21. Measuring the condition of museum collections

Suzanne Keene
Museum of London, London Wall, London EC2Y 5HN, U.K.
Clive Orton
Institute of Archaeology, University College London, 31-34 Gordon Square, London WC1H 0PY, U.K.

21.1 Introduction

The project described below was prompted partly by the National Audit Office Report *Management of the collections of the English National Museums and Galleries*, and partly by the Museum of London's own intention to manage the preservation of its collections more effectively.

The objectives were to develop methods for using the minimum resources necessary to establish the condition of the collections, using the Museum of London as a test bed. The resulting statistics were to be used for planning work and resources for the care of the collections.

Overall strategy and the detailed procedures were developed by Suzanne Keene and the survey design, sampling methods and statistical analyses by Clive Orton. The project had to be completed to a tight deadline (30 March 1990); development of concepts and methods therefore had to proceed in parallel with the surveying. The surveys done last are thus the best planned and designed.

An initial search of the literature showed that most published work falls into two categories: recording the condition of individual objects (e.g. Craft & Jones 1981; Quandt 1986; Raphael 1987) or else general surveys by region or by type of object (Ramer 1989; Storer 1989). Although Ramer used statistical methods to choose which museums to inspect, there is no reference to such methods being used to assess the condition of the collections themselves. Collections surveys of the Horniman Museum relied on surveying every object (Walker & Bacon 1987). Work reported on archive and library surveying seems more promising (Anon 1985, on a strategy for surveying the documents of the U.S. National Archives and Records Service; Anon 1984, refers to a sample survey of the Library of Congress collections); but fuller information on these projects has not yet been located. Library and archive material is however less variable than that in museum collections.

21.2 Definitions of population and data

Because collections differ widely in their level of organisation and internal variability, we need as a preliminary step to be able to classify them according to organisation and variability (section 21.2a). The next step is to define the population and set up a sampling frame from which samples can be drawn in a rigorous statistical way (section 21.2b). Before a sample can be designed, we need to establish the questions that will be asked, and which will be chosen as the 'key' question(s) for which the design will be optimal (section 21.2c). Once all these decisions have been made, the design itself is almost a formality, although a pilot survey is extremely valuable in establishing initial values of the 'key' variables (section 21.3).

21.2a Types of collections

Collections were classified into three organisational types:

(i) *Well organised*. Neatly arranged in store, with all objects inventoried. Total number of objects (population size) known.
(ii) *Partly organised*. Well arranged in store, well described generally, but incomplete inventory and little idea of population size.
(iii) *Unorganised*. Dispersed or disorganised in store, few or no objects inventoried, no factual estimate of population size.

21.2b Description of the population

The following terms had to be defined:

(i) *Collection*. An administrative unit within the overall collection. There can be collections within collections.
(ii) *Store*. A self-contained room in which collections are kept.
(iii) *Storage location*. An important concept, on which the survey design rests. The smallest identifiable grouping of objects within a store, e.g. a shelf, a box on a shelf, a group on the floor, etc.
(iv) *Object*. The unit to which an individual observation relates. An object made up of component parts is counted as a single object.

21.2c Measurement of condition

Two variables were measured:

(i) *Damage type*. Eight broad types of damage were defined. They proved flexible but precise enough to apply to all the types of objects surveyed.
(ii) *Conservation priority*. Four priority ratings were defined:
   1. Urgent. Object at serious risk of further deterioration.
   2. High. Object needs remedial treatment to prevent further deterioration.
   3. Low. Seriously disfigured but not deteriorating; treat before display.
   4. Little. No work needed, or superficial cleaning only.

The second variable was the one used to define 'key' values (usually category 1 plus category 2) for sample design, and will be the one mainly considered here.

21.3 Survey design

In most cases, direct random sampling of objects is not possible, because there is usually no comprehensive list from which samples can be selected. Even if one were available, this might not be the best way to proceed, because of the high overheads associated with locating specific items in a store.

21.3a Principles of design

It was decided to adopt a two-stage sampling procedure with storage locations as the first stage and individual objects as the second. This gives us the flexibility to design samples for different levels of between and within location variability. The notation used is as follows:

	population	sample
no. of locations	N	n
objects at *i*th location	M_i	m_i
total objects	M_o	
proportion of objects in chosen category in *i*th location	P_i	p_i
overall proportion of objects in chosen category	P	p

We sample a proportion f_1 of the locations, with equal probability, and then a proportion f_2 of the objects at each location. Since the proportion f_1 is the same for all locations, this is a *self-weighting sample*. To estimate the overall proportion $\hat{P}$ we use the *ratio-to-size estimate*

$$\hat{P} = \frac{\sum M_i P_i}{\sum M_i} = p \qquad (1)$$

(hence the term *self-weighting*).

We are aware that there are other sampling strategies which may be more efficient (e.g. the probability proportional to size, or pps, method). But we have chosen the above method on the grounds of simplicity in both operation and analysis. A formula for the variance in the general case is given by Cochran (1963: 302). By defining a dummy variable

$y_{ij} = 1$ if object j in location i belongs to the key category,
$= 0$ otherwise, we have $p_i = \bar{y}_i$, and his formula becomes

$$\begin{aligned}\mathrm{var}(\hat{P}) = {} & (1 - f_1) \sum M_i^2 (p_i - \hat{P})^2 \,/\, n\bar{M}^2(n-1) \\ & + f_1(1 - f_2) \sum M_i (m_i/(m_i - 1)) p_i q_i / n^2 \bar{m} \bar{M}\end{aligned} \qquad (2)$$

21.3b Pilot survey

A pilot survey for each collection was undertaken in order to:

(i) quantify the task: number of stores, number of storage locations by store, variation in number of objects per location,
(ii) establish the 'survey rate' for the collection, i.e. the number of objects that could be surveyed per person-hour, and time overheads per store and per location,
(iii) broadly categorise the collections,
(iv) finalise the definitions of damage types and conservation priorities for the collection,
(v) establish the variability of the collection, both number of objects per location and their condition, to assist in design of the main survey.

21.3c Designing samples for specific surveys

The task of design consists of achieving a balance between the sampling fractions f_1 and f_2, depending on the relative variability between and within locations. It is intuitively clear (for example) that if all objects at a location are in similar condition, but the locations differ widely, we would need a high value of f_1 but a low value of f_2. The overall size of the sample depends on the resources available and the time needed to survey. This is best expressed as (i) a fixed overhead c_0 per store, (ii) an overhead c_1 per storage location and (iii) a time per object c_2. Cochran (1963:314) gives the formula for optimal allocation as

$$\bar{m}_{opt} \approx \left(s_2/\sqrt{s_b^2 - s_2^2/\bar{M}}\,\right)\sqrt{c_1/c_2} \qquad (3)$$

where $\bar{m}$ is the average number of objects sampled per location, s_2^2 is the within-location variance and s_b^2 the between-location variance. The overhead c_0 does not enter this equation directly, but must be subtracted from the total time available before the size of the sample is calculated. The factors c_1 and c_2 can be difficult to measure directly. We have found it possible to estimate them from overall work rate figures, using linear regression (see section 21.5a).

21.4 Analysis

21.4a Precision of the results

Equation (2) enables us to attach standard deviations to the estimate $\hat{P}$ of the proportion of objects in any chosen category. These in turn allow us to quote confidence intervals for proportions or numbers in each category.

21.4b Comparative results

The proportions in different categories in different collections, or in different parts of the same collection, can be compared by the use of contingency table analysis. Rather old-fashioned chi-squared tests are adequate for showing differences and high-lighting anomalies. We have found no need here for more advanced techniques such as log-linear analysis.

21.5 Case study — the Social History/Applied Art Collection

This is an extremely heterogeneous collection, consisting of a wide range of classes of objects divided among eight separate stores. We have chosen it as a case study here because more detailed design work was applied to it than to the other surveys. The resources available were two person-months, including the pilot survey and writing-up time.

21.5a Pilot survey

The pilot survey examined 684 objects in 42 locations, from an estimated total of 49,500 objects in 2993 locations. All the objects in each chosen location were examined, taking a total of 29.75 person-hours.

The variances given by the pilot survey are $s_2^2 = 0.062$ and $s_b^2 = 0.075$. To estimate c_1 and c_2 we had the number of locations surveyed (n_i), the number of objects surveyed (m_i) and the time taken in hours (t_i) at each of eight stores. Fitting the equation

$$t_i = c_1 n_i + c_2 m_i$$

gives c_1 in the range 0.4 to 0.5 and c_2 in the range 0.013 to 0.019 (i.e. 25 to 30 and 0.8 to 1.2 minutes respectively) but the fit was bad. Moving to

$$t_i = c_o + c_1 n_i + c_2 m_i$$

gave a better fit but a negative (although non-significant) value of c_2 ($c_0 = 2.77$, $c_1 = 0.30$, $c_2 = -0.01$), giving a new meaning to the saying "I'll do it in no time!" The value of c_0 seemed intuitively too high; we finally settled on the values $c_0 = 1.0$, $c_1 = 0.30$, $c_2 = 0.013$ as a compromise between statistical fit and commonsense. These values lead to $\bar{m}_{opt} = 4$, but owing to an arithmetic error a value of 2 was used, leading to a sample design of $f_2 = 1/8$, $f_1 = 1/4$, with a total of about 1500 objects to be examined at 750 locations. Because of uncertainties in the factors on which the design was based, provision was made for further sampling if time allowed.

21.5b The main survey

In the event, 2449 objects were surveyed from 991 locations, giving a value of m of nearly 2.5. The percentages in the four priorities were

priorities	1	2	3	4
pilot survey	0.7	12.4	40.2	46.7
main survey	5.8	12.4	34.2	47.4

The standard deviation for the proportion in categories 1 and 2 together (18.2% in the main survey) was about 0.85%, compared with an expected value for the survey as designed of about 1.0%, the difference being accounted for by the greater size of the actual sample.

21.6 Discussion

21.6a Design

It is interesting that there is very little difference between the s.d. based on $\bar{m} = 2$ and the one that would have been obtained if a value of $\bar{m} = 4$ had been used instead. The optimum value is the peak of a very flat surface, and there is no need to worry unduly about getting the design exactly optimal. For interest, we looked at the s.d. that would have been achieved if we had just scaled up the pilot survey to the time available. This would have led to a survey of all the objects at 450 locations (i.e. about 7500 objects), with a s.d. of about 1.2%. This shows that there would be no point in examining more than a few objects at each location of this collection.

21.6b Accuracy of observations and data

Any conclusions will only be as accurate as the data on which they are based. The survey teams felt confident that damage factors could be objectively assessed; slightly greater problems were found in allocating conservation priorities. However, reference to the definitions usually allowed decisions to be taken with confidence.

Great attention needs to be paid to defining the priorities and the storage locations. Definitions need to be checked against the type of collection being surveyed, as part of the pilot survey, to make sure they can be applied unambiguously.

There is bound to be a difference in perception between one surveyor and another, and so far we have done no work on this. Two people surveying together helps to average out perceptions. If the methods were to be generally adopted, it would be useful for museums to exchange surveyors for a short time, so that definitions and methods could be uniformly applied.

21.6c Measuring condition over time

One potential use of such surveys is to measure, or audit, changes in the condition of collections over time. Two dangers may arise:

(i) surveyors may adjust their perceptions so that even the best collection tends to have an average proportion of 'priority 1' objects in it,
(ii) the definitions may subtly change over time.

21.6d Survey procedures

It was definitely preferable to conduct surveys in pairs. Decisions could be taken jointly, and the work went more than twice as fast if one person could record while the other examined objects.

21.6e Management

The 'overheads', i.e. the time taken to get to stores, to move around stores, to identify sample storage locations, to collate data and to write up reports all took longer than expected. However, the use of a pilot survey helped to minimise these problems.

A more serious practical problem is how survey staff can be employed in the interval between the pilot and the main survey. There is an inevitable delay while the results of the pilot are assessed and the main sample is designed. One answer might be to use computer software which could help the surveyors to design the main sample themselves, without recourse to specialist statistical advice (see below).

21.7 Conclusions

The methods which have been developed meet all the needs of this project. Using them, six person-months were sufficient to arrive at a statistically reliable assessment of the condition of extremely varied collections comprising in all some 280,000 objects. Only the minimum of simple data were collected for each object surveyed, yet the analysis gives information which can form the basis of a number of different assessments of work required, priorities, causes of deterioration, and even such basic management information as the numbers of objects and quantifications of storage and maintenance costs.

21.8 Further work

The experience gained and the data collected can be used to develop the means for the efficient collection and analysis of data in the future, by

(i) fuller analysis and writing-up of results of these and other surveys, e.g. by cross-tabulating one variable against another. This will involve computer analysis, which if at all possible should be done with standard software, e.g. the Rothamsted General Survey Package (Anon 1989). Only as a last resort should software for analysis be specially written;
(ii) designing standard procedures and recording forms which could be used in the future at the Museum of London and elsewhere;
(iii) developing software for the design of surveys from parameters supplied by a pilot survey. None of the standard survey software that we have yet seen does this.

Acknowledgements

Research time was provided by the Museum of London and University College London. A grant of £7,000 was made by the Office of Arts and Libraries to employ research assistants for five months to undertake surveying.

References

ANON. 1984. "Library of Congress surveys collections", *The Abbey Newsletter*, 8: 5.

ANON. 1985. "Archives' plan stresses storage conditions", *The Abbey Newsletter*, 9: 2.

ANON. 1989. *The Rothamsted General Survey Program (RSGP)*, Rothamsted, Herts, Lawes Agricultural Trust.

CRAFT, M. & S. JONES 1981. *Written documentation AIC 9th Annual meeting*, Philadelphia, American Institute for Conservation.

COCHRAN, W.G. 1963. *Sampling Techniques*, 2nd Edition, New York, John Wiley.

QUANDT, A.B. 1986. *The use of computers in surveying the manuscript collection of the Walters art gallery, New directions in paper conservation*, Institute of Paper Conservation.

RAMER, B. 1989. *A Conservation Survey of Museum Collections in Scotland*, Scottish Museums Council/HMSO.

RAPHAEL, T. 1987. "Collection condition surveys for leather and skin products", *Leather Conservation News*, 3: 2.

STORER, J.D. 1989. *The Conservation of Industrial Collections*, London, Science Museum Conservation Unit.

WALKER, K. & L. BACON 1987. *A Condition Survey of Specimens in the Horniman Museum*, London, Summer Schools Press: 337–341.