DOCUMENT RESUME

ED 070 292 EM 010 585

AUTHOR Larsson, Bernt
TITLE Some Problems in Advising on Research Methodology.
INSTITUTION School of Education, Malmo (Sweden). Dept. of Educational and Psychological Research.
PUB DATE Sep 72
NOTE 22p.; didakometry number 36; Paper presented at the Conference of the Society for Multivariate Experimental Psychology (Nijmegen, September 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Behavioral Science Research; Measurement Techniques; *Methods Research; Operations Research; Research; *Research Design; *Research Methodology; Research Skills; *Statistical Analysis; Statistical Studies

ABSTRACT

By and large behavioral scientists, particularly in education and psychology lack the sophisticated mathematical understanding necessary for adequate statistical analysis, especially multivariate analysis. And incompetent use of measurement theory, psychometry, and statistics cannot but lead to bad research. The remedy for this is better mathematical training of the behavioral scientist and the creation of special well educated advisors on research methodology. The advisor's role is not only one of dealing with attitudes to research methodology and the consequences of inadequate methods, but also to further the methodological education of the client. Preventative measures must be taken to keep researchers in education and psychology from slowly dividing themselves into two non-communicating camps depending on the knowledge of mathematics and statistics. The problems are solvable only if more researchers are willing to change their attitudes to methodology. (MC)

special-topic
bulletin from

DEPARTMENT OF
EDUCATIONAL AND
PSYCHOLOGICAL RESEARCH

SCHOOL OF EDUCATION
MALMÖ, SWEDEN

didakometry

Larsson, B.:

SOME PROBLEMS IN ADVISING ON RESEARCH METHODOLOGY

No. 36 September 1972

ED 070292

SOME PROBLEMS IN ADVISING ON RESEARCH METHODOLOGY

Bernt Larsson

This paper discusses the poor conditions of research methodology in the behavioural sciences due to insufficient consumer knowledge and lack of well-educated advisers in this area. It starts with a definition of research methodology, examplifies the kind of knowledge needed by the adviser and then goes on to the main subject, that is, different aspects of the adviser's tasks. Among other things, it deals with attitudes to research methodology, the consequences of inadequate methods, the methodological education of the client and the adviser's own research. Most examples are taken from multivariate analysis.

Paper read at the conference of Society for Multivariate Experimental Psychology in Nijmegen in September 1972.

INTRODUCTION

Research methodology can be defined very broadly, to include such things as the construction of proper concepts, theories and hypotheses and the usefulness of and similarities between different models. In this paper a more limited view is taken and research methodology refers to generic methods of collecting and processing data. They can usually be categorized as follows:

1. Design
2. Sampling technique
3. Instrument for data collecting
4. Statistical analysis
5. Computer use

The first point concerns the architecture of an examination, e.g. the possibilities of performing an experiment and advantages and drawbacks of different designs. We have partly the general type of discussion to be found in Bracht & Glass (1968) and Campbell & Stanley (1963), partly the more technical considerations as they appear in Cohen (1969) and Winer (1962).

The second point seems to be of different importance to different researchers. I do not intend to argue on this issue, but the fact remains that not many educational reports can boast of random samples, not to mention simple random samples, which most statistical tests assume.

The third point comprises measurement theory and psychometry, e.g. problems of construction and evaluation in connection with questionnaires, tests, ratings and interviews. As far as application is concerned, this is an underdeveloped area and I have argued elsewhere (Larsson, 1971) that several differences of results may very well be the effect of bad measurement theory. However, I do not wish to discuss it here, but instead concentrate on point 4.

This point is intimately connected with the first one and contains a large number of problems which are central to mathematical statistics. This paper will for the most part illustrate its discussions by examples from this area, particularly from multivariate analysis.

Point 5 may concern the development of programs, problems of choice between different facilities of both software and hardware, and the computer use for simulations and documentation.

Research methodology is more or less based on mathematics. The degree of mathematical involvement is dependent on the particular field of interest. Well-defined areas, e.g. most areas within physics, have usually no need for any fancy statistical methods to discover their experimental effects. While this need not necessarily be true, the opposite will certainly hold. One cannot but expect small effects in experiments which are founded on insufficient measurement theory and which are rather uncontrolled, as is often the case with research in education and psychology. No statistical method can eliminate these bad conditions, but while the behavioural scientist tries to make them better, we may hope, he is (or should be) a big consumer of statistics, which helps him to discover the effects, if any.

And this is where troubles start. Unless the behavioural scientist is very specialized, he needs several kinds of statistical ammunition to hit his targets. This requires a rather profound knowledge of both applied and mathematical statistics, a knowledge which he seldom posesses. (I could add knowledge of measurement theory, of psychometry, and so on, but let us mainly limit ourselves to statistics.) The fact is also valid for most postgraduate students. Some readers may smile sceptically and say that his students have sufficient knowledge. Well, I do not know what super-students some readers may have, I can only say that I have not met them. In fact, I do not think it is realistic to let all students have this knowledge. After all, they intend to do behavioural research, not to become statisticians.

It is perhaps convenient to say something about the kind of knowledge I am thinking of before proceeding with the discussion. If the behavioural scientist wants to guide himself properly, as concerns the choice of statistical methods, he must know how and why they are constructed, e.g. to understand their limitations. I doubt whether books like Cooley & Lohnes (1971), Hays (1963) and Winer (1962) are sufficient for this purpose. Parts of the content of e.g. Dempster (1969), Miller (1966), Morrison (1967) and Rao (1965) must also be understandable. As examples of somewhat more specialized books on the same mathematical level, we may mention Cox & Miller (1965) and De Groot (1970). To this we may add books dealing with methods which by tradition belong to the behavioural sciences: Bock & Jones (1968), Horst (1965), Lawley & Maxwell (1963), Lazarsfeld & Henry (1968), Lord & Novick (1968) and

Rozeboom (1966). Good working knowledge of elementary calculus and matrix algebra is usually indispensible in this connection. More advanced knowledge can also be useful, when reading e.g. Dempster (1969).

What is the actual knowledge of the postgraduate student and the researcher of the behavioural sciences? I have limited experience from Scandinavian educational psychology only, but I do not think that things are much different elsewhere. Though there is a great variation among people, the knowledge is, on the average, certainly not on the level exemplified above. Most students and researchers have a dim apprehension about calculus and it is very few who know anything essential about matrix algebra. Unfortunately not too seldom, you meet behavioural scientists whose knowledge of mathematics beyond the square root is almost infinitesimal. It is obvious that statistical understanding will not develop easily under such conditions, and this is particularly valid for multivariate analysis. So there is in general a wide gap between the knowledge which the scientist really has and that which he should have, if he wants to be his own statistical adviser.

ADVISING ON RESEARCH METHODOLOGY

Since I have already said that it would be unrealistic to assume the latter kind of knowledge in every researcher, the reader may have anticipated my view that we need special advisers on research methodology within the behavioural sciences. But this idea will fail unless the adviser can communicate adequately with his client, and whether this is possible or not depends on several factors. I shall discuss some of them, mainly the kind of work which the adviser must do, the knowledge and attitudes of the client needed for efficient communication with the adviser, and some of the consequences of certain assumptions not being fulfilled. The discussion will start from the adviser's different tasks, and for each task I will incorporate my opinions about the client where convenient.

If the cooperation between adviser and client is to give good results, this implies considerable work by the adviser. He should

1. propose and describe methods to the client,
2. interpret output from computers,

3. be co-author of reports containing problems with strong methodological anchoring,
4. communicate with programmers and other advisers of the client,
5. keep himself informed on literature of research methodology,
6. take part in conferences and other contacts of importance for research methodology,
7. do his own research.

Some aspects on the choice of methods

The first task is, of course, important and perhaps the first one which comes to mind when speaking about an adviser for this area. The second task is quite similar to the first one so we can just as well discuss both points simultaneously. I have separated the two points, so that nobody will believe that the client is away for good, once he has been sent on to the programmer. Even if we pretend that the adviser has all the knowledge, which is required for a good adviser, there are other things that must function. The adviser must have time and the right attitude, the client must have sufficient knowledge and be motivated. Also, the client must know his own problems so that he can specify what he wants from the adviser. Of course, this is reciprocal in that the adviser must also understand the client's problem. It must further be underlined that the client must seek advice in time and not wait until data are collected, which unfortunately many clients do.

If the adviser has not time enough to explain methods and/or if the client is not able to understand the 'explanation', two things can occur: either the adviser realizes the facts and proposes a simpler method, which the client can comprehend, or the client leaves the adviser, no wiser than when he came (which may mean that the client uses the proposed method without understanding it). The latter situation can occur if the adviser makes wrong assumptions concerning his clients' knowledge, but it can also occur when the client visits the adviser without having any 'honest' intents. He may have a firm resistance to learning a new method and expects to find ready-made, easily understandable methods, which solve all his problems without compromises or any substantial effort from him. In fact, certain clients may use an adviser only to verify their own prior opinions and work and do not listen to anything which will question these. Such attitudes do definitely not produce good research and persons having them are more or less impossible to advise.

Forgetting such cases which I hope to be exceptional, one may ask 'What are the consequences of proposing methods which the client can easily understand or already knows?' It is extremely difficult to state anything generic about consequences. Perhaps one dares to say that only advising the client about methods from his own repertoire - though creatively - will too often produce research of insufficient quality. These advices typically keep the client away from applications of multivariate methods, as these are the methods about which the behavioural scientists usually have little knowledge. Perhaps somebody has the objection that this is not true with factor analysis, which is widely known among researchers of the behavioural sciences. Personally, I do not believe in such a statement. I agree that factor analysis is widely used, but this is not a criterion for an understanding of the method. In fact, the question of consequences is more complex than that above, because you can add 'What are the consequences of proposing convenient methods which the client is not able to use properly?'

Let us take an example. A client's problem is such that a good advice would be to let him perform a pxq factorial discriminant analysis, for which we suppose a convenient program is accessible. The client has good knowledge of elementary statistics up to 'the t-test level' and he knows something about ANOVA but nothing at all about discriminant analysis. Both the adviser and the client are motivated to do their best, but the adviser has not much time to devote to each client. What is the 'best' advice in such a situation? For the sake of simplification, let us assume that the choice is between four alternatives (we have r dependent variables):

1. Perform $r\,(p^2 + q^2 - p - q)/2$ t-tests.
2. Perform r pxq factorial ANOVAs.
3. Perform one pxqxr factorial ANOVA.
4. Perform one pxq factorial discriminant analysis.

The first alternative implies that the client himself can handle the output, although it can be a tremendous lot of output. Besides the amount of output three other features are clearly negative: no interaction tests, the mass significance problem and the correlations between the dependent variables. As far as I know, it is very difficult to state how they in detail contribute to wrong decisions, only that they contribute to wrong decisions. The second alternative produces interaction tests, mitigates the mass significance problem, but not the third one. The third alternative is equivalent to the fourth, provided that the covariance matrices for the dependent variables are all of a certain form (from cell to cell in the pxq factorial design).

On the other hand, the client is uncertain about some properties about a pxq factorial ANOVA, e.g. how to perform contrast analyses, and still more uncertain about the more complex ANOVA with repeated measures, not to mention discriminant analysis. There are many technical details which must be considered for this situation, but it is evident that the adviser must compromise, which means that alternative 2 or 3 is probably chosen. To put it bluntly: what is the most serious error, to suggest a sub-optimal method which the client can use, or to suggest a good method which the client does not understand? Indeed, I do not know; and there is probably no general answer to this question.

The example above could be multiplied, but I shall only take one further example. A client has a problem which fits factor analysis of variables and he has a determined opinion about what kind of a structure he expects. The client has a very vague apprehension of factor analysis and the adviser has two different programs to offer, again for the sake of simplification:

1. The principal axis method with varimax rotation, see e.g. Dixon (1970).
2. A general program for analysis of covariance structures, as this theory is described in Jöreskog (1970).

Although the mathematical development of both methods are difficult and because of that are beyond the usual client's capacity to understand, the second method is far more complex. But the first methods is not very acceptable in this case, since the client has a special hypothesis about structure, and this cannot be tested by the first method. If he uses the first method he could very well cheat himself to believe that his hypothesis is correct, while the other method had given him a significant result. (The rotated factor matrix seems to be 'correct', but the second method reveals that the hypothesized structure is insufficient to properly reproduce the covariance matrix.) Regarding that the client has, under all circumstances, bad knowledge of factor analysis, I think that the adviser can here just as well choose the more appropriate method. The only real alternative is to frankly assert that the client himself is not capable of doing any factor analysis at all, an assertion which perhaps will not be accepted.

I assume that the reader finds the sketched situation to be unsatisfying. If so, I am satisfied, because that was what I intended to show. It is my opinion that lack of expertness on research methodology in education and psychology makes much research unnecessarily bad. How many results are by this reason not discovered and how many facts are just artefacts? For instance, how many differences of results in the same area mirror methodological differences only? Who has ever measured the consequences of a bad methodology and, even worse, how many decisions can have been made on false grounds, e.g. due to significances based on negligible effects? Such decisions can have great economic consequences, when concerning many people. I am frightened when thinking of it, because my subjective probability of being wrong in this case is too far from 1.0. I can only hope that somebody will convince me that there is nothing to worry about.

Some aspects of the instruction on research methodology

Meanwhile we can return to the earthy troubles, where we left them before the last paragraph. One of the facts is that the few available advisers on research methodology have too little time for each client and an easily suggested remedy is to recommend convenient literature which describes his problems. It is a nice thought but it will not work: many aspects of his problems are so technical that it will be impossible to find literature which matches the client's knowledge. For several problems I think that the literature can be produced, but it is also evident that some problems are so technical and/or some clients' knowledge is so deficient that it is almost impossible to write a book of normal size without assuming prerequisites of the reader. For instance, I can mention three books on multivariate analysis which were published by educational psychologists in 1971: Cooley & Lohnes (1971), Tatsuoka (1971) and Van de Geer (1971). Although these books vary concerning mathematics, none of them is accessible to most behavioural scientists, as far as I can see. Also, books dealing with measurement theory and psychometry have with years become more mathematically involved.

Compare e.g. Torgerson (1958) with Bock & Jones (1968) and Gulliksen (1950) with Lord & Novick (1968).

One part of the formal education of research methodology, at least on the postgraduate level, should be courses in mathematics for those who do not have the necessary prerequisites here. I am aware of some institutes which have taken the consequences of this and offer courses in mathematics up to the level of elementary matrix algebra. However, too many institutes have no such courses, and I cannot tell whether this depends on economics or other factors. With regard to the very shifting mathematical background students in the behavioural sciences have, these courses should be compulsory for those who need them. Otherwise much teaching on research methodology will be meaningless, and this goes for lectures as well as for individual advice. It is quite naive to think that anybody can learn anything, no matter how mathematically involved, without spending a good deal of time on first learning mathematics. It ought to be self-evident, but my experience of many people's reasoning says it is not so.

On the other hand I am no friend of too much formal teaching on research methodology for students, at least not in the almost ridiculous way it is sometimes performed in Sweden. Often the students, and particularly the postgraduate students, are fed with an overwhelming lot of methodological facts within a few semesters. Most of the facts meet the student before he has done any substantial research and are then supposed to last for the rest of his life. Then the only thing that really lasts for many students is a feeling of distaste for everything that smells statistics. I can understand them: it requires an interest in methodology well above average to survive in such a climate.

The only methodological course which I think is necessary at the postgraduate level, except a course in mathematics, is one which gives a broad overview and not details of different methods, etc. It should give syntheses, point to differences and similarities of methods, discuss strategies of choosing methods. The more profound knowledge should be left to the individual adviser to give to the student (researcher), when he is motivated by his own research work to go into details with methods suggested. Of course, this means that the adviser must spend rather a lot of time with each client, and this would be more expensive than the arrangement of formal courses. However, I am quite convinced that the individual way produces better applications of research methodology and does it more rapidly.

It is not at all surprising that the researcher now and then has use for courses on special subjects, including methodology. Particularly if several members of one or more institutes are interested in the same subject, there are good reasons to start a course. I have a feeling that this type of courses is not very common. Maybe there is a lack of motivation to further educate oneself, because few researchers have permanent appointments? (At least so in Sweden, where most people doing research have one year appointments.) I think that formal courses are much more advantageous to the researcher than to the student, because the former's work experience makes it easier for him to assimilate the new knowledge. I also think that such courses can be efficient if you want to introduce a new method, etc. at your institute, but should be followed up with individual advising. The adviser on research methodology has an important role in presenting statistical innovations to other members of his institute.

The time lag between the development of statistical ideas and methods and its applications in education and psychology is often surprisingly great. A typical example is the concept of power. Although the idea of power and statistical power analysis is over forty years old, it seems to have been noticed by few behavioural researchers. Maybe tables such as those by Cohen (1969) will give rise to a more extensive use of this important concept. And even if an idea has been introduced to some behavioural scientists, it can take time before it is firmly settled. As a probably extreme case I can tell that about three years ago I began to give the advice that Hays' ω^2 can be a good complement to the usual ANOVA output. It still happens that people ask me what this index implies. I presume that if I had had a one hour's lecture on this concept for the members of our institute this would have raised the efficiency of the information spread considerably.

One reason for the time lag with the multivariate methods is that they had to wait for the computer revolution. One prominent instance of this is canonical correlation analysis, which, in spite of its introduction by Hotelling (1935) to educational psychologists, had very few applications in the behavioural sciences before 1960. The lack of computers, which prevented many researchers from using multivariate methods, while they were in their most formative years, has produced a generation which still does not use these methods very extensively, because they have got other methodological prejudices, which are not easily moved. Another essential reason for time lag is the mathematical barrier. Multivariate methods are, almost by definition, not easy to understand. There are so many relationships which are difficult to see through, e.g. because we cannot visualize more than three dimensions and that even such a simple concept as variance has no unequivocal extension to more dimensions. It must also be admitted that several theoretical problems still wait for their solutions.

The theoretical problems of multivariate analysis are of two kinds, which are interdependent, viz. inference problems and data-analytic problems. Mathematical statisticians involved in this area are mostly concerned with inference problems. The direction of this research has been criticized (see Aitkin, 1971, p. 233) for dealing almost entirely with methods which discover effects and almost never with methods which examine the nature of the effects. Problems of data analysis are more or less incidentally worked upon and comprise the difficulties of presenting data to human receivers in intelligible ways. Savage (1970) believes that data analysis will one day have its own foundation.

The problems of data analysis are also the problems of the adviser, when he wants to explain methods or ideas to his clients. He must do this in a non-technical way, while nevertheless retaining the vital points of that which he is explaining. How this should be done - when it can be done - must be a judgment from situation to situation. However, some general statements may be made, of which I want to mention two. Try to translate technical or otherwise new concepts into concepts which the client already understands. Try to emphasize similarities between methods rather than differences, at least in the first place. Both recommendations rest on the assumption that you should take advantage of the knowledge which the client already has achieved.

Factor loadings can be regarded as correlations in many applications and eigenvalues can be regarded as variances here. You can inform the client that canonical correlations and, therefore, multiple correlations are special types of product-moment correlations. Likewise, ω^2 is an ordinary squared multiple correlation between the dependent variable and the dummy variables defining the effect. And for a 2 x C contingency table Cramér's statistic (see Hays, 1963, p. 606) is identical to ω and thereby the multiple correlation, while this statistic is not a canonical correlation in the general case. If convenient, prefer Bonferroni t-tests to other tests for contrast analyses (see e.g. Miller, 1966, pp. 67-70). When it comes to multivariate tests about means, Roy's union-intersection principle is illuminating (see Morrison, 1967, pp. 118-119 and several other pages). It has the further advantage of giving a simple view on how discriminant analysis is connected with MANOVA. E.g. if you want to explain the multivariate t-test for two independent samples, the explanation could run like this: 'If these dependent variables are weighted together, you get a new dependent variable on which you can perform a t-test. Let us assume that you choose the weights in such a way that no other dependent variable, produced by weights, has a greater t-test value. If, in spite of this choice, it does not give a significant t-test, then we can say that the two groups are not different, simultaneously for all original dependent variables.' On the other hand, multivariate concepts like the generalized variance are not easily explained. This variance is defined as the determinant of a covariance matrix and is the product of its eigenvalues. Neither the variance equivalence of the eigenvalue nor the connection of the determinant to certain volumes make this concept more understandable to me.

The similarities of methods are not so easily pointed out. That multiple linear regression analysis is a special case of canonical correlation analysis is perhaps one of the easiest similarities to mention. Perhaps also that discriminant analysis is a sort of specialization out of MANOVA. Somewhat harder to show for a client is the connection between multiple linear regression analysis and ANOVA and ANCOVA. Papers like that of Cohen (1968) can be useful here: the idea of dummy variable coding is essential. If we extend it to multivariate analysis this means that canonical correlation analysis is the core of many methods. For latent structure analysis, the axiom of local independence is a common property to all its methods and may be used to emphasize similarities. It is also important to emphasize similarities and differences between different designs, e.g. explaining advantages and drawbacks of the following four alternatives, where experiment group/control group always is one factor and you have measurements both before (x) and after

(y) the experiment: 1) x/y is a second factor with repeated measures; 2) x is categorized and becomes a second factor; 3) only one factor with dependent variable y-x; 4) only one factor with dependent variable y and covariate x. I am strongly convinced that very much remains to be done on similarities and differences between methods. This involves mathematical analyses as well as simulations and experiences from real data. I also believe in a better 'marketing' of such studies, so that they can be at least partly accessible to other than experts on a certain subject.

The adviser's third, fourth, fifth and sixth task

The third point of the adviser's working list (p. 5 above) is important, for two reasons. The first one is that it gives the adviser some possibilities of incorporating methodological studies here and there within the client's projects. While this is not often convenient, the opportunities should certainly not be neglected. The second reason is that some research reports contain rather advanced methods, which can be too difficult for the client to describe and use properly alone. This is in line with what has been said before and to be able to compromise between good methodological advice and faulty consumer knowledge, the adviser should write certain parts of the reports when necessary. It is then important that the adviser know his readers. Although a complex matter cannot be simplified too far without losing its essential properties, much more than what is done today can be done to facilitate understanding. To hide oneself behind a mathematical iron curtain, when writing for a broad audience, is either a sign of ignorance or boastfulness. Of course, this takes time and implies that one adviser cannot serve many clients, but on the other hand it will raise the methodological quality of a number of reports.

In connection with the fourth point, I would like to stress the important role of the programmer. Every sufficiently large research institute ought to have its own programmer. It is my definite opinion that he should have good knowledge from the fields of mathematics, numerical analysis and mathematical statistics and some experience with problems of education and psychology rather than the opposite (be a behavioural scientist with informal experience in programming). This is because he must be able to develop new programs along with his other services to the clients. I consider this as highly important: my experience with some available standard programs is not good and in other instances there are no such programs at hand. (I am conscious of the fact that such experiences can be very different from place to place due to different software facilities.) Again, it takes time and cuts down the number of clients which the programmer can serve.

The last three points of the adviser's list concern his own further education. The art of keeping oneself informed on literature is, as every researcher knows, not easy. Besides different abstracts, review articles of Annu. Rev. Psychol. and Rev. educ. Res. may constitue good starting-points, but continuous scanning of certain journals cannot be avoided. Of the statistical ones I have found Ann. math. Statist., J. Amer. statist. Ass. and J. roy. statist. Soc. most useful. Other journals not belonging to behavioural sciences and which are of more or less permanent interest are Biometrics and Biometrika. Among our own journals you may find items of methodological interest here and there, but I have found the likelihood of this to be greatest in the following journals: Brit. J. math. Statist. Psychol., Educ. psychol. Measmt, J. math. Psychol., Multiv. behav. Res. and Psychometrika. I may also add Behav. Sci. and Econometrica.

Of course, it is also important to read books on research methodology, of which I have already mentioned some. However, there are, with some exceptions, not many new things to be found in textbooks. I have sometimes wondered how books on what is called measurements and elementary statistics can be released in such numbers as is the case today. They are evidently attractive, because I do not think that publishing companies are economic fools. Is there really an increasing number of people who read these books or do the same people who have not managed to grasp the content and therefore, with an admiring optimism, hasten from book to book in the vain hope of eventually understanding the subject? If the last presumption is correct, due to insufficient basic knowledge of the reader and/or pedagogical failures of the books, much has to be reconsidered.

I do not intend to waste many words on point six: to take part in conferences and other valuable contacts. One can say a great deal concerning the construction of conferences, such as the number of members and the limitation of the subject, but most researchers agree that personal contacts are very important. This is particularly true for the isolated specialist, like some advisers on research methodology, because his daily work seldom makes it possible for him to have mutually profitable conversations within his own area of specialization with colleagues.

The adviser's own research

The adviser can do research of his own on many different problems. As I see it, there are three main areas:

1. The organization of research methodology.
2. The use of known methods.
3. The development or introduction of new methods.

The first area will perhaps surprise some people, but I believe it to be essential to examine how research methodology should be organized. The methodological quality is dependent on an efficient system to spread methodological knowledge. Many models for this are conceivable, as are criteria for evaluating such models. For instance, one can have some special institutes doing methodological research or its members can be allocated at different educational and psychological institutes, the degree of specialization of these members has to be discussed and this goes for the teaching of methodology too.

Criteria are not easily produced, but one simple way to get some is just to ask researchers about what they need and what they are missing. Besides the difficulty of defining a researcher, which must be done under all circumstances, there is the risk of getting meager results because some researchers do not see any problems here, for different reasons. Another way could be to do a content analysis of what has been written about these problems or let some experts do the job; both attacks can also have several drawbacks. Anyhow, some criteria are more disputable than others. For instance, few researchers will deny that experiments are preferable to non-experiments, other things being equal, while the proper choice between parametric and non-parametric statistics will certainly divide researchers on many applications. So you may be able to write down some 'unequivocal' criteria, take a sample of reports (which by the way requires a definition of a population) and report your results. Something similar to this has been done by Stanley (1967), whose paper also takes up some other points which I have touched upon.

The second point is most in line with the purpose of the Society of Multivariate Experimental Psychology: to develop methods through feedback from the experience of their use. I have already mentioned something: the problems of data analysis and the problems about similarities and differences of methods. Much is done on measurement theory which is not used in field applications and I suspect that different results in a special area may very well be the effect of different researchers realizing their concepts on different scales. Why not make a study of this to learn the consequences of different scales? Or to take another example: what would be the consequences of doing all the t-tests (or the ANOVAs) instead of the factorial discriminant analysis in the situation mentioned before? And still another one: what Cattell (1952) has once named R and Q analysis are in some sense the same analyses, but from other points of view one of them is better than the other. Why is the individual-space used most of the time? Can we use e.g. Q discriminant analysis for anything?

Perhaps somebody wants to include the answer to the last question within the development of new methods, and the boundaries are indeed very fluent. This reflects reality, for experience with known methods points to certain things, where more research work seems to pay off, that is to develop methods further. (This is valid for programs too.) One principal dividing point, concerning statistical methods, is whether they involve latent classes, clusters, etc. or not. Methods which comprise manifest classes have been examined to a greater extent than the other kind of methods (the split-up methods), but there remains a great deal to be done. For example, methods like canonical correlation analysis are usually based on some optimality criteria which may be quite insensitive: very different weights can still give a correlation in the neighbourhood of the optimal correlation. Many times one is interested in knowing whether weights with special restrictions (e.g. the two weight vectors defining the first canonical correlation are equal) produce a correlation sufficiently near the optimal one. I believe that this kind of hypothesis testing is both useful and underdeveloped. I have seen a test for fully assigned weights of discriminant analysis (Rao, 1965, pp. 482-483), but I do not know if this has been extended to arbitrary restrictions, e.g. with the help of the generalized likelihood-ratio criterion.

The split-up methods are often based on such things as the axiom of local independence, the mixture of multivariate normal distributions, discriminant analysis criteria or combinations of these. They have the common property of producing homogeneous classes which differ from each other. Other types of split-up methods aim at multivariate matching, e.g. individuals are to be matched in quartets on the basis of several variables or items are to be clustered into three parallel tests. Of course, you may also imagine combinations of split-up methods and methods with manifest classes. While I can see some meaning in performing e.g. a 2 x 2 factorial design on factor analysis I cannot see the significance of a latent profile analysis which produces a 2 x 2 factorial design or any other latent design. Anyhow, the split-up methods offer a tremendous amount of problems.

There are several other developments and/or introductions which can be indicated. Such an area concerns checks on agreement between judges, and the area needs to be cleared up. It seems to me that the attacks on this problem are almost as numerous as the authors describing it. Multivariate non-parametric statistics is another area which we should use to a greater extent. As is clear from Aitkin (1971, p. 248) something is already done and 'the systematic development of these procedures should produce important shifts in emphasis in the use and teaching of statistics'. One may also wonder how long time it will take until we dare to use Bayesian statistics, which has produced a metatheoretical issue that has been debated with sometimes great heat. Much has been done in this area, see e.g. De Groot (1970), but we hardly use it. (Strange to say, we instead use several Bayesian ideas in applications as e.g. psychometry and behavioural decision theory, which is clear from Novick & Jackson, 1970, and Slovic & Lichtenstein, 1971, respectively.) Are we scared by the orthodox statisticians' objections, or do we not expect ourselves to be capable of honestly assessing prior probabilities?

FINAL DISCUSSION

Simple problems mean simple statistical methods, complex methods are required for complex problems and when the complexity of the problems grows the behavioural researcher has usually not the knowledge to match the problem with an appropriate method. It is e.g. not unusual to find rather complex designs which use too simple tests and as more and better programs become available you cannot blame those. (I think that the failure of matching complexity of problems with methods goes for the construction of mathematical models as well. However, I have not defined research methodology to include constructions of theories, etc.)

Incompetent use of measurement theory, psychometry and statistics cannot but lead bo bad research, the consequences of which are not easily predicted. The remedy for this is better mathematical training of the behavioural scientist and the creation of special, well-educated advisers on research methodology. This idea is by no means new, because it is pretty close to what Stanley (1967) tells about Laboratory of Experimental Design. Also, several aspects here are in agreement with those in Härnquist (1972), whose paper treats a more general subject. But it is quite a different thing to induce researchers to admit that these bad methodological conditions exist and still another to get funds for improving them. (According to my own experience, the latter seems to be the most difficult task.) I am afraid that researchers in education and psychology are slowly dividing themselves into two camps depending on knowledge of mathematics, statistics, etc. and that these two camps have increasing difficulties in communicating with each other. This process must be stopped and I do not intend to propose that to be stopped by neglecting all statistics, etc.

My purpose with this paper has only been to draw the attention to some methodological problems within education and psychology. I have also given some hints here och there how some of the problems may be solved. However, unless more researchers are willing to change their attitudes to methodology, very little can be done. There are still too many researchers who look upon statistics, etc. with suspicioun and who show resistance to learning and applying better methods. Some will acknowledge good knowledge of research methodology with their tongues but not with their hearts, saying 'multivariate analysis is a fine thing but I have no use for it'. Three possibilities come into my mind: either they will not admit their inadequate knowledge and/or they are not able to see any problems here, or these problems are only figments of my imagination. If the third cause is the correct one, the content of this paper will come in for severe criticism. It will, in such a case, be very interesting to follow the line of proof. If there are no reliable counter-arguments, it is about time to set free resources to improve research methodology.

REFERENCES

Aitkin, M. Statistical theory. (Behavioural science application.) Annu. Rev. Psychol., 1971, 22, 225-250.

Bock, R.D. & Jones, L.V. The measurement and prediction of judgment and choice. San Francisco: Holden-Day, 1968.

Bracht, G.H. & Glass, G.V. The external validity of experiments. Amer. educ. Res. J., 1968, 5, 437-474.

Campbell, D.T. & Stanley, J.C. Experimental and quasi-experimental designs for research on teaching. In Gage, N.L. (Ed.) Handbook of research on teaching. Chicago: Rand McNally, 1963, 171-246.

Cattell, R.B. The three basic factor-analytic research designs - their interrelations and derivatives. Psychol. Bull., 1952, 49, 499-520.

Cohen, J. Multiple regression as a general data-analytic system. Psychol. Bull., 1968, 70, 426-443.

Cohen, J. Statistical power analysis for the behavioural sciences. New York: Academic Press, 1969.

Cooley, W.W. & Lohnes, P.R. Multivariate data analysis. New York: Wiley, 1971.

Cox, D.R. & Miller, H.D. The theory of stochastic processes. London: Methuen, 1965.

DeGroot, M.H. Optimal statistical decisions. New York: McGraw-Hill, 1970.

Dempster, A.P. Elements of continuous multivariate analysis. Reading, Mass.: Addison-Wesley, 1969.

Dixon, W.J. (Ed.) BMD Biomedical computer programs. X-series supplement. Berkeley, Los Angeles: Univer. Calif. Pr., 1970.

Gulliksen, H. Theory of mental tests. New York: Wiley, 1950.

Hays, W. Statistics for psychologists. New York: Holt, 1963.

Hotelling, H. The most predictable criterion. J. educ. Psychol., 1935, 26, 139-142.

Horst, P. Factor analysis of data matrices. New York: Holt, 1965.

Härnqvist, K. Training and career structures of educational researchers. Rep. Inst. Educ., Univer. Göteborg, nr 18, 1972.

Jöreskog, K.G. A general method for analysis of covariance structures. Biometrika, 1970, 57, 239-251.

Larsson, B. Use of mathematical language within educational research reports: Some examples of misuse. Pedagogisk-psykologiska problem (Malmö, Sweden: School of Education), nr 140, 1971.

Lawley, D.N. & Maxwell, A.E. Factor analysis as a statistical method. London: Butterworths, 1963.

Lazarsfeld, P.F. & Henry, N.W. Latent structure analysis. Boston: Houghton-Mifflin, 1968.

Lord, F.M. & Novick, M.R. Statistical theories of mental test scores. Reading, Mass.: Addison, Wesley, 1968.

Miller, R.G. Simultaneous statistical inference. New York: McGraw-Hill, 1966.

Morrison, D.F. Multivariate statistical methods. New York: McGraw-Hill, 1967.

Novick, M.R. & Jackson, P.H. Bayesian guidance technology. Rev. educ. Res., 1970, 40, 459-494.

Rao, C.R. Linear statistical inference and its applications. New York: Wiley, 1965.

Rozeboom, W.W. Foundations of the theory of prediction. Homewood, Ill.: Dorsey Pr., 1966.

Savage, L.J. The shifting foundations of statistics. Stencil, 1970.

Slovic, P. & Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organiz. Behav. human Perform., 1971, 6, 649-744.

Stanley, J.C. On improving certain aspects of educational experimentation. In Stanley, J.C. (Ed.) Improving experimental design and statistical analysis. Chicago: Rand McNally, 1967, 1-60.

Tatsuoka, M.M. Multivariate analysis: Techniques for educational and psychological research. New York: Wiley, 1971.

Torgerson, W.S. Theory and methods of scaling. New York: Wiley, 1958.

Van de Geer, J.P. Introduction to multivariate analysis for the social sciences. San Francisco: Freeman, 1971.

Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Department of
Educational and
Psychological Research
School of Education
Malmö, Sweden

Abstract card

Larsson, B. Some problems in advising on research methodology. *Didakometry* (Malmö, Sweden: School of Education), No. 36, 1972.

This paper discusses the poor conditions of research methodology in the behavioural sciences due to insufficient consumer knowledge and lack of well-educated advisers in this area. It starts with a definition of research methodology, examplifies the kind of knowledge needed by the adviser and then goes on to the main subject, that is, different aspects of the adviser's tasks.

Indexed:
1. Research methodology
2. Advising on research methodology

Reference card

Larsson, B. Some problems in advising on research methodology. *Didakometry* (Malmö, Sweden: School of Education), No. 36, 1972.