Mr Nakken

This report not to be cited without prior reference to the Council*

International Council for the Exploration of the Sea

C.M. 1976/B:14
Gear and Behaviour Committee

REVISED VERSION OF THE SECOND AND THIRD REPORTS OF THE WORKING GROUP ON STANDARDISATION OF SCIENTIFIC METHODS FOR COMPARING THE CATCHING PERFORMANCE OF DIFFERENT FISHING GEAR

Contents

* General Secretary
ICES, Charlottenlund Slot,
DK-2920 Charlottenlund,
Denmark

1. Introduction

The Working Group first met in Hamburg in April 1972 and presented its first report to the 60th Statutory Meeting of the Council. A revised version was submitted to the 61st Statutory Meeting and published in April 1974 (Cooperative Research Report No. 38, pp 1-22). Although much of what it contained was of wider application in the field of comparative fishing, the report had been written with bottom trawling specifically in mind.

Accordingly, the Working Group was invited to continue its study of the standardisation of methods of comparing different fishing gears and in particular to extend its discussions to gears other than bottom trawls.

Two further meetings were held. The first, which took place in Hamburg from 13-15 March 1974 dealt principally with methods of comparing two-boat bottom trawls, mid-water trawls and purse-seines. The participants were:

Dr H. Bohl, Convener	F.R.G.
Mr H.B. Becker	Netherlands
Dr U. Buerkle	Canada
Mr J.J. Foster	UK
Mr S. Lens	Spain
Mr C. Nédélec, Guest	FAO
Mr J.A. Pope, Secretary	UK
Mr M. Portier	France
Mr G. Vanden Broucke	Belgium
Mr J.G. de Wit	Netherlands

The second meeting, held in Ostende, Belgium, on 23-24 April 1975, dealt with passive gears. The participants at this meeting were:

Mr J.A. Pope, Convener	UK
Mr H.B. Becker	Netherlands
Mr R.S.T. Ferro	UK
Mr R. Fontayne	Belgium
Mr J.F. Foster, Guest	UK (WFA)
Mr J.J. Foster	UK
Mr L. Karlsen	Norway
Mr C. Nédélec, Guest	FAO
Mr S. Olsen	Norway
Mr J. Prado	France
Mr H. v. Seydlitz	F.R.G.
Mr J. van Hee	Belgium

Reports were presented at the 62nd and 63rd Statutory Meetings of the Council. These were subsequently revised and combined by Dr H. Bohl and Mr J.A. Pope to produce the present report.

2. Terms of Reference

The Working Group attempted no formal definition of active gears and passive gears, assuming that the allocation of at least the more important gears in the ICES area to one or other of these categories would cause no conflict of opinion. Thus active gears, which imply the movement of the catching device, include trawls and seines while passive gears, which rely on the movement of the animals to effect capture, include gill-nets, traps and static-hooked lines.

In addition to the need for procedural advice on the execution of experiments involving different versions of the same gear, the Working Group recognised the vital need for guidance on the comparison of the catching performances of different gears within categories (e.g. trawls versus seines) and of gears in different categories (e.g. trawls versus gill-nets). The results of such comparisons find their application in, for example, modes of allocating effort quotas in any multi-gear fishery regulated either nationally or internationally in this way. However, although there was some discussion of this important topic it was necessarily brief as the Group's terms of reference specifically excluded it. In the time available the Working Group could not deal adequately with a wide range of gear types and it was agreed that discussion should be restricted to include only gears of importance in the ICES area.

Guidance on the standardisation of procedures for carrying out comparative fishing experiments requires the prior identification of factors which may play a part in determining the actual catches which will be observed. Many of these are common to a variety of gears and methods of fishing and have been discussed very fully in the first report (Coop. Res. Rep. No. 38, Sections 4, 11). In general, therefore, this report deals only with those factors which were considered to be specifically relevant to the gears and methods discussed here. A feature of significant importance in fishing with mid-water trawls and purse-seines is the need for skill in manoeuvring the gear in three rather than in two dimensions. Thus, qualities such as human skill which are difficult or even impossible to alter or control are introduced and may play a dominant role. Special consideration was paid to these distinctive features particularly as they affect the design of comparative fishing studies.

Also included in this report is discussion of two important aspects of comparative fishing experiments - namely design and analysis. Whilst many experimental designs have been proposed and successfully applied to a wide range of experimental studies in many fields of investigation, the use of complicated designs in comparative fishing is not generally recommended because of the often high risk of failure to execute the plan due to uncontrollable factors such as adverse weather conditions, large changes in fish availability, etc. The choice of an appropriate method of analysing the results of an experiment is much more open to debate. Among other things it will depend on the particular functions to be estimated, the existence of previous relevant information, the experimental design used, the amount of data and the statistical properties of the observations. There exists a considerable volume of published material dealing with statistical methods of data analysis and the Working Group realised that it would be unwise, unnecessary and even impossible to attempt to cover other than a few topics of major importance.

3. Two-Boat Bottom Trawls

The vessels forming the pair in a comparative fishing experiment must be of the same type and power and the lengths of the vessels should be as nearly equal as possible. When comparing catches by different pairs the lengths of all vessels should be similar. The vessels of a pair should have similar towing arrangements, either over the side or over the stern, and warp loads must be equal on each side of the trawl. A description of the towing points and the method of gear handling (net drum, gallows) should be made.

Changes of course during towing should be avoided. If course changes have to be made the haul should be considered invalid. The distance between the two ships must be kept constant and this distance measured and recorded. Experiments should not be carried out in sea conditions that do not allow accurate distance keeping.

With regard to the fishing gear itself, the construction and material of the warps and the length paid out should be recorded; also the magnitude and placing of weights and the length, material, diameter and weight of the groundrope. A constructional drawing of the net should be supplied.

The gear factors which are either difficult to control or cannot be controlled independently are less complicated for bottom pair trawling than they are for bottom otter trawling. Differences in the angle of attack of the bridles may occur when the direction of the current at the bottom deviates notably from the direction of the current at the sea surface. Variations of the headline height occur when the weights touch the bottom heavily. This can happen at an upward sloping bottom or when the vessels are towing against a heavy swell.

4. "Aimed" Gears

4.1 General Considerations

There are close similarities between pelagic (mid-water) trawling and purse-seining. Both methods of fishing rely largely on the detection of suitable shoals of fish and the ability to manoeuvre the gear effectively in three dimensions. The accuracy and reliability of the fish detection equipment carried by the vessel and the skill of the skipper and his crew in interpreting the information provided by such equipment as well as in operating the gear itself may well outweigh any gear design factors. Any comparative study of the catching performance of different versions of such gears must recognise and allow for these important non-gear items.

Since the detection of dense shoals is of prime importance to the success of this type of fishing and as the existence of suitable shoals in a given area at a specified time is unlikely to be within the control of the experimenter, comparative fishing experiments of this type may take a considerably longer time to conduct than in the case of bottom trawling where a more even distribution can usually be more readily found.

Whilst comparative fishing with pelagic trawls and purse-seines should give less variable results if conducted at times when the fish are more evenly distributed in space and therefore may provide a possible way of shortening the period of study it must not be overlooked that such conditions do not correspond to those under which the gear would usually be expected to operate most efficiently.

Increasing the length of time required for an experiment can raise the costs to such an extent as to make it prohibitive for a study to be carried out by research vessels and the use of commercial vessels may be necessary in order to provide enough information. Increasing the duration of an experiment also increases the possibility of major changes occurring in conditions many of which cannot be controlled in any way by the experimenter. The need to collect a large amount of information on many factors known or suspected to influence catching performance becomes necessary in such situations. This state of affairs is, of course, not peculiar to comparative fishing experiments and many statistical techniques have been proposed to assist the analysis and interpretation of such observational studies.

4.2 Mid-Water Trawls

With the above general considerations very much in mind the Working Group reviewed specific aspects of comparative fishing with mid-water trawls. It regarded the approach by way of studies of the performances of similar commercial vessels using standard and experimental gears as being the one most likely to provide reliable conclusions. When working with commercial vessels it is very desirable to have experienced observers on board both the experimental and standard (control) vessels to make accurate records of all information to be collected. If enough observers are not available it is essential that properly designed log books be issued to the skippers of all participating vessels with clear and adequate instructions as to how records are to be kept.

Items of particular relevance to mid-water trawling which should be noted are the type of equipment used for locating shoals and the means employed for varying the levels of the net. Notes on fish locating equipment should include the method of operation, manual or automatic, details of scale expander and whether steering is automatic during location. During a tow any changes in level of fishing achieved by changing warp length and/or propeller thrust should be noted and the tactics employed by the skipper should be observed. The use of modern techniques of psychology to study the behavioural tactics of the skipper in response to the information presented to him by his equipment is recommended as a valuable new area of research in this field. It is to be expected that on commercial vessels adjustments to the gear will be made during the course of a trip. However, only changes necessary to permit the application of the best tactics in a given situation should be allowed. Changes in the basic design of the net and otter boards should not be made at any time. It is absolutely essential to note any adjustments made and their effect on the gear.

4.3 Purse-Seines

The Working Group noted that developments in purse-seines had taken place over the years and it was suggested that this was evidence that the fishing industry had succeeded in carrying out at least some kind of comparative fishing evaluation. Over recent years purse-seines had gradually become bigger, both in length and depth, sinking rates had increased through changes in hanging ratio and total lead line weighting, and maximum fishing depths had been considerably extended. The Group recognised that it could be argued that some of these "improvements" were achieved not by comparative evaluation of alternative approaches but simply by recognising the direction in which changes for improvement had to be made (e.g. "the bigger the net the better" principle). But some comparative evaluation must have been made at some stages in the evolution of the purse-seine, even if only on a trial and error basis.

Several "types" of purse-seine exist and each is operated in rather different ways. These types include

(a) shallow water purse-seines (e.g. those used for fishing sardines);

(b) deep water purse-seines (e.g. those used for herring);

(c) purse-seines used in association with fish attraction and aggregation procedures (e.g. those used in conjunction with light attractants).

The latter group is a special case in that it may not matter which design of gear is used provided the size matches the aggregation power of the light. Yet this could be the one case when, by controlling the density of fish available, a direct comparison of the effectiveness of different gears in catching the known concentration could be made.

It was agreed that a number of alternative comparisons should be considered which cover the groups (a) and (b) above. In particular, attention should be given to exploring comparisons of fishing units (vessel, gear and crew) in a similar way to that used by fishing organisations.

Short term comparisons of purse-seines were considered to be unfeasible in general. Comparison of the catching efficiency of different purse-seines in shallow water on scattered, reasonably uniformly distributed fish shoals might be possible but in general circumstances long term studies would be necessary. Two main categories under this heading were considered, namely (1) comparisons between gears on similar vessels when the catch per unit fishing time is used as the measure of efficiency and (2) between fishing units (vessel, gear and crew) when economic profitability over a fishing season or equivalent time is used as the measure.

In comparing different purse-seines used by similar vessels, experiments should be conducted basically along the same lines as described for mid-water trawling and these need not be repeated again here. It is essential that throughout the net shooting and hauling operations the sinking speed and pursing speed should be noted and also the fishing depth. Essential environmental data to be recorded are water turbidity, surface temperature, occurrence and depth of thermocline, general weather conditions, current strength and direction and, for shallow water operation only, the depth and type of the sea bed.

When comparing different purse-seines used by different vessels, separate assessment of the influence of all factors related to either ship or net is not likely to be possible. A "global" comparison could be made by taking into account the economic aspects of the fishery, the results of each fishing unit being recorded and analysed in terms of profitability (i.e. sales minus exploitation and depreciation costs). A sufficient indication of the difference between gears may be given where several vessels of similar capital and running costs are operating on the same grounds. Although the value of such comparisons for profitability analysis is not denied it was considered that techno-economic studies of the extent required are beyond the scope of comparative fishing. No definite technical recommendations are given, therefore, for conducting comparisons of this type.

5. Gill Nets

In its discussions the Working Group did not distinguish between the different types of set nets and drift nets.

For both set nets and drift nets it was considered that the main factors of importance in determining fishing performance were

(i) the hanging ratio (horizontal and vertical) defined as
E = length of line/length of stretched netting

(ii) the material of the netting yarn

(iii) the colour of the netting yarn

(iv) the flotation and weighting.

Catches from comparative fishing experiments with set nets and drift nets should be reported in terms of numbers and weight of fish per unit area of mounted net and per unit length of mounted net.

When comparing the catching performance of different versions of gill nets it is recommended that either entire fleets of the net being compared should be used or, if different nets are used within fleets, the lengths of the different sections of the fleets should not be too small. It is also recommended that the different nets be joined by ropes of say, 2-3 fathoms length rather than tying them closely together. In this way possible interaction between nets of different type which might affect the experimental results may be avoided or at least minimized. An interactive effect might occur, for instance, due to possible differences in visibility of the nets being compared in such a way that fish might be led into less visible nets by the more visible nets acting like a barrier. In this situation the higher catch of the less visible nets could partly be a result of the presence of the more visible nets and the observed difference in performance might be greatly reduced if the more visible nets were not present. If twine colour is not a factor being studied then all nets should be of the same colour.

Set nets may be set in straight or zig-zag lines and nets at the end of a fleet may be set in a curve or in a spiral. The setting procedures should therefore be standardized for all fleets used in an experiment.

In the case of drift netting either the same or similar vessels should be used since differences in vessel size, shape and superstructure could result in important differences in the speed of drift of the nets and the tension on them in the water.

The direction of the setting in relation to currents and tide should be carefully recorded. Observations should be made on the distribution of fish caught both in the horizontal and vertical direction.

6. Hooks and Lines

The Working Group considered the following types of lines:

- hand-lines and pole-lines
- set lines
- drifting lines
- trolling lines.

The possibility of concentrating fish in the area fished by the hooks (e.g. by 'chumming' or by light) was noted, particularly in the case of hand-lines and pole-lines. This operational aspect makes the comparison of these two types of lines difficult. Discussion was therefore mainly restricted to baited long-lines, bottom set or drifting.

6.1 Long-Lines

When reporting the results of comparative trials of hooks and lines, catches should be related not only to the number of hooks but also to the length of line.

The influence on the catch of the type of bait, its shape and size were considered to be particularly important. The way of attaching the bait to the hook, which can affect the holding of the bait on the hook, can also affect the catch.

The hook type and size were also itemised as factors affecting the catch. The influence of the distance separating the snoods, particularly in relation to the density of fish on the fishing ground, was recognised. If it is desired to modify the distance between hooks it is recommended that this be most easily done by removing every 2nd or 3rd (etc) snood along the line.

The time of day when shooting and hauling the long-line takes place and the duration of the set are important factors which should be controlled in any experiment. The hauling phase could have an influence on the catch, especially for the last part of the line (when hauling the anchor), possibly resulting from the additional twist of the line produced by the increasing tension on the line as hauling proceeds. Catches from this part of the line might be discarded from the analysis or treated separately.

In order to reduce the influence of the movement or drifting of the vessel it is recommended that experiments be carried out using vessels of the same or similar size and type. Particular environmental factors which may influence catch rates include tide, current, light, moon phase and the nature of the sea bed.

6.2 Trolling lines

Possible experiments which might be conducted with trolling lines include studies of the influence of the line location on the outrigger or on the vessel stern, or assessments of the differences in catch rates obtained with lines of different types (shape and colour, for instance). Specific environmental factors which should be taken into account include position of sun in relation to the course of the vessel, state of sky, direction of wind. The size of the vessel may also influence catches.

7. Traps

The term traps covers various types of gear, such as pound-nets, pots and fyke nets.

It was agreed to restrict the discussion to baited pots which seemed to be more amenable to comparative fishing. As far as pound-nets and fyke nets are concerned it was the opinion of the group that the influence of their location (topographical factor) makes their comparison difficult and time-consuming. However, results may be obtained by fishing alternately with different nets in the same location.

With regard to baited pots, the distance between the pots was considered to be a factor of prime importance due to possible interaction between the pots and between the various baits used.

There are two ways of setting pots, either in line or separately. In line, attention should be paid to any changes of environment (depth, bottom nature, etc) which could affect the distribution and behaviour of the fish. Because of the influence of the current ('shadowing effect' on the zone of action of the bait) it is recommended that the lines of pots be set across the current. When the pots are set separately, the different types should be located at random. It was agreed that for comparing different baits the pots should be similar, and for comparing different types of pots, the bait should be the same.

As a specific factor in the pot fishery the effect of 'parasites' (organisms feeding on bait) should be taken into consideration.

If the fishing ground is not uniform with regard to environmental conditions, the experiments should be restricted in area and replicated.

The behaviour of the fish or shellfish caught in the pot should be noted, particularly when a possible escape can take place after a certain period of time.

8. Observational Studies

8.1 General Remarks

The Working Group devoted a good deal of its time to discussing the possibility of being able to draw correct conclusions from data gathered from "undesigned and uncontrolled" experiments. The appropriateness of the statistical techniques given in the Working Group's first report to the analysis of data from such studies, was also debated. The main conclusions were that while care must be exercised in collecting, analysing and interpreting data from observational studies, such studies are capable of providing a wealth of useful and meaningful information as evidenced by their increasing use in a wide range of research investigations (e.g. medical research).

8.2 Definitions, Principles and Procedures

An observational study is a survey of a process which seeks to explain the response of a variable quantity in terms of one or more explanatory variables. The process cannot be interfered with by the observer nor can any of the major explanatory variables be controlled or adjusted as in a designed experiment.

The first step in organising such an observational study is to define its objectives and the variables to be measured. At the outset it is advisable to prepare a written statement of the aims of the study. This ensures that all participants have a precise formulation of the hypothesis to be tested and also ensures that, as the study progresses, the objectives do not become altered and that different participants do not develop conflicting views as to what is to be done.

This statement should lead logically to the definition of the response variable to be measured and to how the measurements are to be taken and what comparisons are to be made. In comparative fishing studies the response variable needs to be carefully selected so that all appropriate information can be collected throughout the entire period of study. Possible response variables include (1) total catch in numbers per haul, (2) catch in weight per day's absence, (3) proportion of a given species which is of a certain size caught per haul, (4) value of all catches per vessel over a fishing season. There may be many more possibilities and the one or ones chosen must be stated.

In addition, all other variables to be measured during the study must be decided. Such "explanatory" variables may be truly causal variables (e.g. the number of weights on a purse seine) or simply classificatory variables (such as time of day). Particularly for the latter type, binary variables may be sufficient (e.g. 1 = day, 0 = night). In other situations a multi-pointed scale may be required. Also when a variable is difficult to measure it may be sufficient to record it as an ordered variable (e.g. none, few, average, many).

If the study is likely to continue over a long period and particularly if the effect on the response variable is not expected to be observed quickly it will be necessary to include a large number of explanatory variables in the study.

When the explanatory variables (x_1, x_2, ...) are measured on a continuous scale a comparison of the response variable (y) for the group using the standard gear can be made by first calculating the regressions of y on x_1, x_2, ... for each group and adjusting the mean responses to remove the effects of the regression. When some variables are classificatory the regression analysis may be extended to become an analysis of covariance.

An extreme form of classification is that known as matching in which pairs of units with identical values of <u>all</u> explanatory variables are chosen, one being used as a control and the other as an experimental unit. Clearly the use of sister ships in a study should be an advantage but non-ship factors may still differ appreciably even in such situations.

It will be useful to conduct a pilot study before the main exercise is undertaken. Such a pilot study will serve to identify possible explanatory variables on which information should be collected in order to check the adequacy of log books and to provide a comparison between the vessels of the control and experimental groups <u>before</u> the experimental gear is introduced.

9. Experimental Design

In its narrowest sense the term 'experimental design' means the arrangement or order of the experimental units (e.g. hauls) to be used and the method of assigning to these the different factors (e.g. gear variants) to be studied. The elimination or minimization of the effect of known, or suspected, but uncontrollable variation, the provision of unbiased estimates of the differences between the effects of the experimental factors with the desired degree of precision for the least amount of experimentation and a procedure for calculating the amount of uncontrolled variation are the main advantages of a good experimental design.

Essentially the aim is to group the experimental units into sets which are as alike as possible and assign a different experimental factor (or 'treatment') to each unit within a set. Thus, hauls made by the same gear at the same position are often more alike than hauls made at different positions. A comparison of catches made by different versions of a gear, in which each version has been used at a different site from every other, may reveal differences which merely reflect changes in fish population abundance or behaviour or differences in environmental conditions rather than in the catching performances of the gears involved. On the other hand, if comparisons are made within sets they are not subject to sources of variability which affect the sets differently.

As an illustration, consider an experiment to study the catching performance of two types of baited pot (A and B, say) of which 48 of each type are available. Twelve fleets each of 8 pots might be made up, all pots in a fleet being of the same type. If six different fishing sites were available it would clearly be sensible to use all sites and to arrange to have two fleets at each site, one of each pot type. In this situation the experimental units are fleets, 12 in all, and they are arranged in 6 sets of 2 units each. Results might look as follows, the figures being (hypothetical) catches in numbers.

Total Catch of Fleet

Site	Type A	Type B	Difference (B-A)
1	14	20	6
2	5	10	5
3	8	12	4
4	22	20	-2
5	9	13	4
6	17	25	8
		Mean =	4.2
		s.e. =	± 1.38

Note that the effect of differences between sites (due perhaps to real differences in levels of abundance) is eliminated. Changing the observations at any given site by the same amount leaves the results unchanged.

This technique may be extended to comparisons of more than two pot types in an obvious way giving rise to what are known as 'randomized blocks', the blocks being sites and the relative position of the fleets of different pot types at each site being determined by some randomization procedure. A possible plan for comparing 6 pot types might look as follows.

Position (East to West)

Site	1	2	3	4	5	6
1	A	B	D	F	C	E
2	D	C	A	B	F	E
3	B	F	A	C	E	D
4	C	A	F	E	D	B
5	E	C	B	A	F	D
6	D	B	E	A	C	F

At 4 of the sites (2, 3, 5, 6) pot type D appears at an outside position while types A, C and F occupy such a position at only 1 site each (1, 4 and 6 respectively). If it were considered that the position of a fleet within a set might produce a systematic effect on catches, a second grouping into sets corresponding to the six different positions could be introduced. This would give rise to a Latin Square design in which each pot type appears not only once at each site but also once at each position. Such an arrangement is shown below.

	Position (East to West)					
Site	1	2	3	4	5	6
1	A	B	C	D	E	F
2	B	C	F	A	D	E
3	C	F	B	E	A	D
4	D	A	E	B	F	C
5	E	D	A	F	C	B
6	F	E	D	C	B	A

It may often happen that we wish to study the effect of several different types of factor. Thus we may wish to study pots having three different sizes of eye-opening S_1, S_2 and S_3 (say) and also to study whether the fitting of a particular type of non-escape device affects catches. These factors may be studied within the framework of a single 'factorial experiment' in which the treatments applied are combinations of the various 'levels' of the different factors. In this hypothetical example there are two factors, size of opening and non-escape device. The former has three levels corresponding to the three sizes of eye, the latter has two levels namely non-escape device present (P) or absent (A). There are 2 x 3 = 6 possible combinations of factor levels,

S_1A S_2A S_3A S_1P S_2P S_3P

An experiment of this type is referred to as a 2 x 3 factorial experiment. The treatments may be applied in a randomized block or Latin Square design.

If the effect on catch of fitting a non-escape device is not the same at all sizes of eye-opening we say that there is an 'interaction' between the fitting of the device and size of eye-opening. The advantage of using a factorial experiment is that it allows a test for the existence of interaction to be made and, if it exists, allows its nature to be explored. The possible existence of interaction may be revealed by constructing a two-way table of total catches from each of the six sets of fleets of the six pot types. This might look as follows.

Device	S_1	S_2	S_3
Absent	96	150	61
Present	84	302	102
Difference	-12	152	41

The above table suggests that the fitting of a non-escape device is advantageous at the intermediate eye size and possibly also at the largest size but not at the smallest. Standard errors for each of the above differences may be calculated by established procedures.

While the factorial principle may be extended to the study of more than two factors, each at several levels, the employment of experiments involving a large number of factors is not generally recommended in comparative fishing.

10. Statistical Analysis

Statistical analysis of the results from some simple experiments were given in the previous report (Coop. Res. Rep. No. 38, section 8). In this report some basic requirements for the validity of application of the analysis of variance are presented and the method of carrying out an analysis of variance of a two-factor experiment with interaction described. Familiarity with elementary statistical techniques is assumed.

10.1 Preliminary Considerations

The mathematical descriptions (models) used in the analysis of variance are linear models. That is, the variable being analysed, catch say, is considered to be capable of being represented in a mathematical fashion in the following way.

Catch = (a function of specific systematic factors)

plus

(a random element representing the combined effect of non-systematic, unmeasured and uncontrolled factors)

i.e. $y = f(\mu, \alpha, \beta, \gamma, \ldots) + \varepsilon$

$= (\mu + \alpha + \beta + \gamma + \ldots) + \varepsilon$

where μ is a constant representing the overall mean of the catches and α, β, γ are constants referring to the effects of the different systematic factors.

For example, suppose catches taken by two versions of the same bottom trawl fished from the same vessel are to be analysed. Suppose further that the gears were fished sometimes from the port side of the vessel and sometimes from the starboard side and that hauls were made during both daylight and darkness. Assuming it were possible to carry out 8 hauls in each period of 24 hours a possible experimental design might be one in which 4 hauls were made during the hours of daylight and 4 hauls during darkness hours. Thus the plan for the first three periods might look like

Period	Daylight				Darkness			
1	AP	BP	BS	AS	BP	AP	AS	BS
2	BS	AP	BP	AS	AS	BS	AP	BP
3	AS	BS	AP	BP	BS	AP	BP	AS

where A and B refer to the two versions of the bottom trawl being used and P and S refer to port side and starboard side respectively. Because the effect on catch produced by fishing from different sides of the vessel and by changes in light intensity may be considered to be systematic, this has been recognised in the experimental design by arranging that both gears are fished from both the port side and the starboard side in both daylight and darkness. The mathematical model for the analysis of the results of such an experiment would be

$$y = \mu + \gamma_i + \sigma_j + \lambda_k + \varepsilon$$

where y is the catch (or possibly some function of the catch such as its logarithm), μ represents the overall mean level, γ_i $(i = 1,2)$ represents the effect of gear i, σ_j $(j = 1,2)$ the effect of side j and λ_k $(k = 1,2)$ the effect of daylight or darkness. The terms to be included in the model are determined by the nature of the experiment and its design. Terms which cannot be estimated from the data should not appear in the model. The random term, ε, represents the combined effect of all those factors, known and unknown, which have not been controlled and whose effect is known, or has been arranged by randomisation, to be unsystematic. There is a not uncommon misconception that the validity of such models demand that there be a Normal distribution of the species in space so that when the species are distributed in a highly aggregated fashion, as many are, this assumption fails and brings down the entire model with it.

The spatial distribution of a species may be classed conveniently into one or other of three groups, namely, uniform, random and aggregated. A uniform distribution occurs when the fish are regularly spaced in a symmetrical square lattice in two-dimensions or cubic lattice in three-dimensions. The density per unit area or per unit volume would, in such a situation, be constant at all points. Probably no fish species is distributed precisely in this way although some may approximate to this.

When fish are randomly distributed in space they may be thought of as exerting no force of repulsion or of attraction on each other. The number of fish in any volume is neither influenced by nor itself influences the number of fish in any other volume. In such a situation the number of fish in the different unit volumes into which the total habitat may be divided will have a frequency of occurrence exactly predictable by a statistical probability distribution known as the Poisson distribution. This type of distribution is unaffected by the size of volume chosen as unit although, of course, the average number of animals per unit volume will depend on the size of the unit volume. When the mean number of fish per unit volume is large the frequency distribution of the number of individuals in different volumes will be closely approximated by a Normal distribution. This is a mathematically verifiable property of the Poisson distribution. Further, catches from such a population will have a Poisson, or for large numbers caught, a Normal distribution.

When fish are aggregated into shoals then the occurrence of individuals in any specific unit volume will increase the probability that individuals occur in neighbouring volumes. The entire habitat will be characterised by having a number of volumes with no individuals present and others with a large number present. The frequency of occurrence of different numbers of individuals per unit volume will in this case not be given by the Poisson distribution law but will follow some other statistical distribution such as the Negative Binomial distribution. Catches by fishing at random on such a population will also follow a Negative Binomial distribution, but, if fishing is "aimed" at schools this will no longer be true as zero catches will either not occur at all or be under-represented.

It is theoretically possible to find a mathematical transformation of the random element which will have a Normal distribution and, in practice, some convenient transformation can usually be derived which will be, to a sufficient degree of approximation, Normally distributed. The point of this is to allow standard statistical tests of significance to be made where otherwise they would be invalid. The Normality requirement is a prerequisite of the figures actually used in the computations not of the basic original data. If tests of significance and confidence limits are not required it is not essential to demand Normality of the stochastic term or any transformation of it although for other, analytical, reasons the statistician may prefer to work with transformed data.

Although the simplest response function is a linear one it is not universal and if non-linear responses are known or thought to exist they should be so specified. The form of the function (f) is a matter of judgment and will be based on the experimenter's knowledge and experience of the process under study.

A wrongly specified model will lead to incorrect conclusions. Non-linear response functions may be handled by ordinary least-squares procedures although the existence of non-linearity may result in mathematical complexity. This may be avoided by seeking a linearizing transformation of the response function suitable for analysis. As with the normalizing transformation this linearizing transformation is introduced solely for mathematical simplicity. If enough information is available, both transformations may be deduced from the data.

10.2 Analysis of a Two-Factor Experiment with Interaction

The computations required to perform an analysis of variance of a two-factor experiment, which includes a test for the presence of an interaction between factors, are given in this section.

Suppose the experiment relates to long-lining and the two factors are hook type and distance between snoods. If two hook types are to be tested at three different spacings between snoods, a minimum of 2 x 3 = 6 lines will be required, three composed of hook type A and three of hook type B, one of each group of three being at each of the three snood spacings to be studied. However, one set of catch data for each line would not yield an estimate of the magnitude of the random variability which is necessary for carrying out tests of significance. Thus it is necessary to replicate the experiment. If this were done a minimum of five times a sufficiently accurate estimate of the variance of the random component should be obtained. More replication would, of course, improve this estimate. With such a design it is possible to include terms representing the presence of interaction between hook type and snood spacing.

A suitable model for this analysis would be

$$y_{ijk} = \mu + \alpha_i + \beta_j + I_{ij} + \rho_k + \varepsilon_{ijk}$$

where

- y_{ijk} is the catch per hook from the kth replication of hook type i at spacing j
- μ is the overall mean catch per hook
- α_i is the effect of hook type i (i = 1,2, ..., a)
- β_j is the effect of spacing j (j = 1,2, ..., b)
- I_{ij} is the effect of interaction between hook type i and spacing j
- ρ_k is the effect of replication k (k = 1,2, ..., r)

The terms α, β, I and ρ are not all of the same kind. The hook types and spacings studied are probably the only ones the experimenter is interested in, i.e. they may constitute his total population of hook types and spacings and the results of the experiment are to be applied only to these. The terms α, β and I are, therefore, referred to as _fixed_ effects. The replication terms, ρ on the other hand may be regarded as random quantities, being a sample from a larger population of replications that might have been carried out over a longer period of time or over a much wider area. Without any loss of generality we may then assume

$$\Sigma\alpha_i = \Sigma\beta_j = \sum_i I_{ij} = \sum_j I_{ij} = 0$$

We may also assume

$$\rho_k \sim N(0,\sigma_\rho^2), \quad \varepsilon_{ijk} \sim N(0, \sigma^2)$$

The analysis of variance table for this model is given below.

Source of variation	Degrees of freedom (df)	Sum of squares (SS)	Mean Square (MS)	F-test
Hook type (A)	a-1	S_A	$s_A = S_A/(a-1)$	s_A/s_E
Distance (B)	b-1	S_B	$s_B = S_B/(b-1)$	s_B/s_E
Replication (set)	r-1	S_R	$s_R = S_R/(r-1)$	s_R/s_E
Interaction (AB)	(a-1)(b-1)	S_{AB}	$s_{AB} = S_{AB}/(a-1)(b-1)$	s_{AB}/s_E
Error	(r-1)(ab-1)	S_E	$s_E = S_E(r-1)(ab-1)$	
Total	abr-1	S_T		

$$S_T = \bar{x}_{\cdot\cdot\cdot} = \sum_{ijk}\sum\sum x_{ijk}/abr$$

$$S_A = \bar{x}_{i\cdot\cdot} = \sum_{jk}\sum x_{ijk}/br$$

$$S_B = \bar{x}_{\cdot j\cdot} = \sum_{ik}\sum x_{ijk}/ar$$

$$S_{AB} = \bar{x}_{\cdot\cdot k} = \sum_{ij}\sum x_{ijk}/ab$$

$$S_E = \bar{x}_{ij\cdot} = \sum_k x_{ijk}/r$$

The consequences of non-normality of the ε_{ijk}'s on the significance level of the F-test are not too serious. Only very skewed distributions would have a marked effect on the significance level. The best way to correct for lack of normality is to carry out a normalizing transformation. The most common one is the logarithmic transformation. Right-skewed frequency distributions are often made more symmetrical by transformation to logarithmic scale. Other transformations which are frequently used are the square-root and the arcsine transformation. A fortunate fact about transformations is that very often several departures from the assumptions required by analysis of variance are simultaneously cured by the same transformation. Thus frequently, by making the data homogeneous, we also make them approximately Normal. When a transformation is applied, tests of significance are performed on the transformed data. If none of the transformations can make the data meet the assumptions of analysis of variance, non-parametric methods may be applied. However, we should point out that in cases where the assumptions hold even approximately, the analysis of variance is generally the more efficient statistical test.

11. Amount of Experimentation

A question frequently posed in the planning stage of an experiment is 'how many hauls (sets, etc) will be required to provide a sufficiently accurate answer to the questions being studied?'. This was dealt with in the Working Group's first report. A similar question is 'how is it possible during the experiment to decide whether further work need be done or not?'. The motivation for the second question is obvious. Once an effect has been clearly established the unnecessary expenditure of further resources is wasteful.

Although an important question it is not an easy one to answer. There seems to be very few instances when too much unnecessary experimentation has been done and the simplest answer would seem to be to continue testing for as long as possible. This, of course, is not an entirely satisfactory answer. A full examination of the problem would, however, require a knowledge of decision theory which is far beyond the scope of the present report. Instead some intuitively simple procedures may be suggested.

If the magnitude of the sampling variation to which the observations are subject is not known it will either have to be guessed (always a dangerous procedure) or estimated from a preliminary set of results involving say r_1 observations on each experimental treatment. The standard error of the difference between two treatment means each based on n observations is $\sqrt{2}s/\sqrt{n}$. If it is desired to estimate the mean difference with a standard error of p %, the necessary value of n may be obtained by solving the equation $\sqrt{2}s/\sqrt{n} = p$ or $n = 2s^2/p^2$ where s is the standard deviation estimated from the preliminary experiment and is expressed as a percentage of the mean level. The number of observations still required is then $n - r_1$.

Alternatively the values of the mean difference at each successive stage of the experiment may be plotted on a graph on which the curves $\pm t\sqrt{2}s/\sqrt{n}$ have been drawn, t being an appropriate value of Student's t corresponding to a probability level of 0.05. If the successive mean differences based on all available data consistently fall outside the region lying between the curves $\pm t\sqrt{2}s/\sqrt{n}$ a decision to terminate the experiment may be taken.

These procedures are, however, simply rough (though useful) guidelines.

12. Selected Reading List on Experimental Design

The following is a short list of text books devoted to principles of experimental design and analysis. It is by no means exhaustive nor is it intended to be but it does contain some of the accepted authoritative accounts of the subject.

Cochran, W.G. and Cox, G.M. 1957 "Experimental Designs". 2nd ed. New York: Wiley.

Cox, D.R. 1958 "Planning of Experiments". New York: Wiley.

Davies, O.L. 1954 "The Design and Analysis of Industrial Experiments". New York: Hafner.

Fisher, R.A. 1935 (rev. ed. 1960). "The Design of Experiments". London: Oliver and Boyd.

Quenouille, M.H. 1952 "The Design and Analysis of Experiment". London: Griffin.