traffic accident, say, to a T.S. Eliot-style poem. The other represents the range of volume of writings extending from an encyclopaedist's bulk to a few fragments of a poet. Clearly a pedestrian, widely-ranging encyclopaedist would present more and more manageable material than would the fragmentary remains of an ethereal poet. It would be more meaningful, that is, to do a content analysis of Pliny's *Natural History* than to do one of Sappho. Or take another pair of intersecting lines. This time one represents the range of coding-units from words through interaction-units to themes; the other the range of analyses from analysis of a single, short passage to analysis of trends within a mass of documents. Clearly a trend-study of word usage -- say that of the concept *demokratia* in fifth and fourth century Greek -- will yield better substantiated results than will analysis of themes in a narrow span of context (one reason why studies of themes in Pindar -- not done via content analysis, to my knowledge -- seem to convince no one but their originators). Finally, if one is looking for subliminal patterns -- as in Morton's study of the *Epistles* of St. Paul, say -- or if one is scanning upon material involving an interchange of messages (a case where disingenuousness can be identified), one can be more sure of one's ground than if one is analysing a sophisticated narrator for bias. Even here, however, modern sophistication is such that, with one's enormous semantic and analytical advantages over an ancient writer, it is often possible rigorously to demonstrate that a particular bias did exist.

A common assumption in using the technique is that frequency of occurrence of an item establishes the significance of that item. This is often, but

32. A point which seems not to have escaped the notice of Classicists at the University of Pennsylvania, to judge by the theses which they have been producing: e.g. Ch. H. Herkert, *Historical Commentary drawn from the Natural History of Pliny the Elder for the Years 54–76 A.D.*, 1956, 133pp.; M. A. Th. Burns, *A Commentary on the Reign of Augustus based on the Evidence of the Naturalis Historia*, 1960, 347pp.

33. J. Pelikan, *The Shape of Death*, 1962 (five Fathers of Church on life, death and immortality) would seem to be (I've not read it) an example of a cluster of themes studied as they appear in a mass of documents.


35. It is unlikely, for instance, that a good modern historian would be as unsophisticated in regard to balance in presentation as Woodhead's study shows Thucydides to have been (see n. 22).
not necessarily always, true.\textsuperscript{36} For instance analysis can reveal an \textit{omission} which may be glaringly significant once revealed. One way of getting around this inferential problem consists of looking for clusters of things, or converging trends, pointing up the significance of a theme or issue. But, as the leap from data gathered to final inferences is not subject to the rigorous controls provided by the technique, findings are always liable to debate as to their validity. Consequently a content analysis study should include a check up on its findings. One takes one's (inferential) conclusions, turns them into questions and runs a content analysis based on these questions either on an unexamined portion of the text originally analysed or on a similar or related body of text(s). The objective is to produce findings which confirm or disconfirm one's initial conclusions by reproducing them or coming up with converging evidence – or the reverse.\textsuperscript{37}

Let me sum up by quickly reviewing the advantages and disadvantages of using this technique. Firstly, then, it keeps one's frame of reference and one's focus of attention uniform across time despite distractions. It enables one to process masses of diverse materials with uniform precision. One can easily check the validity of findings reached by content analysis, so it can compel conviction.\textsuperscript{38} It allows of cumulative research in other ways too: one can build upon the findings, or use the technique in several different but consecutive ways, within the one body of material.\textsuperscript{39} Secondly, it lessens what might be termed 'data-processing' problems. When using it, one does not have to produce one's own technique of analysis as well as the actual substantive study upon which one is engaged. Moreover it suggests all sorts of ways of going about the analysis. Thirdly, its very completeness often produces findings which one had not anticipated (the so-called 'Serendipity Effect'),

\textsuperscript{36} The classical study which demonstrates this is A.L. George's 'non-frequency content analysis'; see pp. 7-32 of Pool's \textit{Trends in Content Analysis}.

\textsuperscript{37} For instance I analysed the books and articles cited in \textit{A.J.P., J.H.S., J.R.S} and \textit{T.A.P.A.} for 1945, 1955 and 1965 to see what Classicists were reading; when I found them not to be reading in the Social Science area, I checked for Social Science titles in some contemporary books put out by leading presses: 'Problems and Prejudices in the Humanities', \textit{The Humanities Association Bulletin} (20) 1969, 32–34.

\textsuperscript{38} The findings of an impressionistic study cannot be proved; therefore they need not be accepted, and can be ignored. As it is not linked with any body of theory about man or society, such work can proceed from any kind of viewpoint. Hence, instead of building up a body of knowledge step by step in an integrated fashion with constant checkings on validity, a field in the Humanities can be a mass of random and unconnected observations. The achievements of German scholarship in Classics would seem to be connected with the greater amount of coordination in thesis work associated with German university traditions.

\textsuperscript{39} As, for instance, my series of studies on the picture of Marius at various points in antiquity allowed a survey of the changes in that picture from epoch to epoch (see note 16). But it was also possible to use the data which the content analysis collected on Marius' times as a check on the sources which the individual writers claimed to be using and on the survival of contemporary sources at later points in antiquity: cf. \textit{Rh. Mus.} 105(1962) 304 & 317.
because *all* aspects of the search are conducted over all parts of the documents. It thus often happens that questions which one would not ordinarily think of asking unexpectedly unearth significant findings.

But the basic advantage which the tool confers is this. Ordinarily one conducts one's inquiries in a 'proof or disproof' frame of mind, and one's attitude, in consequence, unconsciously tends to make one sensitive to the data in a partisan fashion. This technique, however, makes the whole enquiry of a neutral or *investigatory* nature. One is not trying to 'score points', merely to elicit all relevant data. The elaborate nature of the analysis prevents one from seeing what the findings imply (individually they are not significant enough to have much meaning, very often), and one is anyway simply too busy for anything other than the task in hand, that of extracting data.

One disadvantage is that, when conducting a content analysis, one cannot start adding to the things one is looking for halfway through. If one did, it would be impossible to say that all (or in fact *most*) of the text had been examined for all of the items involved. The pilot-study, or pre-test, is a way of meeting this difficulty: by analysing a part of the text one should find out what further items there are on which one may need to check. This defect does not mean that impressionistic studies are superior, however. One unconsciously concentrates on only part of the total perceivable range of data, anyway, in the act of focussing (and then this restricted focus shifts imperceptibly, in impressionistic studies).

A more serious disadvantage is that, by making one particularly sensitive to specific sets of things, this technique may impose patterns of emphasis on the data. Hence one has to take great care in setting up one's initial coding-units and categories. This can often be done with reference to some body of work which is known to be relevant to one's problems and in no way influenced by one's own idiosyncrasies and sensitivities.40 Often, too, the whole object of the analysis is to see whether in fact certain things do occur in the source material in certain combinations41: that is to say the test is simply 'Is such and such a pattern there or is it not?' In using this technique, however, one is always conscious of one's frame of reference, so the possibility of thinking critically about it exists (as is not the case in impressionistic studies, where it is unconscious).

---

40. It has been shown, for instance, that the concept of authoritarianism is relevant to classical antiquity: see E. E. Hagen, *On the Theory of Social Change*, Dorsey, Illinois 1962, Part II, esp. chap. 4 (whereas Freud's *libido* may well be of central importance only in modern western cultures: see M. Duverger, *The Idea of Politics*, Methuen 1966, 26–27). Work on authoritarianism (see the works of Rokeach and Adorno, note 12) indicates syndromes of characteristics to look for and sensitive social areas to look at.

41. Consider, for instance, the question: in his treatment of Marius, is it the case that Plutarch only has information on Marius when the latter is in some form of association with one or other of the contemporary memoirists?
Finally, content analysis is only one of the many techniques which a scholar should have in his control and, consequently, may often be only one process among several in conducting an investigation into a problem. As has been shown, one needs to know something about multivariate analysis to make fullest use of content analysis. But one may have to use both of these and other modes of analysis. Let me provide an illustration. Recently, I heard a paper on 'Propertius, the Parthians and Propaganda'. This actually consisted of a study of Propertius' poetry for references to the Parthians. In fact, however, it should have comprised:

(a) a study of Propertius' personality and his relationship with the Augustan Establishment;
(b) a content analysis of his poetry for the whole subject-area of Parthia (including Roman reactions to her experiences with Parthia);
(c) a similar study of the genre in which Propertius' poetry was written (to see what was owed to its requirements, and what was unique to Propertius);
(d) an identification of the target audience of the poetry;
(e) a content analysis of the reactions of that audience.

Only by considering all these things together can one investigate propagandistic aims or effects (the paper didn't specifically identify these different aspects of propaganda).

Content analysis is a tool. Without it there are certain things one cannot do. But it needs to be used with sensitivity, skill and discernment to get good results. A chisel does not make a Michelangelo - but he would never have produced his sculptings with his fingernails. Content analysis will not solve historical problems, for inference will always be involved in their interpretation. But it will move discussion of such problems to a new level by making sure that all the facts are in before inference commences, and by laying down foundations for work upon a problem in a spirit of objectivity and openness. Moreover any criticism levelled against it can also be levelled at the impressionistic 'method', which has certain disadvantages not shared by content analysis.

42. The point is that content analysis relates to one's overall views on human nature, let us say, or social change, as a tactic relates to its overall strategic background: see Young op. cit.96–102. But there is no space in this study to go into the matter of systems approaches, which deal specifically with bringing one's background frame of reference more fully under conscious surveillance. However, it is no mark of superiority in the impressionistic approach that it is as unconscious of problems in respect of background assumptions as it is of the lesser problems in respect of techniques of analysis.

43. For an illustration of content analysis used within a systems approach (as in this example) see Budd & co., op. cit.3–4.
**Matrix 1: Cicero’s Writings**

<table>
<thead>
<tr>
<th>THEME: Marius’ career</th>
<th>Source of information</th>
<th>Late, middle or early stage in writings</th>
<th>Frequency of mention (High, medium, low)</th>
<th>Fullness or skimpiness of detail</th>
<th>Errors or self-contradictions</th>
<th>Verbiage: striking or commonplace</th>
<th>Attitude: pro-neutral contra</th>
<th>Etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parents COMPONENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Childhood SUB-THEMES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adolescence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First military service</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First public office</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Watch for converging findings** — i.e. a sub-theme receiving emphatic treatment in several categories.
DIAGRAM no. 2

sample

measurement

coding unit

category

context unit

process of extracting data

matrix

conclusions

back check
DIAGRAM No. 3

hatred

ACTIVE

LOVE

PASSIVE

slime

COMFORT

PAM

ROSE

KITTEN

LEISURE

30

20

10

0

10

20

30

30

WEAK

168
Acta Classica is published annually by the Classical Association of South Africa. The journal has been in production since 1958. It is listed on both the ISI and the SAPSE list of approved publications.

For further information go to: