POST SEASON MAIL SURVEY TECHNIQUES AND PROCEDURES

W. SCOTT OVERTON, JR., Florida Game and Fresh Water Fish Commission

Proc. Annu. Conf. Southeast. Assoc. Game & Fish Comm. 7:71-81

The problem of obtaining accurate game kill statistics is being given serious consideration by many game agencies at the present time. A number of methods of obtaining this information have and are being used with varying degrees of success. The primary problem involved in the use of surveys of kill success and hunting pressure is the determination of the validity of the results, and many of the methods are difficult or impossible to evaluate. As a source of data of measurable accuracy and as a basis for comparison with experimental procedures such as bag checks, the post-season mail survey is ideally adapted.

Mail surveys have been conducted in Florida for three seasons. The present paper attempts to outline routine procedures found best adapted for use in the Florida surveys in the hope that other game agencies planning similar surveys will be able to benefit from our investigations.

Acknowledgments are due Dr. A. L. Finkner and Dr. R. L. Anderson of the North Carolina State College Department of Experimental Statistics for invaluable guidance in setting up the Florida mail survey and the statistical procedures for its evaluation; to the entire staff of the Game Management Division of the Florida Game and Fresh Water Fish Commission for many suggestions and much collection of data; and to many researchers in other states who have contributed helpful ideas.

It is admitted that many of the procedures here presented are arbitrary, and will probably cause some controversy. Such controversy will be welcome as we are yet attempting to improve these procedures.

PLANNING

Much of the success of a survey depends on the thoroughness of the preliminary planning. Survey designs, office procedures, schedules for interviews, techniques of analysis: all must be planned as fully as possible. In Florida, planning has included personal contact and much correspondence with County Judges in assuring prompt delivery of the carbon copies of licenses, from which the sample is drawn. In addition, effort has been directed towards improving the license design in order to facilitate their legibility. This activity has resulted in a vastly improved source of names and addresses of licenses, although there is still room for considerable improvement.

DRAWING THE SAMPLE

The procedure for drawing the sample used in the Florida surveys is as follows: licenses are filed by county in order of receipt from the license selling agency. The sampling ratio, or interval, is calculated as N_j/n_j (number of licenses in jth stratum/desired sample in jth stratum). The first license to be pulled is determined

by the use of a table of random numbers, and every $(N_j/n_j)th$ license thereafter is pulled from the file. Starting points and sampling ratios are determined for each stratum.

Whatever the source of names — reports from license selling agents, license stubs, or carbon copies of licenses — it is likely that illegible or otherwise unusuable names and addresses will occur. In the Florida surveys, this problem has been of considerable importance and it has been necessary to select a standard procedure for such a situation, as follows: the next succeeding usable name is used, if it falls within five names of the one originally selected. If not, the next preceding usable name is used, if it falls within five names of the original selection. If, after checking five forward and five backward, no usable name has been found, the procedure is repeated for the next five forward and the next five backward, etc. In the event that no usable names are available between one-half the sampling interval forward and one-half the sampling interval backward, the attempt to locate a substitute name is discontinued and the normal sampling procedure is resumed from the point of interruption.

This procedure involves the assumption that each name is most representative of those closest to it; an assumption that is in accord with the use of the sampling interval.

STRATIFICATION

Stratification by ecological or administrative zones is often necessary or desirable in a survey of statewide scope, particularly if individual estimates are desired for these zones. In addition, if considerable difference in activity exists from one zone to another, the standard error of statewide estimates may be minimized by adjustment of sample size within strata. When estimates of s_j are available:

$$n_j = \frac{N_j \, s_j}{S(N_j \, s_j)}$$ (Hendricks 1942)

where: s_j = standard deviation of the individual measurement for the jth stratum.
n_j = sample size for the jth stratum.
N_j = number of licensees in the jth stratum.
S = summation.

The inclusion of eight or ten species in the survey limits the application of this formula and the Florida surveys have been conducted with approximately the same sample size in each of the five administrative Divisions in order to get estimates of comparable accuracy for each Division. This procedure leaves something to be desired, and will possibly be changed in the future.

QUESTIONNAIRE DESIGN

Questionnaire design is probably the most important single controllable factor affecting the accuracy of replies. All questions must not only be concise, but must

have no possibilities of misinterpretation. It is doubtful if a perfect questionnaire will ever be designed; certainly we have not yet designed a perfect one in the Florida surveys.

For example, in the 1950 - 51 survey, the question was asked, "How many days did you hunt the listed species? (Give number of days)." On tabulation it became obvious that many respondents were adding fractions of days into wholes. In the 1952 - 53 surveys this question was asked, "On how many different days did you hunt? (species)?" To this question were received such answers as "Saturdays only," "Legal days only," "All different days," etc. Percentage of occurrence of such answers was low, but indicates imperfect questionnaire design, and complicates analysis of the data.

Another factor to be considered in the design is the ease of tabulation. An attempt to design a questionnaire well adapted to key punching for machine tabulation is as yet unsuccessful. The basic problem involved is in inducing the placement of answers in the desired blanks. Kill entered in the blank for days and days in the blank for kill are very noticeable occurrences. This particular problem can be minimized by requesting the kill data before the question on days. How to prevent answers being placed all over the page is a moot question.

A very important point in questionnaire design is attractiveness and order. There is a tendency in both mail and interview questionnaires to skip questions, or otherwise to fill out the questionnaire incompletely. This may be minimized by liberal spacing and grouping and by the elimination of superfluous questions.

The post card is probably the most popular form of questionnaire for use in kill and hunting pressure surveys because of its advantage of economy and ease of handling. However, the tabular form necessary in order to include questions on kill and days for eight or ten species provides maximum opportunity for entering answers in the wrong position and, it is believed from use of post cards, for overlooking species of relatively minor importance. The 1952 - 53 surveys were conducted with full sized, fully explanatory questionnaires in an effort to eliminate such inaccurate and incomplete answers. A test survey utilizing post card questionnaires (mailed in the same fashion and accompanied by the same letter or introduction as the full questionnaire) was conducted of a comparable sample of licensees in one of the five Divisions. The volume of response for these two surveys was almost identical (Table 1). Analysis for differences in results is not yet complete and cannot be reported at the present time.

MAILING PROCEDURES

Class of outgoing and incoming material, and intervals between successive mailings must be considered in planning the survey. Approximately 90 per cent of the potential returns are received within three weeks after the questionnaire are mailed, and this interval has been adopted as a standard in planning office schedules.

For the 1952 - 53 survey two secretaries were able to address and prepare outgoing material, record incoming questionnaires and pull names of respondents from the mailing list on this schedule. This survey involved stuffing and sealing first class envelopes for an initial mailing list of 8000. The use of double return post cards would require during the mailing period about two-thirds of the office time required by the type of survey used.

Table 1. Response and record of non-delivery of the three annual mail surveys, illustrating effect of different mailing and questionnaire techniques. Percentages calculated as per cent of initial mailing.

		First Mailing				Second Mailing				Third Mailing			
	Number Mailed	Percent Returned	% Returned Undelivered	Quest. Form	Outgoing Class	Percent Returned	% Returned Undelivered	Quest. Form	Outgoing Class	Percent Returned	% Returned Undelivered	Quest. Form	Outgoing Class
1950 - 51[a]	4339	16.9	1.9	Double Post card	Stamped	28.6	9.1	Single Post card	First Class	No Third Mailing			
1951 - 52													
Northwest	890	13.3	1.0	Double	Non-Metered	16.6	0.2	Double	Non-Metered	8.7	0.8	Double Post card	Non-Metered
Northeast	899	10.8	1.6	Double	Non-	13.5	0.7	Double	Non-	7.3	0.2	Double	Non-
Central	898	16.4	0.3	Double	Non-	14.0	0.4	Double	Non-	9.0	2.9	Double	Non-
South	905	13.7	0.2	Double	Non-	17.3	0.2	Double	Non-	12.4	0.0	Double	Non-
Everglades	897	16.9	0.0	Double	Non-	16.3	0.0	Double	Non-	9.5	0.7	Double	Non-
Total	4489	14.2	0.6	Double	Non-	15.5	0.3	Double	Non-	9.4	0.9	Double	Non-
1952 - 53													
Northwest	1005	23.6	6.8	Quest.	First Class	18.9	1.9	Quest.	First Class	No Third Mailing			
Northeast	1121	23.4	8.5	Quest.	First	7.9	1.5	Letter only	First	21.8	1.1	Quest.	First
Central	1066	25.6	7.8	Quest.	First	10.1	1.6	Letter only	First	23.4	0.8	Quest.	First
Central	1066	25.1	8.0	Single Post card	First	10.7	1.1	Letter only	First	22.9	1.0	Single Post card	First
South	1030	29.8	4.8	Quest.	First	11.7	1.2	Letter only	First	21.7	0.3	Quest.	First
Everglades	989	26.5	9.3	Quest.	First	23.5	2.3	Quest.	First	No Third Mailing			

[a] The second mailing consisted of 20% of the non-respondents to the first mailing. Figures for second mailing are estimates of relative returns for comparison to other data. Actual mailing was 636, returns 219, undelivered 70. For other years, each mailing included all non-response to previous mailings.

It can be seen from Table 1 that two first class mailings produced about 47 per cent response, while the returns from the double post card total only 39 per cent from three mailings. The costs of sample preparation, addressing and preparing material for mailing, and recording returns have been estimated as closely as possible. These costs, plus postage, average about $0.40 per return for both the survey involving first class mailing and the survey involving the non-metered — business-reply double post card. This is exclusive of administrative and analytic costs, and does not take into account the reduction in the size of the interview sample by the greater percentage of response.

In an effort to cut mailing costs, second mailings to three of the five Divisions in 1952 - 53 included only a letter requesting that the questionnaire sent several weeks before be returned. As may be seen from Table 1 this trial was a dismal failture, and it was necessary to make a third mailing to these Divisions. It is impossible to estimate from these data the volume of returns that would be received from three first class mailings, each including questionnaires, but it is well illustrated that a larger volume, at the same average cost, was returned over a shorter period of time by the use of first class mailing.

NON-RESPONSE

Non-response is a major problem in mail surveys. Even with improved mailing techniques providing increase returns, the volume of non-response is sufficient to contribute considerable bias if there is any appreciable effect of hunting success or participation on the likelihood of response. Surveys in other fields have illustrated such an effect, and surveys of game kill are no exception (Table 2).

Generally, bias of non-response seems greater in surveys dealing with a single species than in surveys dealing with many species. Also, in the surveys dealing with several species, there seems to be a considerable difference in the effect of non-response on the different species. The analysis of these characteristics has not progressed to the point at which it may be reported, but the evidence that these differences occur is clear.

Detailed analysis of non-response is being made in the attempt to set up historical data on which to base regression equations by which the bias may be corrected. In the meantime, it is necessary to estimate by some means the success and activity of the non-respondent group. A personal interview survey of a sample of the non-respondents is widely used in this capacity, and was adopted for the Florida surveys. The basic theory behind this approach is simple — the sample is considered representative of the entire group of non-respondents, and the estimates are calculated by the use of the following formula:

$$\bar{y} = \frac{r\bar{y}_1 + (n - r)\bar{y}_2}{n}$$

Where: $\bar{y}$ = estimated mean
$\bar{y}_1$ = mean for mail response
$\bar{y}_2$ = mean for interview sample
r = number of mail respondents
n = number in initial sample

Table 2. Arithemetic averages of Divsion Means of successive returns. Florida Post Card Survey, 1951 - 52 Hunting Season.

	Kill per respondent					Days per respondent			
	Mail returns			Interview		Mail returns			Interview
Species	First	Second	Third	Sample	Species	First	Second	Third	Sample
Deer	0.08	0.09	0.05	0.05	Deer	2.4	1.6	1.8	2.6
Turkey	0.17	0.16	0.14	0.10	Turkey	1.2	0.9	1.1	0.8
Quail	12.20	12.20	13.30	13.90	Quail	3.2	3.5	3.8	3.5
Squirrel	7.60	6.20	8.30	6.00	Squirrel	2.8	2.7	3.3	2.6
Dove	8.20	7.80	7.30	5.50	Dove	2.1	2.1	2.1	1.4
Duck	3.20	2.20	2.10	1.00	Duck	1.7	1.2	1.1	0.5
Coot	2.20	1.80	1.80	0.80	Coot	0.7	0.5	0.7	0.3

Recommended sampling ratios for these enumerative sub-samples range from 20 to 33 per cent, 20 per cent probably being sufficient for game kill surveys. In order to allow for the contact potential of 85 per cent (with poor record of non-delivery this potential has been around 75 per cent) and a realized contact percentage somewhat lower, a 25 per cent sample of the non-respondents is drawn. About 75 per cent of the interview sample was completed during the 1952 - 53 surveys. The 1951 - 52 survey was followed-up by an interview of all non-respondents in two sample counties in each Division. This procedure is not as desirable as a random sample drawn from non-respondents in all counties (as in 1952 - 53), but was used because of a personnel and time shortage.

NON-DELIVERY

Return of questionnaires by the Post Office as undeliverable due to insufficient address, persons moved leaving no forwarding address, and other causes, is an important item of consideration. Table 1 includes the record of non-delivery for the three years of mail surveys. It will be noted that the yearly differences in return by the Post Office are due to the class of mailing. It is believed that better record could have been obtained on the non-metered post cards used in 1951 - 52 by a guarantee of return postage printed on the card.

In the usual treatment of this group, the assumption is made that non-delivery is random, and the number returned in this manner simply deleted from the survey sample. However, the large number of transients in Florida, particularly during the winter months, undoubtedly contribute heavily to non-delivery. Therefore, it is likely that this group is best represented in the Florida surveys by the non-respondent group, and is so treated.

UNUSABLE ANSWERS AND INCOMPLETE DATA

The problem of unusable or incomplete questionnaires, in whole or in part, is of considerable importance in mail surveys. As indicated earlier, this problem can be minimized by careful questionnaire design, but it is likely that unusable answers will always occur to some degree. Treatment in any particular instance is dependent on many factors, and the only general rule that may be made is that complete objectivity be maintained at all times during the editing and coding operations and in the calculation of estimates.

Treatment of such cases has been standardized to a large extent in the Florida surveys. This treatment is insofar as possible in accord with the following considerations:

1. The estimation of total kill is of first ranking importance in the surveys.
2. It is assumed that where no kill is reported or indicated none occurred.
3. It is assumed that all kill requires hunting effort and the following situations are treated thus:

1. No information — treated as non-response.
2. No information for any one species, general answer, "Hunted but did not kill anything" — treated as response, non-participation for all species.
3. No information for any one species, general answer, "Hunted but did not kill much" — treated as non-response.

4. Indication of kill for particular species, number not reported — coded as "unknown kill."

5. Days hunted for particular species reported in such a manner as to be unusable — treated in same manner as a report of "no days" and therefore coded as "none."

6. Indication of participation for particular species, but no information — occurrence low, inclusion with "no kill and no days" in accord with assumption that where no kill is reported or indicated, none occurred.

The tabulation of data for any one species should be in a manner similar to the following example:

	n	SK	SD	SK^2	SKD	SD^2
Not hunting species	n_0					
Hunting: no kill and no days	n_1					
Unknown kill and no days	n_2					
Kill and no days	n_3	SK_3		SK_3^2		
No kill and days	n_4		SD_4			SD_4^2
Unknown kill and days	n_5		SD_5			SD_5^2
Kill and days	n_6	SK_6	SD_6	SK_6^2	SKD_6	SD_6^2
Total	r	SK	SD	SK^2	SKD	SD^2

The importance of each of these classes varies with questionnaire form and with species. On the post card questionnaires used in Florida, the indication of kill was very low when information was not obtained, and classes 2 and 5 were included in classes 1 and 4, respectively. Preliminary examination of the 1953 survey data, collected by the use of full page questionnaires, indicates that classes 2 and 5 may be sufficiently large to warrant individual treatment in some cases. Class 4 is relatively small for most of the small game species, and increases in significance for species with lower kill success.

It would seem unnecessary to present formulae for all of the possible combinations of significant classes, but rather to recommend formulae for calculating missing values where such values are deemed necessary.

$$SK_2 = SK_3(n_2/n_3)$$
$$SK_5 = SK_6(n_5/n_6) \text{ or } SK_6(SD_5/SD_6)$$
$$SD_1 = SD_4(n_1/n_4)$$
$$SD_2 = SD_6(n_2/n_6)$$
$$SD_3 = SD_6(n_3/n_6) \text{ or } SD_6(SK_3/SK_6)$$

It has been indicated that in most cases some of these class tabulations will be unnecessary. However, it is much simpler to tabulate unnecessary classes and then combine, than it is to retabulate for special cases, and these tabulations give a positive record of incidence of incomplete reports. It is likely that with standardized questionnaires it will be possible to anticipate necessary tabulations for each species, but until such time, tabulation in the manner described will likely be more satisfactory.

In calculating mean kill and mean days from the tabulate data, it is important to remember that the simpler the calculation — the less frequent the calculation of

missing values — the easier the calculation of confidence limits. Here again applies the rule of objectivity: if the more logical treatment makes little or no difference in the means, the choice should lie solely with ease of treatment. In the calculation of SK_5 and SD_3 the first and the two suggested formulae excel in this respect.

ESTIMATES

Means may be calculated by hunters of the species ($SK/_{r-n_0}$ and $SD/r - n_0$) or by respondents ($SK/_r$ and SD_r). The latter is preferred by the writer as population estimates of this mean are expressed as "kill per licensee" and "days per licensee," and the calculation of total kill and confidence limits is simplified.

Estimates from an unstratified survey are obtained by the formular previously outlined:

$$\bar{y} = \frac{r\bar{y}_1 + (n - r)\bar{y}_2}{n}$$

and

$$s^2_{\bar{y}} = \frac{s^2}{n} + \frac{(n - r)(k - 1)s_2^2}{n^2}$$

where: $k = \frac{n - r}{r_2}$

n = number in sample.
r = respondents to mail survey.
r_2 = number of schedules completed in sample of non-respondents.

and

$$s^2 = \frac{rs_1^2 + (n - r)s_2^2 + r(\bar{y}_1 - \bar{y})^2 + (n - r)(\bar{y}_2 - \bar{y})^2}{n}$$

The formulae for s_1^2 and s_2^2 are dependent on the methods of calculating $\bar{y}_1$ and $\bar{y}_2$ respectively, and a statistician should be consulted before these estimates are attempted. If no corrections for missing values are made, and if r is sufficiently large to be assumed equal to $(r - 1)$,

$$s_i^2 = \frac{r_i(SX_i^2) - (SX_i)^2}{r_i^2}$$

and if r is small,

$$s_i^2 = \frac{r_i(SX_i^2) - (SX_i)^2}{r_i(r - 1)}$$

where: s_i^2 indicates s_1^2 or s_2^2
s_i indicates r or r_2.

For a stratified survey of 5 strata,

$$E = \sum_{j=1}^{5} N_j \bar{y}_j$$

$$\bar{y} = \frac{1}{N} \sum_{j=1}^{5} N_j \bar{y}_j$$

$$s^2_{\bar{y}} = \frac{1}{N^2} \left[\sum_{j=1}^{5} \left(N_j(N_j - n_j)s_j^2/n_j + N_j^2(n_j - r_j)(k_j - 1)s_2^2 j/n_j^2 \right) \right]$$

where: N = number of licensees in state.
N_j = number of licensees in the jth stratum.

ACCURACY

In Table 3 are presented the means and standard errors of means calculated from the first two mail surveys in Florida. It will be noted that sampling error is very large.

Table 3. Estimates of sampling error of mean kill per licensee as determined from the Post Card Survey. 1950 - 51 data calculated from two mail returns. 1951 - 52 data calculated from three mail returns and interview sample of non-respondents.

	1950 - 51		1951 - 52	
Species	Mean	$s_{\bar{y}}$	Mean	$s_{\bar{y}}$
Deer	0.0659	0.0145	0.0650	0.0127
Turkey	0.1604	0.0303	0.1074	0.0184
Quail	10.4900	1.1300	13.6300	1.5000
Squirrel	7.0800	0.5900	7.3300	1.5600
Dove	5.6400	0.63	6.2100	0.5300
Duck	1.4830	0.2330	1.5630	0.1880
Coot	1.5090	0.3830	1.2520	0.1910
Goose	0.0155	0.0107	0.0060	0.0018

The accuracy of the individual answer is also in many cases open to question. In this respect, the time element is very important; the greater the time lapse between the hunting season and obtaining the answer, the greater the probability of serious error of memory. Little is known at the present time about the bias resulting from inaccurate replies. Although studies are being conducted elsewhere, no effort is being devoted in Florida to study of the bias itself.

In 1952 - 53, in an attempt to set up procedure eliminating delay after the hunting season, and theoretically reduce bias of memory, a test survey was conducted to the respondents to the 1951 - 52 survey. This "permanent mailing list" was conducted simultaneously with the random sample in order that results be comparable. About 40% returns were received from the first mailing of the

PML, in contrast to 25% received from the random survey. Second and third returns were about the same for both surveys. Tabulations are not yet complete, and comparisons of results cannot be made at the present time. It is hoped that the PML will prove useful, at least in auxiliary capacity, in future surveys.

REFERENCES

Cochran, W. G. 1950. Sample Survey Techniques. Mimeo Series #7, Institute of Statistics, North Carolina State College, Raleigh, NC.

Hendricks, Walter A. 1942. The Theory of Sampling. Mimeo Series #1, Institute of Statistics, North Carolina State College, Raleigh, NC.

Hendricks, Walter A. 1948. Mathematics of Sampling. Spec. Tech. Bull., Virginia Agri. Exp. Sta., Blacksburg, VA.