PEARL

04 University of Plymouth Research Theses

01 Research Theses Main Collection

1979

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND REPERTORY GRID STUDY

KEEN, TERENCE R.

http://hdl.handle.net/10026.1/2175

http://dx.doi.org/10.24382/3213 University of Plymouth

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

PEDAGOGIC STYLES III PHYSICS TEACHING: ALI ATTUTUDE SCALING AND BEPEBTOBY GBID STUDY





x70523668		
PL LEAR	YMOUTH POLYTECHNIC NING RESOURCES CENTRE	
ACCH. No.	5 366278 FHS 809	5,5
CLASS No.	T-530.07	KEE

.

55.0278

CONTENTS

PAGE NO.

Acknowledgeme	nts
Abstract	••••••••••
CHAPTER 1.	Introduction 6
CHAPTER 2.	Setting the scene with respect to the literature
CHAPTER 3.	The Research Hypothesis 47
CHAPTER 4.	The Attitude Inventory 60
CHAPTER 5.	Factors associated with attitudes to Physics: an evaluation of the instrument designed, and the method by which sub-groups may be identified for the investigation of the hypothesis
CHAPTER 6.	The Repertory Grid as developed for use in this research
CHAPTER 7.	Selection of sample, use of instruments 175
CHAPTER 8.	Analysis of results with respect to the hypothesis
CHAPTER 9.	"Newsworthy results" : a summary of findings
CHAPTER 10.	Conclusion: Extension studies, etc 285
Appendix	•••••••••••••••••••••••••••••••••••••••
Bibliography .	••••••••••••••••••

ACKNOWLEDGEMENTS

Without the help of many individuals and institutions, this research could not have been conducted.

Particular thanks are extended to both academic advisers and supervisors;

Roger Bullock; Dartington Hall Social Research Unit Roy McHugh; Open University

Adrian Thomas; Open University

All provided crucial guidance and stimulation in the early years whilst my special thanks are extended to Prof. Brian Lewis (Open University) for his support and encouragement in the latter stages of the research as principal supervisor.

The Polytechnics and Universities, through providing me access to their staff and students, made the data gathering possible and particular thanks must be extended to Plymouth Polytechnic for the award of a staff fellowship which enabled me to pursue the research full time for one year.

Finally, but by no means least, I would like to acknowledge the tolerance of my wife, Pru, and children, Nicky and Martin, who have so often cheerfully accepted that 'Dad' will be yet again burning the midnight oil!

Thanks.

PEDAGOGIC STYLES IN PUYSICS TEACHING : AN ATTITUDE SCALING AND BEPEBTOBY GBID STUDY

ABSTRACT



ABSTRACT

The need for a greater understanding of the teaching-learning process in higher education is clearly evident from the rapid increase in research activity in this area over the last decade. Economic and political pressure often defines the rationale for such research in terms of increased efficiency; it is the author's contention that such an outcome might still result even if the emphasis is shifted towards effectiveness. Indeed it is the body of academics who are concerned with staff development that have redefined the problem in these terms. Much research has concentrated on the outcomes of the teaching-learning process in order to try and measure teaching effectiveness. Whilst it cannot be denied that these researchers have recorded some interesting observations, the author has shown that this perspective on the problem presents a less than complete picture.

In the work reported in this thesis an adaptation of a Repertory Grid methodology based on the principles of Personal Construct Psychology was used primarily to examine in detail the ways in which different teachers and students perceive the roles of themselves and others in the teaching-learning process and to explore their perspectives on the range of pedagogic styles commonly utilized by the educators.

The author has been able to demonstrate that groups of students and teachers who share common attitudes towards the discipline of physics are more likely to demand a common pedagogic style in order to facilitate effective, and therefore efficient learning. The development of an attitude scale enabled these groups of common attitude to be identified on four attitudinal components, ranging from exam orientation to pleasure from physics. The author investigates a number of hypotheses relating attitude and pedagogic style. The conclusions drawn from the data enable the author to make comment on the validity of a number of commonly held misconceptions within the arena of physics teaching in higher education.

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND BEPERTORY GRID STUDY

INTRODUCTION



- 1.1. Doctoral Theses can originate, it seems, in two rather different ways. One increasingly-common method is for the researcher to subject some already-established body of literature to an "in-depth" scrutiny, with a view to identifying questions and issues which look as if they ought to be examined more closely. Typically, the aim of such an exercise is to uncover possible paradoxes and anomalies, hidden assumptions, suspect methodologies, doubtful looking chains of reasoning, counter-intuitive assertions, alternative explanations, seemingly-suspect findings, and so on. The ensuing thesis work then consists in attempting to resolve or (at the very least) illuminate some of the queries that the literature search managed to reveal. The point to be noted about research of this kind is that it is a body of literature - or, more precisely, the <u>deficiencies</u> in such literature - that is the main source of inspiration for whatever hypotheses are tested. In other words, the hypotheses that get tested are a reaction to the literature, rather than a reaction to problems that present themselves in the real world.
- 1.2. The work reported in this thesis did <u>not</u> originate in this way. The originating ideas did not come from some already-existing body of literature. Rather, they were the outcome of several years of personal reflection on the problems of teaching undergraduate level physics in the FE/HE area.
- 1.3. To the conscientious teacher, teaching poses a wide range of difficult and challenging problems. This seems to be

particularly true of teaching which is aimed at first year undergraduates - because it is during the first year that student attitudes (toward the subject matter being taught) are likely to "harden." To make matters worse, there are certain subjects - and physics (like mathematics) is generally agreed to be one of these - that seem, for reasons which are by no means well understood, to be "extra difficult" to teach effectively. The teaching of physics to first year undergraduates is, therefore, a doubly hazardous venture. It is hazardous by virtue of the students being in their first (and probably most impressionable) year, and it is hazardous by virtue of certain special difficulties that seem to inhere in the very nature of the subject matter. However conscientious and well intentioned the teacher might be, it is altogether too easy for students to be "turned off" - even to the point of dropping out - by what they (i.e. the students themselves) regard as bad teaching.

1.4. This, then, was the problem that first presented itself to the present writer back in 1971 - shortly after he had been appointed, as a qualified physics teacher, to teach physics to first year students in a College of Further Education. The problem, in its most general form, is that of discovering ways in which such teaching might be rendered more effective. Clearly, this particular problem did not emerge from any reading of the professional literature. Rather, it manifested itself as a pressing and recurring problem in a real-life teaching situation. As a corollary, it follows that the relevant

literature (i.e. literature relevant to solving the problem) had to be sought <u>after</u> the problem had appeared.

- 1.5. There are obviously many possible reactions to the kind of problem just described. For example, an expert in curriculum design might seek to redesign the first year physics curriculum. An expert in human motivation might look for novel incentives. An expert in micro-teaching or audio-visual aids might experiment with different ways of presenting the normal first year material. An expert in individualised teaching might try some variant of, say, the Keller Plan. And so on. Since the author did not at that time consider himself to be an expert in anything at all, he reserved his judgment on what might usefully be done to facilitate the teaching and learning process. Instead of looking for some "off the peg" technique to try out on the first year students, he embarked on a series of informal discussions - with the students themselves, and with other physics teachers, and with other non-physics teachers of first year students. The general aim of these discussions was to get a feel for how these different kinds of persons "saw" the problem as a whole.
- 1.6. These discussions went on, intermittently, for several months. As they progressed, the present writer became increasingly aware of the fact that, within each of the three groups of discussants (students, physics teachers and non-physics teachers) there were marked differences of attitude: (a) towards physics, per se, and (b) towards the ways in which physics might best be taught.

For example, some students seemed to view the learning of physics as a somewhat burdensome task that they had agreed to go through in order to pass an examination. Other students seemed to be wanting to learn the subject for its own sake. Yet others seemed to be primarily interested in the practical use that they could later make of physics, in the real world. A little reflection suggests that these "attitude And so on. to physics" are roughly synonymous with "reasons for learning In other words, they say something about the physics." perceived goals of the learning process. For one student, the goal is a piece of paper stating that he has satisfied the For another, the goal is the satisfaction that examiners. comes from studying a subject which one really enjoys. For yet another student, the goal is the promise of practical realworld applications of the knowledge being imparted.

1.7. As might be expected, these different ways of viewing the teaching/learning process tended to be shared, to varying extents, by the teachers themselves. Thus, some teachers seemed to be primarily concerned with pushing their students through the requisite examinations. Others seemed to be trying to communicate their own love and enthusiasm for the And others seemed to attach special importance, in subject. their teaching, to the practical uses to which physics could be To the extent that a teacher pursues one general goal put. (e.g. the goal of getting his students through the mandatory examinations) rather than another, he will obviously tend to adopt a PEDAGOGIC STYLE which strikes him as being most

appropriate for the achieving of that goal. Such a style is then likely to <u>irritate</u> any student who has a strong preference for a somewhat different goal!

1.8. To judge from the kinds of discussions that the writer originally held, a failure to share common goals was by no means the only source of frustration and friction among students and teachers. It is sometimes glibly suggested that people learn most from other people who are different from themselves. This may be true of some people. But the majority seem more likely to reject people whom they perceive as being different from themselves. Conversely, they seem to "warm" towards people whom they perceive as being essentially similar - in values, ideologies, world views, and the like - to themselves. There is little doubt that students can be quite quick to reject teachers whom they perceive (perhaps rightly, perhaps wrongly) as having incompatible values and outlooks. Some teachers are equally quick to show disapproval of students who do not seem to share their values and outlooks. The whole problem is exacerbated by the fact that teachers and students often seem to be unaware of the values and outlooks that they tacitly subscribe to and/or evince. For example, a teacher who pays lip service to the need to learn physics "for its own sake," might actually teach as if his sole objective in life is to get his students through their examinations. And a teacher who fancies himself to be clear and incisive might come across to his students as being woolly and muddled.

- 1.9. Considerations of this kind have led in recent years to a variety of "Aptitude-Teacher-Interaction" experiments, in which attempts have been made to achieve some sort of match (e.g. of personalities and/or cognitive styles) between teachers and their students. However, the discussions which the present writer held, back in the period 1971 to 1973, persuaded him of the need to cast the net somewhat wider and to embark on a general investigation (with respect to first year undergraduate physics) of the interaction between PEDAGOGIC STYLES of teaching and ATTITUDES toward the subject matter. This, in a nutshell, is what the present thesis is all about.
- 1.10. Any attempt to explore the complex relationships between Pedagogic Styles and Attitudes obviously calls for a discussion of what these words are taken to mean. More will be said about this shortly. For the moment, it may be sufficient to note that, irrespective of the way in which these words finally get defined, there is clearly a need to construct a data-eliciting procedure which will reliably exteriorise the attitudes in question, and which will also exteriorise (i.e. make available for public scrutiny) the ways in which people perceive and judge For reasons that will later be made clear, it pedagogic styles. was accided to use the Repertory Grid Technique as the The first stage of the appropriate exteriorising instrument. research program was accordingly devoted to the careful construction and validation of such an instrument. The instrument was then applied to the three groups of persons which have already been briefly defined.

1.11. As subsequent chapters try to show, the repertory grid technique elicited a rich crop of data about people's perceptions and assessments of different pedagogic styles, and about their attitudes to physics, per se, and to the "goals" of physics teaching. As a result of the research it was possible to classify pedagogic styles and attitudes in promising-looking ways. It was also possible to compare and contrast the somewhat different perceptions of the three groups of subjects (students, physics teachers and non-physics teachers) studied, and to explore the ways in which pedagogic styles and attitudes More precisely, it was also possible to probe the interact. validity of no less than twelve hypotheses to do with attitudes and their influence on the perceived effectiveness of different pedagogic styles. Although the entire project is written up as if it were a piece of "pure" research, frequent attention is drawn to some possible practical implications of the findings. Of central importance, here, is the author's conjecture that attitude matching may prove to be more effective than aptitude In other words, if a teacher has the or ability matching. choice of matching pupils (in small groups) on the basis of shared attitudes toward the subject matter being taught, then this might prove to be pedagogically superior to matching them on the basis of, say, similarity of I.Q. Another way of expressing the point would be to say that it is a mistake to assemble, in a single group, all students possessing a high I.Q., if the resulting high-I.Q. group contains students who have radically different attitudes towards the subject matter being taught. However, this is only one specimen example of the way

in which the results of the present research might influence physics teaching in the future. More generally, it is the author's contention that the criteria used to form student groups, and the methods of instruction advocated by teachers of undergraduate physics, the students' perceptions of their own needs in terms of pedagogic style, and teachers' perceptions of their students' needs all require further investigation.

Furthermore the author intends to provide evidence which will not only permit an investigation of the areas noted but which may substantiate the claim that teachers of physics, teachers of disciplines other than physics and physics students all tend to use different criteria when assessing a teacher's effective-It is common practice in most institutions of further or ness. higher education to use a measure of aptitude or ability to determine the composition of student learning groups. (The author expands this argument later, when the concept of 'perceived academic potential' is introduced). Furthermore the academic teacher in Great Britain, unlike his American counterpart, is not regularly assessed and he is given considerable autonomy in choosing his preferred pedagogic style, which, it will generally be agreed, is chosen intuitively by the individual staff member without reference to student feedback or Whilst the emergence of staff development: peer appraisal. units (S.C.E.D.S.I.P., 1973) directly concerned with the improvement of teaching effectiveness, has occurred in the carly 1970s such units have been preoccupied with the behavioural assessment of teaching against criteria externally specified by their 'experts.' (Proceedings of the 3rd. International

Conference on the Improvement of University Teaching 1977). The author contends that both the methods used for student group composition and the strategies adopted to assess teaching efficiency (rather than appraise learning effectiveness) are less than ideal, generating a need for an investigation of the areas noted in section 1.11. above.

1.12. No excuse need be, nor indeed is, offered for this initially negative and critical view, if it is recognised that such expressions constitute a very subjective set of hypotheses arising from extensive personal experience, both as a teacher of physics in FE/HE and in the role of one who endeavours to analyse and improve the teaching of others.

The opinions expressed represent a starting point, an orientation for constructive criticism, analysis and development of an alternative system. As Claude Bernard stated in 1865, "A hypothesis is the obligatory starting point of all experimental reasoning. Without it no investigation would be possible and one would learn nothing; one could only pile up barren observations. To experiment without preconceived ideas, is to wander aimlessly." (Bernard, 1865).

1.13.1. The author has already used terms, which have definitions specific to this research, these require explanation. "Pedagogic style." This term is taken to embody all of what a teacher 'does' when exercising his teaching function. Teaching is seen at one level (roughly, the level of "actions")

as a practical activity wherein the teacher spends a large proportion of his time writing on blackboards, marking books, preparing lessons, arranging classroom equipment, listening to Whilst at a second level his students talking, etc. (roughly, the level of "goals") it may be seen as facilitating Many problems of a practical nature develop in learning. teaching, most being associated with the teacher's basic task that of making available certain kinds of knowledge and enabling the student to relate the un-commonsense knowledge of the subject to the commonsense knowledge already internalised This aspect of teaching within his cognitive structure. emphasises the teacher as a practitioner. Additionally, the teacher is an exponent of applied knowledge working within Whilst this two part model of professional traditions. teaching could be criticised for its over simplification, it does provide a fragework upon which the research may be based. The autonomy of teachers in Great Britain has meant that each individual can evolve his own particular set of skills and techniques which he uses as he considers appropriate in order to The author uses the term facilitate learning in his students. 'pedagogic style' as a means of specifying a complex set of techniques, behaviours or understandings which characterise one individual's teaching role in a specified setting. Thus, a 'pedagogic style' is both a description of how teacher 'A' manages a given learning situation and additionally, how a hypothetical teacher, 'B', might manage the same learning situation.

- 1.13.2. "Perceived academic potential." Students of physics are frequently grouped by 'ability' (based on a previous year's performance) before they are exposed to new learning situations. Logically, if they are to learn new material, they cannot be classified according to their real ability to learn in the new setting as this cannot be evaluated until the learning Thus, their performance on a test of past experience is over! learning experiences is used to predict their performance in the It is postulated, therefore, that any test of new setting. previous learning can only provide a measure of "perceived academic potential" for performance in a new setting. The author explains the concept of "perceived academic potential" further in a joint paper with Reid (keen & Reid, 1977).
- 1.13.3. Effective teaching. Educators regularly debate the merits of evaluating effective teaching. In this research, the author has chosen to simplify the problem by consciously avoiding the temptation to quantify effectiveness, which would, of necessity, have meant using some form of teacher appraisal system (e.g. IDEA, Biles 1975). Such systems, and there are many described in the literature, are, without exception, behaviourly based. Thus the teachers performance is compared with some set of supposedly good attributes of teaching, usually teacher behaviours. (Such as 'adequate use of questioning,' 'uses set induction,' etc.). The author determined from the beginning of the research to avoid the imposition of criteria of 'effectiveness' specified by some external 'expert.' Instead of quantifying effective teaching the author chose to 'qualify'

the term by placing each and every subject who provided data in the position of a personal 'expert' who knows what he perceives as effective teaching. A methodology was developed which enabled these individual perceptions of effective teaching to be made explicit to the external observer. Thus, the term 'effective teaching', as used by the author, accepts the semantic differentiation from 'Efficient teaching' which was considered at length by Sayer (Sayer, 1977) and others, and defines a pedagogic profile (see paragraph 1.13.1.), perceived by either an individual respondent or a group of respondents, which he or they believe promotes maximum learning in the student groups exposed to that pedagogic profile.

1.14. An investigation which purports to question the areas identified, demands a methodology which can measure attitude towards certain prespecified aspects of physics, and which can make explicit pedagogic styles perceived as representing effective teaching by respondents.

The first methodological problem, namely attitude measurement, may be resolved by a review of existing measures of attitude. A detailed discussion relating to such a review together with the rationale for the development of a new attitude scale is undertaken later, for the purposes of this introduction it is sufficient to state that a new inventory was seen to be required and was, in consequence, developed. The second methodological problem, that of specifying pedagogic style, presented a rather more serious hurdle. Having chosen to define 'pedagogic style'

as indicated it became necessary to explore the possibilities of a <u>multi-dimensional "profile"</u> which could be seen to cover all the recurrent criteria seen as important by the respondent. Many behaviourist style techniques are available for such profile formulation but all suffer from the serious drawback of specifying supposedly important criteria from some external 'expert' source. Two techniques existed, however, which were not behaviourally oriented, these were Kelly's repertory grid technique and Osgood's Semantic Differential. For reasons specified in later chapters the repgrid technique was adopted and subsequently developed by the author in a unique way relative to previous attempts to use cognitivist techniques for the production of 'pedagogic style' profiles.

1.15. Early in the developmental stages of the research the author became aware of the magnitude of the task he had undertaken. A clear need to specify the scope of the study was evident, the nature of the research problem being seen by the author as The techniques guite different in each sector of education. adopted, and the findings described, cannot necessarily be transferred from FE/HE to secondary or secondary to primary education. Therefore, a decision had to be taken with regard to the sector of education which was to form the basis of the A variety of factors contributed to the decision research. made to restrict the study to FE/HE. The major factors included the experience of the researcher which was mainly in that sector, the relative ease of access to a representative sample of respondents and, not insignificantly, the demand

being made by physics faculty staff and students to investigate physics teaching generally.

The FE/HE restriction limits the research only in so far as is necessary, in order to complete the project within the limits imposed by time and resources.

1.16. Having crudely specified the areas of interest, considered the methodological requirements and determined the scope of the study the author was able to embark on a review of the literature and specify the particular hypotheses which the research data would support or refute.

The reseach was conducted by adopting, what some writers describe as a phenomenonological approach. (To these writers, an approach is phenomenological if it calls for an exteriorisation or verbal reporting of one's current subjective experiences). In its extreme form a phenomenonological approach demands that a researcher becomes one of the subjects in his field of study and conducts the research from 'within.' The author had experienced being a member of each of three categories of respondents, and the specification of the hypotheses was made from his position of former experience. This, together with the intent to place the onus for 'expert' opinion on the respondent justified the phenomenological label.

No strong prescriptive comments can be necessarily made from the conclusions, but the raising of perceptional awareness of the

problems, of pedagogic relevance, together with an objective comment on one static picture of physics teaching, may lead to the identification of useful areas of subsequent research specifically designed to be of a prescriptive nature.

•• -4

PEDAGOGIC STYLES IN PHYSICS TEACHING : AN ATTITUDE SCALING AND BEPERTORY GRID STUDY

SETTING THE SCENE WITH RESPECT TO THE LITERATURE



- 2.1. People (like the present author) who have gained their initial qualifications in the "hard" physical sciences, tend to take it for granted that researchers should always give (or be capable of giving) precise definitions of the key terms that they use. It therefore comes as something of a surprise to discover that researchers in the "soft" human and social sciences are so often casual about the terminology that they For example, if one examines any half-dozen papers deploy. (chosen at random) on the subject of "prejudice," it can be quite difficult to decide whether, and to what extent, such papers are actually talking about "the same thing." Moreover, what one researcher talks about under the heading of "prejudice," another writer might talk about under the seemingly-different heading of, say, "ethnic aggression." Under such circumstances, words such as "prejudice" and "ethnic aggression" would seem to be little more than labels of convenience - tag words which signal the fact that the paper "has something to do" with what commonly (and vaguely) tends to go under the name of prejudice or ethnic aggression. Such words are simply not intended to have any very precise meaning.
- 2.2. Since many highly-regarded theorists and researchers in the soft sciences seem quite sanguine about the deployment of loose and ill-defined terminology (and since, moreover, their <u>ancillary</u> supporting terminology often tends to be even more loose and ill-defined), it is natural to ask whether precision really matters in the field of human and social research.

Several comments are in order here. First of all, it may be noted that definitions are, by their very nature, unlikely to do justice to anything as complex as a psychological phenomenon. Definitions typically consist of no more than one or two sentences. At the very most, they can highlight some feature of the phenomenon defined - and also, perhaps, briefly indicate the way in which the phenomenon compares and/or contrasts with some relevantly related phenomenon. If anything more detailed is required, more than one or two sentences would be called for - and that would move the researcher into the realm of "explications" or "elucidations," rather than succinct definitions.

Secondly, it was Kant who first remarked that definitions are the <u>end</u>-points, rather than the starting-points of intellectual enquiry. Of necessity, one might start with a provisional or tentative definition. But, as one's enquiries proceed, the initial definition will be seen to be inadequate and in need of modification. Perhaps after many years of enquiry, one eventually will be able to "capture the essence" of a phenomenon in just one or two sentences. But, even then, the one or two sentences that are used will mean very much more to the creator of the definition, then they will to the reader of the definition. This is so because the casual reader will have no conception of the numerous "candidate definitions" that were considered, and rejected, along the way.

A third point to be noticed is that the Human Race has actually

managed to come quite a long way on fuzzy thinking! Precise thinking may well be desirable, but it cannot be convincingly argued that precise thinking is essential to scientific As a matter of fact, some theorists hold that the progress. quest for exactitude in human affairs is always a mistake - and the most that can be said in favour of the quest for exactitude is that it sometimes turns out to be an instructive mistake! It is fairly clear, for example, that mathematical models of human behaviour do scant justice to the way in which humans actually behave (except under special laboratory conditions in which the subjects agree to behave in accordance with the experimenter's model). There may, however, be something to be learned from the ways in which the model fails to predict what actually happens.

Finally, it may be pertinent to remark that attempts to secure greater precision, in the use of psychological terminology, do not seem to have been particularly fruitful. It is not clear what is gained by lengthy discussions as to what a particular term or word might better signify

2.3. What, then, should be the psychological researcher's stance toward the main terms that he uses in his research reports? If the foregoing considerations have any validity at all, there would seem to be no <u>necessity</u> for the researcher to clarify, with any great precision, the terms in question. But it would obviously be <u>helpful</u> if he gave a reasonably adequate characterisation of what he personally takes the terms to mean. This is

particularly the case if (as sometimes happens) the researcher uses a set of terms which seem to him to be interrelated in ways that need to be pointed out. It is also the case if he believes that he is seeing more significance in a particular term than the reader is likely to see. By way of example, suppose that a social psychologist is conducting research into the nature of Character. Now most people are conditioned by society to respect so-called "Men of Character." But it can be pointed out that a Man of Character is actually a somewhat predictable and, to that extent, rigid, man. Since this facet of Character is unlikely to be noticed unless the researcher explicitly points it out, there is clearly a need to point it out.

2.4. In keeping with the above considerations, the present author devoted some time to considering what he personally meant by the word 'attitude.' In particular, he wondered what he ought to say that might not be self-evident to professional researchers working in the attitude assessment area. He finally decided that an appropriate course of action would be to give a brief account of what he took attitudes to <u>be</u>, and what he saw them as <u>doing</u>.

2.5. The Concept of Attitude.

Textbooks of elementary psychology nowadays tend to define Psychology as "The Science of Behaviour" - by which is normally meant the science of <u>observable</u> behaviour. Typically, they then go on to assert that behaviour is determined partly by internal factors and partly by'external factors.'

At first glance, there seems to be nothing objectionable about such views. However, it is the business of science to enquire more closely into the acceptability of seemingly-unobjectionable statements. If we do this, a whole cluster of difficulties can be seen to arise.

For a start, it turns out that the word 'factor' (in the kinds of textbooks cited) is something of an evasion. To some authors, an internal factor is essentially an internal <u>physical</u> state - such as the state of the stomach, or the hormonal state, or the neural state of the brain. To other authors, an internal factor is essentially a <u>mental</u> state - such as a state of hunger, or anger, or fear. To say that behaviour is partly caused by "internal factors" is therefore to avoid the issue of whether behaviour is "caused" by physical events or mental events.

Similar problems arise in connection with the notion of an "external factor." In some textbooks, an external factor is some sequence of physical events impinging on one of the sense organs. In other textbooks, an external factor is a sort of "presenting experience" - such as the sight of a snake, or the cry of someone in pain. There is a tendency in some textbooks to "gloss" these different conceptions by referring to internal factors as "some set/sequence of neural or mental events" - and by referring to external factors as some set/sequence of physical or perceptual experiences, etc. However, this simply raises the question as to how Psychology itself is to be defined. There was a time when Psychology was defined as the Science of Mind (or Mentation), rather than the Science of Behaviour.

Both definitions have their limitations. Psychologists who concern themselves solely with mental happenings (thinking, feeling, desiring, willing, and so on) might legitimately be accused of dealing with Disembodied Minds. Psychologists who restrict themselves to the domain of observable behaviour might similarly be accused of dealing with Diseminded Bodies.

It is further to be noted that not everyone agrees that behaviour is due partly to internal factors and partly to external factors. Some theorists have argued that internal factors - the state of one's body and brain and mind - are themselves determined by external factors. It may be <u>convenient</u> to "explain" behaviour in terms of (say) a person's attitudes and prejudices. But if these attitudes and prejudices are themselves the result of external conditioning, then, ultimately, only external factors are responsible for that person's behaviour. This argument is in fact one version of the argument in favour of Determinism (and against Free Will).

At the other extreme, there are some people - especially some Eastern Mystics - who hold that behaviour is caused solely by internal factors. The view here is that the entire World is nothing but a projection of the beholder's mind. This being so, the world is in the mind (rather than the mind being in the world or - even more restrictedly - in the brain!) and the mind is consequently responsible for everything that we do. No doubt almost everyone would dismiss this latter view out of hand, but it is worth noting that it does exist and that it represents the polar opposite of the Deterministic argument.

In practice, most people seem to hold some intermediate position. Even the most rigid believer in Determinism tends to chastise his children for disobedience - thereby implying that perhaps they do have some free choice in the way that they behave! The overwhelming view seems to be that our everyday behaviour is, indeed, caused partly by external factors and partly by internal factors, and that these "factors" can legitimately be construed either in physical or mental terms.

In the case of human behaviour, there seems to be a bewilderingly large number of "internal factors" that function as partial causes of everyday behaviour. To note this fact, the English language contains such words as need, greed, intelligence, ability, aptitude, disposition, attitude, prejudice, taken-for-granted-assumptions, commitment, character trait, presupposition, ideology, and so on. Any Thesaurus will immediately reveal a whole lot more of such terms.

What all these words have in common is that they all refer to what might be described as "internal controlling influences" on everyday behaviour. One aim of Social Psychology is to see whether such words - and the phenomena that they denote - can be systematised and theorised about in fruitful ways.

Clearly, the kinds of words just listed all have different nuances of meaning. For example, prejudices tend to be somewhat resistant (often irrationally resistant) to change. And they also tend to be situation-specific or context specific.

Thus a person who is anti-Semitic is a person who tends to behave in a particular way toward Jews. But a person who is suspicious tends to be suspicious "right across the board." Suspicion is a character trait which is less "situation specific" - and that is why it tends to be called a character trait, rather than (say) a prejudice.

What, then, is an <u>attitude</u>? An investigation of everyday usage supports the view that it is a fairly general controlling influence on behaviour. It is also one that people can to some extent articulate and be conscious of - although some theorists do speak of people possessing unconscious attitudes. However, it is important to recognise that there is no uniquely correct conception of attitude waiting to be discovered. It is rather the case that, within limits which do not do too much injustice to commonsense usage, we are free to define an attitude in any way that is theoretically promising.

For the purposes of this thesis, there is a textbook definition of 'attitude' which 's well-regarded by attitude researchers, and which happens to accord quite well with the views that have just been expressed. The definition is due to Milton Rokeach ': leading American authority on the subject), and it appears in his book, "Beliefs, Attitudes and Values" (published by Jossey-Bass Inc., San Francisco, 1968). It reads:

"An attitude is a relatively enduring organization of beliefs around an object or situation predisposing one to respond in some preferential manner."

This definition has several merits. First of all, it is

consonant with the commonsense notion that attitudes are (or can be) important determiners of human action. Secondly, it leaves open (in a highly desirable way) certain questions to do with the relationship between behaviour and action. In particular, it does not rule out the possibility of attitudes affecting behaviour, or behaviour affecting attitudes, or each affecting the other. Thirdly, it points to a crucial connection with belief, and it additionally implies that attitudes exist in clusters, rather than in isolation from one Some further implications of the definition are that another. attitudes make people more predictable to one another. In effect they act as blinkers, or as constraints/limitations on Finally, the definition is entirely psychological. thinking. Unlike a better-known definition of Attitude by Allport (which refers to "neural states of readiness to respondent") the definition is not contaminated with references to neural or physiological processes. The definition is accordingly very well suited to the type of experimental enquiry described in later chapters of this thesis. In the present author's opinion, the Rokeack definition, when taken in conjunction with the remarks that introduced it, give an entirely adequate account of how this thesis treats the concept of 'attitude.'

2.6. Having outlined the research problem in Chapter 1, two particular areas have been examined, namely: attitudes and pedagogic style together with the means by which the latter may be recorded. Although the intention is to review specific areas, it is recognised that some relevant research will span two or more of them. One such piece of research is that of Paul Gardner (Gardner 1975) which highlights the relationships perceived to exist between students' attitudes to physics and certain aspects of the teacher's personality and behaviour. Tenuous though these links appear to be, they raise questions about the relationships between attitude and perceived aspects of pedagogic style which the author will endeavour to answer by his research. Gardner concludes that teacher behaviours (aspects of pedagogic style) may enter into lawful relationships with student variables in general and student attitudes in particular.

Gardner's work spans a five year period from the presentation of his doctoral thesis "Attitudes of Physics" through a series of recent papers which consider some of the areas of interest After spending a year on study leave in for this research. England during 1973, Gardner published three papers identifying some relationships relevant to the U.K. arena of physics Some of the earlier papers consider data collected teaching. in Australia and are, for a variety of reasons, (not the least of which is the cultural difference) of dubious applicability in the United Kingdom. The three papers identified make specific reference to physics education in the U.K. and enabled the author to focus the attention of this research on a specific range of guestions within his area of interest and to which either no adequate answer had been previously provided or to which little research effort had been devoted. The papers concerned were: "Research on Teacher Effects: Critique of a traditional Paradigm" (Gardner, P.L. 1975 - The Bournal of Social Psych. 95 pp. 91 - 97).

"Attitudes of PSSC Physics Students: Relationships with Personality and Teacher Behaviours." (Gardner, P.L. 1975 -The Australian Science Teachers Journal vol. 21 (i) pp. 75-85).

These papers provide support for some of the views, and concerns, expressed by the author in Chapter 1; consequently a brief report of their content follows to indicate how the questions framed as hypotheses in Chapter 3 arose from the subjective comments in Chapter 1 in the light of this (and other) research.

2.7. The decline in enrolments in undergraduate physics courses was worldwide in the early 1970s. In Britain 2400 university science places were unfilled in 1973/74, and the Daily Telegraph (10th. November 1973, p.5.) reported a predicted and complete halt in all university science building programmes for In America the 1960 predictions (American Institute a decade. of Physics) of a large increase in female physics students did not materialise and in Australia even those students who had opted to read physics displayed a sharp decline in enjoyment of the subject during their course. (Gardner 1973). At the same time the physics enrolments in Australia showed an absolute decline from 1971 onwards. Gardner concludes that in many different countries, students are voting with their feet against the study of physics despite an enormous expenditure during a 20 year period on the development of innovative physics curricula.

Various writers have tried to explain this state of affairs in

terms of the poor 'image' of physics (Ahlgren & Wallberg 1973), the counterculture's distaste for science and technology (Rodzak 1969), the failure of physics curricula to include social aspects of science which are of interest to adolescents, particularly females (Ahlgren & Wallberg, 1973), and the operation of economic forces outside the control of the science educator (Reitz, 1973).

The emergence of a number of "combined studies" degrees in the late 1970s (Smethurst, 1978) appears to be a consequence of the educational technologists of the time accepting the points made in the papers cited without questioning the possibility of a superordinate component or factor which may subsume some of the valid, but perhaps relatively insignificant, explanations given. Gardner focused his attention on two areas, student attitudes and teacher behaviours. (He uses "teacher behaviours" in much the same way as the author uses "pedagogic style" and the phrases may be, within this thesis, interchanged).

First, a study was conducted to identify relationships between pupil needs and pupil attitudes to various aspects of their physics course. The results suggested that there is not a strong tendency for students to project their own personal attitudes onto their classroom environment. Such a conclusion might be in error due to weaknesses of the Stern (Stern 1958) statistical approach which calculated correlations between individual students for the total sample. It is conceivable that there might be a strong correlation between a personality factor score for a student with a specified classroom
environment (note: a classroom environment is part of a pedagogic profile as defined) within, but not between, student groups. This might mean that certain types of student (classified by attitude) respondent well to a given teacher's pedagogic style but because groups are usually arranged such that they consist of mixed attitude common ability groups the classrooms would mask such relationships.

Gardner concludes that students who are compulsively orderly tended to rate their teacher higher on the corresponding scale, nurturant students perceived their teachers as more nurturant, students who see themselves as well-organised tended to describe their teachers in similar terms and those who are interested in developing their understanding of a field tended to see their teachers as promoting this objective.

This last observation led to the second kind of study relating to the effects of teacher behaviours to student outcomes. Assuming that the instruments are reliable a zero correlation between the teacher behaviour (i.e. aspect of pedagogic style) and the student outcome implies that the teacher behaviour being studied does not affect the students in any meaningful way. The research concludes that teacher behaviours and pupil outcomes (for which one might read: pedagogic styles and changes in the affective and cognitive domains of the learners) may enter into lawful relationships with one another, the nature of the relationships being different for different kinds of student. The results strongly suggest that instead of grouping students in terms of ability (however measured), one might bring

together (in a single group) all students who share a similar preference for a particular kind of pedagogic style and that such a strategy might prove beneficial.

- 2.8.1. Having identified the general areas of interest as specified in Chapter 1, the author was able to direct attention in the literature review to two specific regions which enabled a research programme to be specified which was based on a foundation of existing knowledge and sought to answer the questions seen as unanswered. These two regions were, attitude measurement and recording of pedagogic styles. Methodological strategies in connection with these two areas would also be reviewed.
- 2.8.2. In addition to Gardner's work substantial quantities of research have been conducted on attitude to physics and attitude measurement in general.

Walberg (Walberg, H.S. 1967) developed an inventory specifically related to physics, whilst many other researchers are interested in attitudes to science. (Laughton & Wilkinson 1973; 1978 & 1970; Mayer 1963 & 1959; Aiken & Aiken 1969; Themas 1970; Ahlgren & Walberg 1973; Wicke & Yager 1966; Selmes 1971; Perrodin 1966; Ormerod 1973; Ormerod 1971). These researchers all used attitude as one variable in their studies, many developed their own instruments whilst some adopted or adapted the instruments of others. The important observation that the author would wish to highlight from all of

the work reviewed is not methodological considerations of attitude measurement, but the importance attached to attitude, as a variable, by a wide range of respected researchers. Quite clearly, the last decade has provided increasing awareness of the need to consider attitudes when investigating the teaching-learning process.

Turning now to attitude inventories, it was considered important to review literature outside the specific research interest of attitude to physics. Attitude measurement represents an area of psychological research from which lessons may be learnt and subsequently applied to the specific problem in hand, namely adopting an attitude inventory for this research. The standard references for practitioners make a useful approach Amongst those Scott (1968), Oppenheim (1966), to the problem. Waite (1961) and Edwards (1957) provide the intending evaluator of student attitudes with methodological bases for further work. A major weakness of such standard works is the rather superficial treatment given to the philosophical problems associated Sherif, Sherif and Nebergall with attitude inventory design. (1965) and McGuire (1969) provide a rather more scientific appraisal of the problems inherent in attitude inventory design while other researchers look at specific areas of importance. Amongst the latter category are: OsGood and Tannenbaum (1955) who consider the principle of congruity in the prediction of attitude change, Heider (1946) whose interest focuses on cognitive organisation, Chein (1948) and Doob (1947) who debate the behaviourist-cognitivist dichotomy and Cook and Selltiz (1964) with their multiple indicator approach to attitude

The literature review was not restricted exclusively to 2.8.3. attitude measurement in the discipline of physics. The review aimed at identifying the kind of inventory most suitable for an investigation of the research hypotheses rather than endeavouring to find an inventory containing suitable elements for direct To achieve this end it was found that existing adoption. inventories could be classified into four categories. Firstly. those designed for particular types of student, e.g. Harrison (1971), Brown & Davis (1973). Secondly, those wishing to measure attitude to a particular specified change in the learning environment, such as Finchs (1969) instrument to measure student attitude toward individualised and laboratory Thirdly, those considering only some aspects of instruction. the total attitude as exemplified by the scales of Laughton & Wilkinson (1973) and Bollen (1972). Finally, the inventories purporting to measure the 'total attitude' of the respondent towards the 'total discipline' concerned, e.g. Coon (1969), Skurnik and Jeffs (2971) and Gardner (1974), of these, Gardner alone was responsible for an inventory specifically designed for the discipline of physics. The author, presented with the choice of either adopting the Gardner PAI (Physics Attitude Inventory) or designing his own, decided to develop a new The rejection of Gardner's instrument did not instrument. imply a lack of confidence in the manner in which the PAI was The developed, in fact the contrary view is more appropriate. rejection was based on the problems associated with using an instrument in Great Britain which had been developed using

respondents from an alien culture (Australia), and of an age group different from respondents identified for this research. Adoption of PAI would have necessitated anglicising and piloting the instrument, in fact, the whole process would have represented little saving of resources over the chosen alternative of developing a new scale.

- 2.8.4. Of the four types of inventory found, it became increasingly clear from the research cited that an inventory measuring a limited number of pre-specified attitudinal factors which may be combined to form a composite score would be the most appropriate type of instrument in order to investigate the areas identified in Chapter 1. As will become evident later, four such factors proved to be important in this research: Exam orientation; Practical Orientation; Intrinsic motivation; and Personal Pleasure from Physics.
- 2.9.1. The second area of literature review was that of pedagogic style and how it may be recorded.

Whilst the term pedagogic style has not been widely used by other researchers, there have been many studies looking at the teacher role in the teaching-learning situation. Many of these are relevant to this study. Armidon and Flanders (1963) consider the role of the teacher in the classroom whilst Bales (1950) concentrates on the interaction process occuring in the teaching-learning situation. Wherever 'pedagogic style' or similar terms are used, almost without exception some form of evaluation is implied. Some researchers following the lead set

by Stufflebeam (1968) watered down the concept of 'evaluation' by calling it "information for decision making." This could be interpreted either as containing a judgement or of being noncommittal, e.g. a descriptive account. However, those who support descriptive 'evaluations' (amongst whom the author would be named) are not so naive as to suppose that judgements are no part of their work, but they are more interested in illuminating possible judgements rather than recommending them.

With some notable exceptions, such as Mollet (1977), most 'evaluations' of the teaching situation are attempts to specify good or bad examples, whatever they may be, with the descriptive illuminative approach finding little support amongst academics. Clearly, the expressed intention of using just such an approach in this research project will need to be justified if it is to be academically credible. Such discussion is included in later chapters illuminated by the following extension of the literature survey to include rather more specific methodological studies.

2.9.2. The problem raised by the literature reviewed in paragraph 2.9.1. was related to the competing paradigms within educational research. Power (1976) considers this problem and identifies three paradigms. The 'Agricultural-Scientific' paradigm is analogous to the 'evaluative experimental' approach. The idea is that one can utilise the powerful intellectual and statistical tools of the sciences in studying educational, as well as natural, phenomena. This is a logical outcome of the success story of science. This paradigm implies that, to be worthwhile,

all educational research should be objective, empirical, nomothetic, and value-free; and that well-designed empirical studies will ultimately uncover the laws governing human The contrary view implies that descriptive behaviour. evaluations stemming from a phenomenological approach can provide equally worthwhile outcomes. Power agrees, but divides the view into two different paradigms. Firstly the anthropological paradigm encompassing ethnographic approaches to The first research and, secondly, the philosophical paradigm. of these is the true phenomenological approach which at its extreme, demands that the researcher does extensive field work in the problem setting of a kind which allows him to become a participant observer rather than a detached manipulative The researcher, controlling and measuring people and events. second is that which attracts the author's support and which lends credence to the approach proposed and adopted for this The philosophical paradigm claims that to explain is study. to analyse the incomprehensible into simpler, more understandable components, and to show how these components are inter-Such research Kaplan (1964) and Scriven (1966). related. provides support for the author's approach to the problem. The commonsense everyday knowledge of the research subjects is the prime source of data and, it is the cognitive structure of the respondent to which one has to direct investigation without the imposition of external criteria or beliefs in the superiority of Having established the Young (1971). academic knowledge. credibility of such an approach from the literature, it still remains to consider research based upon such a philosophy for the investigation of pedagogic style within the discipline of

Such investigation is rewarding, for both the depth physics. and breadth of such studies enables the perceptive reader to identify or explore their weaknesses and strengths. Six studies can be identified as representative of this area, namely: Gardner (1974), research on teacher effects; Reed (1961), with teacher variables of warmth, demand and utilisation of intrinsic motivation related to science interests; Horsby-Smith (1973), styles of teaching and their influence upon the interest of students in science; Shavelson (1970), some aspects of the relationship between content structure and cognitive structure Elliott (1971), perceptions of high in physics instruction; school physics and physics teachers; and Mackay (1971), changes on affective domain objectives during two years of physics study. Each of these studies considered alone provides valuable insights into pedagogic styles adopted by physics teachers. Considered together, the importance of pedagogic style relative to the learning of physics students becomes very evident but at the same time, these researchers have shown no commonality of technique in the identification of particular pedagogic styles. In the latter respect, these reports may appear to contradict one another, but in practice it is not contradiction, but a question of apparent conflict which stems from the use of different starting points and inconsistent definitions of pedagogic style between researchers. The hypotheses specified in Chapter 3 are framed to clarify these conflicts and to provide a means of viewing physics teaching in a systematic and (See also Rothman, Welch and Walker (1969), consistent manner. Arbib and Hanscombe (1972), Handley and Bledsoe (1967/68)).

The literature reviewed in paragraph 2.9.2. enables a 2.9.3. philosophy to be adopted, but also creates the need to develop a methodology appropriate to the problem, concerned with the tenets of the chosen philosophy, and acceptable in terms of practical This problem has already been encountered in application. attitude measurement, but now needs to be reconsidered with respect to recording pedagogic styles. In order to achieve this end, relevant literature can be consulted to ascertain how similar problems have been solved by other researchers and hence to deduce an appropriate strategy for this research. The teaching-learning process is described by Fenker (1975) in terms of a simple communication model with four components; the sender, the encoding structure, the decoding structure and the Clearly, the 'pedagogic style' of the teacher destination. encompasses the first two aspects, but the teacher's behaviour (and hence his pedagogic style) will be affected by his knowledge In simple terms this states that in of the latter two. adopting a particular style, a teacher will either overtly or covertly assess his expectations of anticipated student Such relationships are identified in the work of responses. Weick (1968), Festinger (1957), and Shavelson (1972). Many useful techniques have been developed which fall within this area and taken into consideration the total cognitive process. A number of these are mentioned in a book entitled "Evaluating teaching in Higher Education" (U.T.M.L. 1976). Whilst this considers research conducted in Great Britain it does rather Bailey (1977) at Kansas State neglect American research. University developed a means of recording pedagogic style which did not impose criteria upon the teacher, but having used this

technique, he then adopted the emergent factors as criteria to be behaviourably applied in subsequent attempts to record pedagogic style:

2.9.4. Two techniques emerge from the literature as providing an appropriate means of recording pedagogic style.

These are Kelly's Repertory Grid and Osgood's Semantic differential. Frankly, close examination of some studies using one or other of these techniques would result in the reader being unable to decide which was being applied, as both systems have been extensively adopted, and often approach one another very closely. Largely for the reasons specified by Kelly (1955) the rep-grid technique would seem to be best suited to the demands of this project. A consideration of the rep grid technique is made in Chapter 6 and its use in education is discussed below.

The work of Tho mas at Brunel University has done much to extend the scope of the technique. The literature raises some rather important questions, the first of which is the applicability of repgrids to certain types of respondent. Children do not respond well to the technique and researchers, such as Cashdan and Philps (1978) using types of repgrid with young respondents have had to abandon the technique in favour of more 'traditional' free-flowing interview techniques. However, support for the use of repgrids can be obtained from the work of Applebee (1976) working at Goldsmiths College on the development of childrens response to repertory grids. The evidence shows the post 16

hh

year old students respond well to the technique.

A second question relates to the use of repgrids in the respondents' semantic competence in distinguishing the concepts 'same' and 'different.' It should be noted that Glucksberg (1976) has investigated this problem and concluded that the generally expressed concerns are unfounded. A contrary view is expressed in a book entitled 'Opposites.' No review of repertory grid literature could exclude Bannister and Mair (1969) who have provided a first and most useful handbook for the potential user of repertory grids, together with follow-up publications (e.g. Bannister & Fransella 1977).

Only one technique appears to have ever used repertory grids as a means of identifying and recording pedagogic style. This is the "Tuckman teacher feedback form" developed by Tuckman (1976). Kelly's theory of personal constructs explains the underlying Both teacher and observer use personal purpose of the TTFF. constructs, conceived as bipolar adjective pairs, to interpret or construe the reality of the classroom. Thus, using a set of relatively unstructured dimensions, observers can report that construction of reality to represent a picture of how the teacher is behaving, in other words the TTFF accepts Kelly's argument that reality is a construction of the observer and so presents a means of eliciting and describing an observer's Tuckman whilst starting construction of a teacher's behaviour. from a useful base ends up with specification of 28 behavioural characteristics which can then be directly applied.

- 45

The second international conference on Personal Construct Theory held at Oxford in 1977 included a variety of attempts to record pedagogic style. Apart from Pope (1977) and Keen (1977) all the other attempts used the repgrid as a means of producing criteria which may be behaviourally applied when appraising teachers. (See proceedings of 2nd. International Conference on Personal Contact Theory). The author has chosen to develop an application of the repertory grid which is not so constrained, and depicts the pedagogic style as perceived by the respondent, free from external criteria.

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND BEPERTORY GRID STUDY

THE RESEARCH HYPOTHESIS



- 3.1. Chapter 1 includes a comment of Claude Bernard, stating that, in his view, a hypothesis is the obligatory starting point for all experimental reasoning, Without it, no investigation would be possible and one would learn nothing; one could only pile up barren observations. To experiment without preconceived ideas, is to wander aimlessly. It is, however, important to be aware of what one's preconceived ideas are in order to avoid the injection of unconscious bias. In this chapter, it is the author's intention to refine the views expressed in paragraph 1.1. in order that they may be stated as specific hypotheses, which will be investigated in this research project.
- 3.2. One possible approach to this research would be to use an 'experimental' design where a direct relationship between the independent variable, pedagogic style, and the dependent variable, attitude, could be investigated, i.e. a change in the latter could be attributed to a change in the former. Although, in theory, such an experimental approach is possible, it is not seen by the author as appropriate for a variety of reasons. Perhaps the most important objection is the covert inference of a direction of causality which is by no means well established in the literature (Cronbach and Snow 1977). A scoond and practical problem is the nature of pedagogic style. By definition this is a multivariate factor of such a nature that it would be difficult to identify groups of respondents each exhibiting the same pedagogic style.

The chosen research approach, therefore, is a comparative study

conducted in the present dimension where respondents may be classified into groups by either an attitude score, over all factors, or by their score on each of four factors of their attitude to physics. Thus, although the attitude inventory developed is in fact multivariate (four factors), the contributing factors are easily identified for each respondent who will have a score on each. This is not the case with pedagogic style where the factors contributing to a respondent's perceived pedagogic style may be unique to him.

- 3.3. Pedagogic style may then be recorded for each individual respondent irrespective of whether they are students or teachers in one or more of the following forms. (It can be argued, and indeed is in Chapter 5, that even although a respondent may be devoid of all knowledge of 'teaching skills' he still has a perception of what a given teacher, or a hypothetical ideal teacher 'does.' This formal description of what a teacher 'does' is defined as a perceived pedagogic style).
 - The pedagogic style I believe I would adopt when teaching physics in a given setting.
 - ii) The pedagogic style I believe to be the most effective in the same teaching-learning setting as (i) above.
 - iii) The pedagogic style I believe would be the most ineffective in the same teaching-learning setting as
 (i) above.

The pedagogic styles identified can then be compared between respondents classified previously as similar by attitude.

Similarities and differences can be identified and quantified, and composite group profiles computed for "between-group" analysis. Details of the mechanism for such analysis is given in Chapters 4, 5 and 6.

3.3.1. It needs to be made explicit that the author considers the investigation to be of value in terms of providing a means by which practising teachers of physics in higher education may view the arena in which they operate. In his progression Q^{*} through the process of research the author adopted a methodology which in itself proved to be a mechanism for raising perceptual awareness and as a consequence this 'means' aimed to achieve in some measure the 'ends' specified for the research. Indeed, one of the many strengths of the repertory grid based methodology is its relative protection from 'unconscious bias.' Work by other researchers has shown that the intended consistent biasing of grid scores proves to be almost (Reid 1976). impossible to achieve. Respondents to a grid intending to elicit reasons for marital problems often attempt to conceal real contributing factors. Such respondents freely admit in feedback sessions that their attempts at concealment were fruitless. Whilst this relates to a conscious attempt at biasing, research evidence is available from the wide range of activities researched using repertory grid based methodologies that bias unconsciously introduced is made explicit by the analysis, and respondents are compelled to admit that the feedback data makes explicit facts they now realise to be true, but which had hitherto not been apparent to them, thus unconscious

bias and conscious bias, whilst possible, almost always becomes apparent from the analysis. Thus the methodology emerged as a strong instrument for achieving the research objectives as it collected data in the respondent's terms unclouded by unconscious bias and in itself raising perceptual Such was the emergent strength of awareness of the problem. the instrument developed, the author used it, almost unmodified, as a teaching appraisal tool called by the acronym TARGET. (Teaching Appraisal by Repertory Grid Elicitation Techniques). This is described in Chapter 9. (The author was the originator of the TARGET project and then operationalised the system as a part of his employment with Hopwood, who contributed to the national operations aspect). However, at this point, the author would wish to stress that this thesis describes an investigation of pre-specified hypotheses many of which are clearly 'newsworthy' (to use Poppers terminology) in that they do not reject the views popularly held by practising teachers of physics in higher education (see Chapter 9).

3.4. The research aims to provide a static picture of physics teaching in institutions of further and higher education in the period 1975 to 1979. Hence the use of the phrase 'in the present dimension' in paragraph 3.3. Furthermore, it will be possible to observe the perceptions of both teachers and students with respect to what pedagogic style represents the most effective approach in a given teaching-learning situation. As a critical comparative survey the research does not intend to prescribe changes in pedagogic style, but by identifying

possible areas of conflict between teachers and learners, the findings may assist practitioners in the field to perceive their teaching as others see it. In other terms it will make explicit those peer group values which so often remain undisclosed to those outside of the culture from which they were generated.

The major research aim can be specified, namely that of producing a research report relevant to, and accurately representative of physics teaching in FE/HE during the 1975 to 1979 period. The research problem therefore is to look in detail at the way in which different teachers of physics play their role in facilitating learning in their students, and to identify the ways in which different observers (teachers, physics teachers and physics students classified by attitude factors) perceive these pedagogic inputs.

The research is not intended to be prescriptive and so any conclusions may be open to the criticism 'so what?' The author would defend his stance by stating his personal intuitive belief, which, (due to the fact that no research has been undertaken in the area), is not supported by research evidence, that many teachers of physics are skilled practitioners with a wide repertoire of pedagogic skills. Their pedagogic style, does not, in his experience, often change once established, perhaps due to a lack of information regarding his student and peer perceptions of what he is doing. If, however, an analysis of the data collected in relation to the hypotheses enables any

practitioner to identify differing pedagogic preferences, it may well be within the competence of many teachers to explore the effects of changing their own pedagogic profile towards that perceived as more effective by their students who can be seen to exhibit similar attitudes to those in the data producing sample. It should be noted that a perceived effective pedagogic style may be LESS effective than one considered rather ineffectual, but if this is so (and it will depend on the way in which effectiveness is defined, viz. Sayer 1977 & Gilbert 1978) the teacher concerned will be able to question his strategies from a position of raised perceptual awareness with a communication channel available between himself and his students based on the overt statement of their differing perceptions of what constitutes an 'effective pedagogic style.' Irrespective of the way in which an effective pedagogic style is defined or interpreted by differing groups there may be compelling reasons for assigning students to learning groups by different criteria than academic potential (however measured) which is commonly used at the present time. Thus the hypotheses have been framed in such a way as to facilitate the production, after data analysis, of a descriptive report concerned with the questions raised in Chapter 1 and in relation to the foundation of knowledge extracted from the review of the literature in Methodological, and sampling strategies are Chapter 2. considered in detail in following chapters.

3.4.1. Cronbach and Snow (1977) in their ATI studies (Aptitude and Treatment Interaction) considered some similar problems from a rather different perspective. Their concern was matching

aptitudes to treatments in three ways, all assuming individual students and ignoring social-psychological factors operating at group level: capitalization of strengths; compensation and remediation. The author's approach differs in two major respects from the ATI approach Firstly, the group factors are not ignored, indeed they are promoted to a high level of significance as it is believed by the author that group interactions do effect learning and secondly, the emphasis is placed not on individual learning differences (which are acknowledged to exist) but on common group perceptions of effective pedagogy, thus demanding changing of pedagogic procedure on the part of practitioner/teacher in order to achieve increased effectiveness. The author does not, however, refute the importance of individualized learning and has expressed a triadic distribution not too dissimilar to that made by Cronbach and Snow (1977) in his paper with Reid (1976) entitled Guided Learning. His classifications were (a) remedial, (b) supportive and (c) extending, and this matches the Cronbach terms (a) 'remediation', (b) compensation' and (c) 'capitalization of strengths' respectively.

3.5. The hypotheses

3.5.1. Attitude to physics, measured by the attitude to physics inventory total score if appropriate, (and on each independent factor if composite scoring is inappropriate), will be more positive for teachers of undergraduate physics than for first year undergraduate students of physics.

- 3.5.2. Attitudes to physics, measured by the attitude to physics inventory total score if appropriate (and on each independent factor if composite scoring is inappropriate) will be more positive for first year undergraduate students of physics than for academic teaching staff from disciplines other than physics.
- 3.5.3. First year undergraduate students of physics will be more positively orientated towards the requirements of an examination, measured by Factor A (examination orientation) of the attitudes physics inventory, than will teachers of undergraduate physics.
- 3.5.4. Teachers of undergraduate physics will exhibit a more positive orientation towards practical work, intrinsic motivation and obtaining pleasure from physics instruction, measured by the three respective factors from the attitude to physics inventory, than will first year undergraduate students of physics.
- 3.5.5. There will be a significant difference between the positive attitude and negative attitude respondents of all three categories, and on all four factors, in the way in which they corporately perceive and categorise observed teaching acts on criteria associated with effectiveness of teaching, identified by a statistical comparison of element vectors and subjectively by the respondents' verbal constructions associated with the principal components in the construct space

of the appropriate respondent sub-group.

- 3.5.6. There will be a greater positive correlation between the perception and classification of observed teaching episodes by teachers of physics and students of physics when respondents with similar attitudes are compared than when differing attitude groups are compared.
- 3.5.7. The corporate perceptions and categorisations of observed teaching episodes, against effective teaching criteria, by each of the classes of respondent will show greater similarity between students of physics, than any similarity between any other two groups. The comparison will be made objectively by a statistical comparison of element vectors and, where significant relationships are seen to exist, by listing construct labels as used by each class of respondent respectively.
- 3.5.8. First year undergraduate students of physics formed into groups by their score on each component factor within the attitude to physics inventory (four factors) will exhibit a commonality of perception, when categorising observed teaching episodes against pedagogic effectiveness criteria, between common attitude groups across factors to a greater extent than between differing attitude groups within factors.
- 3.5.9. Positive attitude respondents from each component factor, will exhibit a greater 'acceptance range' measured by a lower rate of decline of Eigen values generated from a principal

component analysis of the respondent's grids with common element samples, than will the negative attitude comparison group.

- 3.5.10. Teachers of undergraduate physics will exhibit a greater
 'acceptance range' measured by a lower rate of decline of
 Eigen values generated from a principal component analysis of the respondents' grids with common element samples, than will the student category of respondents, who will themselves exhibit a greater acceptance range than teachers of subjects other than physics.
- 3.5.11. An appraisal of the constructs used by respondents in completing their grids will show that for teachers of undergraduate physics the ratio of discipline orientated constructs to pedagogy orientated constructs will be greater than for either of the other two groups of respondent.
- 3.5.12. First year undergraduate students of physics formed into learning groups by their total attitude score, measured by the attitude to physics inventory, will have common perceptions of the pedagogic style they associate with effective teaching of undergraduate physics with less variance than any mixed attitude grouping.
- 3.6. A full discussion of these hypotheses together with interpretive notes appears in Chapter 8. At first reading the hypotheses may appear to have tenuous links, they do, however,

fall into four basic categories connected as illustrated in the figure 3/1.

FIGURE 3/1 ABOUT HERE

Fig. 3/1 may be read from left to right and top to bottom simultaneously! Thus starting from the top box, the author realised that attitudes and their influence on the effectiveness of pedogogic practice gave rise to a variety of possible relationships. These in turn stimulated thought in four specific areas; (attitudes, factors of attitude, etc.) which by means of a feedback mechanism refined the four areas as shown in the diagram to facilitate the writing of specific hypotheses. Finally, the research identified certain new relationships as indicated in the right hand box.

FIGURE 3/1



PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND BEPERTORY GRID STUDY

THE ATTITUDE INVENTORY



- 4.1.1. Having delineated the research hypothesis in Chapter 3, and considered the range of possible methodology, it became imperative for the practical problems associated with the methodology to be confronted and solved if the research was to progress.
- 4.1.2. Notwithstanding the existence of a substantial body of literature describing repertory grid techniques, applied in a variety of settings, (Bannister and Mair 1969 et al) the research hypothesis demanded a sensitive and unique application of Kelly's original grid technique in order to identify the pedagogic profiles of effective and ineffective physics teachers, as defined and described by each class of respondent. In achieving this objective the author was compelled to reconsider repertory grid theory in order to ensure that both the application of the theory, and the subsequent analysis, provided outcomes which would prove to be an appropriate means of investigating the stated hypotheses. These outcomes are described in Chapter 7.
- 4.1.3. Unlike repgrids, which date their origin as recently as 1955 (Kelly 1955), attitude scaling has been an area of intense activity for psychologists for much longer. Such progress appears to have been made that the reader might anticipate the selection of an attitude inventory suitable for the research would prove relatively straightforward. Initially the author shared this view until a detailed analysis of the literature proved such optimism ill founded, for, upon closer examination,

the literature proved to be far from conclusive in its appraisal of existing instruments, many of which appeared to exhibit some weaknesses. Such a review of the literature associated with attitude measurement provided the evidence which rendered it necessary to develop a new attitude to physics inventory for the research. The need for the development of a new instrument is based on the criticism of other instruments in paragraph 4.2.

- 4.2.1. Investigation of the research hypotheses demands that two sets of people be identified from each category of respondent, consisting of persons with negative and positive attitudes to physics respectively. It follows, therefore, that whatever attitude inventory was used it had to be capable of identifying persons with negative attitudes, as well as persons with positive attitudes, towards the discipline of physics.
- 4.2.2. It might appear that any scale capable of identifying respondents who have a positive (suppose these respondents score high) must of necessity identify those with a negative attitude, (i.e. those respondents who score low on the scale). Whilst this <u>may</u> be the case, one cannot be certain that it is so, as the distinction is in fact only between respondents with a positive attitude to physics and those respondents who do not exhibit a positive attitude. It does not follow that those who do not have a positive attitude have, of necessity, a negative attitude. An example illustrating this point is the merchant navy cadet who is compelled to study physics in order to pass

his Second Mates Examination. Most students of this category fail to see the relevance of such a requirement, (perhaps due to inappropriate curriculum content or poor teaching), and would certainly fail to score highly on a scale where high scores indicate positive attitudes to physics. For most of these students, however, it would be equally unlikely for them to score highly on a scale where high scores indicate negative attitudes to physics. They exhibit a neutral attitude frequently summed up by comments like: "Physics is a necessary evil, I don't see why I need to do it, but I equally don't have anything against it." A scale designed to fulfil the requirements of this research would positively identify some persons who have a positive attitude to physics and at the same time positively identify some persons who have a negative attitude to physics.

4.2.3. Within the constraints imposed by the considerations described in the preceding chapters, the author was able to review existing measures of attitude, extending such an inroad into the literature beyond attitude to physics alone to include general attitude inventories and attitude to science scales. Clearly the adoption of an existing scale would have resource advantages, and so the review of the relevant literature was undertaken with the specific objective of identifying a scale suitable for adoption.

An appendix is attached reviewing existing inventories.

- 4.3. The task confronting the author, having generally reviewed attitude scaling and specifically attitude to physics inventories, was to select or design an instrument suitable for all three categories of subject from whom data was to be collected during the research. The scale must reflect the style and effectiveness of Gardners scale (i.e. not exhibit any of the three kinds of defect discussed in appendix 1) and demonstr ably have high validity and reliability. Needless to say it must enable an investigation of the hypotheses described as the basis of this research to be undertaken. Such a scale was devised as described in paragraph 4.4.
- The development of the attitude inventory reflected the 4.4.1. strategy adopted by Gardner in developing the PAI (Physics Attitude Inventory), namely the adoption of Thurstone type The development was based on Thurstone techniques techniques. after undertaking a three phase appraisal of what was required. Firstly, interviews were held with lecturers, students and other interested persons. These were recorded and later analysed to identify implicit or explicit objectives considered Secondly, these objectives were translated to be important. into clearly defined constructs. Thirdly, a check was instigated to ensure that the effects of any treatment being studied were reflected in the manner in which the instrument would measure outcomes. It followed, therefore, that before these specific techniques were considered some fundamental decisions had to be made relating to the factors to be

identified and the element sample included to measure these factors.

4.4.2. The factors to be investigated are both limited and defined by the hypotheses, which, it will be recalled, were based upon a subjective appraisal of the general relationship between students' and teachers' attitudes to physics and perceived pedagogic style.

Specifically four attitudinal factors are seen as important by the author having considered his preliminary interviews:

(a) Exam orientation. (A high score on this factor would identify respondents who believe that pedagogy should primarily aim towards the student group passing examinations rather than their learning physics for physics' sake).

(b) Practical bias. (A high score on this factor would indicate that the respondent believed that the pedagogic experience should result in learning based on demands alternated with the practical use of physics in the 'real world' rather than a purely theoretical basis).

(c) Intrinsic Motivation. (A high score on this factor would indicate that the respondent was intrinsically motivated to succeed at physics - for whatever reason. Such success may be achieved by passing an exam so that he need never study Physics again, or reaching personal achievement goals).

(d) Personal pleasure from physics. (A high score on this factor would indicate that the respondent would derive enjoyment from his exposure to the pedagogic experience in

physics, irrespective of performance criteria).

4.4.3. Having specified the attitudinal factors to be considered the author considered the feasibility of combining these factors to produce an overall attitude score. If the correlation between the factors proved to be high, and the reliability of each factor likewise proved to be high also, then a direct addition could be made to provide a fifth general attitude to physics score for each respondent. This analysis has been undertaken and is discussed in Chapter 5.

Such an instrument would need to be designed in order to identify both 'positively' and 'negatively' orientated respondents, as opposed to 'positively' and 'not positively' orientated respondents.

4.4.4. The element sample from which the instrument was constructed was extensive. Firstly, all of the eight inventories reviewed in detail were considered. Every individual element from each inventory was written on a card.

Twelve judges were then identified, consisting of six teachers of undergraduate physics and six first year undergraduate students. No attempt was made to make the selection of judges in any way representative, the contrary was in fact the case. Specific teachers known to be interested in the research were approached, and students were invited to volunteer in the knowledge that extensive work would be

required of them in their capacity as judges for Thurstone type inventory design. Having established the panel of judges they were invited to undertake their first task.

At meetings consisting of four judges (two from each category) the pile of cards were considered in turn, each for its relevance to undergraduate students and attitude to physics. If the four agreed the card was appropriately marked; if not the statement was either re-worded (on the same card) or rejected as irrelevant (and coded accordingly). At the end of the exercise the panels of four had viewed the original elements together with the amendments made by the previous All of the cards receiving unanimous approval meetings. were placed in a pool of 'potentially acceptable elements'; similarly, all those cards which were not unanimously accepted were withdrawn and destroyed. The remainder, (which it will be recalled may have been modified) were considered again at a single meeting (at which ten of the twelve judges attended) and again modified to an 'acceptable' form by the majority of those present, and placed on the pile of 'potentially acceptable elements,' or finally rejected and destroyed.

The author undertook an exercise to increase the number of 'potentially acceptable elements' by visiting one University and one polytechnic to address a group of first year undergraduate students from each institution. The author asked each group to undertake two tasks. Firstly, to list the five things they liked most about Physics and the five things they liked

least; and secondly to write a few paragraphs in a time limit
of ten minutes saying why they <u>either liked or disliked</u>
Physics.

From the completed scripts the author's wife, (who is not a Physicist and therefore unlikely to be biased in identifying and selecting attitude comments) carefully constructed a pack of cards, each one carrying an attitudinal comment which had appeared at least three times from the 124 scripts obtained. These cards were included in the pile of 'potentially acceptable elements.' This pile of cards consisted of 203 separate elements, each possibly related with attitude to physics.

The judges were then recalled individually and invited to post each card into one of five boxes. The first four boxes were labelled with the attitude factors listed in pawagraph 4.4.2. and the final box labelled 'inappropriate.' The instructions given to the judges demanded that they consider each card in turn as potential element in an inventory intended to measure one of the four attitude factors noted on the boxes.

After each 'sort' a code known only to the author (to preclude the possibility of other judges following previous 'sorts') was entered on each card to identify the box into which it was placed.

At the end of twelve 'sorts' all of those cards which had been posted into the same box ten or more times were selected as the

'core' element cards for each factor. Each remaining card was considered by the author, and a score assigned to it according to the consistency of sorting by the judges. The highest scoring elements were added to those already chosen. This produced four sets of elements; fourteen in factor A, (Exam Orientation): eleven in factor B, (Practical Orientation); twelve in factor C, (Internal Motivation); and eight in factor D (Pleasure from Physics). The score cut-off point was chosen in an arbitrary manner given that the score must be consistent across factors and that the minimum number of 'shortlisted' elements for each factor should be in excess of the number ultimately required for the instrument.

The precise number of elements to be included was determined according to a number of criteria. Clearly the instrument, which was to form only a part, albeit an essential part, of the administration procedure must be capable of completion in a reasonable time, say less than thirty minutes. An average time for completion was identified as twenty minutes in order to fulfil this requirement . Josephs (1973) had shown that great length and complexity was not a prerequisite of validity or reliability and six of the eight measures identified as representing acceptable design characteristics used as few as three elements to measure individual factors. A subjective decision was made by the author to identify a total of twentyfour elements using Thurstone techniques, six for each factor.

Each of the twelve judges was then invited to consider each

element in turn, decide whether agreement with the element constituted a positive or negative attitude to the factor, and to assign a score out of ten to indicate how well they considered the element measured the factor concerned in the direction to which they had assigned it.

Reduction from 45 to 24 elements was made firstly by rejecting all those elements where there was not 100% agreement on the direction of attitude by the twelve judges, and secondly, by selecting the six element cards with the highest composite score.

4.4.5. Having derived the content of the inventory, pilot testing was undertaken. The format chosen for the inventory was to code each of the 24 elements for source (from whence the element was derived), and for factor (each of the four factors) and for direction of polarity, (whether agreement indicated positive or negative attitude). The inventory was printed with the 24 element sequence randomly ordered, and for each factor set of elements two subsets of three were randomly identified to enable split half reliability testing to be undertaken. Likert type scoring was adopted with the respondent indicating 5, 4, 3, 2 or 1 for each element. In every case 5 indicated the strongest agreement.

Feedback from the piloting, conducted with a sample of twenty respondents <u>not</u> representatively selected, proved acceptable, (paragraph 4.4.6.) and with minor modifications to the rubric,
the final inventory was printed ready for use (Appendix 2 includes an example). Chapter 6 discusses the analysis of data collected from respondents.

4.4.6. Having designed the instrument to measure attitude to physics.its reliability and validity needed to be considered.

An indication of the reliability of the inventory could be obtained by use of the split half technique. The test elements were divided into two categories by random selection, with the proviso that each factor was equally represented in each 'half.' It would have been possible to administer the test to a group of respondents and obtain a numerical value for However, such an approach was rejected. the reliability. The objection to such an approach lay in the realm of sample A quoted reliability of better than + 0.87 might selection. look impressive, but unless the sample were chosen with the same degree of care, and within the same constraints as the research sample, then one could not assume that the same degree of reliability would be evident in the sample used for the main data collection. Thus when data was collected, a reliability coefficient was calculated using the split half method. If the reliability had been less than acceptable, any conclusions that may have been drawn would have had to be limited in their usefulness by the reliability of the test. Whilst this technique left the author with a certain amount of uncertainty in his mind until the analysis stage, it did ensure that any reliability quotient stated was relevant to the sample chosen

for the investigation.

Validity, however, needed to be considered in detail prior to test administration. The face validity of the new inventory was high. Not only did it look as though it measured what it should, (the refining techniques described included a face validity check) but it will be recalled that as the panel of judges were required to decide on the aptness (face validity) of each element, only those elements with high agreement between judges were used. Thus, face validity was supported by the design methodology.

The content validity likewise was ensured by the selection procedure. If any element received a wide band of responses from the judges it was rejected. Thus each element, and consequently the complete measure, is able to adequately measure four factors associated with attitude to physics. The validity of combining factors must await the correlation coefficients between factors; this is discussed in Chapter 5.

If, when the data was considered, the principal factors identified were to consist of the elements grouped together in the design, when results were factor analysed, evidence would support the decisions which had been taken. If these decisions were not supported by a factor analysis then the reliability would have to be recalculated from a different split half with the newly identified factors equally represented in each half.

The predictive validity is difficult to state without a test-In many respects the predictive validity retest format. However, if the content relates strongly to reliability. validity and reliability were to be acceptable then it is highly likely that a subject would respond similarly on two different occasions, (If his attitude in the interim period has not altered). In the context of this research the . predictive validity was of less import due to the fact that an experimental situation was not to be used, and the measure was intending only to ascertain the respondents' attitude to physics at one moment in time, and not to predict how his In the pilot attitude would alter given certain stimuli. study a very small set of four respondents were re-administered the inventory two weeks after the first session, the correlation was + .88).

Concurrent valility is often the only kind of validity considered by researchers. It represents the way any one technique correlates with another purporting to measure the same thing. The author's detailed consideration of the difficulties of content validity of other measures of attitude to physics, makes him unwilling to consider such a direct comparison worthwhile.

The construct validity of the measure was guaranteed by the use of a panel of judges to select the elements, whilst a 'belt and braces' safety precaution was built in by virtue of the fact that nearly all the elements had been elicited from other

measures of observable construct validity, or from the sample of subjects drawn from the constituents who would ultimately respond to the inventory.

The inventory was designed in such a way that it could 4.4.7. easily be completed by first year undergraduate students. The research hypothesis demanded, however, that it should be equally suitable for their lecturers. A similar strategy to that described in the preceding paragraphs was adopted to develop a parallel inventory for teachers of undergraduates. However, considerable difficulties arose, not the least of which was the limited number of members of that constituency which could be called upon to assist with the development. An alternative approach was tested which involved asking teachers to complete the inventory as they think they would have completed it when a student themselves. Follow up interviews with the teachers concerned supported the adoption of this technique and led the author to adopt this approach. The attendant advantage accrued from adopting this strategy was the possibility of directly comparing different groups. Chapter 5 pursues this application of the inventory in relation to the data obtained.

PEDAGOGIG STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND REPERTORY GRID STUDY

FACTORS ASSOCIATED WITH ATTITUDES TO PHYSICS: AN EVALUATION OF THE INSTRUMENT DESIGNED, AND THE METHOD BY WHICH SUB-GROUPS MAY BE IDENTIFIED FOR THE INVESTIGATION OF THE HYPOTHESES.



The research samples from three constituencies; teachers of 5.1. first year undergraduate physics students, first year undergraduate physics students and teachers (of undergraduates) who teach disciplines other than physics. An investigation of the hypotheses demanded that each constituency be appropriately sampled, and from each sample subsets of respondents be identified, each of which exhibited a common attitude to physics (or a factor of attitude to physics) in either a positive or a negative direction. The importance of the instrument designed to measure attitude is clear, for if its use provided unreliable or invalid common attitude groups, then no matter how clearly conclusions could be drawn from the data, the validity and reliability of the whole research findings would be limited by these considerations when applied to the attitude inventory.

The sampling from a universe of potential respondents was of equal importance: it was not adequate to merely sample randomly, for random sampling only ensures the absence of bias and does not ensure representativeness. The sampling strategy is discussed at length in Chapter 7, and so in this chapter, the discussion is restricted to the attitude inventory on the assumption that the sampling techniques utilized were adequate.

5.2.1. Reliability of the inventory is first considered. Appendix λ includes a table of every respondent's score on each factor of the inventory, together with a total score corresponding to the sum of the four factors. (The theoretical implications of such an addition are considered in paragraph 5.3.). The six

elements, identified by the Thur stone techniques described in Chapter 4, relating to each factor were randomly split into two groups of three in order to permit a split half reliability coefficient to be calculated. It will be recalled from Chapter 4 that all twenty-four elements were ordered on the response document in a random form.

5.2.2. The test-retest method of evaluating reliability yields information about the stability of rank orders of individuals A high correlation indicates that over a period of time. respondents have changed little in relation to the other members of the sample and that the test measures the same A low correlation, functions before and after the interval. however, may mean that individuals have changed in different directions, or in the same direction at different rates. Whilst changes in means and standard deviations may assist the researcher in deciding which kinds of systematic changes might be taking place, such a method places considerable demands on him to choose appropriate statistical strategies. Unless there are compelling practical reasons for knowing the stability of scores over a time period, this technique is less than ideal as a measure of test reliability. A further objection to the use of testerctest procedures in this research was the likelihood of attitude change occurring directly as a result of involvement in the project, as all respondents were required to complete a repertory grid, thereby exposing them to the likelihood of experiencing a raising of perceptual awareness and perhaps a consequential

change in attitude. Having rejected the use of a test-retest for these reasons, consideration was given to the value of using the split-half method, provision for which had been included in the inventory design.

5.2.3. The reliability of any instrument can be defined as the proportion of the variance exhibited by a set of measurements (in this case obtained by using the instrument with a given set of respondents) with the true variance. Such a reliability has restricted applicability in that a high value for reliability for one class of respondent does not imply that the test will remain equally reliable when used with a different class of respondent. Error theory, which is well documented in a variety of standard statistical handbooks, (Hays 1973: Thorndike & Hagen 1969) when applied to reliability as defined, leads to the conclusion that a correlation coefficient provides a most satisfactory means of measuring reliability. As the data are independent, then the adoption of Pearson r as a means of computing the coefficient Whilst it is interesting to calculate a is acceptable. reliability coefficient for this test as applied to all the respondents selected for this research, the coefficient so obtained is of limited value. The limit of applicability is determined by the different classes of respondent utilized in the study; the inventory must prove reliable for each group, (Teachers of Physics, students and other teachers), given that all three will exhibit different characteristics.

The results of the computation are tabulated in Table 5/1.

TABLE 5/1

. . . .

SPLIT-HALF RELIABILITY - FACTORS OF ATTITUDE TO

PHYSICS INVENTORY

CLASS OF RESPONDENT	N	SPLIT-HALF RELIABILITY COEFFICIENT	LEVEL OF SIGNIFICANCE
ALL RESPONDENTS COMBINED	69	+ .624	.001
TEACHERS OF UNDERGRADUATE PHYSICS	19	+ .609	.01
FIRST YEAR UNDERGRADUATE PHYSICS STUDENTS	33	+ .701	.001
UNDERGRADUATE TEACHERS OTHER THAN OF PHYSICS	. 17	+ .509	.05
	CLASS OF RESPONDENT ALL RESPONDENTS COMBINED TEACHERS OF UNDERGRADUATE PHYSICS FIRST YEAR UNDERGRADUATE PHYSICS STUDENTS UNDERGRADUATE TEACHERS OTHER THAN OF PHYSICS	CLASS OF RESPONDENTNALL RESPONDENTS COMBINED69TEACHERS OF UNDERGRADUATE PHYSICS19FIRST YEAR UNDERGRADUATE PHYSICS STUDENTS33UNDERGRADUATE PHYSICS STUDENTS33UNDERGRADUATE TEACHERS OTHER THAN OF PHYSICS17	CLASS OF RESPONDENTNSPLIT-HALF RELIABILITY COEFFICIENTALL RESPONDENTS COMBINED69+ .624TEACHERS OF UNDERGRADUATE PHYSICS19+ .609FIRST YEAR UNDERGRADUATE PHYSICS STUDENTS33+ .701UNDERGRADUATE PHYSICS33+ .701UNDERGRADUATE PHYSICS17+ .509

Line 1 is included for interest, but it must be noted that even such a high level of significance (better than .001 level) is meaningless except where all respondents are pooled: as such pooling is not used by the author, this value cannot be cited The final three lines in support of instrument reliability. The inventory proves are, however, of great value. acceptable for use with all three classes of respondent at a The implication of level of significance better than 0.05. these significance levels is best observed in terms of probability of error in producing specimen groups, i.e. there is less than 0.1% chance of a student of physics being classified into the 'wrong' attitude group, less than 1% for teachers of physics and less than 5% for other teachers. The research has specified the .05 level of significance (5%) in the grid aspects (see Chapter 6) and so an application of the same criteria renders the attitude inventory acceptable for all classes of respondent used in the research. (Note: The Spearman - Brown prophecy formula may be applied with a resultant <u>increase</u> in apparent reliability. The figures quoted, therefore, may be considered conservative in nature).

5.3. In calculating the values displayed in Table 5/1, the score on each half of the total inventory was used. Such computation is acceptable because no assumption is being made about the homogeneity of the inventory. The arguments of Chapter 4 demand that great care is exercised in determining whether the four factors of the inventory do measure some super-ordinate

construct (which may be called 'attitude to physics') but, as yet, no evidence has been presented to the reader allowing him to determine whether or not this is so. In the absence of such evidence, the test would appear to be homogenous, but if one or more of the factors is quite independent of the others, the inventory would be hetergeneous in nature. Such evidence can, however, be extracted from the data. Α correlation coefficient (Pearson r has been used) may be calculated for each factor compared with every other factor for each of the three classes of respondent. Such computations having been made, no significant differences were detected between the matrices for each class of respondent, so one correlation matrix will suffice to indicate all of the Table 5/2 indicates the values obtained when relationships. the scores for all 69 respondents were used in making the calculation.

TABLE 5/2 ABOUT HERE

The null hypothesis concerning homogeneity may be expressed as:

"There is no significant similarity in the factors measured by sub-sections A, B, C and D from the attitude to physics inventory."

TABLE 5/2

.

CORRELATION MATRIX SHOWING THE RELATIONSHIP BETWEEN FACTORS OF THE ATTITUDE TO PHYSICS INVENTORY

	Preference for examination orientation	Practical bias to Physics course teaching	Intrinsic Motivation	Pleasure from Physics
	Â	В	С	D
A	+ 1	178	+ .112	+ .231
В		+ 1	* 1 + .338	008
c			+ 1	* 2 + .503
D	-			+ 1

*1 Significant at 0.01 level

*2 Significant at 0.001 level

After Fisher (1965), critical values of r which have to be exceeded in order to reject the null hypotheses may be seen to have the following values for the three specified levels of significance:

Significant at 0.001 level r must exceed 0.380 Significant at 0.01 level r must exceed 0.310 Significant at 0.05 level r must exceed 0.235

Only two cells from Table 5/2 exceed these values, namely C and D at 0.001 level, and B and C at 0.001 level. One must conclude, therefore, that the attitude to physics inventory is an heterogenous scale consisting of four factors which cannot, with two exceptions, be considered as measuring the same concept. The exceptions indicate that, for all respondents considered together, the constructs identified as 'pleasure from physics' and 'intrinsic motivation' are likely to be construed as measuring some common, undefined, super-ordinate construct as are 'intrinsic motivation' and 'practical bias in physics course.'

Care must, however, be exercised as a specification of a second hypotheses, namely:

"There is a significant difference between the factors measured by sub-sections A, B, C and D from the attitude to physics inventory,"

fails to be supported in any cell of the matrix, indicating that, whilst the factors must measure independent constructs, these constructs are associated with one another by some quality which, whatever one choses to call it, is certainly not quantified by the direct addition of the scores obtained on each of the four factors. (The name one might assign to either the super-ordinate constructs or the linking quality will have implications for the validity of the scale, but remains irrelevant to a discussion on reliability).

5.4.1. The validity of the inventory cannot be considered in isolation, as much of the evidence provided in paragraph 5.3. is relevant to such a consideration.

The validity of the attitude inventory can be investigated by searching for the answer to one question, namely:

Is the inventory a valid means by which the research respondents may be categorized for further investigation?

As the answer to this question is crucial to further investigation of the hypothesis, evidence is presented in the following paragraphs which will lead to a justification of the use of the inventory for the identification of appropriate sets of respondents.

5.4.2. No test or inventory may be more valid than it is reliable, nor may be validity be expressed other than for a specified application of the test or inventory. These two limitations must be imposed upon this consideration of validity. Firstly, the following comments apply only to the inventory when used with the sample of respondents selected for this research and, secondly, the validity cannot quantitatively exceed the reliability quotients expressed in Table 5/2.

The weakest type of validity is face validity, which is a subjective qualitative assessment of whether or not the inventory measures what it purports to measure. Most respondents in this research were asked two questions in order to investigate face validity:

- (i) Do you think the attitude inventory you have just completed can measure your attitude to physics?
- (ii) Do you consider that your attitude to one or more of:
 - a) Examination orientation in physics teaching
 - b) Practical bias in physics teaching
 - c) Intrinsic motivation in physics teaching
 - d) Pleasure from physics teaching

could be at all relevant to how favourably disposed you might be to studying physics?

Of the 37 respondents asked these questions, 32 answered 'yes' to (ii), and 26 answered 'yes' to (i). One may conclude, therefore, that between 70% and 86% of all respondents considered the attitude inventory to exhibit face validity.

The questions were deliberately general in nature, as validity is not an either/or criterion, but a characteristic possessed by an inventory to a lesser or greater extent.

5.4.3. The next consideration was content validity. In common with face validity, being data-free, it may be conceptualized by considering it an estimate of the representativeness of the content of the inventory as a sample of the universe of possible content. Clearly such a definition demands that the author justifies the factors which have been included, as well as those which have been omitted, and this in turn relies on a detailed analysis of the literature resulting in a clearly specified rationale for the content. Such consideration has been described in Chapter 4 and so it remains only to ensure that the content validity remains satisfactory for each factor. The adoption of the Thurstone type techniques described in Chapter 5 provides the best guarantee of content validity, as the team of judges have to make individual and corporate decisions which must be both consistent and under-pinned by theory in order to provide elements for inclusion in the The author contends, therefore, that the inventory. theoretical rationale of Chapter 3, followed by the careful application of Thurstone techniques as described in Chapter 4, provides sufficient evidence to ensure the content validity of the inventory.

5.4.4. Construct validity is more difficult. Frequently researchers quantify construct validity (which is not datafree) by comparing respondents' performance on the test inventory with their performance on another instrument known to measure the same quality. (Often called concurrent Previously cited arguments precluded this validity). Firstly, had there been another instrument of strategy. acceptable validity the author would have used it in Secondly, evidence preference to developing a new inventory. has been provided to question the validity of some instruments formerly accepted as valid by earlier researchers (Chapter 4).

Alternative approaches are possible. The whole inventory may be factor analysed, and the element groupings thus obtained compared with the predicted element groupings. The proportions of variance of each element grouping can then be used to identify a rank ordering of factors. Ideally, such an approach leads itself to an homogeneous inventory or task, which, on the reliability evidence in paragraph 5.3., this inventory clearly is not!

The author had intended to utilize the factorial validity technique until the emerging heterogeneous nature of the inventory appeared to render such an approach inappropriate. Prior commit ment to a technique found to be inappropriate placed the researcher in a less than ideal position, as data relevant to the computation of construct validity was not available for most respondents. Fortunately, however, a

final group of ll respondents remained and each was asked to place a mark on each of four scales (corresponding to positive and negative attitudes) for each of the four factors included Of the eleven respondents, only four were in the inventory. eventually located in positive or negative sub-groups, but for every one of these four, on every factor for which they were identified as exhibiting a significantly strong attitude, perfect agreement was obtained between their actual responses on the inventory and their overall opinion expressed on the (A subsequent post-data collection study has factor scales. replicated these findings with an agreement in over 90% of Interesting though this may be, it cannot be cited as cases. evidence for construct validity in the research, as the sample was different, although intuitively one is able to accept findings based on a sample of four with a little more confidence in the knowledge that a subsequent study has replicated these results).

A final consideration of construct validity as relevant to this research relates to the manner in which the inventory is used. If an instrument purports to make fine discriminations between respondents (i.e. distinguishing between students' I.Q's. so that students may be assigned to eight I.Q. groups, each 10 points long in the range 60 to 140), then, of necessity, the construct validity must be very high. However, high construct validity is less crucial when crude discriminations are to be made, (i.e. To identify those students, in the previous example, who have significantly high or low I.Q.). In this study the author has designed an

inventory which appears, (on the basis of the very limited evidence from four respondents) to exhibit high construct validity, thus rendering it suitable for making fine discriminations. Yet, as paragraph 5.5. describes, the instrument is in fact used only to make crude distinctions identifying those respondents who exhibit significantly positive or negative attitudes on each factor.

- 5.4.5. The predictive validity of the instrument will have to await replication of the research, although some evidence is becoming available from an application of the methodology of this research (see Chapter 11) to suggest that the inventory can be used to identify certain groups, behaviour patterns for which may be predicted in the area of preferred pedogogic practices.
- 5.5. Having established that the inventory designed to measure certain factors of attitude to physics is both reliable and valid, the way in which the inventory is to be used must be considered. Chapter 7 contains a description of how the respondents were selected; this section describes how the performance by these respondents on the attitude inventory was considered, and how sub-groups of the respondents exhibiting common attitudes were identified.

5.5.1. Prior to the application of statistical techniques to the data in order to categorize respondents, certain checks have to be made to ensure that the statistical techniques chosen are appropriate. In order to simplify subsequent operations the data were considered to be normally distributed. Clearly, such an assumption needs to be justified. The author used the \mathcal{V}^2 Pearson goodness-of-fit test for each factor and for each class of respondent as a means of determining whether the group in question were distributed normally. The null hypothesis may be stated:

There is no difference between the sample score distribution under test and a normal distribution at the 5% level of significance.

In order to reject the null hypothesis, the computed value of \mathcal{V}^2 must exceed the tabulated value of \mathcal{V}^2 for a specified level of significance (0.05) and an appropriate degree of freedom.

Although the distribution of scores is continuous, it is necessary to think of the population as grouped into a finite number of distinct class intervals when applying Pearson \mathcal{J}^2 test. Furthermore, it is necessary for the expected number of respondents in each interval to be relatively large, certainly greater than, or equal to, five. It should be noted that an arrangement which specifies interval length in order to secure one common expected frequency of respondents refers to the

population distribution, (not the sample population) which is assumed to be normal under the null hypothesis, and that the choice of class intervals was made before the data are seen. The arrangement described is, however, quite arbitrary and any Whatever number be number of class intervals may be chosen. chosen the intervals will be of unequal size in order to give equal probability of expected frequencies per interval. It would be perfectly acceptable to select some arbitrary classinterval size in z score terms, and allow the probabilities to The method used exhibits two advantages: it be unequal. makes derivations from normality either in the middle or end of the range to be more easily detected as well as simplifying the computations.

The formula used is: $\mathcal{T}^2 = \frac{(0-E)^2}{E}$

with (J - 1 - x) degrees of freedom

where:

0 is observed frequency in interval j E is expected frequency in interval j J is number of intervals chosen x is the number of parameters estimated. (In this case M the mean, and I the standard deviation of the sample, are used as estimates of M & I for the population, and so x will always be two in this analysis).

The first stage is to compute means and standard deviations for each group of respondents on each factor. Table 5/3 contains the results of such calculations.

TABLE 5/3

The means and standard deviations of groups of respondents on each factor of the attitude to Physics inventory

		FACTOR A	FACTOR B	FACTOR C	FACTOR D
TEACHERS	·И	21.90	20.95	20 . 53	18.53
PHYSICS	6	2.85	4.38	3.70	2.74
STUDENTS OF PHYSICS	М	20.03	21.55	21.46	18.91
	6	3.59	4.02	4.62	4.74
TEACHERS NOT OF PHYSICS	М	19.82	21.12	17.88	17.29
	6	3.73	4.09	4.33	4.70

Within the limits imposed by the criteria of group size and number values of ψ^2 may be calculated. Table 5/4 indicates the critical values of ψ^2 and whether or not the null hypothesis is rejected.

TABLE 5/4

Cells marked with an asterisk represent group responses which may be considered to be distributed normally, i.e. the null hypothesis is rejected.

		FACTOR A	FACTOR B	FACTOR C	FACTOR D
PHYSICS TEACHERS	critical x ² d.f	2.65	2.65	2.65	2.65
N = 19 _.		4	4 *	4	4 *
PHYSICS STUDENTS	critical f ² d.f	6.3	6.3	6.3 3	6.3
N = 33] 	6	6	6	6
OTHER TEACHERS	lcritical I _f ² d.f	2.65	2.65	2.65 [.]	2.65
N = 17	3	4	4 *	4	4

The null hypotheses fails to be <u>rejected</u> in any cell when working at the .05 level of significance. Thus all the sample distributions may be considered normal.

One example calculation follows to illustrate the method adopted.

Respondent No.	Score on C.	Respondent No.	Score on C.
2018	21	4058	30
2020	18	4059	20
-3021	12	4060	25
3022	20	5061	23
3023	18	5062	18
3035	25	5063	16
3036	20	5064	18
4044	23	5065	26
4049	· 25	5066	26
4050	25	5067	11
4051	21	5068	19
4052	24	5069	21
4053	22	5070	15
4054	28	5071	22
4055	27	5072	16
4056	21	5073	30
4057	22		

Students responses to factor C.

= 21.46 MEAN 4.62 =

б

= 33 И

No. of groups J = 6 (i.e. average N₁ = 5.)

Degrees of freedom = 3

Value of ψ^2 to be exceeded in order to reject null hypothesis is = 7.8 (0.05 level of sig.)

Having decided (an arbitrary decision) that there will be six intervals, they may be diagrammed. (See fig. 5/5).

FIGURE 5/5

Z scores and their conversions



Intervals and frequencies table

Interval	Expected frequency	Observed frequency	$(0 - E)^2$	$\frac{(0 - E)^2}{E}$
	E	0		
0.00 - 16.840	5.5	. 5	• 25	.05
16.841 - 19.430	5.5	5	.25	•05
19.431 - 21.460	5.5	7	2.25	•41
21.461 - 23.490	5.5	5	.25	.05
23.491 - 26.080	5.5	6	.25	.05
26.081 – 👁	5.5	5	.25	.05
	N = 33	N = 33		E 0.66

 $x^2 = 0.66$ d.f = 3

In common with all other similar calculations, the value of X^2 falls short of the critical value and so the distribution may be considered normal.

5.5.2. Due to the nature of the repertory grid usage planned (Chapters 3 and 5 refer) only small groups of respondents are required who exhibit common, and strong attitudes to each factor. Four to six respondents in each class was considered to be an ideal group size. However, the same criteria must be applied to every class of respondent, and to permit cross group comparisons. Whichever criteria for selection were adopted, the numbers of respondents falling into each sub-

group would be expected to vary, particularly as N for each classification of respondent was different. Such variation may be minimised by basing the cut off points in the normal curve having first shown each group may be considered to be normal. An arbitrary decision to try 20% from each end of the distribution proved most acceptable. In fact the 21.19 and 57.62 percentiles were chosen (after Fisher 1972). Figure 5/6 indicates the percentile points with corresponding scores for each group on each factor together with the number (N) of cases identified as falling into the group thus formed.

FIGURE 5/6 ABOUT HERE

Appendix includes a table itemizing the actual respondents contained in each attitude grouping.

5.5.3. By adopting the techn ique described, the attitude inventory proved suitable as a means of identifying a total of twentyfour groups of respondents, each group exhibiting a strong (the strongest 20% of all respondents of the same class) pusitive or negative attitude to a factor attitude to physics. These groups became the comparison groups used, together with the repertory grid described in Chapter 6, to investigate the hypotheses previously specified.

FIGURE 5/6

The percentile cut off points and corresponding scores for each group of respondents on each attitude to physics factor.

G R O U P	F A C T O R	21.19 percentile 21.19%.of most negative attitude respondents		57162 percer 21.19% of posi atti responde	ntile most itive itude ents
		Negative attitude Range of scores	N	Positive attitude Range of scores	N_+
·	A	< 19.62	3	> 24.18	4
PHYSICS	В	< 17.45	5	> 24.45	4
TEACHERS	С	< 17.57	4	> 23.49	5
	D	< 16.34	5	> 20.72	4
	A	< 17.16	9	> 22.90	7
PHYSICS	В	< 18.33	7	> 24.77	9
STUDENTS	С	<17.76	5	> 25.16	6
	D	< 15.12	6	> 22.70	9
	Α	< 16.84	2	> 22.80	3
OTHER	В	< 17.85	3	> 24.39	5
TEACHERS	С	< 14.42	4	> 21.34	3
	D	< 13.53	4	> 21.05	3

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND REPERTORY GRID STUDY

THE REPERTORY GRID AS DEVELOPED FOR USE IN THIS RESEARCH



6.1. This chapter considers grid methodology from a user's perspective. Chapter four delineated the rationale for adopting a grid based methodology; in this chapter that foundation is built upon to illustrate how the author has used the grid as a means of gathering data relevant to the hypotheses.

Given the general description of the kind of data which may be elicited using repertory grids, (previous chapter), one can now direct attention to the way in which that task may be undertaken. For a description of a typical way (it will become evident that grids may be used in an infinite variety of ways) of using grids one can do no better than quote Kelly's own words.

6.2. "Methods of Assessing Personal Constructs"

"Perhaps the best place to start the discussion of methodology is with the description of a particular technique. Then, later, I can attempt to describe the broader methodology of which this technique is a particular example.

Suppose I were to give one of you a card and ask you to write on it the name of your mother. Then I would give you another and ask you to write the name of your father. On a third you might write the name of your wife, and on a fourth the name of the girl you almost married - but didn't! We could continue until you had as many as twenty or thirty cards, each showing the name of a person important in your life.

Then suppose I should select three of these cards, perhaps the ones of your father, your mother, and your boss or supervisor. Suppose I should ask you to think of some important way in which any two of them seem to be alike and in contrast to the third. What will you say?. Perhaps you will say that your mother and your boss have always seemed to know the answers to the questions you asked but that your father hesitated or told you to seek out your own answers. Now, if this is a distinction you can apply to your father, your mother, and your boss, can you extend it also to the other persons you have named? You probably can. The important fact is that as you apply it to person after person you are not only characterising those persons but you are also providing an operational definition of what you have in mind. Applied to enough persons this operational definition provides a more extensive definition of a particular channel of your thought than do words you may use to symbolise it.

Now, suppose I select another three cards, perhaps the ones with the names of your mother, your wife, and the girl you did not marry. What about them? Is there an important way in which two of them - any two - differ from the third? Perhaps you will respond immediately by saying that your wife and your mother are loving but that the girl you did not marry turned out to be harsh.

And how will you extend this personal construct to the other persons who are important in your life? Now let me suppose - for the sake of this discussion - something which I doubt would be true of anyone in this audience. Let me suppose that each person you characterise as 'loving' is a person you have previously characterised as ready to answer your questions, and each person you characterise as 'harsh' is one you previously characterised as sending you off to look for your own answers. Suppose this were true in case after case, on out to infinity. What could we say then? Would we then be ready to say that the two constructs were identical in everything but name?

Not quite! In our illustration the two constructs have been applied only to persons as whole entities. There is still the question of whether the constructs are applied identically to the separate acts of persons. To go even further, this suggests that, in general, the equivalence of constructs is determined by their similar application to all types of events, not merely to human events

Moreover, we need also to make sure that both constructs occupy exactly the same range of convenience. That is to say, can the first construct in my illustration - the response-rejection construct - be applied to all the events to which the second construct - the loving-harsh construct - can be applied; and, of course, vice versa? If there are some events that can be classified by the person as responsive or rejecting but which he cannot treat in terms of lovingness or harshness, then the range of convenience of the two constructs are different and the constructs themselves are therefore not quite the same.

All of this is a mathematical or logical problem and it leans to the formulation of one of the theorems underlying personal construct theory. Since, however, this paper is more concerned with the methodology of personal construct theory than with its mathematics, I shall limit myself to pointing out merely that such propositions exist.

Let us return to our deck of cards. We can represent the data produced so far in a flat matrix with events - in this case the names appearing on the cards - ranged along the top from left to right, and with the constructs ranged along the side from top to bottom. The entries in the matrix are single digit binary numbers, indicating simply whether the event is regarded one way or the other in terms For example, if you regarded your of the construct. mother as loving, this particular datum would be represented in the matrix by the numeral 'l' in the first cell of the second row - below 'mother' and opposite If you regarded your father as harsh the 'loving-harsh.' numeral '0' would be entered in the next cell, etc.

Now we may go on to expand the matrix until it is large enough to give us a stable idea of how the person construes his world. Starting with different triads of cards we can successively produce row after row of matrix entries.

This is not an interminable undertaking. Experience shows that only persons with the most complex or schizoid outlooks require more than twenty or thirty rows to express their repertory of constructs. Repertoires used in everyday affairs are generally quite limited, and, especially so it appears, among those who prefer to act rather than reflect.

As you can see, the matrix can be factor-analysed to see to what extent the person is employing a variety of constructs on only a few constructs masquerading under different names. We can examine the columns in the matrix to see which figures in his life are viewed as similar to others, or whether, indeed, there is any great variety perceived among them

But let us turn away from the particular kind of matrix we have described - which, after all, is only one example of the application of the methodology - and look at other kinds of personal construct matrices. Suppose, instead of asking you to write the name of a person on each of the cards I gave you, I would ask you to list an important experience you had had. Suppose, for example, I asked you to think of your wedding and make a note of it on the first On the second card you might note the occasion when card. you had a serious quarrel with your parents, on the third the time when you believed you were near death, on the fourth the ceremony at which you were awarded your university degree, then the meeting when a paper you presented was most severely criticised, and so on. .Then suppose you were to construe these events, three at a time as you did the persons in your life, and then extended the constructs to all the other events you had mentioned. This would generate another kind of matrix whose columns and rows, as well as its verbal content, cculd be analysed.

Some researchers have used the methodology to come to an understanding of how a young person, confronted with making a vocational choice, views the different occupations and professions open to him. Others have used it to analyse personal factors in job dissatisfaction. Some have studied changes in the construing process during a year of university training, and others have studied similar changes during psychotherapy." (Kelly, 1955)

6.3. The adaption of Kelly's method described in his own words in paragraph 6.2. into the format used by the author, falls into a number of separate areas. These are listed below. (The numbers in parenthesis indicate the paragraph number in which the point is elaborated).

Purpose and function of the grid	(6.4)
Size of grid	(6.5)
Element Sclection	(6.6)
Construct Selection	(6.7)
Scoring system	(6.8)
Analyses	(6.9)
Administration procedure, Reliability and	
Validity	(6.10)

6.4. In order to investigate the specified hypothese (Chapter 3), an instrument was required which permitted aspects of pedagogic practice to be identified and classified into 'teaching profiles,' for both individuals, and groups of respondents identified as exhibiting a common attitude to one or more factors of attitude to Physics. In deciding to adopt a methodology based on attitude scaling and the repertory grid, the author chose a cognitive stance which did not impose onto

the populations being studied any behaviourally specified pedagogic practices. The onus was, instead, placed on the representatives of those populations to generate their own criteria perceived as relevant to the classification of teaching. An example drawn from the medical profession might clarify the way in which this stance placed the onus on the respondent rather than the researcher.

A medical doctor, upon first meeting a potential patient in his consulting room, may adopt one of two strategies (or some amalgam of both). He might 'pass the time of day' with idle conversation whilst systematically applying tests to the client in order to ascertain what is wrong with him. Alternatively, he may ask the direct question 'what appears to be wrong with you?' In the latter case, the client is identified as the expert, for only he "knows" what is 'wrong,' the task for the doctor is then to translate a non-technical non-specific description into a class of illness to which a standard remedy may be prescribed. Clearly the manner in which the doctor in my example frames has questions will expedite diagnosis.

The author used the repertory grid as a systematic way of asking questions about physics teaching in general and specifically regarding perceived effectiveness of that teaching. The grid method can be used in order to provide composite 'pictures' of effective physics teaching as perceived by the classes of respondents identified in Chapter 5.

The particular decisions which had to be taken in order to use a grid based method to achieve this end are described in the following paragraphs.

6.5. A construct must not be confused with the verbal label which may be used to name it, some constructs may not be verbally symbolized whilst others may be very inadequately named by Provision must, however, be made for the respondent to words. write some verbal representation of his construct. The grid format used reflected common practice amongst many users of the technique by consisting of columns of elements and rows for constructs with a space mode available for the construct The size of grid is the first description in each row. variable needing to be established. There is evidence from Kelly (1955) and other researchers to suggest that respondents rarely need more than 20 spaces for construct. The author therefore adopted 20 as the number of constructs spaces (no. of rows) on his grid. In the subsequent data collection phase of the research many respondents found it difficult to progress beyond twelve to fifteen rows and there is evidence to suggest that those who did 'fill' the grid frequently used verbal descriptions for constructs towards the end of the grid which were on the surface quite different to those used earlier, but which correlated very highly with previously One can conclude, therefore, that in specified constructs. selecting 20 as the number of constructs permitted in the grid the author was unlikely to be limiting any individual respondent's exploration of his total component space with

1.05

respect to the element sample. There is less guidance from the literature in selecting the number of elements for inclusion. Clearly the elements need to representatively sample the construct space of each respondent (see paragraph 6.6), but equally important is the need to keep the grid completion time within reasonable limits. Such conflicting demands may be considered in relation to other research conducted using grids. An arbitrary choice of 20 elements seemed appropriate for a number of reasons. Such a number produced a symmetrical 20 x 20 grid which could be completed within a $2\frac{1}{2}$ hour time limit by the majority of respondents. The chances of 20 elements representatively sampling the component space of each respondent was high, and in any event a check could be made for each subject, and if the sample proved to be unrepresentative for him that data could be excluded from the final analysis. The grid format thus decided, the graphical layout could be finalised and grids printed. Appendix 2 includes a copy of the grid format used.

6.6.1. Adopting a grid based methodology permits considerable autonomy to the researcher in selecting the elements he proposes to use. Some examples of different classes of elements which have been used are: photographs of people (Bannister 1962), standing models (Salmon 1976), situations (Fransella 1972), occupations (Shubsachs 1975), and many others. The author, in his endeavour to identify aspects of pedagogy, decided to choose teachers, and as the pedagogic practice was to be specifically related to physics teaching, these teachers were to be of that discipline. (Subsequent
research has shown such discipline based concern to be less important than the author thought at that time: Keen & Hopwood Certain criteria must be applied to element selection 1978)). which limits the researcher's autonomy. If constructs related to physics teaching are to be elicited, then the elements must facilitate such elicitation by enabling discriminations to be made between triads selected from the element sample relevant No matter how carefully an element to pedagogic practice. sample is drawn from the universe of potential elements, there will be some elements which fall outside the range of convenience of some constructs used by some respondents. There must therefore be provision for any individual respondent to indicate when a construct is inapplicable to a specified element, and such provision was included in the grids used by the author (paragraph 6.8.).

The elements, as has already been implied, need to be representative of the universe from which they are drawn. Thus the twenty physics teachers identified as elements in the grid must be representative of the universe of physics teachers. This apparently difficult task can be undertaken in a variety of ways; the author adopted a strategy which selected representative elements in the manner described in paragraph 6.6.2.

6.6.2. The pedagogic practices which required identification by the grid were those related to perceived effectiveness. The first three elements were, therefore, easily specified as role

titles.

A. Self (me as a teacher of physics).

- B. The most effective teacher of physics I now know, or have ever known.
- C. The most ineffective teacher of physics I now know, or have ever known.

Both the 'effective' and 'ineffective' elements were required as a means of checking that pedagogic practices which distinguished between the self and the effective teacher were also criteria which were related to effectiveness. Evidence will be presentd later to show this to be an essential check, as it was not uncommon for subjects to specify constructs (such as good diction - poor diction) which, when applied, clearly distinguished themselves as being quite different from <u>either</u> the effective or ineffective elements, but which generated no apparent distinction between these latter two elements. (The implication being that diction is not a criterion which may be used to assess effectiveness; See Chapter 9).

Further role titles were then used to sample each respondent's experience of physics teaching. In order to generate these a group of lecturers were asked to nominate physics teachers who were memorable to them for any reason and to add to the list produced any teachers known to them and whose teaching is familiar to them. Twelve teachers undertook this task; all

the people named or described on all twelve lists could be classified within the following thirteen categories:-

1. Myself.

- 2. The teacher I now believe to be the most effective teacher of physics I now know or have ever known.
- 3. The teacher I now believe to be the most ineffective teacher of physics I know or have ever known.
- The most senior physics teacher in my own institution excluding myself.
- 5. A teacher who taught me physics at school.
- A teacher who taught me physics at a post school institution.
- Another physics teacher, (other than those listed), that I used to know or work with.
- 8. Yet another physics teacher, (other than those listed), that I used to know or work with.
- 9. My colleague A.
- 10. My colleague B.
- 11. Up to three nondescript
- 12. persons that could not
- 13. be otherwise classified.

These thirteen categories became thirteen elements in the proposed grid in the following slightly modified form:

- A. Myself (as a teacher of physics).
- B. The teacher I now believe to be the most <u>effective</u> teacher of physics I now know, or have ever known.
- C. The teacher I now believe to be the most <u>ineffective</u> teacher of physics I know or have ever known.
- D. The most senior physics teacher in my own institution excluding myself.
- E. Any teacher who taught me physics at school. (Excluding any people named on previous cards).
- F. Any teacher who taught me physics at any post school educational or industrial institution (excluding any people named on previous cards).
- G. Any other physics teacher whem I know or have known (excluding colleagues working in the same institution as myself and any teacher already named).
- H. As for G.
- I. My colleague (excluding any teacher already named).
- J. As for I.
- K. Add anyone, or leave blank.
- L. Add anyone, or leave blank.
- M. Add anyone, or leave blank.

These elements refer to the column headings on the grid referred to in paragraph 7.5. and illustrated in Appendix 2.

Adopting such an approach to role title element elicitation assists the user in his attempt to produce representative elements. However, representativeness has two levels of interpretation. The strategy described is likely to,

1.10

(although the reader has not been provided with any evidence that it will). produce a set of elements representative of the respondents construction of physics teaching. It is likely that differing respondents will have been exposed to quite different samples of physics teaching themselves - it would not, for instance, be unreasonable to expect a practising teacher and graduate of physics of many years experience, to have been exposed to a greater sample of the universe of physics teaching than a first year undergraduate reading physics. The element sample elicited by the thirteen role titles listed would therefore be quite different in the extent to which it may sample the total universe of physics teaching. Furthermore, each individual respondent will have named different persons for each role title, rendering any comparison between respondents guite inappropriate.

Thus to improve the element sample some mechanism was required which increased the representativeness of the sample in relation to the total universe of physics teaching and yet was common to all respondents, thus permitting comparisons between individuals and groups of individuals. Elements consisting of video-taped episodes of physics teachers in 'action', which could be viewed by all of the respondents, would achieve the required objective providing care was taken to ensure the episodes chosen were themselves covering a wide range of pedagogic styles.

A total of six hours of video-tape was made at a variety of

locations in Great Britain, in both Universities and Polytechnics. The teachers video-taped were initially approached by the author and given a guarantee of anonymity. Artificiality was minimised in a variety of ways. The author obtained a range of convenient dates and times to visit and video-tape episodes from each contact who had agreed to 'volunteer.' Filming was then undertaken by 'surprise' with no prior warning given. This approach proved costly, as timetables are amended rather more frequently than one might expect, and on a number of occasions the author arrived to find the arranged teaching episode to have been cancelled! Where video-taping could proceed, the author alone stayed in the room with portable television camera/recorder using normal lighting levels. This reduced the tension caused by technicians and lights, but introduced an inevitable loss in quality. Audio recording was facilitated by using a radic microphone carried by the teacher in his pocket and freeing him from static microphone positions or trailing leads. Finally, each session was observed from beginning to end (with times varying between one hour and four hours), yet only 20 minutes of film was taken of each encounter. This selectivity introduced editing subjectivity, but increased validity as neither the teacher nor the class knew when recording was actually happening, and in consequence normal relationships were usually maintained with the author being forgotten or ignored soon after commencement. Finally, at the end of each session the class were informed of the nature of the research, told how the video-tapes would be used and asked if any objected to being included. The usual guarantee of anonymity was given. No group objected, and in

consequence all the video-tape which was taken could be utilised.

In addition to the six hours of video-tape recorded specifically for the research, the author had met colleagues from other institutions who placed at his disposal nearly four hours of video-taped material relating to physics teaching. From the ten hours of video-taped material available, seven (to add to the 13 role title elements in order to reach the chosen 20 element sample size) episodes of such a duration that all seven could be viewed within a time limit of 30 minutes were required. The time limit was set partly to ensure all seven episodes could be seen consecutively on one 30 minute tape, but more importantly to limit the time for the total administration session (already 30 minutes attitude inventory plue 2 hours 30 minutes grid) to a maximum of three and a half hours.

Before describing the method used to produce the seven episodes, and to check their representativeness, a digression is required into grid theory to explain how representativeness may be tested.

In paragraph 6.10. the method of grid analysis is considered at length, but for the purpose of this exercise it is sufficient to consider what a construct system is, and how an element may be located within it. For the reader who requires a more thorough mathematical treatment paragraph 6.10. should now be read before continuing.

One can describe an individual's component space as an n - dimensional hyposphere in which all constructs and all elements have a unique location. An example may be used to illustrate the point:

Consider a hypothetical respondent called Fred, who, no matter what the topic of conversation, endeavours to relate it to his particular interest. We commonly say such an individual has a 'one track mind,' or, to use the parlance of grid technique, he could be considered uni-dimensional in his outlook. If such a character completed a grid where all the elements were people in his acquaintance, and all the constructs related to the conversations he had had with them, subsequent analysis might show only one component to be significant for him. Thus his n-dimensional hyposphere becomes a one dimensional component or, in pictorial terms, a straight line.

Figure 6/1 illustrates Fred's component space.

FIGURE 6/1

POSITIVELY POSITIVELY IRRELEVANT RELEVANT e b а TO MY TO 14Y îd 'SPORTS' 'SPORTS' INTEREST. INTEREST. 'ORIGIN' OR POINT OF TOTAL UNCERTAINTY.

If Fred is primarily interested in 'sport,' his component may be defined as shown in the figure. Every element (person) and every construct will have a unique location somewhere on the line.

Point 'b' may have been someone who plays golf and is classified by Fred as being at this position on the component in comparison with point 'a' which is Fred, i.e. where he places himself (another element) on the component. Point 'e' may have been a construct arising out of a conversation on cooking which Fred considers irrelevant to his unidimensional interest. Point 'd' corresponds to total uncertainty.

Fred's brother, Peter, is not unlike Fred. He and Fred share their sporting enthusiasm, but Peter has a two-dimensional component space illustrated in Figure 6/2.



One can look at the same elements and constructs in this new two-dimensional space.

Peter places himself (a) a long way from the origin and 45^o between the two components. He clearly perceives himself as both interesting and 'sporty', and he is very confident about this classification because the plot is a long way from the origin. The golfing person 'b' is situated 45^o between the components, but nearer the origin. Peter perceives him as 'sporty' and not interesting, but he's less certain (nearer the origin) than he was for himself.

The third person (d) he is uncertain about in relation to the 'sporty' component, as was Fred, but quite sure on the 'interesting' component. The converse is true for the construct on cooking (e).

The reader will perceive that the exercise may be continued into 3, 4, 5 or n components, although conceptual difficulty arises when more than three components have to be visualised.

Provided that the elements used in a given grid may be plotted onto the component map, the distance of the element from the origin will provide a good indication of how clearly it was constructed, whilst the distribution of plots will indicate representativeness. This describes the technique used to check the representativeness of the video elements used by the author. The technique is, however, not without flaw. The

element plot bears some strong resemblances to Heisenberg's Uncertainty Principle in Physics.

The Heisenberg uncertainty principle is concerned with the measurement of the velocity and location of a particle or The wave theory of light describes the motions of photon. photons statistically, but fails to specify in detail how a The wave theory of matter describes given photon will move. the probable location of particles, but fails to specify precise orbits of the kind described by Böhr. In simple terms Heisenberg states that we may, with precision, either know the location or velocity of a particle but not both. The Parallel the author wishes to draw is concerned with element locations Statistical computations may be made on in component space. grid data, and graphical plots of the kind illustrated in Figure 6/2 can be produced with ease. They clearly give the precise location of every element in the respondent's component space, but such precision necessitates the introduction of some uncertainty regarding how the observer interprets the If this uncertainty is reduced by respondent's element. lengthy, and detailed, follow up discussions with the respondent, then the plot ceases to be valid as the discussion will have refined the respondent's classification of the element and it will have moved either in his component space, or alternatively, his component space may be re-defined. Thus a researcher using grid methodology may either know the location of an element in a respondent's component space, or his construction of the element, but not both.

To return to the main argument, it must follow that, given the limitations perceived within the element plot technique, there remains the possibility of using these plots as a means of assessing representativeness. The author has used such a technique. It should be noted, however, that such a strategy improves the chances of selecting representative element samples but, for the reasons cited, cannot guarantee representativeness until the grids have been completed by the respondents and subsequently tested in the manner described.

The ten hours of video-tape was viewed many times by the author together with two colleagues skilled in the use of microteaching. (The author, together with Hopwood, developed a microteaching clinic at Plymouth Polytechnic, where video analysis of practiging teachers was used to assess and improve effective-From their observations, twenty-seven 'clips' were ness). identified as exhibiting specific but different teaching styles. These 27 'clips' became the elements of a grid which was completed by eleven volunteers (who were the same people who had been judges in the Thurstone process used to develop the attitude inventory). All the grids were placed end to end to form one mammoth grid 27 elements wide, and 189 constructs long. This grid was analysed and an element plot made. The resultant component space (which was a 3-dimensional space at the .05 level of significance) was then considered. Seven video 'clips' were selected from the sample of 27 such that they were spread as evenly as possible throughout the combined component space of the eleven respondents. In order to achieve this end

a rather cumbersome method was used. (Since then a rather more sophisticated mathematical analysis has been developed). The 3-dimensional component space was represented by a sphere. The projections of all the elements onto the surface was obtained from the 3 two-dimensional plots (Components 1 vs 2; 2 vs 3 and 3 vs 1) and pieces of paper glued onto the surface The surface was then divided into to represent each element. seven equal area divisions and one element selected from each division. Had any of the seven areas been devoid of elements, the experiment thus far would have had to be repeated, as the 27 original video-tapes would have proved devoid of certain In practice, every area had at least one pedagogic styles. element within it. Where more than one element was available, then the one which was construed with greatest certainty was (As decided by its actual distance from the origin). chosen. Thus seven video-tapes were chosen as the remaining elements for the grid design used in the research. It will be observed that the described method of selecting the video-tapes was based on the combined component space of cleven volunteer The actual representativeness could, therefore, respondents. not be guaranteed for all the respondents. A check was built into the analysis which computed the distance between elements for these seven specific elements. Where this proved the element sample to be unrepresentative for a given respondent, his grid was not included in the research data. Of 75 respondents, three had to be rejected against this criteria. Paragraph 6.9. explains how the video-tapes were used in eliciting constructs from respondents.

6.7.1. The instrument, if to be of value, must be useful; it must be able to be used to embrace the future as well as pigeon-hole The method must therefore elicit permeable the past. Permeable constructs are those to whose context constructs. new elements can be added; they are still being developed by the person as a means of organising and summarising new evidence or events. The first constraint on the design, therefore, is the need to elicit these permeable constructs. This, of course, precludes the possibility of using only supplied constructs, but does not preclude the possibility of having a mix of supplied and clicited constructs. It must be remembered also that repertory grid techniques are only an efficient and effective way of collecting the data that could otherwise be obtained by skilled interview. (Bannister and The same caution relating to applicability must Mair 1968). therefore apply, namely the applicability generally of one person's view, or indeed the inference that a man's view today will remain constant. Kelly's theory, however, is one which acknowledges the possibility of changes within an individual at a variety of levels. Within some parts of a person's construct system, minute-to-minute or day-to-day fluctuations and changes may be common, while other constructions may remain stable over longer spans, and some may show little fluctuation across a period of years.

> In the Rep. grid, the researcher cannot assume that the words used by the subject, to describe the construct discriminations he is making, mean much the same to the subject as they do to him. This is an assumption underlying all questionnaire and

rating scale approaches to data collection.

Sometimes a subject will provide the same word to specify one pole or another of constructs derived from different triads. The researcher has to assess whether the same word covers a slightly different construct, or whether the same total dimension is being repeated. In the former instance, the construct (including both emergent and implicit poles) would be retained; while in the latter instance, the subject would be asked to think of some other dimension of discrimination between the figures involved. One of Kelly's main concerns was with the importance of eliciting personally significant constructs; this must be considered when supplied constructs are being used rather than elicited constructs. Although it seems guite legitimate to supply constructs or rather labels which, for the subjects, may represent constructs (after all we do it every day when we hold a conversation), exclusive use of supplied words seems to negate the important ideas in construct theory noted in the individuality corollary (Kelly 1955). Interpretation of word labels used by respondents, or indeed an understanding of what they perceive as an interpretation or a supplied word label, can often be aided by reference to certain mathematical features of construct links demonstrated in the Where some constructs are supplied and others elicited, grid. then some understanding of the meaning ascribed to the

construct labels supplied by the examiner, in terms of the subject's system, can be gained from examining the mathematical relationships between supplied and elicited constructs.

This particular discussion can be summarised in the statement that grid methods do not assume that the subject means what the experimenter means by particular verbal labels involved in the test - on the contrary, the method is designed to help ascertain what the subject means by particular verbal labels. With grid method, it is possible to attempt an assessment of the particular and personal definitions which a subject attaches to the words he uses, by examining the element selections made on the basis of his constructs, and by quantifying the interrelationships between these constructs.

One of the original features of grid form, which has been ignored in more recent studies, is the insistence that the subject specify both poles of each construct. This problem particularly arises when constructs are supplied. While the supplied emergent pole of a construct may be translated by the subject into his own terms, the supplying of a contrast pole may saddle the subject with what, for him, are poles of two different constructs, causing him to work his way round a psychological corner when ranking or allotting his elements. It would seem advisable to check on the subject's view of the contrast when constructs are being supplied. It follows. therefore, that in conceptual analyses, single dimensions, as commonly used in grid format, may introduce unnecessary artificiality. Kelly was certainly aware of the importance of avoiding the limitations imposed by too concrete a use of his methods, but the tendency to use single-word descriptions of constructs is still common in grid work, although less so than was the case a few years ago.

1.22

In summary then, it would appear that there are advantages to be obtained from using a mixture of supplied and elicited constructs; however, when bi-polar constructs are provided the mis-match between the researcher and respondents' definition of the construct can lead to a loss of validity. Overcoming this problem by the provision of only the emergent pole of a construct seems a less than adequate solution.

In this study, the author has adopted a technique, which seems to have been little used previously. Constructs have not been provided but instead the respondent was given a list of construct labels which he may have glanced through if at any stage he found difficulty in formulating the pair and singleton from the triad. These supplied constructs were, of course, bi-polar but only one pole was stated, and it could have been either the emergent or implicit poles as the subject decided. The list, whilst not exhaustive, deliberately related to those aspects of pedagogic style often cited by educationalists as being important attributes of teaching. (The author always discussed with the respondent his "agreed" interpretation of each construct if requested. The "agreed" interpretation was based on the notes following each construct, but the respondent did not have a copy of these notes unless he specifically asked the administrator for an explanation or definition).

Experience gained in administering the grid to the seventyfive respondents in the research has indicated that for all but a few, such a list was not required until after the first ten

constructs had been elicited.

6.7.2. The minimum context card form of construct elicitation was used, each respondent being presented with triads of elements and asked to determine some important way in which two were alike, and yet, by the same token, different to the third. The triads were identified by shaded cells on the grid illustrated in the appendix. Whilst there are no rules appertaining to triad selection, a description of how the author designed the triad sequence is important.

The selection was partly systematic, partly intuitive and partly random! The stages were as follows.

FIGURE 6/3 ABOUT HERE

- a) Each row required a triad. There are 21 provided rows and so 63 cells were available for distribution.
- b) Logically each element should be contained in a triad the same number of times and so each element was allocated three cells.
- c) The three remaining cells were allocated to the 'myself' element to encourage respondents to consider constructs most relevent to themselves.
- d) Elements R, S and T were likely to be filled by fewer respondents and so one cell was subtracted from each and

FIGURE 6/3. A reduced copy of grid format with triads marked x.

.



allocated to F, I and J, being the three elements common to <u>all</u> twelve of the initial panel used in selecting the element descriptions.

- e) The final distribution became:-A B C D E F G H I J K L M N O P Q R S T 6 3 3 3 4 3 3 4 4 3 3 3 3 3 3 3 2 2 2
- f) In order to facilitate the generation of new constructs, no element should be used more than three times consecutively and preferably not more frequently than twice in three rows. The exceptions to this general rule is where between-element comparisons were to be encouraged (e.g. with the video-tapes) or where groups of elements could be identified as having similarities, or falling closely together, and which could therefore be used to force the respondent to make finer distinctions.
- g) Bearing in mind the comments in the previous six stages, the actual distribution was undertaken in a systematic manner with random changes of emphasis after rows 5, 12, 15 and 18.
- h) Finally, it should be emphasised that it really doesn't matter how the triads were selected as they are only used as a means of aiding construct formation. It would have been equally valid to either deal three cards at random, or to have let the respondent select <u>any</u> three cards from choice. The only advantages of the system adopted here arei) Each element is used on average three times.
 ii) The procedure is common to all respondents.

- 6.8. In the initial Kelly grid the subject provided his
 - discrimination between the elements in the first triad, then the administrator put a tick in the appropriate grid cells for the two elements which are subsumed by the emergent pole of the construct, (the emergent pole being the end which describes the pair) and leaves the cell relating to the singleton element He then asks the subject to look at the people named blank. on all the other cards and classify them as being one or other The administrator inserts ticks or end of his construct. If the subject finds it leaves blank as appropriate. impossible to place the element on either the emergent pole or the implicit pole, then the administrator inserts a cross. The choices for the second row are recorded in a similar fashion on the prepared grid matrix.

This whole process is then repeated a further 15 or 20 times, with groupings of three figures being used as contexts for the clicitation of constructs - constructs from the repertoire (hence the term repertory grid) which the subject uses to structure his inter-personal world. The examiner then has before him a grid of ticks and blanks which formalises and represents the intersection of various bi-polar construct dimensions with various external events (in this instance physics teachers) in the subject's life.

When this style of grid completion is used, a simple form of analysis, called matching scores, can be used. A matching score matrix is similar to a correlation matrix in which low

scores represent positive associations. One of the first variations in the method of grid completion originated from In the variation he suggested that, when a construct Kelly. dimension had been elicited from a subject, the person should then be allowed freedom to classify as many or as few of the other elements involved, as he saw fit, under each pole of the construct. This procedure permits the appearance of very lopsided constructs - constructs where very few elements are subsumed by one pole and many by the other. The problem with this is that two constructs may show a high matching with each other (matchings calculated by counting the numbers of ticks coinciding with ticks, and blanks with blanks in a comparison involving any two rows) suggesting a positive association between the two constructs. However, this matching may be misleading due to the lopsidedness of the response.

Because of the possibility of deriving distorted estimates of construct relationships when the subject is allowed complete freedom in his choice of elements (to fit one or other pole of any construct), and because of the mathematical difficultics involved in correcting these biases, Bannister (1959) suggested an alternative form for element allotment. In this form the subject is required to place half the element sample at the emergent pole of each construct and the other half at the implicit pole. This became known as the split-half form. (Not to be confused with any split-half reliability consideration which is not applicable to grid theory). As already stated, in Bannister's split half form, the subject is

required to place half of the elements at the emergent pole of each construct, and those which are left are supposedly Thus, constructs, which Kelly allotted to the implicit pole. assumes to be basically dichotomous, must be used in a scalar fashion, since the subject is required to grade his elements to achieve an equal apportionment. He must decide which elements (people) show the various characteristics specified by When the constructs are being each construct most markedly. elicited, this procedure for element allotment necessitates that some constructs be discarded, since some cannot be used readily in scalar fashion (e.g. male-female). This procedure is likely to eliminate grossly lopsided constructs. This prior selection of constructs seems beneficial from the point of view of gaining less distorted estimates of construct relationships, but may well rule out the possibility of exploring the various features of construct systems which may be associated with lopsided constructs.

One immediate effect of the split-half method (with 20 elements) is to restrict the range of matching scores to 11 points (as opposed to 21 in the original form) going up by steps of 2, from 0 to 20.

In a split-half grid with 20 elements, matching scores of 4 and 16 are significant at the 1 per cent level, while matchings of 6 and 14 have a probability of association at the 12 per cent level. For most practical purposes, matching scores between 8 and 12 can be regarded as within the chance range.

This variation demands that the subject makes decisions on a black/white basis and is not given the freedom to perceive shades of grey. As such demands are likely to colour responses it becomes difficult to justify the use of this variation.

A further variation which overcomes many of the problems inherent in the split-half method is the rank order form. Here, the subject is asked to use his constructs in a more overtly scalar fashion, and to rank the elements from the one which he considers shows the particular characteristic (indicated by the emergent pole label) most markedly to the one which shows it least (i.e. that which is most clearly characterised by the contrast pole). For example, he might rank the elements from the most intelligent to the most stupid.

Any rank order correlational method can then be used to estimate the degree of similarity in element placement on any two constructs. Since correlations are not linearly related, the raw correlation estimates cannot serve as scores. Construct relationship scores are calculated by squaring the correlations and multiplying by 100, thus providing an estimate of percentage variance between the two constructs represented by two grid rows. Here, significance levels are calculated in the usual manner for rank order correlations, and, again, the same significance limits (degrees of freedom equalling number of elements minus one) can be set to apply to the whole grid. Where n = 10, the 5% level of statistical significance (one-tail) is reached with a correlation of

rho = + or - 0.564. However, in grid work, there seems no good reason why the conventionally sacred levels of significance should be given particular importance, since it seems unlikely that people, when using their construct systems, only make decisions when the probabilities of errors are 5 in 100 or less!

This variation, whilst representing a distinct improvement over the former, still fails to produce an improvement over the original form which is significant enough to merit its adoption. The principle objection in terms of this research comes from the <u>demand</u> to rank the elements when it is conceivable that on some constructs there may well be a bias which should not be hidden by forced ranking. Whilst the respondent has the freedom to elect twenty equal first, it is highly unlikely that he would do so. Where this variation really scores is in its capability to greatly increase the available range of scores in any matching between element selections.

There are an infinite number of variations to scoring repertory grids; those described are perhaps the most widely used. The third and final one to be considered has been less used than either the split-half or rank order forms and yet, as will become evident, exhibits the greatest improvement in flexibility of the three! It is called the rating form. When this method is used for element allotment, the subject is asked to rate each element on each construct dimension. The subject may be asked to rate each element from "extremely" kind

to "extremely" cruel. If a seven point scale is used, the numbers 1 to 7 may arbitrarily index the amount of the characteristic attributed to each element. Alternatively, if only one pole of the construct is presented, the subject may be required to rate each element on a scale ranging from most kind to least kind. The mid-point rating can be used either as an average position, or (if appropriate instructions are given) as a rating to be given in cases where the construct dimension does not properly subsume the element, i.e. the element may be outside the range of convenience of that construct.

The rating method, on logical grounds at least, would seem to offer a number of advantages. The subject is allowed much of the freedom of Kelly's original method, in that he can nominate any number of elements he chooses for either pole of any He is given the opportunity of making distinctions construct. between people who, in the original methods, might receive only a uniform tick or an equally undifferentiated blank. Although some fineness of grading is thus introduced, if five-or-sevenpoint scales are used, the amount of differentiation required of the subject is not as great as that demanded by the ranking Furthermore, the subject may give the same rating to method. elements which might be artificially separated by the ranking method, where generally no ties are allowed, though there seems no good reason why the ranking method should be limited These two features together make it possible in this way. for a subject to place all the elements at one pole of a construct and yet differentiate between them.

1.32

As has been illustrated the problem of range of convenience can be specifically acknowledged and partially solved within this . Further, lopsided constructs need not be type of format. excluded from the grid examination. The range of possible scoring methods for this grid form is almost certainly greater than for any of the other forms already discussed. Matching scores of the type obtained in Kelly's original method can still be derived, and most correlational techniques can be used to provide relationship scores of the type obtained with the The aim throughout is to give the subject rank order method. as much freedom as possible to express his judgements and to throw the onus of formalising and quantifying on to post-test statistical processing.

Any size of rating scale may be acceptable, though five or seven point scales would seem to be most satisfactory. The primary difficulty with a seven point scale is getting respondents to use the whole of the scale; whilst 7, 4 and 1 are easily identified as the emergent, centre and implicit pole of the construct respectively, the apportioning of 6, 5, 3 and 2 is difficult. It is for this reason that the author adopted a five point scale

- 5. The emergent pole of the construct.
- 1. The implicit pole of the construct.

 The mid-point between emergent and implicit poles.
 and 4, easily identified as less extreme tendencies to one or other of the two poles.

0. The score inserted when the element cannot be included on the construct scale.

(The inclusion of 0 does not make this scale a 6 point scale as '0' implies that the scale is inappropriate).

It should also be noted that respondents will be encouraged to define the emergent and implicit poles differently, and only revert to the "black - not black" type of distinction when no other definition is clearly evident.

This can be illustrated by an example:-Suppose the triad consisted of:-

i) himself

ii) the person to whom he is responsible

iii) a person responsible to him.

Two possible constructs are

Either:

PAIR (emergent) (i) + (iii) because their primary function

is teaching. (Score 5 each).

SINGLETON (implicit pole) (ii) because his primary

function is not teaching. (Score 1).

Or:

PAIR (emergent pole) (i) + (iii) because their primary

function is teaching. (Score 5 each).

SINGLETON (implicit pole) (ii) because his primary

function is administrative. (Score 1).

In this case the latter would be a better construct, because not only does it define the similarity between the pair, it clearly shows why the singleton is different.

A variety of statistical techniques are available for the analysis of the 'rating variation' of grids; the method adopted for analysis in this research is described in paragraph 6.9.

6.9.1. A number of techniques are available for the analysis of grids. All the "cluster" approaches to the analysis of grid data have as a main aim the extraction of the more simple formal structure which is obscured by the detail of the original matrix - they yield a simpler picture, with the inherent virtues and vices of simplicity. A number of computer programmes are available which allow for the factorial analyses of the grids of individuals or the averaged grids of groups. Amongst these are programmes for principal component analysis. This analysis delineates significant orthogonal structure both of constructs in relation to elements, and of elements in relation to constructs.

In adopting a five point rating scale for the grids obtained in this research, the author included the facility to use zero as a sixth point for those elements which could not, in the view of the respondent, be classified on the construct. Binner (1958) and Gottesman (1962), in studying Kelly's notion of construct permeability, also encouraged subjects to mark a zero in the Rep test when neither emergent nor contrast pole

could be applied to a figure. They used the sum of all zeros in a grid as a measure of the permeability of the constructs involved. This method of element allotment is similar to that used by Fjeld and Landfield (1961) and Landfield (1967), but these writers, like Hess (1959), thought that the measures derived were more relevant to Kelly's concept of range of convenience than to permeability. The author argues, as Binner and Gottesman did, that when the zeros in the whole grid are totalled, it is indeed permeability which is being investigated; however, if only single constructs or small groups of convenience which can be estimated by the number of zeros.

Of the programmes available to analyse grids, the author selected a package developed by Patrick Slater with finance provided by the Medical Research Council (Slater 1972). This programme has weaknesses, some of which may be considered serious. Particularly weak is the adoption of the Pearson 'r' statistic to compute correlation coificients from data not wholly independent. The acceptance of a programme with perceived weaknesses is justified in a variety of ways. At the time the decision had to be made, alternatives exhibiting fewer weaknesses were not available. Since then programmes such as Pegasus (Thomas 1977) offer a more acceptable alternative. Resource limitations precluded the possibility of developing a new grid analysis package; and even if such an attempt had been made there could be no guarantee that the production did not exhibit weaknesses in other areas.

Finally, and perhaps most important, was the strength of the Slater package. Without doubt it did have weaknesses, but these were of less importance than may be implied from the criticism levelled at the package by researchers in the last This is exemplified by the experience gained when few years. the author presented a paper on the methodology to an international conference (Keen 1977). One delegate severely criticised the conclusions drawn on the grounds that the analysis, based on Slater's Ingrid package, was invalid. The critic requested the raw grid data which he ran against his own grid analysis package. There was no significant difference between either the element groupings nor construct loadings. Conclusions drawn from the alternative analysis were identical in every respect with those drawn by the author in his original paper.

6.9.2. The Slater grid analysis package offers seven major analytical programmes. The author used two only: called INGRID and PREFAN.

Grid data is obtained by the evaluation of a series of attributes. So that raw data is in the form of a matrix (a_{ij}) where any entry a_{ij} refers to the evaluation of stimulus j and attribute i. In all of the programmes the stimuli are referred to as elements and the attributes as constructs in common with the termology used throughout this thesis.

The first programme developed, and central to the whole package, is INGRID. This programme is an adaptation of a

standard principal components analysis, but, in addition, gives information about the relationships between the elements, between the constructs and their inter-relationships. It can be applied only to single grids.

A group of grids may be formed by one individual completing . grids on a number of occasions, or by a group of individuals The suitability of the programmes available completing grids. for such analyses depends upon the degree of alignment of the In this research the grids were aligned by element arids. (Only 10 of the 20 elements were but not by construct. common to all respondents and so a variable format card was used to select only the aligned elements when groups of grids The programme used for groups of grids is were compared). The appendix contains specific information called PREFAN. on the format in which data is presented.

These programmes can be used to analyse grids with not more than 30 elements or constructs per grid, so the 20 \times 20 format adopted by the author is acceptable for the package.

The programme will analyse any number of individual grids one after another in sequence. The elements may be ranked or graded in terms of the constructs. Grading may be dichotomous, i.e. on an all-or-none scale, as favoured by original form of Kelly grids; or on a 7-point scale, (as favoured by users of Osgood 'semantic differential') or on percentage scales, or on 5-point scales as in this research. Whatever the rating system adopted for evaluating the elements,

it should be maintained for all the constructs in the same grid, and all the constructs should include all the elements in their range of convenience. For instance, if ranking is used all the elements must be ranked in terms of each of the constructs. Ranking reduces to grading when ties are allowed; that is to say, grids with tied ranks are analysed as graded.

Each grid is introduced by a pilot card giving, among other information, the number of constructs and elements, the rating system used and the options selected for the output. So grids with different numbers, systems and options can be included in the same sequence. Principal component analysis is incorporated. There is a set of tables at the end of the output, defining the relationships among the elements, and between the constructs and the elements in terms of direction cosines mathematically equivalent to correlation coefficients. Thus all the associations among the constructs and the elements are expressed in comparable terms and can be assembled and presented in a single table.

The mean and the total variation, i.e. the sum of squares of deviations from the mean, are calculated for each construct; the grand total of the variation for all the constructs (V) is accumulated; and the percentage contributed by each construct to V is derived.

If the respondent is applying the same grading scale consistently with all the constructs, the means, totals and

percentages per construct will not differ greatly. Theoretically it may seem reasonable to suppose that the constructs in terms of which the respondent can discriminate better between the elements will be the ones which will have the larger totals and percentages; but in practice the ones with the larger totals may turn out to be those where the respondent discrimination is cruder. Elements may then be pushed out to one extreme of the construct scale or the other and no finer distinctions made - seen for instance as black or white with no intermediate shades of grey. Cross differences between means, totals and percentages are evidences that the same grading scale is <u>not</u> being applied consistently with all the constructs.

The researcher must decide whether to retain such differences between the constructs in the later stages of the analysis or to eliminate them by having the constructs normalised, i.e. rescaled so that they each have their total variation put equal to unity. The choice to normalise all scores, at this stage, in this research was based on piloting experience which indicated increased face validity when normalising was under-Respondents who used the grading scale consistently taken. for all their constructs differ from one another in the way they used it. Some avoided both extremes on every scale; some gravitated towards one pole to avoid the other; some favoured the extremes and avoided the midpoint. Such differences between respondents may be described by measures of bias and variability.

<u>Bias</u> increases when more elements are referred to one pole of a construct than the other. The difference between the mean for the construct and the midpoint of the grading scale measures the amount and direction of the bias in it. The direction cannot be treated as the same in all the constructs, since they distinguish between the elements in different ways, but the amount of the bias can be accumulated for all of them and expressed as a standard deviation.

A variance ratio may be calculated from this measure and the measure of variability, to test whether the amount of bias is significant. Using <u>b</u> for bias and <u>v</u> for variability

 $F = \underline{m} (\underline{b}/\underline{v})^2$

where <u>m</u> is the number of elements. F is entered in a table of variance ratios, taking its degrees of freedom as m and <u>n</u> (<u>m</u>-1) respectively, where n is the number of constructs.

The test may not throw much light on the psychological interpretation of the observation, as evidence from another closely related measure indicates that a significant degree of bias is normal.

<u>Variability</u> increases the more widely the elements are contrasted on the grading scale. It reaches its maximum when the elements are evenly balanced at the opposite poles. The amount that actually occurs in a grid is measured by the standard deviation of the grades about the construct means. If there is any construct in the grid where every element has been put in the same grade, it is discarded.

The breakdown of the total variation about the construct means, V, into its subtotals per construct is the first of many given in the print-out. A breakdown by element is given, as is a breakdown by component. Indeed, the results of every stage in the analysis serve directly or indirectly to provide breakdowns of V, by construct by component, by element by component, etc...

When ranking is used, all the constructs have the same mean, namely $(\underline{m} + 1)/2$ where \underline{m} is the number of elements; and the same variation, $(\underline{m}^3 - \underline{m})/12$. So there is no need for their values to be repeatedly listed for each construct and no individual differences in bias or variability are observable. The output for ranked data and the total variation about the construct means, V, is available together with the total per construct.

Before the analysis is carried any further, the original grid is replaced by a table of deviations from the construct means, D, which is not printed out. It may be visualised as a table with a row for every construct and a column for every element. The sum of the entries in every row is 0.0. The sum of their squares per row is displayed in the print out when ranking has been used or the option to normalise the constructs has been Otherwise the sum of squares for each construct, i.e. taken. for each row of D, remain unchanged. For the correlations the variances of the constructs must be normalised. Thus, geometrically speaking the constructs are all assigned locations at an equal distance from a common origin, and differ
only in being placed in different directions away from it. They lie on the surface of a hypersphere, and the difference between any two of them can be expressed as an angular or circumferential distance: the angle they subtend at the centre.

An angle of 0° corresponds with a correlation of + 1.0. It implies that the constructs are located at the same point on An angle of 90° corresponds with a the hypersphere. It implies that the constructs are correlation 0.0. independent of one another. An angle of 180° corresponds with a correlation of - 1.0. The two constructs are located diametrically opposite one another: one provides the same scale of measurement as the other, but in reverse. In some contexts it is an advantage to consider the angular distances between constructs rather than their correlations: the average of a set of angles is itself an angle, whereas the average of a set of correlations is not itself a correlation. So the angle corresponding to each correlation is printed out Such measurements can be used for comparing alongside it. grids. For instance it would be possible to compare the average angular distances between the constructs 'like me as I am' and 'like I would like to be' in the grids of respondents from different classes, without necessarily using a standard set of elements for every grid or keeping the other constructs the same.

The entries referring to each element in the table of deviations, i.e. the entries in each column of \underline{D} , are summed, and so are their squares. The results are listed. The

cumulative total of all the sums of squares is also printed out giving the value of V, if it was not given previously because ranking was used or if it was redefined by normalisation.

All the constructs have the same total variation under ranking or normalisation, namely V/\underline{n} , where \underline{n} is the number of constructs. The totals per element are also expressed as percentages of V, and listed with the element totals, opposite the element number.

The sums of squares for different elements may vary widely. Α small sum of squares implies that the informant's attitude towards the element is indifferent: he has rated it neither high nor low but near the mean on all the constructs. Conversely if the sum of squares is large the element must be an important one in the subject's construct system, whether his attitude towards it is consistently favourable or consistently unfavourable, or favourable in some respects and unfavourable in others. A large range of positive and negative quantities among the totals indicates a simple construct system where all the constructs tend to give convergent results; all running, for instance, from high to low along a common evaluation scale. But a narrow range does not necessarily indicate a much more complicated system: the constructs may all still relate to a common scale, only with some running in the opposite direction In either case, the elements will differ from to the others. one another mainly in one dimension, which will be the major Their distribution along the axis of the construct system. axis must always be balanced about a central point, but other-

1.44

wise it may take on any form: it is not at all unusual to find many elements clustered at one end of an axis balancing a few or only one at the other. In a more complicated construct system the elements may spread out from the central point in several directions, the less important individually remaining closest to the centre and the more important spreading further And the distance of an element from the centre is a away. function (the square root) of its sum of squares. If one salient element is sharply distinguished from the rest, the contrast between it and them may well form the most important axis in the construct system. For better or worse it sets the scale or standard according to which the rest are judged; thus trend-setting has been used to describe it. Where such a phenomenon occurs, comparing the elements will evidently provide a simple and clearer interpretation of the grids than comparing the constructs. However, in every grid both modes of interpretation are admissable and may often be combined with advantage.

The distances between pairs of elements are worth examining as well as their individual distances from the centre. It is here that evidence of clustering or isolation will be found. The expected distance between two elements drawn from a construct system at random can be shown to be the square root of (2V/(m - 1)). This quantity is displayed in the printout, and the observed distances between all possible pairs of elements are compared with it and displayed in a table of distances between elements. Observed distances expressed proportionately to the unit of expected distance will vary

about 1 for a lower limit at 0 to an upper limit at the square root of \underline{m} - 1. As this exceeds 2 when \underline{m} is over 5, one might expect their distribution to be skewed, and progressively more so as \underline{m} increases; but the effect is scarcely noticeable. The upper limit can only be reached when the entire grid is concerned with the contrast between two elements, leaving the rest in the middle.

Distances can be used for comparing grids in the same way that angular distances were formerly used. 'Myself as I am' and 'Myself as I would like to be' can, for instance, be used as elements instead of the corresponding constructs mentioned above. And then the average distance between the two elements could be used for comparing groups, without necessarily standardising all the other specifications of the experimental grids.

In the course of calculating the sums of squares for the elements and the distances between them, the squares and products of the deviations in <u>D</u> are summed by element to form an <u>m</u> by <u>m</u> covariance matrix <u>D'D</u>. It is of no direct interest for interpreting the grid, but it is technically important as a central part of the analysic. As well as being the source of some of the previously stated results, it is the matrix to which the principal component analysis is actually applied.

Like the grid from which it was obtained, the table of deviations, \underline{D} , has a row for every construct and a column for every element. The typical entry in it, say the one in row 1,

column J, is the difference between the rating (grade or rank) assigned to element J in terms of construct I, and the mean rating of all the elements in terms of I. If the grid shows the ratings of <u>m</u> elements in terms of <u>n</u> constructs, <u>D</u> is a table with <u>n</u> rows and <u>m</u> columns; and in every row under every column there is an entry, which is a number expressing the difference between the mean rating and the rating found for the element on the construct concerned. It is positive if the rating for the element is above the mean, negative if below. The sum of the entries in each row is 0.0., as explained carlier, and the sum of the squares of all the entries in all the rows is V.

In terms of Cartesian geometry, the column of entries for an element gives its location in a space where there is an axis for every construct, so the complete table defines the dispersion of the elements as a scatter of \underline{n} points in the construct-space, which has \underline{n} dimensions. The table can also be read by row: the entries for a construct locate it as a point in a space with an axis for every clement, so the complete table also defines the dispersion of the constructs as a scatter of \underline{n} points in an \underline{m} - dimensional element-space. The two views are strikingly different although they are built of the same data.

Principal component analysis is consistent with both views of the data. It provides a common co-ordinate system for the two dispersions and thus establishes the connection between the two techniques. Its most important advantage, however, is

different: the components form an ordered series, each accounting for an independent part of the total variation from the largest to the least. In this respect principal component analysis is unique. No axes, other than those of the components, can be used to analyse the total variation in this orderly way; any rotation of the axes sacrifices the advantage.

If the elements are given similar rating on a large number of the constructs, the main differences between them can be shown Their measurements on it can be found by on a single scale. adding their ratings on the constructs in certain proportions. The scale which shows the greatest amount of variation is the axis of the first component. The amount of variation shown on it is given by the LATENT ROOT, which is a sum of squares accounting for part of the total variation about the construct The proportions in which the ratings for an element means, V. on the constructs should be combined to obtain its measurement on the scale of the component are given by a set of coefficients, one for each of the constructs, and called the construct vector. The measurements themselves are listed as element loadings.

This is not the only way of considering the results. The importance of distinguishing between certain elements or groups of elements may govern the informant's choice and use of the constructs included in the grid. So it is just as reasonable to relate a component directly to the elements and define it in terms of an element vector, from which construct loadings can be derived.

A principal component is completely defined by its latent root, its construct vector and its element vector.

<u>The latent root</u> is a single numerical quantity which must be positive or zero; and the sum of the latent roots of the components is equal to V.

<u>The construct vector</u> is a set of coefficients, one for each of the constructs. It is listed in a column. The coefficients are normalised, that is to say, scaled so that the sum of their squares equals 1.0.

<u>The element vector</u>, similarly, is a normalised set of coefficients, one for each element. It is also listed in a column in the print-out, though there are contexts in which it needs to be treated as a row of numbers.

The element loadings for a component can be obtained from the element vector by multiplying the coefficients by the square root of the latent root; and multiplying the coefficients in the construct vector by the same quantity gives the construct Although vectors and loadings are related in this loadings. simple way they are both printed out as they are each particularly convenient for special purposes. If the first component accounted for all the variation in D the entry in row 1, column J would be exactly equal to the product of three terms: the coefficient of construct I given in the construct vector; the coefficient of element J given in the element vector; and the square root of the latent root of the So the differences between the observed and the component. computed values of the entries in \underline{D} , that is to say the

residual deviations, would all be 0.0. And the sum of the squares of the computed values would be exactly equal to the sum of squares of the observed values, namely V. In general, of course, the fit is not exact. Although the specifications of the first components are chosen to account for as much as possible of the observed variation in \underline{D} , some non-zero residuals are almost sure to be left. They form a table, $\underline{D}(1)$, with \underline{n} rows and \underline{m} columns as before, including any zero values. The sum of their squares, denoted V(1), is the residual amount of V not attributable to the first component.

A second component may then be computed to account for the residuals in $\underline{D}(1)$. Like the first, it is specified by an element vector, a construct vector and a latent root. Element loadings and construct loadings are also obtained by re-scaling the vectors so that the sums of their squares equal the latent root. The second component reduces V(1) as far as possible, namely to V(2), which may still not be zero. In that case a second table of residual deviations D(2), will still be left, including some non-zero entries; and the analysis will continue.

Three, four or a good many more components may be needed to complete an exhaustive analysis. But it is unusual to find much variation left in an individual grid after three components have been extracted. Perhaps because people see the world about them in a three-dimensional space their construct systems extending over the same range of convenience do not usually afford much more than three dimensions for the

relative evaluation of one set of elements. Osgood is deeply committed to a particular formulation of the opinion that evaluative systems are three-dimensional (Osgood 1963).

The total number of positive components, \underline{t} , is limited by the number of constructs and the number of elements in the grid. When the grid is replaced by the table of deviations from the construct means, D, the dispersion of the elements is balanced about its central point. If there are only two elements their relative positions will be defined by two points on a straight line with the central point midway between them. The relative positions of the points for three elements can be shown on a surface of two dimensions at most, and so on. Thus the maximum number of dimensions into which a dispersion of m elements can extend is m - 1. And on the other hand as the number of constructs is n, the construct-space cannot have more than n dimensions. So the dispersion of the elements in the construct-space cannot extend into more than \underline{n} dimensions, even if m is greater than n.

The construct vector of the major component specifies the dimension within this space where most of the variation between the elements occurs; the vector of the second, the dimension where most of the remaining variation occurs outside the dimension of the first; and so on. The amount of variation is given by the latent root. Dimensions where no variation occurs will be specified by components with zero latent roots. So \underline{t} , the number of components with positive latent roots, cannot exceed $\underline{m} - 1$ or \underline{n} , whichever is the less.

M. S. Bartlett has developed a test for principal component analyses in general, to decide whether the remaining variation after a given number of the major components have been extracted is scattered at random over the remaining dimensions. His test is applied to the data to determine the number of components significant at the .05 level. The results are not The test works backwards from the smallest always helpful. roots to the largest; and there may be some dimensions in a grid where the variation is restricted as well as some where it Indeed the two effects are likely to be is notably extensive. If there are a few exceptionally small roots the concomitant. test may indicate that all the larger ones are significant, and confront the researcher with a perplexing problem in interpretation. What needs to be considered may be why the variation along some axes is so small, and the explanation may be quite a simple one - the informant may have failed to distinguish between some of the constructs or some of the elements.

Starting with the largest roots provides an alternative approach. We may enquire whether the first component, first two, first three, etc., account for an unexpectedly large proportion of the total observed variation. This alternative has not been used by the author in conducting this research.

The dimensionality of a grid, as illustrated, cannot exceed <u>n</u> or <u>m</u> - 1, whichever is the less. It may be further reduced, for instance, if the informant has given identical ratings to two of the elements, or if one of the constructs has been cut

out because he has rated all the elements the same on it. If, for a given grid there are t dimensions (components) then the complete set of <u>t</u> latent roots is listed in order of magnitude. Their numerical values and their proportionate size as percentages of the total observed variation, V, are also given. The results of the Bartlett test follow in the printout. Each successive application of the Bartlett test is followed by the appropriate value of chi-squared, with the appropriate degree of freedom.

The test does not apply to the first component. Having identified that a given number of components may be considered significant at the prescribed level, certain detailed information is listed in the printout for each component.

Loadings have great interest from many perspectives. The total variation of a component, that is to say its latent root, is the sum of the squares of its element loadings, and also the sum of the squares of its construct loadings. It can be analysed in both ways: either into the amounts due to constructs 1, 2, <u>n</u>, or into the amount due to elements 1, 2, <u>m</u>; and both alternatives are equally valid. A component is in fact a measurement of one way in which the constructs and the elements interact: the way in which it concerns the constructs may be easier to understand that the way in which it applies to the elements, or vice versa; but both need to be considered for a complete interpretation.

The sum of squares for an element is the sum of squares of its

component loadings, and it can be analysed in this way. The residuals listed in the print-out simplify such an analysis. The residual from the first component can be compared with the original sum of squares; the residual from the second with the residual from the first, and so on. A large proportional drop indicates that the evaluation of the element is largely in terms of the component concerned. Similarly the variation of a construct, is the sum of squares of its component loadings, and can be analysed as such. Again the residuals in the print-out simplify the breakdown. A large drop from one component to the next shows that evaluation in terms of the constructs tends to coincide (positively or negatively) with evaluation in terms of the component. Leadings can also be used for representing the results of a grid graphically. Although there are of course a limitless number of ways in which that can be dene, two deserve special consideration as exact geometrical equivalents, the first has already been considered as a means of checking the representativeness of an element sample.

The first is the dispersion of the elements in the component space within the construct space. As already explained, the entries in the same column of \underline{D} , which all refer to the same element, together specify a point for the element in a space where the reference axes are defined by the constructs. The entire array of numbers in \underline{D} , corresponds geometrically with a dispersion of \underline{m} points in this space, which is \underline{n} -dimensional. The distances between the points in it are the distances

between the elements. The centre of the scatter of the elements in this space is made zero when the original grid is Its dispersion is not equally wide in all replaced by D. directions within the construct-space. The construct vector of the first component specifies the axis of the dimension where the variation between the elements is widest. The constructs with the highest positive coefficients contribute most to defining it positively; the ones with the highest negative values to defining it in the opposite direction. (Or both directions may be defined by the emergent and the latent poles of the same construct). The orthogonal, i.e. independent dimensions with successively smaller variation are defined by the construct vectors of the successive components similarly. Altogether the components define a sub-space of not more than t dimensions within the construct-space. The locations of the elements in this component-space are given by their component loadings; and their dispersion can be mapped conveniently in two dimensions at a time. To make a map one sets out two lines at right angles on an ordinary piece of graph paper to represent the axes of two components. Use their construct vectors to characterise them in their positive and negative senses. Next one plots the position of the It is natural to consider the elements from their loadings. two major dimensions first. As the map cannot show how far the scatter extends outside them, generally it can only be How close the approximation is depends on how approximate. much of the total variation is attributable to the first two (See figs. 6/1 and 6/2 for example). This is components. frequently as high as 90%. If the distance between two

elements given in (D) is much greater than the distance mapped, it must occur in dimensions defined by the construct vectors of one or more of the components not on the map. So it must show up in some other map, constructed in the same way, referring to some pair of components in the construct-space. Although it is natural to consider the major components first, components which show relatively small amounts of variation are not necessarily There must be some reasonable explanation devoid of interest. if variation is particularly restricted along some dimension in It may not be at all difficult to detect. the construct-space. Two constructs may be effectively the same or almost the same, e.g. the ratings of the elements may differ scarcely if at all in terms of two constructs such as "Good bad" and "Like ... Or two elements, e.g. "Hyself as I am" and ... dislike." "Hyself as I would like to be" may be virtually indistinguishable in terms of the constructs used in the grid. In either case, little or no use may have been made of one of the dimensions available in the grid for distinguishing between the constructs or the elements. Why the informant fails to make such a distinction may be a question of interest. The most useful results to examine in search of explanation, apart from the graphs described already, are the table of correlations between the constructs, and the table of distances between the elements. Sometimes the explanation may be more obscure and involved; and sometimes practical considerations What is unjustifiable is a may make it not worth seeking. general assumption that dimensions where variation is small can have no interest.

The second geometrical equivalent for the numerical entries on the grid, which can be shown by graph, is the dispersion of the constructs in the component-space within the element-space. The entries in one row of a grid, or of \underline{D} , can be treated like Taken together, all the entries the entries in one column. referring to one construct specify a point for it in a space where the reference axes are defined by the clements, or, in The entire array of numbers other words, the element-space. in D corresponds geometrically with a dispersion of n points in this space, which is m-dimensional. The variances of the constructs in a grid always tend to be the same. They must be exactly the same if the elements are ranked in terms of the constructs, or graded dichotomously with a 50:50 split: they are equated if gradings of any other kind are used and normalised; and they still tend to be the same for other, nonnormalised gradings, provided the same grading procedure is used consistently throughout the grid. Thus the points for the n constructs in the element-space are all either exactly or approximately equidistant from a central point, and consequently must lie on or near the surface of a hypersphere. Its dimensionality depends on the number of components with positive latent roots; that is to say, it occupies a tdimensional component-space within the m-dimensional element-Constructs with high loadings on the first component space. only will cluster together in one region on the hypersurface if their loadings are positive, and in the diametrically opposite Such distributions region if their loadings are negative. are easy to examine when they are represented geographically. If the positive pole for the first component is set on the

equator of a geographical globe at longitude 0° , its negative pole will be on the equator where the international date-line, longitude $180^{\circ}E$ or W, intersects it. Constructs with high loadings on the second component only will be located on the surface around points 90° away from the two poles of the first component, e.g. on the equator at longitude $90^{\circ}E$ for positive loadings and longitude $90^{\circ}W$ for negative loadings. Constructs with high loadings on the third component only must then be located around the points 90° away from those already occupied, e.g. at the North pole for positive loadings and the South pole for negative ones.

The geographical model can only include three components of course; but if the first three components account for over 80 per cent of the total variation in grids of the customary size, the model for them will usually give a very useful indication of the relationships between the constructs.

Polar co-ordinates of the constructs can be computed and these are displayed on the printout as H, V and R. The values of H and V (horizontal and vertical measurements in degrees) are the ones to be used for plotting the positions of the constructs; and the convention intended for adoption is the one already described. A point on the equator is selected as the origin, with $H = 0^{\circ}$ and $V = 0^{\circ}$. Positive values of H are reached by moving to the right, i.e. eastwards around the globe for the given number of degrees, and negative values by moving westwards. Positive values of V are reached by moving upwards, i.e. northwards, negative southwards. The radial measurement

R. is not used in map-making. To be precise, a construct can only be located on the surface of the sphere if its value of ${\ensuremath{\mathsf{R}}}$ Otherwise, it should be located beneath the surface, is 1.00. proportionately nearer the larger its R. R. has other It defines the multiple correlation interesting properties. between the construct and the first three components; and R^2 is the proportion of its total variation they account for. Α construct with a very small R could perhaps be treated as if it had disappeared beneath the surface completely, for most of its variation must occur in dimensions unmapped. (To refer to an earlier argument, it approaches the point of total uncertainty, i.e. the origin). The advantages of using an actual geographical globe for plotting the relationships between the constructs are very great, as it shows which points are diametrically opposite one another. All the results obtainable by any method of rotation in three dimensions can be seen by picking the globe up and turning it around in one's fingers.

There is no necessary relationship between the number of clusters found in the dispersion of the constructs and the number of components found in the analysis. The description given for explaining the use of the polar co-ordinates, above, is highly schematic. Clusters are seldom, if ever, neatly separated by angles of 90° . One rather loose cluster extending into many dimensions is sometimes all there is to be found; if so, it is just as likely to be around the point H = $+ \text{ or } - 180^{\circ}$ and V = 0° as around H = 0° and V = 0° . But sometimes, neither region is occupied: they may both be relatively empty spaces around one of which, at least, several

small, perhaps trailing groups of constructs are gathered. Among grids in general, the absence of any simple regularities When they both refer to the same components, a is perpetual. map of the dispersion of the elements in the construct-space and one of the constructs in the element- space are two different views of the same variation, for the variation of a component in the construct-space is identical with its variation So if suitable conventions are adopted, in the element-space. either map may be projected onto the other. But it would be very misleading just to superimpose one onto the other, and map the constructs and the elements together as n + m points in a common component-space. The constructs should be represented as direction-lines in the component space; and the elements as direction-lines in the component space within the element-space where the constructs appear as points. As the direction of a line from an origin can be indicated by a single point at any distance from it, the difficulty can be overcome; but the convention being adopted must be clearly understood.

Such composite map: take advantage of a unique property of principal components, that they provide a stationery coordinate system in both spaces, allowing the two dispersions to be aligned. They may reveal interesting and unexpected relationships between elements and coincident or diametrically opposed constructs. Diametrically opposite the point for a construct, another point can be marked on the globe for its contrast. And since the projection for an element should be regarded as a hypothetical construct referring to it rather than as the element itself, it too may be regarded as bipolar

and represented by two points, the pro-element and the antielement. To obtain the polar co-ordinates for the opposite pole of a construct or a pro-element; when H is positive, one can subtract 180°; when H is negative one can add 180°; and change the sign of V.

For the reasons explained earlier, the relationships between the constructs and the elements and of the elements with one another can all be expressed in terms of direction cosines. These are mathematically equivalent to correlation coefficients and serve like them to describe how closely two variables are associated. The printout concludes with four tables listing results of this kind. All these measurements refer to the relationships of the elements and the constructs with one another in the whole component-space, and enable a mathematical model to be used in place of the physical model proviously described.

- As is clearly evident, not only is rep grid methodology a powerful tool, but the analytical techniques available assume maximum advantage from the data format.
- 6.10.1. The reliability and validity of grids is a special consideration quite different from the consideration given to these concepts in Chapter 5. In many contexts, reliability estimates will themselves be a test of the validity of particular grid measures; and it may often be the case that particular validities hinge on the finding of <u>low</u> reliability. (This was recognised by Cronbach and Heehl (1955) in their

discussion of construct validity). In terms of grid method, people who respond similarly to the same "stimuli", or elements, will be considered similar in that respect, i.e. in relation to their element allotment and construct relationships in that They will not necessarily be considered particular area. alike in other areas of construing, as yet unexplored. If two people respond differently to a set of elements in terms of a number of constructs, they need not necessarily be considered as different people as regards their construct system for In construct theory terms, dealing with these elements. similarity between persons is seen in terms of similarity of the constructions they have placed upon experience, not in terms of similarity of the experience they have undergone. This conception is embodied in grid measurement, in that it is possible for two people to allot the same elements quite differently to the poles of the same constructs; yet when the grid is analysed they may be shown to be operating very similar construct systems.

The data provided Ly 2 grid are in the form of a matrix (of correlations or matching scores), which is assumed to represent the pattern of relationships between the constructs used in the grid, and to be an index of the network of implications between the obtained sample of constructs for the individual subject. This matrix of relationships (or any specified part of it) can itself be the object of "reliability" assessment. Thus, for any given subject, a test-retest coefficient can be calculated, by rank ordering each of the matrices, which his grid

performance has yielded on two separate occasions, from the highest positive relationship through zero to the highest negative relationship, and then calculating the Spearman rho Such a rho is, in one sense, between these two rank orders. an index of factorial similarity between the two matrices, and represents an estimate of the degree to which the pattern of the construct inter-relationships of the subject has remained Again, in terms of such a stable across the time interval. derived measure, no general estimate of reliability can be given, because, as will be shown, it will be grossly influenced by factors such as the elements and constructs used, individual and group differences, methods of administration, and However, as a kind of statistical platitude, it can so forth. be said that by using elements such as people known personally to the subject, with supplied constructs of a conventional type and with either a rank order of split-half matching administration, normal subjects, doing repeat grids, on either the same or different elements, tend to yield coefficients of reliability which fall largely within the range 0.6 to 0.8. Field and Landfield (1961) concluded that (i) given the same elements, the subjects, after a two week interval, produce very similar constructs (Pearson r = 0.79); (ii) when allowed to take the test entirely afresh and considering new elements, subjects equally reproduce their earlier constructs (Pearson r = 0.80.).

The author had reservations about such a technique as one might expect an individual's construction to change with time and by

virtue of having been involved with completing a grid. In an associated piece of work, the author found (Hopwood and Keen 1977) that where feedback was witheld from a group of respondents (who had completed a grid as a means of having their teaching appraised) test - retest correlations were as high as those stated. However, when feedback was provided to the respondent between the two elicitations, the correlation plunged to near zero.

This illustrates the interlocking of the usual concepts of validity and reliability when applied to grid methods, in that they can be looked on as experimental explications of the dichotomy corollary, which states that "a person's construction system is composed of a finite number of dichotomous constructs" These studies suggest that the grid tests are "reliable" in the sense that where constructs about people are concerned (and, one would tend to assume, where other types of construct are concerned), normal sampling procedures apply. Grids are delving into a limited repertoire of constructs which the subject has available, and there is no fear of being confronted with the everlasting pages of an infinite personal dictionary.

This highlights the fact that users of grid methods are faced on each occasion with an explicit sampling problem, in that they must select elements, constructs, methods of administration and methods of analysis appropriate to the context in which they are attempting to make predictions. This certainly does not relieve them of sampling problems, nor does

it guarantee that they will adequately solve them: but it does mean that they have explicitly to face them time and again, and they are less likely to ignore the dangers of sampling error by resting content with a pseudo, once-for-all solution to the problem.

Returning to the main theme of this paragraph, namely the reliability and validity of grid methods, one has to consider the implications of supplying constructs. The problems which can arise when the constructs are supplied by the experimenter, rather than elicited from the subject, are obvious - any assurance that the constructs are a meaningful part of the subject's repertoire is forfeited, and some of the grid matrix variance will inevitably reflect a degree of failure by the subject to translate them into his own terms (Isaacson and Langfield, 1965, and Cromwell and Caldwell, 1962). Yet in experiments with specific hypothesis and in those involving group comparisons, construct supplying seems useful. Tactics which reduce, though they do not eliminate, distortion effects in supplied construct grids, are suggested.

In any grid or grid series involving supplied constructs, additional elicited constructs from the subject can be used. The presence of such individually varying elicited constructs does not affect the calculation of relationships between the standard constructs, but it does provide a means of checking something of the meaningfulness of the supplied constructs. A markedly lower level of intercorrelation between supplied than between elicited constructs could indicate that the

.165

supplied constructs were largely verbiage for the subject. An individual examination of supplied/elicited construct relationships in the grid matrix might yield information as to any unusual meanings which had been attached to the supplied The strategy adopted by the author and described in labels. paragraph 6.9.1. and 6.9.2. of eliciting constructs by triads and providing 'prompt' words arose from the experience of pretest interviewing of single subjects and groups on the topics covered in the grid to yield information as to the common vocabulary available, and increase the experimenter's chances of couching supplied constructs (or key words) in terms meaningful to his subject. This type of methodology is less prone to the dangers inherent in providing constructs, and yet it still does permit the analysis between applied and elicited constructs noted above.

In this research, the selection of the 20 elements was undertaken to try and maximise the validity; although, as was indicated in the previous paragraph, no guarantee could be made until the data were available. The effect of element selection on validity is two pronged. Firstly, the number of elements should reflect the subjects discriminating capacity; but this is really only of real importance where ranking rather than rating of elements is demanded, and, secondly, that all the elements used for sorting in the grid must be within the range of convenience of the constructs supplied and elicited. This rule has been constantly in mind in designing the particular grid for this research. It cannot be denied, however, that as so much of construct theory hinges on the

notion of the bipolarity of constructs, that the reliability or unreliability, in different individuals, or under different conditions or between different constructs, of the distribution of elements between the poles must be a matter of Any discussion about lopsidedness of a major interest. construct implies that the relative allotment of elements to the two poles of the construct for any given individual is an indicator of stability, in the sense that we would expect the proportions in each pole to be roughly the same when an individual uses the same construct on two random samples of Care, therefore, has to be taken in interpreting elements. the results indicating lopsidedness, this problem is minimised by the adoption of a rating scale as opposed to the traditional binary response.

Another aspect of grid method is relevant to the concern with the predictable stability, rather than the general stability, of scores. The grid is not a test, but a variable technique: it can be cast into many different forms, involving any number of different types of constructs and elements, and many kinds of scores can be derived. Since there is no such thing as <u>the</u> grid, there can be no such thing as <u>the</u> reliability of the grid.

Any consideration of a specific reliability coefficient quoted for a particular instance of the application of grid method, must take into account, firstly, the particular measure extracted from data supplied by the grid; secondly, the type of experimental situation within which repeat grid data were

obtained, and thirdly, the general parameters which affect reliability coefficients in any grid context.

In conclusion of this discussion on the reliability of grids, the author is forced to accept that since this is a procedure relating to a theory which affirms that "man is a form of motion", it is necessary to challenge the orthodox notion of high reliability as an invariably desirable characteristic of However, even bearing in mind the slender evidence tests. available, it would seem that constructs which account for little of the variance in grid matrices are likely to be those and this restriction with many near-zero relationship scores: of variance, plus the tendency of low correlations to wobble randomly around zero, may, of itself, reduce the reliability Consequently, it would seem wise to assume that coefficient. reliability will be, in part, a function of grid form and type of administration.

Moving on to validity, as distinct from the previous discussion, which looked at the interrelationship of validity and reliability, one is forced again, in grid methodology, to reject traditional ideas. As stated by English and English (1958):

"there is no such thing as general validity. Nor is there absolute validity - we determine the degree of validity. And the validity index has no meaning apart from the particular operations by which it is determined."

These structures apply to any consideration of the validity of

grid technique in a highly specific way. Just as with reliability, to ask what is the validity of the grid is like asking what is the validity of the questionnaire. The questions would be, what questionnaire?, in what context?, The number and manner of grids is, used for what purpose? for practical purposes, infinite. Perhaps all that can be done is to consider whether the basic rationale of grids makes sense, whether their application gives results which tally with the underlying assumptions and whether the use of grid methods to date justifies their continued elaboration. It need not be stated that the author considers that grid methods to date are valid and so justify further elaboration!

: .

In terms of construct theory, these modes of validation seem Construct theory envisages each individual as incomplete. developing and operating an elaborate, even though sometimes poorly articulated, construct system, designed to deal with Measures of such a system can only be many situations. minimally validated in artificial and restricting contexts. It has been argued that the appearance of statistically significant internal relationships in grids gives validational support to the central tenet of construct theory - that people construe in an organised way - and this suggests in turn, that the grid is a useful means of objectifying this theoretical Although the fate of any theory need not turn on contention. the outcome of a particular experiment or the validity of a particular prediction, its validity cannot be maintained if a succession of hypotheses, when put to adequate experimental test, do not turn out to be palpably true. Thus, the theory

1.69

can only become valid when someone is able to make use of it to produce verifiable and largely verified hypotheses. Hypotheses can be erected and tested for a single subject, in a manner not possible with conventional normative tests, and a range of conventional population statistics (including forms of cluster analysis) is potentially available for use in analysing individual grids. The individual can be studied, not merely in the sense of taking subjects one at a time, nor in the sense of making the content of test items relate to a particular individual, but also in the sense of developing a complete experimental design for a single subject. Specific hypotheses can be prepared before the grid is administered, or indeed, during the analytical stage . Since the grid allows internal tests of significance, no prior validation specific to the problem in hand is necessary, even though the basic assumptions underlying the method have to be met.

6.10.2. The precise administration procedure for the whole data collection exercise is described in Chapter 7; however, this chapter would be incomplete without a description of the manner in which the grid elicitation was possible.

In piloting the grid, as designed for use in this research, the author found group administration to be acceptable only where the group size was small enough for the administrator to fully interact with all the respondents. An optimum size proved to be three respondents. Thus all the data collection was undertaken either on an individual (one to one) basis or by small groups of two or three respondents. In every case

the author was the grid administrator in order to control any variance of instructions which may have resulted from the use of additional administrators. Each session was planned to take three and a half hours on one day, (although this was frequently spread over a lunch break). Only one respondent required more time, insisting on using 21 constructs (although subsequent analysis indicated these not to be independent, and requiring nearly five hours of uninterrupted time to complete the whole exercise). For the majority of respondents, $2\frac{1}{2}$ - 3 hours proved adequate, and some (about twelve) completed the whole session in less than $2\frac{1}{2}$ hours.

Before administration, a guarantee was given to each and every respondent that total anonymity would be preserved. The respondent's name was not recorded, although each subject was given a personalised number which he was requested to quote should be wish to seek any 'follow up' information. The second part of the quarantee of anonymity concerned the names of persons, other than the respondent, who may be used in the administration session. Respondents were informed that they would be required to cite names of others on cards, but that upon completion of the exercise the cards would be destroyed Effore the author and subject left the administration room. The personalised number was randomly drawn from a set of numbers and against this number the author recorded only the class of respondent, i.e. Teacher of physics, Student of physics or teacher other than of physics.

The attitude inventory was then administered. For student respondents the inventory rubric was explicit about the way in which the inventory should be completed. For the remaining two classes of respondent a further verbal instruction was given, namely that they were to complete the questionnaire, as they believed would be appropriate if they were now a student of physics in their first year of undergraduate study. Some difficulty was experienced with many of the non-physicist respondents who claimed that they, never having studied physics, To these respondents could not undertake such a task. encouragement was offered, and they were asked to attempt the exercise with the proviso that where any difficulty was experienced, the administrator would be available to help. In none of these cases was further aid sought, suggesting that these respondents could, in fact, complete the inventory with relative ease, within the constraints imposed.

Thirty minutes was allocated to this section (attitude inventory) including the introduction and collection of completed forms. This timing proved adequate for all but a few respondents. Any variation in timing was relatively unimportant as administration was always undertaken either individually or in small groups (2 - 3 persons). It should be noted that whilst lunch, coffee and tea breaks were a feature of the administration session, these were always taken between the administration of the attitude inventory and the grid and at no other time.

The second phase was concerned with the grid format and the The respondents were introduced to the grid with a elements. brief (2 - 3 minute) explanation in simple terms, of what they It was made clear that they would would be called upon to do. be guided through the process and should, therefore, only concern themselves with each immediate task as it was presented An envelope containing 20 cards corresponding to the to them. They were told that 13 of the 20 elements was then provided. cards contained role descriptions to which they should associate a named person who should be cited in the appropriate The same name was not permitted to be space on each card. repeated, although some cards could be left blank if the respondent found it quite impossible to relate a name to the This part of the exercise was variable in its role title. consumption of time and frequently exceeded the 30 minutes The remaining seven cards had nominal time allocation. passport size photographs of physics teachers, the letter code (corresponding to the grid column) and a key word or phrase related to what the teacher was doing on the videotape. The respondents were all given the same precise instruction, namely to watch a videotape which lasted 29 minutes. They were told that each of the seven teachers on the cards would be seen clearly identified by the photograph, key word and lever; furthermore they were invited to make any comments they wished, related to what they saw, on the appropriate cards. The exercise was then undertaken in one continuous viewing. Questions were not permitted whilst the videotape was running.

Ł

Χ

173

 \checkmark

The triad method of eliciting constructs was then explained, as was the scoring system and, very slowly, the first line of the grid completed with the administrator giving continual, The respondents were then invited to positive, guidance. continue using new constructs for every row, seeking advice wherever they felt it necessary. The administrator kept involved by asking questions where appropriate, in the usual manner, to check such items of appliability of elements to a construct, bipolarity of constructs, etc. After six to ten constructs had been elicited, the optional list of items was provided for reference. Few respondents used the list to formulate actual constructs, but reading it often stimulated involvement and seemed to facilitate new construct elicitation.

On eliciting his 12th. construct, a respondent was informed that, for many people, 20 constructs were more than were required and that he was to feel free to cease completing the grid if it became apparent to him that he had exhausted his repertoire of constructs.

The same procedure was used for all respondents.

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND REPERTORY GRID STUDY

SELECTION OF SAMPLE,

USE OF INSTRUMENTS



- 7.1. In this, as in any research project, sampling is an important mechanism for obtaining data from a restricted group of respondents which may be considered representative of the total universe of respondents. Clearly it would be quite impossible to seek information from every student and, or, teacher of physics in Great Britain within the limitations of a Ph.D. research project. Apart from the National Census, which does seek information from every member of the population (or universe of respondents) all data gathering techniques rely to a greater or lesser extent on various sampling techniques. Before the strategy for sample selection adopted by the author is considered, some terms require definition.
 - 7.2. The <u>Universe</u> of respondents is defined as all possible respondents within the appropriate category.

The portion of the universe to which an author has access is called the <u>population</u>. In this research the author negotiated access to teachers and students within seven universities and twenty polytechnics (although it should be noted that not all of those institutions providing access to students were used in the research). The population of respondents therefore were the students and staff fulfilling the appropriate group defining criteria who taught or studied in those institutions to which access had been gained.

The invited sample is defined as all clements of the population to which an invitation to participate in the research was

extended. The <u>accepting sample</u> was that portion of the invited sample that accepted the invitation and agreed to participate. Finally the <u>data-producing sample</u> was that portion of the accepting sample that actually produced data.

The research hypotheses relate to the universe of respondents and yet their investigation is based on the data-producing sample. Care has to be exercised in selecting a sample such that the ultimate group of data-producing respondents are not unlike the universe from which they are drawn.

- 7.3. Representativeness of the sample is the major consideration. It is not sufficient to select a representative population alone nor indeed a representative invited sample, it is the dataproducing sample which needs to be representative, and so the question of representativeness is a recurring one.
- 7.3.1. At every stage of the sample selection from population to data-producing sample there is a requirement for representativeness, but representative in terms of what? Clearly representativeness is required in terms of those variables known to be related to the phenomenon under investigation. This implies that to achieve representativeness a researcher must know the characteristics which are related to the phenomena and/or behaviour that he is to study. Previous research and theory can provide this knowledge, the former of the two being the most useful in that it provides empirically demonstrated relationships whilst the latter simply suggests that a relationship should

exist. Additionally a researcher is likely to have subjective ideas of additional characteristics which he hypothesises might be relevant although neither previous research nor accepted theory supports his 'hunch.' Often, as with this research, the 'hunch' list of characteristics is longer than the objective list!

- 7.3.2. Three distinct class of respondent had been defined for this research, namely
 - a) Teachers of first year undergraduate physics
 - b) First year undergraduate students of physics
 - c) Teachers of first year undergraduates from disciplines other than physics

Thus a stratification of the total universe of respondents had been made to produce three classes of respondents who would produce data relevant to the specified hypotheses. This stratification ident: fied three 'characteristics' of respondents in that each sample group would have to exhibit some 'qualifications' which enabled them to corporately fulfil the class title. In other words respondents in the class of 'teachers of first year undergraduates physics' had to be currently employed for at least some of their time in the task of teaching physics to first year undergraduates!

From the literature, some characteristics can be identified as important. Of these the sector from which the respondents were drawn seemed crucial. There is some unpublished evidence to suggest that the public sector (Polytechnics) and the private sector (Universities) would produce different types of respondent
for a variety of reasons. Representativeness in this respect could be ensured by selecting equal numbers of respondents from each type of institution. (An analysis of the data collected has indicated that there was no significant different between these groups on any of the hypotheses).

The two characteristics hypothesised as being of great importance were attitude and pedagogic practice. Whilst evidence was available from the literature to support this, these variables were crucial to the investigation of the hypotheses and necessitated the development of the instrumentation described in Chapters 3 and 6. Representativeness on these variables could not be ensured before the research was undertaken and so random selection techniques were utilised. Other characteristics considered worth ensuring representativeness on were:

- i) Age (a,c)
- ii) Sex (a,b,c)
- iii) Experience; length of service (a,c)
 - iv) Physics main vs. Physics subsidiary (a)
 - v) Teacher trained vs graduate only (a,c)
 - vi) Volunteer vs Conscript (a,b,c)
- vii) Unknown characteristics (a,b,c)

(Note: The letters in parenthesis correspond to the classes of respondent listed at the start of paragraph 7.3.2.). In order to ensure representativeness on all these variables one needs to know which characteristics are related to the phenomenon under study, and have the ability to measure each characteristic as well as having population data on the characteristic to use as

the basis for comparison. Where such knowledge was not available random sampling was used. The flow chart of Fig. 7/1 illustrates how the research sample was identified.

FIG. 7/1 ABOUT HERE

Some elaboration is necessary, particularly with regard to the activities associated with the box asterisked.

Having identified seven universities and twenty polytechnics which it was possible for the author to visit for collecting data, five of each were invited to participate. (Proportionally about half of all students of physics are in each type of institution). The ten were chosen to cover a geographical area from Plymouth to Newcastle with both large and small institutions included. Nine agreed to participate. Whilst the flow diagram indicates that there was a subsequent invitation extended to another institution to replace the one which refused, that acceptance arrived too late in the time schedule for it to be used and so data was collected from four universities and five polytechnics.

In selecting the data producing sample the characteristics listed had to be considered. For many (such as sex, age and experience) measurement was straightforward, for others more difficult. The author was particularly concerned that he may have omitted some important characteristic which neither he nor

1.80



previous research had identified. To minimise the risk of such an omission all the potential respondents from the nine institutions were listed and the data producing sample identified by randomly selecting individuals from these lists in such a way that representativeness was maintained across the criteria known to be important and which could be measured. Sex may be cited Less than 5% of physics teachers are female. as an example. The author's sample reflected this imbalance of the sexes. It is important to note that the issue is not how large the population is, nor how large a proportion of the universe is represented in the population. It is concerned instead with whether or not, in terms of factors important to the research problem the data producing sample can be considered representative of the universe and hence whether generalisations may be made about the universe for data obtained from the chosen sample. This ability to generalise is not dependent on large numbers of respondents.

The sample from which data was collected consisted of 37 students, 21 physics teachers and 18 teachers of subjects other than physics. Of these some failed to produce valid data and so the data producing sample was somewhat reduced as indicated in Table 7/1

TABLE 7/1 ABOUT HERE

It should be noted that the number of respondents invited to provide data was chosen from both those agreeing to participate (the accepting sample) and those who had declined to participate

TABLE 7/1. Sizes of data producing sample

	Number invited to participate	Number from which valid data was collected	Percentage of accepting sample producing data
Teachers of Physics	21	19	90.48
Students of Physics	37	33	89.19
Teachers other than of Physics	.18	17	94.44

183

.

(in proportion) in an attempt to control for any bias introduced by selecting only from respondents agreeably disposed to provide This was achieved by approaching the sample identified and data. asking them if they would participate. Those who agreed were included. (the accepting sample) and those who declined were approached individually and a small number coerced into participation. The numbers involved were small (8 out of 76) and in proportion to the acceptance/rejection ratio identified when the invited sample was determined (i.e. 10%). Although the anonymity of the data precludes the possibility of testing the idea, it may well be this group generated different relationships. All data producing samples were better than 89% representative of the groups from which data were collected.

7.4. 'Random sampling' is a term which has been used frequently in paragraph 7.3.2. The author has used it as a general term to embody systematic selection. In its pure form one would use a table of random numbers to select individuals from a pool previously numbered in some way, it implies that every item has an equal chance of being selected. Such a technique was used having first stratified the population against the criteria noted, thus each 'draw' of a number from the pool was constrained by the limits imposed by such groups characteristics. In common parlance the draw could be said to be loaded in favour of those respondents exhibiting particular characteristics. Thus the age, sex, experience, graduate/non-graduate, etc., characteristics were represented in the sample used in the same proportion as they were found to exist in the universe of respondents.

7.5. The reader has not, as yet, been provided with the reasoning behind the numbers of respondents chosen. The author believes that size is less important than representativeness, but having determined the minimum number of respondents required for representativeness there is the secondary question related to the numbers required to minimise the introduction of type 1 or type 2 errors with respect to the hypotheses.

The problem is circular in that the latter type of consideration might demand an increase from 10 to 15 respondents rendering it difficult to maintain the same degree of representativeness with The author chose to determine 15 as had been achieved with 10. minimum numbers of respondents in each class to limit the chances of errors to less than 5% and then increase this number to the nearest multiple of the representativeness criteria. (The .05 level of significance had been chosen as the level at which the investigation of the hypotheses would be undertaken it was therefore imperative that representativeness errors in excess of 5% were not introduced, hence 5% was detormined as the In practice, the sample selection was undertaken cut off point. with representativeness nearer (though not below) an error This gave 21, 40 and 18 as the values of N for margin of 1%). Physics teachers, physics students and other teachers respect-Due to a lack of females in the student population, the ively. middle number was necessarily reduced to 37 as indicated in Table 7/1.

7.6. From the very nature of the instrumentation respondents were able to deduce that the research data could be used to identify personal attitudes and personal pedagogic strengths and weaknesses. Teachers and their professional associations are very conservative about assessing teaching and there was an implicit element of assessment embodied in the instrumentation.

The author considered the scenario of undergraduate teaching and felt that many respondents who had agreed to participate may withdraw their consent when, and if, the threat of an assessment element became apparent. For this reason anonymity guarantees were given in the usual way, but supported by an overt methodological attempt to ensure that no individual's response could ever Such caution proved beneficial in that no be attributed to him. respondent withdrew consent (although a few respondents provided In contrast a negative aspect became apparent invalid data). during the data analysis. These were respondents who exhibited special, and interesting, profiles with areas of uncertainty. In such cases the author would have liked to return to the respondent to explore these areas further, however, the anonymity guaranteeing system proved so effective that such follow up enquiries proved impossible to conduct, because individual analyses could not be attributed to a specific individual, nor for that matter could an individual's institution be identified!

7.7. The author visited the nine institutions during the last quarter of 1976 and the first quarter of 1977. Individual administration sessions were arranged for most respondents although small groups of 2 or 3 respondents were occasionally the

subjects of joint administration. The average completion time for the instrumentation was three hours and a total of 186 hours Resource limitations precluded the was spent in collecting data. possibility of collecting the data in a systematic way as subsistence and travelling costs were so high that the author collected data when his other employment commitments took him near to the location of a respondent. The majority of respondents appeared to enjoy the experience, and whilst follow up discussions were not formally offered, twenty respondents have telephoned or written to the author to seek feedback information. Unfortunately such feedback could not be given as the individual's data became anonymous as soon as the original grid was processed. Such interest prompted the development of an associated project described in Chapter 10.

Pedagogic Styles III Physics teachilig: All attitude scaling And bepebtoby grid Study

ANALYSIS OF RESULTS WITH RESPECT TO HYPOTHESES



- 8.1. This chapter considers each hypothesis in turn in relation to the data.
- 8.2. Before moving to a consideration of each hypothesis the format of the data may be considered. For each individual respondent two individualised sets of data are available; firstly information from the respondents completion of the attitude inventory and secondly the analysis of the individuals grid. Additionally respondents have been grouped according to their attitude score on each factor of the attitude inventory and for these groups a corporate analysis (using PREFAN, see Chapter 6) has been undertaken. New variable format cards were produced to select for analysis only those elements common to all respondents. All subjects identified as members of a common attitude group had their individual grids 'stacked' to produce one very long grid having common elements. The analysis of this composite grid may be used to provide some information of group perceptions.

8.3. Hypothesis 1

"Attitude to physics, measured by the attitude to physics inventory total score if appropriate, (and on each independent factor if composite scoring is inappropriate), will be more positive for teachers of undergraduate physics than for first year undergraduate students of physics."

8.3.1. Evidence has been presented in Chapter 6 to show that a composite score produced by adding together an individual's score on each factor of the attitude inventory is an invalid

process. The hypothesis has, therefore, to be considered against each factor. Table 8/1 identifies the actual scores for all the relevant respondents whilst 8/2 specifies the relationships between them.

TABLE 8/1 ABOUT HERE

TABLE 8/2 ABOUT HERE

8.3.2. Before differences between the groups may be compared, a number of factors require consideration. Firstly, the null hypotheses has to be stated. The hypothesis stated in paragraph 8.3. implies a one tailed test (as the direction of any difference is made explicit). The null hypothesis may therefore be stated as:

"The mean score computed from all respondents who were teachers of first year undergraduate physics will not differ from the mean score computed from all first year undergraduate physics students." The hypothesis will apply to each of the four distinct attitudinal factors.

Secondly, limits need to be specified for accepting or rejecting the null hypothesis. Following the discussions of earlier chapters which dealt with the methodological development, it will be recalled that the 0.05 level of significance has been adopted throughout this research. Therefore the 0.05 level of significance will be used in considering the null hypothesis. Such an adoption of a level of significance prescribes the limits

TABLE 8/1	SCORES	OBTAINED	BY	RESPONDENTS	ON	ATTITUDE	INVENTORY

Te	Teachers of Physics				udents o	f Physic	S
A	В	С	D	A	В	С	D
22 20 27 20 20 23 24 18 24 24 20 16 25 25 20 19 25 22 22 22 22	17 18 27 18 24 22 29 25 14 18 18 16 17 22 24 26 16 22 25	20 22 24 20 23 20 27 25 14 17 20 19 24 19 23 24 14 20 15	19 19 16 19 21 20 20 24 16 15 12 21 18 18 16 20 21 20 17	28 18 17 21 23 22 21 26 26 17 20 17 21 18 17 22 21 18 19 22 18 25 15 22 13 22 15 16 25 18 23	21 19 19 20 15 27 21 17 22 21 26 22 28 7 25 28 21 21 24 17 26 22 28 7 25 28 21 21 24 17 26 22 28 7 25 28 21 21 26 22 22 28 7 25 26 22 22 28 7 25 26 22 22 28 7 25 27 21 26 22 22 26 22 22 26 22 22 26 22 22 26 22 22	21 18 12 20 18 25 20 23 25 25 21 24 22 30 25 23 18 16 18 26 11 19 21 15 22 16 30	23 16 11 16 25 21 14 23 20 25 16 16 28 24 19 21 22 25 13 19 20 18 17 17 23 15 15 16 16 22 6 25

TABLE 8/2

Means, standard deviations and relationships between comparison groups

Teachers of Physics	Mean	S.D.
Attitude factor A	21.90	2.85
Attitude factor B	20.95	4.38
Attitude factor C	20.53	3.70
Attitude factor D	18.53	2.74
Students of Physics		
Attitude factor A	20.03	3.59
Attitude factor B	21.55	4.02
Attitude factor C	21.46	.4.62
Attitude factor D .	18.91	4.74

N Teachers = 19

N Students = 33

within which a type 1 error might be made. (Type 1 errors being made when one rejects the null hypothesis by marking a difference significant, although no true difference exists). There is therefore a 5% chance that a type 1 error may be made.

The adoption of a 0.05 level of significance raises the chances of making a type 1 error to 5%, however it limits the probability (Type 2 errors being made when one of making a type 2 error. accepts a null hypothesis by marking a difference not significant, However, in this when a true difference actually exists). particular case there is an additional possibility of making a If, upon examination of the data, the null type 2 error. hypothesis is accepted as a one tail test one is able to say, with 5% certainty of being correct, that teachers do not have more positive scores on a given factor than do students, but without investigating whether they have a significantly lower score (more negative) a type 2 error might be introduced. This danger may be removed by first investigating the null hypothesis with a two tailed test to determine if any difference exists, and then checking with a one tailed test to determine if the direction of the difference coincides with that predicted by the hypothesis under investigation.

8.3.3. In determining the significance of differences between means, a variety of statistical approaches may be made. The best choice of statistic is most likely to be made having considered the population and the sampling procedure. In this case numbers of respondents are relatively small ($N_1 = 19:N_2 = 33$)

but are independent and not correlated (as the samples are drawn from different populations). Had both samples been large (i.e. > 30) the formula

$$\delta_{D_1} = \sqrt{\frac{\delta_1^2}{N_1} + \frac{\delta_2^2}{N_2}}$$

could have been used where:

- σ_{D} = Standard Error of the difference between the two samples means

Direct adoption of this formula would be unwise with one sample of less than 30 respondents. However, the danger of using the formula above is negated if both samples may be considered to The author has computed \mathcal{Z} ² values be normally distributed. (Chapter 6) and found all of these distributions to approximate to a normal curve and so the formula 8/1 would appear to be satisfactory. A second check on the normality of the distribution was made by applying the Kolmogorov-Smirnov test (K-S test, Guilford 1965) which indicated some small deviation from normality. The K-S test itself is more appropriately applied to large samples (30+) and so the result may be spurious. However, sufficient doubt remained in the author's mind to lead him to apply the K-S test as an additional way of determining the standard error of the difference between group means. There

being no difference in the results the adoption of formula 8/1 was made and all the calculations leading to the results in table 8/3 have been based on this formula.

The author was more interested in differences which occurred at the ends of the distribution rather than at the centre and the adoption of this formula accentuated such differences thus reducing the possibility of error introduction.

TABLE 8/3 ABOUT HERE

8.3.4. The null hypothesis is rejected for Factor A (Examination orientation). One may deduce, with at least 95% assurance, that the differences noted have

not appeared by chance, i.e. that teachers of first year undergraduates exhibit a more positive orientation towards examination success than do their students. The factor entitled 'exam orientation' refers to the extent to which a respondent believes that the pedagogic experience in first year undergraduate physics courses ought to be based on the demands of an examination. The mean scores for the teacher group and the student group significantly exceed the neutral attitude value (15), but the teachers score significantly higher than do their students, this implies that both categories of respondent believe exam orientation to be important, but the teachers value examination orientation in their pedagogic practice significantly more than their students.

$N_{s} = 33$ $N_{t} = 19$	Мт	^M s	D = M _t -M _S	x = 6 _D	t = D/x	l tail test at 0.05 level of significance	2 tail test at 0.05 level of significance
FACTOR A	21.90	20.03	+ 1.87	0.91	+ 2.05	REJECT aull hypothesis	REJECT null hypothesis
FACTOR B	20.95	21.55	- 0.60	1.23	- 0.49	ACCEPT null hypothesis	ACCEPT null hypothesis
FACTOR C	20.53	21.46	- 0.93	1.11	- 0.84	ACCEPT null hypothesis	ACCEPT null hypothesis
FACTOR D	18.53	18.91	- 0.38	1.04	- 0.37	ACCEPT null hypothesis	ACCEPT null hypothesis

TABLE 8/3. The significance of differences between means for Teachers of Physics and Students of Physics

Mt = Mean of teacher sample. Ms = Mean of student sample.

6 D = Standard error on the difference between means. D = Difference between means.

t = Fishers 't' statistic which is a critical ratio (although not all CRs are 't's).

Statistics in this table refer to an investigation of the null hypothesis arising from research hypothesis number one.

On all the other factors the null hypothesis is accepted, implying that there is no difference between teachers and students in their attitudes to these factors. Whilst the values are by no means significant it is interesting to comment on an apparent trend, contrary to the result anticipated by the original hypothesis, namely that on all other factors the students exhibit a slightly more positive attitude than do their teachers.

8.4. Hypothesis 2.

"Attitude to physics, measured by the attitude to physics inventory total score if appropriate (and on each independent factor if composite scoring is inappropriate), will be more positive for first year undergraduate students of physics than for academic teaching staff from disciplines other than physics."

The relevant null hypothesis is as follows:-

"The mean score computed from all respondents who were first year undergraduate students of physics will not differ from the mean score computed from all respondents who were teachers of subjects other than physics."

The arguments specified in paragraph 8.3. may be applied to this hypothesis with the consequential production of Table 8/4.

TABLE 8/4 ABOUT HERE

8.4.1. A subjective consideration of the hypothesis might lead the

TABLE 6/4.The significance of difference between means for teachers of disciplines other than physics and'students of physics.

N _s = 33 N _{NPO} = 17	и _{мро} (б)	^M s (6)	D = ^M NOP ^{- M} s	$JD = \frac{6}{\frac{6}{N_{NOP}} + \frac{6}{N_{s}}}$	t = D/ <mark>G</mark> D	l tail test at .05 level of significance	2 tail test at .05 level of significance
FACTOR A	19.82 (3.73)	20.02 (3.59)	- 0.21	0.58	- 0.36	ACCEPT null hypothesis	ACCEPT null hypothesis
FACTOR B	21.12 (4.05)	21.55 (4.02)	- 0.43	0.60	- 0.72	ACCEPT null hypothesis	ACCEPT null hypothesis
FACTOR C	17.88 (4.33)	21.46 (4.62)	- 3,58	0.63	- 5.68	REJECT null hypothesis	REJECT null hypothesis
FACTOR D	17.29 (4,70)	18.91 (4.74)	- 1.62	0.65	- 2.49	REJECT null hypothesis	REJECT null hypothesis

d.f. = $(N_1 + N_2) = 48$. $M_{NOP} = Mean of non-physics teachers$. $M_s = Mean of student sample$. Otherwise coded as Table 8/3.

Statistics in this table relate to the null hypothesis arising from research hypothesis 2.

reader to the conclusion that it states the obvious. As teachers of subjects other than physics will evidently have chosen not to pursue that subject at some time in their past, one might deduce that such a decision would imply that their attitude to the subject (physics) as measured by the four attitudinal factors, would be less positive than for students who had made a choice in favour of studying the subject. This subjective view was found to be correct for factors C and D. Thus both intrinsic motivation to study physics (C) and pleasure from physics (D), are more likely to be positive attributes of the student of physics than of an academic teacher of a subject other than physics. However, the data fails to reject the null hypothesis on the first two factors. Therefore. although the trend of the data (the difference between means) supports the hypothesis, the statistical analysis indicates that there is no difference between first year undergraduate physics students and teachers of subjects other than physics in their attitude towards 'exam orientation' (Factor A - the importance of physics teaching as a means to passing examinations) and 'practical orientation' (Factor B - the importance of the practical aspects or usefulness of physics).

8.5. Hypothesis 3

"First year undergraduate students of physics will be more positvely crientated towards the requirements of an examination, measured by Factor A of the attitude to physics inventory, than will teachers of undergraduate physics." The implication of this hypothesis is that students will see 'overcoming the hurdle' of an examination more important than reading physics for physics sake.

The data suggests the contrary view is correct. Table 8/5 is an extract of the relevant data from table 8/3.

Table 8/5. A comparison of student and teacher attitudes to an examination orientation in physics teaching.

Teacher mean score =	21.90;	$6^{T} = 2.85$
Student mean score =	20.03;	$\sigma_{\rm S} = 3.59$
Difference between means =	+ 1.87;	$\sigma_{\rm D} = 0.91$
t-score =	+ 2.05	

It is quite clear from the data in table 8/5 that the teachers, not the students, prize examination orientation (within the teaching of physics) more highly. The students are less willing to see their teaching directed principally towards examination success than are their teachers. As will become evident in Chapter 9, this result is contrary to popular opinion amongst those concerned with physics education.

8.6. Hypothesis 4.

"Teachers of undergraduate physics will exhibit a more positive orientation towards practical work, intrinsic motivation and obtaining pleasure from physics instruction, measured by the three respective factors from the attitude to physics inventory,

than will first year undergraduate students of physics."

The relevant data may be extracted from table 8/3, and is reproduced as table 8/6.

TABLE 8/6. The difference between the mean scores of teachers and students on Factors B, C and D of the attitude to physics inventory.

	Teachers	Students	Difference
Mean; factor B	20.95	21.55	- 0.60
Mean; factor C	20.53	21.46	- 0.93
Mean: factor D	18.53	18.91	- 0.38

- 8.6.1. None of the differences between means were found to be significant and so one must conclude that there is no difference between the way in which students and teachers respond to these three attitudinal factors (B, practical orientation; C, intrinsic motivation and D, personal pleasure from physics). Although not significant, it should be noted that the trend in all three differences was in the opposite direction to that hypothesised.
- 8.7. Three categories of respondent provided data for the research. These were first year undergraduate students of physics (referred to in future as 'students'), teachers of first year undergraduate physics (referred to as physics teachers) and

teachers of undergraduate subjects other than physics (referred to as non-physics teachers). As was described in preceding chapters, every respondent's score on each of four attitudinal factors was recorded, and from these scores sets, or groups, of respondents were identified as exhibiting a significantly positive or negative attitude. Twenty-four such sets of respondents have therefore been identified, i.e. six categories of respondent on four factors of the attitude inventory. The remaining hypotheses are concerned with the relationships between these sets or groups of respondent. Table 8/7 identifies the respondents numbers included in each group, identified by their code number.

TABLE 8/7 ABOUT HERE

8.8.1. Hypothesis 5.

"There will be a significant difference between the positive attitude and negative attitude respondents of all three categories, and on all four factors, in the way in which they corporately perceive and categorise observed teaching acts on criteria associated with effectiveness of teaching, identified by a statistical comparison of element vectors and subjectively by the respondents verbal constructions associated with the principal components in the combined construct space of the appropriate respondent sub-group."

8.8.2. This hypothesis requires twelve comparisons to be made

TABLE 8/7. Membership of comparison groups

(The numbers refer to the code numbers printed on each respondent's reply package).

		Physics Teachers	Students	Non Physics Teachers
	+	4; 15; 16; 24.	18; 23; 44; 49; 64; 71;	26; 29; 41
e fo. on.		(A)	/3. (J)	(S)
FOR A Ltive Lcates erenc n entati	-	9; 13; 19.	21; 50; 52; 55; 56; 65;	27; 43.
FAC1 Posi indi pref exan oric		(B)	67; 69; 70. (K)	(T)
ы v)	+	4; 8; 9; 19.	35; 51; 54; 56; 57; 63; 65: 70: 72	27; 32; 37; 38; 47.
a c fo bia		(C)	(L)	(U)
TOR B Itive icates ferenc ctical	-	1; 10; 13; 15; 24.	23; 44; 55; 61; 62; 67;	26; 28; 40.
FAC Pos ind prep		(D)	(14)	(V)
	+	4; 8; 9; 15; 19.	54; 55; 58; 65; 66; 73.	27; 37; 47.
s the nt it of the call of the definition of the call of the definition of the defin		(E)	(N)	(₩)
CTOR C sitive dicate sponde trinsi tivate	-	10; 11; 24; 30.	21; 63; 67; 70; 72.	26; 32; 40; 43.
FA FA FA FA		(F)	(P)	(X)
sure	+	6; 9; 13; 24.	18; 23; 44; 50; 53; 54;	26; 29; 38.
s plea		(C)	(0)	(Y)
CTOR D sitive dicate rsonal	-	4; 10; 11 12; 17.	21; 36; 59; 67; 68; 72.	32; 41; 42; 43.
FA(int fr		(H)	(R)	(Z)

between the positive and negative sub-groups of each of 3 categories on 4 attitudinal factors. These will be considered in turn. Firstly Teachers on Factor A. (Cells A and B of Table 8/7). Table 8/8 contains the relevant data.

TABLE 8/8 ABOUT HERE

The null hypothesis asserts that there is no significant correlation between the groups under comparison and in order to reject this hypothesis the correlation has to exceed \pm 0.754 (at 0.05 level with (N - 2) degrees of freedom). As this value is not exceeded in table 8/8 the null hypothesis is accepted. However, trends may be as important as statistically significant relationships, but before an investigation of any apparant trends a digression is required to consider the statistics used.

The INGRID computer analysis package described elsewhere normalises all scores. Thus when computing a correlation coefficient it is acceptable to use Pearson r. (Spearman rho could be used with ranked values but Pearson r has been selected as it reduces the number of operations which need to be applied to the raw scores).

To return to a discussion on trends it should be noted that any comments must of necessity be subjective in nature, as

TABLE 8/8. A comparison between positive and negative attitude respondents on Factor A (exam orientation) drawn from the sample of physics teachers.

		POSITIVE ATTITUDE PHYSICS TEACHERS ON FACTOR A	NEGATIVE ATTITUDE PHYSICS TEACHERS ON FACTOR A
NUMBER OF COMPONENTS SIGNIFICANT		1	1
	1.	Give reinforcement	Explicit
CONSTRUCT	2.	Good at explanation	Practical orientation
DESCRIBING AN EFFECTIVE	3.	Good general manner	Effective for weaker students.
ILACHER	4.	Good spoken presentation	Introverted
PEARSON 'r' BETWEEN ELEMENT VECTORS		- 0.626 (Not sign	nificant)

(This table indicates that there is no difference between the number of components found significant, that a non-significant negative correlation exists between element vectors and some construct labels used by respondents). statistically the evidence clearly states that there are <u>no</u> significant relationships between the sub-groups under consideration.

The statistical computations are embellished by a number of observations; firstly for each class of respondent one component is found significant. The hypothesis under investigation implies that this will be a different component for each class. This may be checked by considering the common element vectors between groups. An element vector locates an element somewhere along a component and if two distinct groups locate the <u>same</u> elements in the same positions one may conclude that the components are identical. In this case the correlation Had the value been more negative than - 0.754 one is negative. could have concluded that the two components were indeed related, one being reversed in polarity to the other, however, that value was not exceeded although - 0.626 is a large correlation indicating that there is a tendency for the two groups to construe along a similar principal component with one group reversing the polarity.

Thus on the evidence presented the substansive hypothesis (para, 8.8.1.) proves to be correct, the two groups do indeed construe effective teaching quite differently.

Trying to interpret construct labels is a dangerous procedure, as it is impossible to know what the respondent meant by the verbal label he assigned to his construct. However, providing

this danger is acknowledged the listing of those constructs known to best describe the component under consideration often provides a powerful, though subjective, reinforcement of the mathematical comparison.

By scanning the construct vectors to identify those with the highest numerical value (irrespective of sign) the four constructs most strongly associated with the component may be The sign then informs the researcher which pole of identified. the construct to identify. Having followed this procedure, table 8/8 lists the four constructs best describing the principal The reader has to attempt the component for each sub-group. impossible task of looking at four phrases (which may or may not be interpreted by him in the same way as the originator of the construct used the phrase) and tease out an underlying component Impossible though the task really is, there is description! often an apparent contrast between the two descriptions. The author hesitates to add further confusion by specifying his construction of the component descriptions in this example but would suggest, however, that if the reader chooses to approach the task it may emerge that these two groups are indeed construing the common elements along different components.

8.8.3. The process of paragraph 8.8.2. may be repeated for the second of the twelve comparison groups. Remaining with teachers of physics the comparison is now made on Factor B. (Cells C and D). The relevant data are tabulated in Table 8/9.

TABLE 8/9 (a).A comparison between positive and negative
attitude respondents on Factor B (Practical
bias) drawn from the sample of physics teachers.

		POSITIVE ATTITUDE PHYSICS TEACHERS ON FACTOR B	NEGATIVE ATTITUDE PHYSICS TEACHERS ON FACTOR B	
Number of components Significant		J.	6	
	1	Effective for weaker students		
Construct labels used by respondents to describe principal component	2	Hatch work to students ability	See Table 8/9 (b)	
	3	Practical orientation		
	4	Clarity of diction		
Correlation between common element vectors		+ .754 (Significant 0.05)		

(This table highlights a significant positive correlation between the groups specified and attitude Factor B)

TABLE 8/9 (b).

As six components were significant for negative attitude teachers on Factor B, this table supplements 8/9 (a).

COMPONENT		The four constructs with highest vectors for each component.				
	1	Reinforcement				
1	2	General manner				
·~ +	3	Audibility				
	[°] 4	Use of blackboard				
	1	Teachers other subjects in addition to physics				
. 2	2	Effectiveness				
1-	3	New ideas				
	4	Research interests				
	1	Chalk and talk				
з	2	Get down to students level				
.)	3	Student orientated				
	4	Confidence of presentation				
	1	Cleanliness of chalk board				
4	2	Ambition .				
	3	Theoretical emphasis				
	4	Experience as teacher				
	· 1	Entertaining				
5	2	Clarity of drawings				
,	3	Likeableness				
	4	Logical approach				
	1	Student involvement				
6	2	Warmth of teacher				
5	3	Clarity of thoughts				
	4	Personality strength of teacher				

In this case the null hypothesis is clearly rejected as both groups do construe effective teaching along the same principal components. However, negative attitude teachers are far more discriminating in classifying observed teaching acts, construing in six dimensions as opposed to the unidimensional approach of the positive attitude teachers on this attitude factor. The substansive hypothesis is therefore not accepted in this case.

8.8.4. Moving to Factor C, the physics teacher groups become E and F from Table 8/7. The data appears in Table 8/10.

TABLE 8/10 ABOUT HERE

Once again the negative attitude teachers construe in more dimensions than do the positive attitude teachers, (3:1) however, there is no significant relationship between their construction of the first principal component/when the substansive hypothesis is supported.

8.8.5. The last physics teacher comparison is between the positive and negative attitude groups on Factor D (Personal pleasure from physics) and these respondents are contained in cells G and H of Table 8/7. Table 8/11 contains the comparison data.

TABLE 8/11 ABOUT HERE

Intuitively one might expect personal pleasure from physics to be construed most differently by groups exhibiting positive and negative attitudes to this factor respectively. The data

TABLE 8/10 (a).A comparison between positive and negative
attitude respondents on Factor C (Intrinsic
Motivation) drawn from the sample of physics
teachers.

		POSITIVE ATTITUDE PHYSICS TEACHERS ON FACTOR C	NEGATIVE ATTITUDE PHYSICS TEACHERS ON FACTOR C		
Number of components Significant		1	3		
Construct labels used by respondents to describe principal component	1	Match work to students ability			
	2	Good	See Table 8/10 (b)		
	3	Student orientated			
	4	Practical orientation			
Correlation between					
common element + .754 (Significant 0.05)					
vectors					

(This table highlights a significant positive correlation between the groups specified and attitude Factor C)

TABLE 8/10 (b).As three components were significant for
negative attitude teachers on Factor C, this
table supplements Table 8/10 (a).

COMPONENT		The four constructs with highest vectors for each component	
	1	Students participate	
1	2	Logical approach	
-	3 Component		
	. 4	Effective reinforcement	
2	· 1	College	
	2	?	
	3	Moves around	
	• 4	Progressive	
3	1	Uses questions and answers	
	2	Mathematical	
	3	Teaching aids	
	4	Fair speed	

TABLE 8/11.A comparison between positive and negative attitude
respondents on Factor D (Personal pleasure from
physics) drawn from the sample of physics teachers.

		POSITIVE ATTITUDE PHYSICS TEACHERS ON FACTOR D	NEGATIVE ATTITUDE PHYSICS TEACHERS ON FACTOR D			
Number of components Significant		1	1			
Construct labels used by respondents to describe principal component	1	Clcar simple message	Good blackboard use			
	2	Clear diction	Good subject knowledge			
	3	-Slow methodical, clearly illustrated	Formal approach			
	4	Brash	Aggressive			
Correlation between						
common element		+ 0.0095 (not significant)				
vectors						

(This table shows that teachers of physics of differing attitudes on Factor D construe effective teaching in unrelated ways). supports this view in that there is no correlation between the first two components (r = 0.0095). The verbal labels indicate the kind of distribution which is made.

The hypothesis stated in paragraph 8.8.1. is therefore supported by the data in all but two attitudinal factors, B and C. Teachers of physics exhibiting a positive attitude to each of the remaining two factors have a significantly different view of what an effective pedagogic style will be than do their colleagues exhibiting a negative attitude.

Thus two hypothetical teachers who have differing attitudes to any one of the two differentiating factors (examination orientation or personal pleasure from physics) may believe themselves to be effective because they adopt their attitudinal peer group description of effective teaching. To the external observer these chosen pedagogic styles will be significantly different. (They may, of course, both be effective depending on the perceptions of the students but such conclusions can only be drawn after an investigation of the other hypotheses being considered in this research.)

8.9.1. The same hypothesis may now be investigated for students. Cells J and K of Table 8/7 identify the appropriate attitude groups for Factor A. Table 8/12 contains the relevant data.

<u>21</u>4
TABLE 8/12. A comparison between positive and negative attitude respondents on Factor A (Exam orientation) drawn from the sample of students.

		POSITIVE ATTITUDE PHYSICS STUDENTS ON FACTOR A	NECATIVE ATTITUDE PHYSICS STUDENTS ON FACTOR A		
Number of components Significant		1	5 *		
Construct labels used by respondents to describe principal component	1	Well qualified	Speaks well, sure and well planned.		
	2	Competent in field	Good talker		
	3	Puts over a reasonable amount of information	Uses 'props' well		
	4	Explains	Clear diagrams		
Correlation between common element	•	+ .754 (Signif	icant 0.05)		
vectors					

* Unlike his treatment of the teacher sample, the author has not listed constructs used beyond the first component unless more than one component is significant for <u>both</u> groups. Clearly the students exhibiting differing attitudes to this factor do construe effective teaching differently although there is a trend (not significant) to suggest that some small relationship might exist. The verbal labels may help to indicate, as before, the kind of difference in perception which exists. As with the teacher groups it is the negative attitude students who have a greater discriminating capacity (5 significant components as opposed to one) indicating that they are more capable of seeing strengths and weaknesses in a variety of pedagogic practice.

8.9.2. Progressing onward to Factor B (practical orientation) the student groups are L and M from Table 8/7. The relevant data is included in Table 8/13.

TABLE 8/13 ABOUT HERE

A correlation of - 0.8216 exceeds the critical value for rejection of the null hypothesis. One may conclude therefore that the two groups considered do construe effective teaching along the same component although the negative correlation indicates that for one group the component is rotated through 180° . The words used to describe the components appear to describe the s ame component also, thus strengthening the statistical evidence. This relationship whilst contrary to that anticipated in the hypothesis under investigation is a replication of the findings obtained from the physics teacher group in paragraph 8.8.3.

A comparison between positive and negative TABLE 8/13. attitude respondents on Factor B (Practical orientation) drawn from the sample of students.

		and the second	a second s
		POSITIVE ATTITUDE PHYSICS STUDENTS ON FACTOR B	NEGATIVE ATTITUDE PHYSICS STUDENTS ON FACTOR B
Number of components Significant		1	1
Construct labels used by respondents to describe principal component	1	Clear teaching	Inspires confidence
	2	Interesting	Seems to hold $\begin{bmatrix} E \\ Q \\ a \\ \end{bmatrix}$
	3	Good presentation	Good L explanations S
	4	Interested in physics for its own sake	Teacher N D
Correlation betwee	en		
common element		8216 (Significar	nt relationship

(Significant relationship if one component reversed)

01.75

vectors

(This table indicates that students of differing attitudes on Factor B (practical orientation), construe effective teaching along a common component save that students at opposite poles of the attitude scale reverse the scale descriptions).

8.9.3. Factor C student groups are identified in Table 8/7 as N and P and these are compared in Table 8/14.

TABLE 8/14 ABOUT HERE

Although the null hypothesis cannot be rejected with a correlation of - .7161 (i.e. one cannot assert that the correlation is not zero) the magnitude of the correlation is indicative of a trend. Thus although the hypothesis stated in paragraph 8.8.1. is confirmed the difference in construction of an effective teacher by these particular groups is not as great as for most previous comparisons. The negative sign indicates that the tendency to correlate is between one groups principal component and the other groups principal component rotated through 180° . The subjective appraisal of the construct labels once again supports the statistical evidence as the words used seem to be related although at the same time distinct differences may be detected.

8.9.4. The final comparison for the student group is on Factor D (Personal pleasure from physics) between groups Q and R from Table 8/7. Table 8/15 contains the relevant data.

TABLE 8/15 ABOUT HERE

In common with the physics teacher sample on Factor C the student sample produces a near zero correlation. Of all four factors for the student sample the fourth (personal pleasure from

TABLE 8/14.A comparison between positive and negative attituderespondents on Factor C (Intrinsic motivation)drawn from the sample of students.

		-		
				NEGATIVE ATTITUDE PHYSICS STUDENTS ON FACTOR C
Number of components Significant		1		
1	Down to carth	Easy to listen to		
2	Neat, well sorted out and ordered	Inspires confidence		
3	Speaks clearly and holds attention	Good presentation		
4	Logical thinking	Good approach		
Correlation between common element vector		ignificant)		
	s 1 2 3 4	POSITIVE ATTITUDE PHYSICS STUDENTS ON FACTOR C 1 Down to carth 2 Neat, well sorted out and ordered 3 Speaks clearly and holds attention 4 Logical thinking - 0.7161 (not s		

TABLE 8/15. A comparison between positive and negative attitude respondents on Factor D (Personal pleasure from physics) drawn from the sample of students.

		NEGATIVE ATTITUDE PHYSICS STUDENTS ON FACTOR D
Number of components Significant		1
1	Not too maths minded	Inspires confidence
2	Professional teacher	Interesting
3	Puts over a reasonable amount of information	Seems to hold class
4	Very interesting to listen to	Puts over well
Correlation between common element vector		significant)
	1 2 3 4	POSITIVE ATTITUDE PHYSICS STUDENTS ON FACTOR D11Not too maths minded22Professional teacher3Puts over a reasonable amount of information4Very interesting to listen to- 0.0650 (not

physics) shows the greatest difference in perception of what an effective teacher of physics might be like. The hypothesis (paragraph 8.8.1.) is therefore supported by the data from physics students for 3 of the four factors of the attitude inventory. The one where the hypothesis is not supported is Factor B (practical orientation) where the student samples, in common with the teacher samples, do not differentiate between styles of effective teaching.

8.10.1. The third, and final, category is that of non-physics teachers. This group was specifically included in the analysis as the elements of the group (teachers of undergraduate subjects other than physics) were less likely to specify discipline related criteria of pedagogic effectiveness than were the other two groups.

The hypothesis of paragraph 8.8.1. is now investigated for this category of respondent on each factor in turn. For Factor A groups S and T from Table 8/7 are used to provide the data in Table 8/16.

TABLE 8/16 ABOUT HERE

A correlation of + 0.8482 indicates a significant relationship. Thus the two non-physics teacher attitude groups on Factor A perceive an effective teacher along a common component, the null hypothesis is therefore rejected.

TABLE 8/16.A comparison between positive and negative
attitude respondents on Factor A (Exam
orientation) drawn from the sample of non-physics
teachers

		the second s		
		POSITIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR A	NEGATIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR A	
Number of components Significant		1	1	
Construct labels used by respondents to describe principal components	1	Good teacher	Liberal	
	2	Clear presentation	Abstract in teaching approach	
	3	Class participation explicit	Perceptive	
	4	Interesting lesson	Interested only in exam results	
Correlation between				
common element		+ 0.8482 (Signific	cant)	
vector				

(This table indicates that irrespective of attitude score on Factor A, teachers of discipline other than physics construe effective physics teaching in the same way). 8.10.2. The non-physics teacher groups for comparison on Factor B (practical orientation) are identified from Table 8/7 as U and V. Table 8/17 lists the relevant data.

TABLE 8/17 ABOUT HERE

A near perfect correlation coefficient of + 0.932 indicates that for non-physics teachers there is no difference between the positive and negative scoring groups on Factor B_{j} (practical orientation) in the way they perceive an effective physics teacher.

8.10.3. Factor C, intrinsic motivation, is the third factor for investigation. The non-physics teachers constitute groups W and X from Table 8/7. Table 8/18 contains the comparison data.

TABLE 8/18 ABOUT HERE

(From Table 8/18 it will be apparent that once again the correlation is significant so both attitudinal groups construe effective teaching along a common principal component). However, it should be noted that the positive attitude teachers (on this factor) are more discriminating than the negative attitude teachers as 5 rather than 1 components were found to be significant.

8.10.4. Factor D, personal pleasure from physics, is the last comparison to be made, in this case between groups Y and Z in Table 8/7. Table 8/19 contains the relevant data.

TABLE 8/18. A comparison between positive and negative attitude respondents on Factor C (Intrinsic Motivation) drawn from the sample of non-physics teachers.

		POSITIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR C	NEGATIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR C	
Number of components Significant		5	1	
Construct labels used by respondents to describe principal components	1	Perceptive	Use of concrete or pictorial material	
	2	Establish rapport	Interested in under- standing rather than fact learning	
	3	Liberal	Correct teaching speed	
4		Interesting to listen to	Clear enunciation	
Correlation between common element vector	1	+ 0.8775 (Signific	cant)	

TABLE 8/17.A comparison between positive and negative
attitude respondents on Factor B (Practical
orientation) drawn from the sample of non-
physics teachers

		POSITIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR B	NEGATIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR B
Number of components Significant		. 1	1
Construct labels used by respondents to describe principal components	1	Perceptive	Good teacher
	2	Could effect interest in learning	Class particip- ation explicit
	3	Liberal	Good lesson
	4	Rapport established	Good teaching method
Correlation between common element + 0.9321			ficant)

(This table replicates the findings of Table 8/16 save that the attitude factor is changed from A to B).

.

TABLE 8/19. A comparison between positive and negative attitude respondents on Factor D (Personal pleasure from physics) drawn from the samples of non-physics teachers

	-	a second a second s		
		POSITIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR D	NEGATIVE ATTITUDE NON-PHYSICS TEACHERS ON FACTOR D	
Number of components Significant		1	1	
1Construct labels used by respondents to describe principal components3	1	Good teacher	Clear, logical explanation	
	2	Class participation explicit	Use of concrete or pictorial materials	
	3	Good lesson	Very interesting teaching	
4		Clear presentation	Teaching speed correct	
Correlation betwe common element vector	en	+ 0.8439 (Signif	icant)	

(Non-physics teachers classified as either positively or negatively orientated to Factor D (Personal pleasure from physics) construe effective teaching along a common component). 8.11. Having considered the twelve comparison groups, some general conclusions may be drawn in relation to the hypothesis under investigation.

The hypothesis states that there will be a significant difference between the positive attitude and negative attitude respondents of all three categories, (physics teachers, students and nonphysics teachers), and on all four factors, (A: Exam orientation; B: practical orientation; C:Intrinsic Motivation: D:Personal pleasure from physics), in the way in which they corporately perceive and categorise observed teaching acts on criteria associated with effectiveness of teaching.

The results may be summarised in a correlation matrix.

TABLE 8/20 ABOUT HERE

Table 8/20 identifies an interesting pattern. One may conclude that respondents associated with physics, be they teachers of physics or students, will perceive different styles of teaching as effective dependent on their attitude to the factors; exam orientation; intrinsic motivation; and personal pleasure from physics as was hypothesised. However, differing attitudes to Factor B (practical orientation) does not alter any class of respondents perception of an effective teacher of physics.

The implication of accepting the hypothesis lies in the region of matching teachers perceptions of effectiveness with that of

TABLE 8/20. A correlation matrix summarising the significant relationships identified.

	FACTOR A	FACTOR B	FACTOR C	FACTOR D
PHYSICS TEACHERS		SIG.		
STUDENTS		SIG.		
NON-PHYSICS TEACHERS	SIG.	SIG.	SIG.	SIG.

The blank cells indicate regions where the hypothesis was supported whilst the cells indicated by 'SIG' show where the two attitude groups share common components).

(See Paragraph 8.11.)

their students. This data does nothing to investigate such matching (or mismatching) but it does suggest that either two students, or two teachers exhibiting different attitudes on any one of three factors will consider quite different styles of physics teaching as effective. Later in this chapter, teacher/ student matchings will be considered.

The second general conclusion arises from the rejection of the hypothesis for non-physics teachers across all four factors. The data therefore suggests that it is only an awareness of 'physics curriculum content' which causes differing attitude groups to identify different pedagogic styles as effective.

Thirdly Factor B (practical orientation) does not distinguish attitudinal groups who perceive effective teaching of physics differently.

Finally, with one exception (non-physics teachers on Factor C) where attitude groups construed effective teaching in more than one component the more flexible component space was always found in respondents who exhibited a <u>negative</u> attitude. These variations are not significant in the statistical sense although one might subjectively conclude that respondents who have a negative attitude are more likely to exhibit greater discrimination when categorising an observed teaching episode than are their colleagues exhibiting a positive attitude.

8.12.1. Hypothesis 6.

"There will be a greater positive correlation between the perception and classification of observed teaching episodes by teachers of physics and students of physics when respondents with similar attitudes are compared than when differing attitude groups are compared."

If this hypothesis is found to be correct then matching students and teachers by attitude will ensure common staff-student perceptions of effective teaching.

The hypothesis may be investigated by establishing a correlation matrix containing all the relevant data. In each case the correlation between the common element vectors (from the grids of attitudinal groups on each attitude factor) is computed. A high positive value (+ 0.754) indicates that on the first principal component the elements (observed teaching episodes) are construed in a similar way, a correlation of + 1.0 would indicate that the component were identical for each class of respondent being compared. The null hypothesis states that there is no similarity between each pair of element vectors and to be rejected at the .05 level of significance, the critical value of r (which has to be exceeded to reject the null hypothesis) is 0.75.

8.12.2. Each factor may be considered in turn. Thus Factor A provides four sets of respondents, positive and negative attitude teachers and students. A correlation matrix can be

drawn as in Table 8/21.

TABLE 8/21 ABOUT HERE

One cell in the matrix was found to contain a value large enough to reject the null hypothesis. That cell compared positive attitude teachers with negative attitude students, both groups construe observed teaching acts along a common component (r =+ 0.78). This clearly opposes the view expressed in the sixth hypothesis which anticipated that such a comparison would have yielded the lowest correlation. No other pattern emerges from the matrix.

The hypothesis under investigation makes no statement about similarities or differences existing only in the first component and yet the matrix considers only this first principal component in making comparisons. One might question the validity of conclusions drawn from such data when no check appears to have been made to ensure that significant differences do not exist in lower order components.

The Bartlett test has been used to identify how many components are significant at the .05 level of significance and where were than one component is significant for <u>both</u> comparison groups the element vectors across <u>all</u> significant components are compared. In Table 8/21 only one component was significant for each comparison pair and so any relationships existing in lower order components could not be significant at the 0.05 level. However, to test this two specimen groups of respondents were chosen at

TABLE 8/21. Factor A, exam orientation, a comparison of teachers and students (classified by attitude to this factor), perception and classification of observed teaching episodes.

		A	В	. C	D
Positive attitude teachers	A	1.0	- 0.01	×	+ 0.78*
Positive attitude students	в		1.0	51	У
Negative attitude teachers	с			1.0	66
Negative attitude students	D				1.0

random from Table 8/21 (actually B vs C) and correlations between the first six components computed. These correlations may be weighted by the Eigen value and their effect on the computed value ascertained. Such a process effected the correlation entered in Table 8/21 by less than 2% although unweighted correlations exceeded the value computed in the first component on some lower order components. This process cannot guarantee that some relationship does not exist in dimensions unmapped, but it does ensure that any significant relationship existing in mapped dimensions will emerge.

In summary one can cite an example. If negative attitude physics teachers on Factor C were compared with positive attitude non-physics teachers on the same factor, three components would be compared, because for the former group these components were significant and for the latter five components were significant. Repeating the exercise described earlier in this paragraph for that group, significant relationships may be identified in the 2nd. and 3rd. components although no significant relationship is evident in the first. The general rule applied throughout the research therefore is to consider, for comparative purposes, the highest number of components found significant by the Bartlett test common to both comparative groups.

The exercise of paragraph 8.12.2. may be repeated with Factors B, C and D, the data for which is tabulated in Tables 8/22; 8/23 and 8/24 respectively.

TABLE 8/22. Factor B, practical orientation. A comparison of teachers and students (classified by attitude to this factor) perception and classification of observed teaching episodes

		A	В	С	D
Positive attitude teachers	A	1.0	+ .53	x	17
Positive attitude students	в		1.0	+ .90*	У
Negative attitude teachers	с			1.0	57
Negative attitude students	D				1.0

(Note:- cells x and y are considered in a separate hypothesis)

TABLE 8/23. Factor C, Intrinsic motivation. A comparison of teachers and students (classified by their attitude to this factor) perception and classification of observed teaching episodes.

		Α	В	С	D
Positive attitude teachers	۸	1.0	+ .25	x	35
Positive attitude students	в		1.0	89*	У
Negative attitude teachers	с			1.0	+ .80*
Negative attitude students	D				1.0

(Note:- cells x and y are considered in a separate hypothesis)

TABLE 8/24.

Factor D; Personal pleasure from physics. A comparison of teachers and students (classified by their attitude to this factor) perception and classification of observed teaching episodes.

		A	В	с	D
Positive attitude teachers	A	1.0	10	x	29
Positive attitude students	В		1.0	+ .77*	У
Negative attitude teachers	с			1.0	+ .01
Negative attitude students	D				1.0

(Note:- cells x and y are considered under a separate hypothesis)

8.12.4. The significant relationships are asterisked in Tables 8/21 to 8/24 respectively. They are

> Exam orientation: Positive attitude teachers) Negative attitude students) Practical orientation: Positive attitude students) Negative attitude teachers) Intrinsic motivation: Positive attitude students) Negative attitude teachers) Negative attitude teachers) Negative attitude teachers) Negative attitude teachers)

Personal pleasure

from Physics:

Positive attitude students)) Negative attitude teachers)

It should be noted that the sign of the correlation is irrelevant for this consideration, as a negative correlation indicates a positive relationship between the components with the polarity of one reversed; in other words the same component is being used although its orientation in construct space is reversed.

Three general observations may be made from the data. Firstly with one exception (Teachers and Students, both with a negative attitude to the intrinsic motivation factor) all the significant relationships show the hypothesis to be false. There is, in other words, no guarantee that students and teachers who exhibit a common attitude score on all factors except intrinsic motivation, will perceive an effective teaching style in a common way.

Secondly the data provides evidence to suggest that the contrary view to that expressed in the hypotheses is the case, namely, teachers and students who exhibit opposing attitudes are more likely to perceive an effective teaching style in a common way. Such was the case in four of five significant relationships.

Thirdly, the emergence of one cell (positive attitude students, negative attitude teachers) as containing a significant relationship in three of the four attitudinal factors. Thus students who exhibit a positive attitude to a practical orientation in physics instruction, exhibit a positive score on the intrinsic motivation factor or score positively on the personal pleasure from physics score are most likely to describe an effective teacher of physics in the same way as a teacher who exhibits a negative score on any, or all, of these factors. The commonality of perception might become more evident if the construct labels used to describe the components were listed. Thus, with the same reservations as previously noted the author has specified the six constructs from each which most closely describe the component compared. This information is included in Table 8/25.

TABLE 8/25 ABOUT HERE

So although the two groups exhibit quite opposite attitudes a subjective look at the constructs used tends to support the statistical evidence recorded. It would be a brave reader who attempted to specify why such a result may have been obtained.

TABLE 8/25. The words used by positive attitude students and negative attitude teachers (Factors C and D) which best describe their common component.

Positive attitude students	Negative attitude students
Clear teaching Interesting	Reinforcement General manner
Down to earth	Students participate
Neat, well sorted out and ordered	Logical approach
Not too maths minded	Good blackboard use
Professional teacher	Good subject knowledge

The result is contrary to popular opinion amongst teachers of physics, but in informal discussion with such persons several explanations have been postulated. Of the range of suggestions, the one most often cited hypothesises \tilde{k} that a student with a positive attitude on these three factors is more likely to seek, as ideal, a pedagogic style which maximises learning and . enjoyment without pressures imposed by examination curricula. Teachers who exhibit a negative attitude may recognise their personal attitude and perceive an ideal pedagogic style in a similar manner because of their own awareness of their own chosen Some evidence is available to support this view teaching style. as more teachers in these categories perceive themselves as 'ineffective' than do teachers in any other class. (This data is accessed by looking at the individual INGRID analysis for each

teacher in the group and comparing the 'self' pedagogic style with both 'ideal effective' and 'ideal ineffective.' The 'self' may be placed on an effective-ineffective continuum and these teachers are generally found to place themselves nearer the 'ineffective' end). The preceding discussion extends beyond the framework of the hypothesis which is clearly rejected, but was included for two reasons, firstly to suggest to potential readers the possibility of exploring the predictive nature of the instrument, and secondly to highlight the fact that the observed relationships do not imply that negative attitude teachers will exhibit the teaching style they proclaim to be most effective.

8.13.1. Hypothesis 7.

In the discussion of hypothesis five, the author suggested that both students and teachers of physics might tend to use criteria associated with the discipline of physics when appraising a teaching experience. Such criteria would be outside the range of convenience (to use Kelly's term) for teachers other than of physics and so their criteria are more likely to be discipline free, and associated with the pedagogic content. This seventh hypothesis attempts to investigate these criteria by these groups and, if supported, by the data, will confirm the subjective observations made by the author in his discussion of hypothesis 5.

The hypothesis is: "The corporate perceptions and categorisations of observed teaching episodes, against effective teaching criteria, by each of the classes of respondent will show greater

similarity between students of physics and teachers of physics, than any similarity between any other two groups. The comparison will be made objectively as in hypothesis 5, using common element vectors, and where significant relationships are seen to exist, by subjectively listing construct labels as used by each class of respondent respectively."

As respondents exhibiting differing attitudes to each of the four attitude factors have already been shown to have differing perceptions of effective teaching, the analysis must be replicated for each of the eight attitude groups. Table 8/26 records the correlations, The hypothesis, if it is correct, asserts that the correlations in column 1 will exceed those elsewhere in the table.

TABLE 8/26 ABOUT HERE

A suitable null hypothesis would be that there is no difference between the columns of table 8/26. In order to investigate the null hypothesis the columns of table 8/26 may be compared using Fishers t formula

$$t = M_1 - M_2$$

$$\sqrt{\frac{x_1^2 + x_2^2}{N_1(N_1 - 1)}}$$

TABLE 8/26. A table of correlations between Physics Teachers, Students and Non-Physics Teachers across all attitude factors and groups

		Physics Teachers vs Physics Students	Physics Teachers vs Non-Physics Teachers	Physics Students vs Non-Physics Teachers
		Column l	Column 2	Column 3
Factor A	+	01	68	32
Exam Orientation	-	66	+ .88*	88*
Factor B	+	+ .53	+ .71	+ .83*
Practical Orientation	-	57	+ .70	58
Factor C Intrinsic Motivation	+	+ .25	62	39
	-	+ .80*	+ .56	+ .71
Factor D	+	10	72	+ .19
Personal Pleasure from Physics	-	+ .01	16	+ .42

(Note: for the hypothesis to be supported, the correlation in column 1 will exceed those in other columns for each row).

where

 M_1 = mean of one distribution

 M_2 = mean of second distribution

 x^{L} = sum of 'deviations squared' from the mean

N_i = sample size

It should be noted that the figures being compared are correlation coefficients and the mean of two or more correlation coefficients is itself not a correlation, this does not preclude the use of the Fisher t statistic as the null hypothesis assumes no difference between columns in table 8/26 in other terms it is concerned with the consistency of differences between the The sign of the correlation is also individual correlations. of little importance as a correlation of - 1.0 would imply perfect matching of respondents components but with one group reversing the polarity. This research is concerned with similarities in construing common events not with the way those systems of construction are applied and so a high negative correlation is equivalent to a high positive correlation with respect to the hypothesis under investigation.

The t statistic was chosen by the author rather than Z or any other alternative because of its relative independence of kurtosis, and its acceptability for small samples. As with all other computations the level of significance in this case was the 5% level.

Table 8/27 contains the results of the calculations.

TABLE 8/27. The results of a comparison between the columns of Table 8/26.

	Mean	Standard deviation	Fisher t with other groups	
Column 1			2	2.12 *
Physics Teachers vs Physics Students	.42	.31	3	1.48
Column 2 Physics Teachers vs Non-Physics Teachers	.72	.21	1	2.12 *
			3	0.64
Column 3 Physics Students vs Non-Physics Teachers	.63	.25	1	1.48
			2	0.64

The null hypothesis is rejected in the cells asterisked.

8.13.2. A consideration of tables 8/26 and 8/27 is quite revealing in relation to the seventh hypothesis as specified in paragraph 8.13.1.

Firstly, very few of the correlations in table 8/26 are significant, in fact only four from the 24 cells. However, with the exception of one line all the correlations in column 2 are greater than those in column 1, indicating that physics teachers and teachers other than of physics tend to construe effective teaching in the same way more often than do physics teachers and physics students. The same could be said of the comparison of column 3 with column 1 where the same observation can be made, for in this case too all but one correlation is greater in column 3. There is no such trend evident between columns 2 and 3 where no regular pattern emerges. The contents of table 8/27 enable these casual observations to be quantified.

The null hypothesis is rejected by the relationship between columns 1 and 2. In other words teachers (whether they be of physics or of another discipline) do construe effective teaching in common terms significantly more often than do physicists (whether they be students of physics or teachers of physics). Furthermore, although the relationship is not significant, physics students tend to construe in terms more frequently associated with non-physics teachers than they do with physics teachers.

Some caution needs to be exercised in accepting these comments as the sample of data from which the conclusions were drawn

includes respondents with either strong positive or strong negative attitudes on the four attitude inventory factors. It did not include those with neutral, or undecided attitudes. However, the kinds of construct labels used appear to be independent of attitude as will become evident from a perusal of table 8/28.

TABLE 8/28 ABOUT HERE

Although many of the construct labels listed in table 8/28 arise from the grids of negative attitude respondents even the most discerning reader would find it difficult to identify which should be attributed to the negative respondents. Given that the mathematical analysis confirms that the lists in table 8/28 describe a common component with better than 80% overlap an interesting exercise might be to re-order the construct labels from the two columns to form the best common description. The author has refrained from including his attempt at this task, as any such record would illustrate the imposition of his interpretation of the construct labels which may be different to that intended by their originators, and indeed different to that of any reader.

8.13.3. The hypothesis under investigation is clearly incorrect in its assertion. Not only do the students and teachers of physics fail to construe effective teaching along common

TABLE 8/28. A comparison of construct labels used by respondents in columns 1 and 2 (Table 8/26) to describe their common components

Note:- Construct labels are only drawn from respondents within cells identified as containing significant relationships in Table 8/26.

		COLUMN 1	COLUMN 2
Physics Teachers		Students participate	Explicit
		Logical approach	Practical orientation
		Competent	Effective for weaker students
		Effective reinforcement	Introverted
	υ	Easy to listen to	Liberal
tudents	Teacher	Inspires confidence	Abstract in teaching approach
sics S	uysics	Good presentation	Perceptive
Phy: Non-Pi		Good approach	Interested only in exam results

dimensions but there is a significantly more frequent relationship between the way teachers of physics and teachers of disciplines other than physics construe effective teaching.

The implication of such a result may be illustrated by an example. If a physics student, a physics teacher and nonphysics teacher all observed and assessed a teacher of physics 'in action,' the framework of appraisal used by the latter two observers would be more alike than either framework when compared to that of the student observer. In simple terms students and teachers see effective teaching in quite different ways,

8.14. Hypothesis 8

There is evidence from learning theory to suggest that learning occurs more effectively when the learner and the teacher share common objectives, also the author has evidence to suggest that learning is more efficient when the learner and the teacher share common perceptions of what pedagogy constitutes an effective teaching style (Keen and Hopwood 1977). It is usual in Great Britain to divide a year intake (of say 60 students) into learning groups of common ability (say three groups of 20 determined on some criteria of perceived academic potential -Keen and Reid 1977) and mixed attitude.

If, as the author suggests, perceptions of effective teaching styles are related to a student's attitude, rather than ability, then learning might be improved by dividing the year intake into

common attitude, mixed ability groups. This hypothesis aims to explore the validity of such a strategy.

The hypothesis is "First year undergraduate students of physics formed into groups by their score on each component factor within the attitude to physics inventory, (four factors) will exhibit a commonality of perception, when categorising observed teaching episodes against pedagogic effectiveness criteria, between common attitude groups across factors to a greater extent than between differing attitude groups within factors."

The evidence, extracted from the data listed in tables 8/29 and 8/30.

TABLES 8/29 & 8/30 ABOUT HERE

An application of the 't' statistic is of dubious validity (due to high S.D.) but such an attempt produces a t value of 1.41 which is not significant at the .05 level. Thus a null hypothesis stating that there is no difference between the data in tables 8/29 and 8/30 fails to be rejected. However, a casual glance at the data is sufficient to note that 75% of the same attitude groups do correlate significantly whilst only 25% of the mixed attitude groups correlate significantly. The high standard deviations arise from extraordinary distributions of results as is shown in Fig. 8/1. TABLE 8/29. Correlations between differing student attitude groups within factors of the attitude inventory

Student attitude groups compared	Correlation coefficient (Pearson r)
A+ with Λ-	+ .51
B+ with B-	82 *
C+ with C-	72
D+ with D-	07
Mean	.53
Standard deviation	.33

* Significant at .05 level

TABLE 8/30. Correlations between common student attitude groups between factors of the attitude inventory

:

Student groups compared	Correlation Coefficient (Pearson r)
A+ with B+	+ .36
. A+ with C+	11
A+ with D+	83 *
A- with B-	92 *
A- with C-	95 *
A- with D-	84 *
B+ with C+	+ .81 *
B+ with D+	+ .14
B- with C-	+ .95
B- with D-	+ .95 *
C+ with D+	+ .63
C- with D-	+ .87 *

Mean = 0.70 S.D. = 0.32


:

MIXED ATTITUDE GROUPS

KEY:



COMMON ATTITUDE GROUPS

So although it is not possible to specify statistical evidence to support the hypothesis under investigation, it would appear that for most respondents the view expressed in the hypothesis is correct, namely, that mixed ability common attitude groups are more likely to have an 'agreed' common idea of what constitutes effective teaching than are common ability mixed attitude groups. At the time of writing there is no institute of further or higher education using such a strategy, but mixed ability teaching is often encouraged in the secondary sector of education in Great Britain. It would be interesting to replicate this particular part of this study in such institutions by 'setting' the mixed ability groups by their attitude and attempting to evaluate any improvement in learning efficiency.

In the discussion of repertory grids in earlier chapters, an 8.15.1. example was given of a uni-dimensional component space (Fred) and a two dimensional component space. If a hypothetical respondent used x constructs in a grid of y elements and each construct was independent, then he would be construing in (x-1) dimensions if $x \langle y, or (y-1) \text{ dimensions if } x \rangle y$. (Slater 1977). One might say that a person who construes in a higher number of components exhibits a greater 'acceptance range' of for the elements under In normal parlance the uni-dimensional (oneconsideration. component) respondent might be said to have a "one-track mind." This ninth hypothesis is concerned with the acceptance range of classes of respondent and suggests that it is more likely for certain classes of respondent to exhibit more dimensions of Measuring this is difficult. The Bartlett test construction.

(Slater 1977) computes chi-squared values from the Eigen values and determines the number of components significant at the .05 level of significance. Such a technique is looking at the rate of change of slope of the Eigen value graph, as such it might classify as insignificant, numbers of components closely associated with one another merely because a big drop in Eigen value occured between the first component and the secondary Thus whilst one can be fairly certain that if the group. Bartlett test identifies x components as significant they are indeed significant one cannot be so certain that other components, as yet unmapped, are not themselves significant. To overcome this difficulty the author chose to ignore the Bartlett test in measuring 'acceptance range' and instead took the profile of the Eigen value graph. The rate of decline of which will be less when the acceptance range is greater. Having looked at the general problem and determined a way of measuring it the hypothesis may be stated.

8.15.2. Hypothesis 9.

"Positive attitude respondents from each component factor, will exhibit a greater 'acceptance range' measured by a lower rate of decline of Eigen values generated from a principal component analysis of the respondents grids with common element samples, than will the negative attitude comparison group."

There being four factors and three classes of respondent, twelve graphs may be drawn. These are included as Figures 8/2 to 8/13 inclusive.



FIGURE 8/2. Factor A, Positive vs Negative Physics Teachers







FIGURE 8/4. Factor A, Positive vs Negative Non-Physics Teachers





FIGURE 8/7. Factor B, Positive vs Negative Non-Physics Teachers







FIGURE 8/10. Factor C, Positive vs Negative

Non-Physics Teachers





FIGURE 8/12.



FIGURE 8/13. Factor D, Positive vs Negative Non-Physics

Teachers



8.15.3. A divergence between the lines of the graphs indicates differing acceptance ranges whilst the absolute slope is a measure of the acceptance range. A steep slope to a graph suggests low acceptance range (i.e. a tendency to "one-track mindedness") whilst a near horizontal line indicates a strong ability to perceive strengths and weaknesses in a variety of pedagogic styles classified by a multi-dimensional framework. Fig. 8/11 is an example of a great difference in acceptance range and here the positive group are much less able to construe clearly in as many dimensions as the negatively orientated group. Figs. 8/5, 10 and 11 indicate the negative respondents have a Figs. 8/2, 6, 12 and 13 indicate the greater acceptance range. positive respondents have a greater acceptance range and Figs. 8/3, 4, 7, 8 and 9 show no real difference. Even a cursory and subjective consideration of the data is sufficient to enable one to conclude that there is no regular pattern emerging. In some cases (8/2, 6, 12 and 13) the data supports the hypothesis, whilst in others (8/5, 10 and 11) the data rejects the From the data available one must conclude that hypothesis. there is no apparent relationship between attitude to any factor of the attitude inventory and acceptance range as defined. However, some interesting observations may be made. Where divergence does occur on the graphs it occurs almost always in the early components, thus whilst many classes of respondent do not differ greatly in the lower order components, many do differ There are other apparent considerably in the first 2 or 3. patterns but these will be discussed in paragraph 8.16. when hypothesis 10 is being considered.

8.16.1. Hypothesis ten considers the acceptance range of physics teachers, students and non-physics teachers and suggests that the acceptance range will diminish from group to group in the order specified, irrespective of attitude.

Hypothesis 10.

"Teachers of undergraduate physics will exhibit a greater 'acceptance range' (as measured in the investigation of hypothesis 9) than will the student category of respondents, who will themselves exhibit a greater acceptance range than teachers of subjects other than physics."

If supported, this hypothesis implies that if a member of each category of respondent were asked to view and assess the effectiveness of a number of episodes of physics teaching, the physics teachers would exhibit the greatest awareness of a range of different ways of being effective than would the students, who would themselves be more able to see merit in a wider range of pedagogic practices than would the non-physics teachers.

8.16.2. A review of Figs. 8/2 to 8/13 does not provide compelling evidence to support or refute the hypothesis. A recourse to a statistical comparison of mean slopes fails to detect any difference between the physics teachers and physics students although both groups have a slightly larger acceptance range than the non-physics teachers. Certainly such relationships which do exist are to small as to be completely buried in the

variation which is evident from the data. The one conclusion which the author is able to draw is that there is an unexpected variety of acceptance ranges amongst the twenty-four sub-groups So great is the variation in of respondents compared. acceptance range that the author wondered if the individuals who constitute a sub-group exhibited as much variation as Fortunately this can be tested by between group comparisons. considering the individual INGRID (Slater 1977) analysis from The task is, however, quite each respondent within groups. daunting as the data needs to be extracted from 119 separate The author chose to analyses and grouped twenty four ways! select at random just four groups to look for variation. The results were surprising; much less variation existed within groups than between groups and the kind of data exhibited in Figure 8/14 is typical of the four sample groups.

FIGURE 8/14 ABOUT HERE

It will be clear from the figure that after the divergence between the 1st. and 2nd. component the lines tend to become parallel so not only is the divergence less than one might have anticipated by the between group comparisons, but that any difference in construction resides primarily in the first component. This is subjectively verified by the construct labels used. In Table 8/31 the construct labels of respondents whose grids form the data of Fig. 8/14 are compared and the similarity between component 2 descriptions may be seen to be greater than for component 1. Such similarity is, of course,



subjective and the reader's interpretation of the labels may differ from either the author's or respondents' interpretations.

TABLE 8/31 ABOUT HERE

- 8.16.3. In relation to hypothesis 10, the author can only remark on his data, which suggests that common attitude groups have similar acceptance ranges exhibited between their members, particularly after the first component has been extracted. There is no evidence to suggest a relationship between groups of respondents as hypothesised and indeed the variation between groups appears unexpectedly high. The author has discussed these results with both respondent groups and with other researchers and is unable to suggest any reason for a wider variation between groups than within groups, in what appears to be a random way.
- 8.17. In completing the grid, the respondents were required to consider teachers of physics and to use construct: which arose from an appraisal of either observed or remembered teaching against criteria of effectiveness. The author recalled his experience in initial teacher training where student teachers frequently requested a supervisor for their teaching practice who had knowledge of their own discipline. To such students the author had always maintained that a teacher who knew little of the topic being taught was more likely to fairly assess the student on pedagogic criteria rather than discipline based criteria. (Arising from the supervisor questioning whether he

TABLE 8/31. A comparison of construct labels used by student respondents to describe components one and two.

.

	Student 1	Student 2	Student 3	Student 4		
Component 1	Competent	Talks to class	Well prepared	Effective		
	Obvious clarity of development	Regional accent	Keen on subject	Logical approach		
	Good spoken prementation	Has experience of lower levels	Commands attention	Good general manner		
	Good blackboard use	Audible	Interested in subject	Reinforcement		
Component 2	Research activity	Steady voice	Interests extend beyond physics	Generally competent		
	Level of teaching	Young	Clear chalkboard	Clear idea of level and content and method		
	Class participation	Interesting	Commands attention	tention Student participation		
	Well spoken	Audible	Clear drawings, etc.	Logical approach		

would use the same strategy to teach the topic under consideration if he were also a teacher of the same subject).

Early in the research planning it became evident that data would be available which would permit this question to be investigated in the discipline of physics, and so hypothesis ll was formulated.

Hypothesis 11

"An appraisal of the constructs used by respondents in completing their grids will show that for teachers of undergraduate physics the ratio of discipline orientated constructs to pedagogy orientated constructs will be greater than for either of the other two groups of respondent."

The constructs used by all respondents may be classified into three groups;

- a) Discipline related.
- b) Pedagogy related.
- c) Others including personality factors, etc.

The hypothesis states that the ratio a/b will be larger for physics teachers than for either of the other groups because physics teachers will use a higher proportion of physics based constructs.

In order to test the hypothesis, five judges were invited to read all the original grids. (Four of the five had been judges

previously in the Thurstone development of attitude inventory). They did not have any means of knowing whether a grid originated from any particular class of respondent, and they were asked to identify the three types of construct and write the fraction equivalent to a/b on a table of grid reference numbers. The author then computed the mean value of a/b for each class of respondent as indicated in Table 8/32.

TABLE 8/32. The ratio of discipline based constructs to pedagogy based constructs for each class of respondent.

Teachers	Students	Non-Physics Teachers
0.33	0.40	0.27

It would be inappropriate to embark on an exhaustive statistical comparison of the values in table 8/32 as they arose from subjective interpretation of construct labels by five judges who frequently disagreed with construct classification. The student group use discipline based constructs most often for a given number of pedagogy based constructs. As one might anticipate the non-physics teachers use physics related constructs least frequently.

The hypothesis is not supported by the data.

Returning to the example given in the introduction of paragraph 8.17., the comment made to the trainee teachers appears to be substantially correct as the non-physics teachers use pedagogy related constructs most frequently. It has been a surprise to most teachers of physics who have commented on this data to find that the students are more strongly discipline orientated in their choice of constructs than the teachers. The numerical data is supported by a perusal of the actual construct labels used. Frequently students cite constructs such as:

"Likes physics,"

"Well qualified in his subject,"

"Knows his stuff,"

etc., when describing an effective teacher. Teachers, on the other hand use constructs such as:

"Well structured lesson,"

"Clear plan,"

"Involves students,"

etc., when describing an effective teacher.

8.18.1. The final hypothesis sought to identify a 'word picture' of students perception of an 'ideal' or 'most effective' teacher. It will be recalled that the hypothesis demanded classification of respondents by <u>total attitude</u>, a classification which, it has been shown, would be quite inappropriate for the attitude inventory devised for the research. It has not been possible, therefore, to investigate the last hypothesis (12) although individual 'word pictures' of an effective teacher have been produced for groups of students identified as positively or

negatively orientated to one factor of the inventory. These profiles have already been considered.

8.18.2. Although it has not been possible to investigate the 12th. hypothesis it is repeated below as the author considers its investigation in the future to be of value. The evidence already presented to support the view that common attitude groups do seem to exhibit shared perceptions of effective teaching is of sufficient significance to merit the later development of a better means of measuring 'total attitude.' This is an area of work already being undertaken by a research assistant employed by Plymouth Polytechnic to continue with the development of the methodology devised and recorded in this thesis. (Further research plans are discussed in Chapter 10).

8.18.3. Hypothesis 12

"First year undergraduate students of physics formed into learning groups by their total attitude score, measured by the attitude to physics inventory, will have common perceptions of the pedagogic style they associate with effective teaching of undergraduate physics with less variance than any mixed attitude grouping." PEDAGOGIC STYLES ILI PHYSICS TEACHING : AN ATTITUDE SCALING AND REPERTORY GRID STUDY

'NEWSWORTHY RESULTS' :

A SUMMARY OF FINDINGS



9.1. A common criticism of psychological and educational research is that it too often wastes time gathering empirical support for the self-evident. Similar comments sometimes get made about the theorising that educational and psychological researchers engage in. At the end of the day, it looks as if an elaborate theoretical framework has been erected to explain what should have been obvious in the first place.

When one comes to the end of a lengthy research project, it can be quite discomforting to be told that one's hard-won findings are "obvious" or "lacking in news value." Such criticisms, once made, are not easy to slough aside. If all one's results could indeed have been predicted in advance (e.g. from commonsense considerations and/or already-existing knowledge) the onus is upon the researcher to explain why he spent valuable research time "discovering" them.

In an attempt to safeguard themselves against criticism of this kind, experienced researchers sometimes pay special <u>advance</u> attention to the "newsworthiness" (to use a term of Karl Popper's) of the hypotheses that they are proposing to test. One way of doing this is to describe the hypotheses to reasonablyinformed colleagues and outsiders - with a view to securing their opinions about the likely outcome. In effect, the aim is to discover whether there is significant agreement about the way in which the results will work out. If everybody agrees that the results are likely to take a certain form, the researcher might then think twice about the necessity of doing the experiment.

Of course, there may be circumstances in which it is desirable to do an experiment, just to make sure that it <u>does</u> work out the way in which everyone expects. But some quite strong justification would surely be needed if a whole succession of experiments failed to yield any surprises at all.

9.2. The hypotheses investigated in this research project were framed by the author from a perspective of one who had been a student of physics, a teacher of physics and was currently an interested observer from outside the arena of physics education. The results have proved interesting to the author as they have enabled him to cite evidence to support or refute views which were based on intuition when the research began. However, to the external observer there is little guidance to the "newsworthyness" of the results, to use Popper's terminology.

By using the term 'newsworthyness' the author is seeking to identify those results which are counter intuitive to the practitioner. One might illustrate a 'newsworthy' or 'counter intuitive' finding by citing Albert Einstein's clock hypothesis, which when validated by experimental procedures not available at the time, Einstein hypothesized the time discrepancy associated with real time travel. There appeared, therefore, to be a need to illustrate in some simple way whether the research findings were congruent with, or in conflict with, the views popularly held by physics education practitioners.

Some kind of 'straw-vote' conducted amongst physics teachers

seemed a suitable means of finding out what they b ieved to be the current view on the hypotheses.

The author made two attempts to evaluate this opinion. The 9.3. first was unsuccessful, for the reasons specified whilst the It should be stressed, however. second proved most rewarding. that this was not intended to be a 'scientific' sampling/data collecting mechanism, merely a simple way of finding out to what extent the author had reflected the intuitive views of physics teachers when framing the hypotheses. The first, and unsuccessful, attempt at gauging the opinion of physics teachers wasin the form of a questionnaire upon which all the hypotheses appeared with questions asking whether or not the respondent agreed with the hypothesis, and if so, to what extent. The questionnaire failed partly because respondents could not easily interpret the hypotheses when these were couched in the jargon of the research project and partly because a seven point response scale proved inappropriate for the response demanded. The attempt was completed for about 20 respondents but such difficulty was experienced that the data collected was rather suspect, and consequently was rejected.

The second attempt endeavoured to overcome the difficulties experienced previously by expressing the hypothesis in common parlance (which, it should be noted, introduced some ambiguity) and by asking respondents to estimate the probability of each statement being true. The method had been used before and the author had been sceptical of its usefulness, but having tow

applied the technique is 'converted' to an advocate of such a method. (For 'straw-vote' opinion gathering the method is ideal, in a serious study the author would still have reservations as the questioner in the interview can 'manage' the situation in such a way as to introduce bias).

Only twelve physics teachers were questioned, six from a single university, four from one polytechnic and two ex-physics lecturers now employed in the area of educational technology. Arbitrary levels were decided for cut-off points for the average probability level of 40% (do not agree) and 60% (agree). Table 9/1 indicates the way in which each hypothesis was supported, refuted or no conclusion drawn from the research data and from the sampling of intuitive views of other physics teachers. The third line indicates those hypotheses where findings are contrary to these intuitive views or, in other words, it identifies those hypotheses where a "newsworthy" result became apparent from the research data.

TABLE 9/1 ABOUT HERE

9.4. It will be noted that on 45% of the hypotheses the research data indicates that the results are in conflict with the popular views of practitioners. The previous chapter has discussed each hypothesis in relation to the data and so it is intended only to highlight in this chapter, those eight sets of results where conflict appears to exist. The data suggests

TABLE 9/1. The hypotheses in relation to data and probability estimating by a group of practising physics teachers.

ABCD are for	Hypotheses																	
factors of att.	1	1	1	1	2	2	2	2	3	4	5	6	7	. 8	9	10	11	12
inventory	A	В	С	D	A	В	С	D	A	В	С	D	A	В	С	D	A	В
Research data	*						*	*	*		*	*	*				*	
/ Supported	1	x	x	x	7	1	7	7	x	x	7	x	x	1	Ō	0	x	0
x Refuted				4													•	
0 No result																		
Straw vote													•					
/ Supported	1	x	/	0	x	/	1	/	1	x	x	1	7	x	1	·0	1	0
x Refuted																		
0 No result						•												
Findings in conflict with				•														
opinion S			S	-	S				S		S	S	S	S			S	

* Significant at 0.05 level.

that teachers of physics prize the fostering of intrinsic motivation in their students less highly than enabling their students to overcome the hurdle of an examination. The students, however, it would appear from the data, do demand a pedagogic style which places the development of 'intrinsic motivation' and 'pleasure from physics lectures' higher in importance than overcoming the examination hurdle. Most physics teachers are in conflict with the data in two respects: firstly, it does not acknowledge that teachers and students may have different priorities in these areas, and secondly, it postulates that students will want to 'pass an exam' rother than be intrinsically motivated to enjoy physics instruction. Irrespective of which attitude factors are considered the data indicates that teachers or students who exhibit different attitudes will perceive different teaching strategies (pedagogic styles) as being effective. Most practising teachers clearly do not see such a distinction and consider an effective teacher of one type of student, (classified by attitude) to be equally effective for any other type of student (classified by attitude). These observations are replicated in two separate hypotheses which consider this problem from different perspectives The author considered that physics (Hypotheses 5 and 6). teachers and physics students would tend to use discipline based criteria to assess an observed teaching episode, whilst other teachers, who were devoid of subject knowledge would tend to use more pedagogic orientated criteria. In fact the evidence available suggests that students are very much aware of aspects of pedagogic competence and are more like teachers of

other disciplines when observing and assessing a er. Physics teachers do, however, use discipline based iteria. The probability estimating technique suggests that physics teachers do not acknowledge that they may be biasing an assessment towards discipline based criteria although evidence is available to show that they do. Additionally, the author's sample of physics teachers do not acknowledge that their students are able to make competent pedagogic criticism in effecting an appraisal of an observed teaching episode.

In his former teaching training activities, the author frequently found himself having to convince students that they would be fairly assessed on teaching practice by someone they (the students) knew had no knowledge of their subject. In almost every case students would prefer to be assessed by a teacher of their own subject. Evidence collected then, as well as in this research data, clearly shows that a physics teacher viewing a student teacher of physics is likely to make statements akin to: "Well it was O.K. but I would do it this way"

The non-physics assessors would make similar comments but the difference centres on how the sentence is completed. For physicists there are usually physics based comments like ".... you could have related the wave theory of light to the propagation of an electro-magnetic wave by considering the field effects of an oscillating electron in free space."

The non physicists' comments tended to be discipline free as examplified by : " could you not have used a water tank. some demonstration or at least a diagram to illustrate your point about the wave theory of light?" (Both comments are transcribed from actual assessor reports on student teachers Whilst this did not form a major part of this of physics). research the examples are included to clarify how the intuitive statements of practical teachers is clearly quite different to the actual strategies adopted. Whilst the data is not statistically significant the trend of results supported the author's belief that mixed ability common attitude groups are more likely to demand a common pedagogic style in order to achieve what they perceive as an effective teaching strategy than are mixed attitude, common ability groups. Popular opinion did not acknowledge that attitude (however measured) had an effect on the perception of an effective pedagogic style.

9.5. Hopefully this chapter has highlighted some of the differences between the research data and the views held by some practising physics teachers. The author firmly believes that any attempt to improve the overall efficiency of physics teaching of undergraduate students would first have to resolve these conflicts and establish a common area of understanding before specific changes of strategy were recommended.

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND BEPERTORY GRID STUDY

CONCLUSION:

EXTENSION STUDIES, ETC.



10.1. On initiating this study the author's contenti that the criteria used to form student groups, the methe of instruction advocated by teachers of undergraduate provides, students' perceptions of their own needs in terms of pedagogic style and teachers' perceptions of their students' needs all required investigation. Furthermore, the author intended to provide evidence which permitted an investigation of the areas specified above, but which would also substantiate his belief that teachers of physics, teachers of disciplines other than physics and physics students all use different criteria when assessing a teacher's effectiveness.

The previous chapter has already considered how the hypotheses, when considered in the light of the data, provide some clarification of these areas of expressed concern and highlighted how these results relate to popular opinion amongst those involved with physics education. There has never been any intention to make prescriptive comments about physics education, the conclusions drawn are intended to be informative for physics teachers or educationalists who are interested in improving the effectiveness of learning amongst physics students. The 'conclusions' are really 'observations' of what was happening in physics education in the period 1975 to 1977. The reader, who is interested in improving student learning may find that these observations are illumative, as did many of the people asked for their opinion in the 'straw vote' survey of the previous chapter. Whether the research will be useful cannot be stated, but for some readers the findings will identify a
different perception of a problem which may have been approached from one perspective only. Thus, the making explicit of differing perceptions of a common problem may enable the practitioner, who is designing new and unique teaching strategies, to consider how different approaches might be construed by those who will ultimately have to experience them as both learners and teachers.

At the outset the research was not intended to be a methodological study and yet, in retrospect, the major contribution this research may be seen to have made may well be the adoption of an established technique, namely repertory grids, into an instrument for appraising teaching effectiveness. Throughout the research period the author was employed in a polytechnic with the mandate to improve the quality of teaching within the institution. Such demands resulted in the emergence of the TARGET project (Teaching Appraisal by Repertory Grid Elicitation Techniques) (Keen & Hopwood 1978) which has been developed by the author together with his colleague, and is based exclusively on the methodology developed by the author for the investigation of the research hypotheses of this thesis. The author perceives TARGET as the link between his academic research (during 1974 - 1978) as described in this thesis, and the practical arena of educational technology. Paragraph 10.2. describes the TARGET project (the whole of paragraph 10.2. has been published as a paper (Hopwood & Keen 1978) in Programmed Learning and Educational Technology 15, 3, 187 - 195) and its relationship to this research.

10.2. TARGET is an acronym for Teaching Appraisal by Repertory Grid Elicitation Techniques. The project is under development and will, ultimately, become operational at four levels, each successive level exhibiting a greater degree of sophistication. This paper is confined to a discussion of Level 1, which is already fully operational.

THE TRADITIONAL APPROACH

Teaching and learning are activities of an extremely complex nature, activities to which considerable research effort has been, and continues to be, applied. A major purpose of such research is to illuminate the nature of each and to point up relationships between them. Even so, there exist few operational guidelines available to teachers seeking ways and means of creating effective learning situations for their students. In other words, associative links between acts of teaching and modes of learning are far from numerous at the practitioner level.

Notwithstanding the tenuous nature of such links, it is commonplace, in formal courses of professional training for teaching, to evaluate teachers and to categorise their ability to teach by means of applying numerical grades, or by the use of such adjectives as bad, good, weak, strong, pass, fail and distinction, to name but a few. More often than not, such evaluations exclude a teacher's own perception of teaching effectiveness and the consumer's opinions, those of his students, are seldom canvassed.

The methods by which teachers and teaching are complex rised are too numerous to mention. They range from a writt mage of subjective comment, through check lists of "approprime" teaching behaviours, to complex and mystifying grading systems employing multiple evaluative criteria. Almost without exc eption, these instruments are used by an observer in the classroom during, or immediately after, a teaching episode.

Although the highly subjective nature of this approach is well known, there seems to be no reliable way of appraising teachers or teaching other than by observing teaching behaviours, and placing apparently arbitrary values upon their appropriateness.

Within a given teacher training unit, there may be no objection to such a process, provided that the student teachers, and their observers, accept criteria to be used for assessment purposes, especially if steps are taken to ensure compatibility of assessments made by different observers. Major problems arise, however, when attempting to compare "good" or "bad" teachers, so defined within one training unit, with their counterparts so defined within another such unit. In many ways, such problems might be expected as, in each unit, trainee teachers are invited to emulate models of "good" teaching based on criteria recommended by different groups of "experts," criteria which are often totally external to the trainees themselves.

In the United Kingdom, and also in the United States of America, there is, as yet, no statutory obligation upon teachers in

establishments of Further and Higher Education to ceive formal training for teaching. In the U.K. those : ll groups of FE/HE teachers who do present themselves for training, do so on a voluntary basis, usually following formal in-service courses for which remission of teaching duties is allowed. Such teachers accept that some quantification of their teaching is, invariably, a course requirement. For the vast majority of their colleagues, however, the rewards attached to teaching accomplishment are perceived to be less than those attracted by high levels of academic achievement in their respective subject areas. In any event, were overt attempts made to raise the "teaching climate" (Eraut 1975) of an institution, British academics and their professional associations, would be likely This resistance could take the form of to resist such efforts. drawing attention to the low reliability and doubtful validity of many of the instruments used for teaching appraisal purposes, and pointing up the plight of those of their American academic colleagues whose careers are often threatened by compulsory evaluation of their teaching ability by the use of instruments of similar dubiety.

Nevertheless, it would be wrong to suggest that teachers in FE/HE are totally unconcerned about their effectiveness in the classroom. This is probably just as well, for institutions of FE/HE are, currently, finding the need to respond to both internal and external pressures to maximise the effectiveness of teaching. It is the opinion of the author , however, that for reasons already discussed, those within these institutions

mandated to improve teaching effectiveness will a dot no credibility in the eyes of their academic colleagues by further elaboration of existing instruments, or by the new development of similar instruments, purporting to measure effective teaching.

Additionally, should professional training for teaching become a statutory requirement for new entrants to the FE/HE teaching service, those responsible for providing the formal training courses could find themselves in impossibly vulnerable positions once the element of voluntary course attendance no longer applies.

PERSONAL CONSTRUCTS

It is against the background of this somewhat depressing scenario that the emergent TARGET system could be seen to offer an alternative and novel approach to teaching appraisal. TARGET is an elaboration by Keen, Hopwood and Reid (Biles, Biles, Keen and Hopwood 1978) of methodology devised by George Kelly (Kelly 1955) and supplemented by Bannister & Mair (1968), Patrick Slater (1972) and others. The system reflects Kelly's philosophical position of constructive alternativism, that is to say, an acceptance that an individual's interpretation of the universe is subject to replacement or revision, rather than presenting a static, once-for-all construction of the universe, or of any part of it.

The TARGET system utilises Repertory Grid techniques, allowing a

respondent to provide plain language statements t interpret events in terms of discrimination, organisation and the anticipation of future possibilities. Statements of this kind are known as personal constructs, and represent interpretations imposed on events, (events associated with teachers and teaching in the case of TARGET), by the respondent. Linkages between constructs illuminate aspects of his construct system, an understanding of which allows speculation concerning his approach to the ordering and anticipation of events.

In general, a construct provides a means of identifying similarities and differences. Furthermore, it offers a continuum along which an aspect of an event may be scaled. A construct is, therefore, two-ended, or bi-polar, involving a particular basis for considering likenesses and differences and, at the same time, for excluding things as irrelevant to the contrast involved.

For example, if the construct "good diction - bad diction" is a feature of a respondent's construct system, he may use this to identify good diction and bad diction in the course of, say, observing a teacher. Furthermore, having now defined the poles of a scale relating to diction, the teacher under observation may be positioned upon it, as may be other teachers observed in the future. It is important to realise, however, that this, and other constructs, are freely elicited by the respondent, and thus take account of individual perceptions of acts of teaching.

As seen by Bannister, and by Kelly, constructs and construct systems exert a controlling influence on the individual:-

"The system of constructs which a person establishes for himself represents the network of pathways along which he is free to move. Each path can be viewed as a two-way street, and while the individual may choose either of these directions, he cannot, so to speak, strike out across country without building new constructions, new routes to follow.

When a person must move, he is confronted by a series of dichotomous choices - each choice being channelled by a Each construct represents a pair of rival construct. hypotheses, either of which may be applied to a new element Thus, just as the which the person seeks to construe. experimental scientist designs his experiments around rival hypotheses, so each person is seen as designing his daily explorations of life around the rival hypotheses which are yielded by the constructs within his system. Moreover, just as the scientist cannot foresee possibilities that he has not, in some manner, conceptualised in terms of hypotheses, so any individual can prove or disprove only that which his construction system allows him to see in terms of possible The construct system sets the limits beyond alternatives. which it is impossible for a person to perceive, and in this way constructs are seen as controls on a person's outlook and also, in an ultimate sense, as controls on his behaviour.

Kelly suggests that the controlling influence of constructs becomes particularly interesting when a person begins to use himself as an event in the context of the constructs he is When a person uses himself as a developing or operating. datum in forming new constructs, he finds that the constructs formed operate as a tight control on his own behaviour. In forming a set of constructs which include the self as an element within their range of convenience, the person plots the dimensions along which it will be possible to organise his own behaviour in relation to others. Thus a person who includes himself in the context of the construct, say, powerful-weak, binds himself to assess his own behaviour in relation to that Whether he sees himself as powerful or weak is of dimension. interest to a psychologist, but it is secondary to the fact that the person has ordered his world and himself with respect to the powerful-weak dimension.

From the point of view of personal construct theory, any person, when viewed by another, is regarded as the point of intersection of a number of personal constructs used by the observer. Just as the North Pole is defined as the point of intersection of a number of geographical constructions, so also the viewed person is to be identified as the meeting point of a series of dichotomised, categorical and personal interpretations."

Aspects of the methodology used in the adaption of grid theory to teaching appraisal, as used in TARGET, are discussed in the second part of this paper.

At the practitioner level, a respondent to TARGET receives, as an end product, three profiles in the form of bar charts. These show:

a) His perception of an effective teaching style.

b) His perception of his own teaching style.

c) His perception of an ineffective teaching style. The derivation of these profiles from grid data will be discussed in some detail later, but for the present it is sufficient to remark that not only are the respondent's own constructs printed on the profiles, but those constructs are organised into groups which best describe each dimension of the profiles. In other words, an attempt is made to depict that part of the construct system used by the respondent to categorise teachers and teaching episodes.

A significant difference between the TARGET system and other systems and methods mentioned earlier lies in the fact that criteria of effectiveness, own teaching style and ineffectiveness, respectively are elicited by the respondent himself, In this way he is freed from the burden of conformity to criteria developed by "experts," criteria often alien to his own personal construct system.

It is not claimed that the use of TARGET will produce "better" teachers but, nevertheless, respondents to a TARGET grid may well experience heightened perceptual awareness of teachers and teaching. Although there is no one-to-one relationship between perceptual competence and a complex motor skill such as teaching, there is some evidence to support the view that motor stills

cannot be acquired without perceptual enhancement, and that repertory grids are appropriate means to this end.

CONSULTANCY

It is almost inevitable that a teacher, on receiving his profiles, will immediately compare one with another. On comparing his own teaching style profile with that perceived by him to be representative of effective teaching, he may conclude that there is some inadequacy or deficiency in his own teaching style. It is at this point where the need arises for the provision of a consultancy or counselling service for, as pointed up by Thomas (Thomas & Harri-Augstein 1976)

"When a habitual skill is disrupted, performance drops and the learner becomes emotionally vulnerable. He requires support and discipline from inside or outside to continue through the trough of this process of change and to enter into a positive reconstructive phase of learning-to-learn."

Further insight is provided by Reid (Reid 1977)

"Miller's work suggests that to construct a more effective plan the learner must dissemble existing plans, often at many levels of complexity, before recombining them into an alternative structure. He will, as it were, be at some stage caught between plans. Invariably this leads to sudden and marked drops in performance, a period of depression, disinterest, anxiety, and feelings of hopelessness. At the point where the distance forwards is equal to the distance back, there are very

strong temptations to reinstate old plans and to reject new methods on the basis that they can be seen not to work. The only solution to this situation is perseverence in the learner, and tolerance and support from the trainer."

Built into any such consultancy service should be provision for skills training and, at this stage, there is a meeting point between interpretive and behavioural approaches. There is, however, a significant different between a training programme of this kind and, as mentioned earlier, programmes designed to train teaching skills considered relevant to acts of good teaching in a local setting.

In the context of training associated with TARGET, it is the teacher himself who is the "expert," and it is he who may be motivated to change, or to modify his behaviour in the light of his own perception of his teaching style relative to his own perception of effective teaching.

There are reasons to believe that, through the experience of TARGET, the teacher's potential to reconstrue himself as an effective teacher will be enhanced and, should this be the case, the consultant or counsellor is now in the position of being able to negotiate with his client an individual training package.

Such a package may assume many forms and utilise various resources, including microteaching facilities. What little experience may be cited to date suggests that, far from such facilities being irrelevant to TARCET, exciting possib.

are beginning to suggest themselves, all of which point to a more vital and realistic approach to in-service training and to in-house induction courses.

At this point, however, having considered some aspects of the origins of TARGET, its underlying philosophy and perceived potential, let us now turn to a more detailed examination of the system, insofar as it provides feedback to the respondent.

COMPLETING THE GRID .

A pack of fourteen "cue" cards is presented to the respondent who inserts names or descriptions of teachers on the first three according to printed instructions. The remaining eleven cards correspond to videotaped excerpts of teaching-learning situations which the respondent then views. The videotape is carefully constructed in order to representatively sample the construct space of the respondent, a test being incorporated into the analysis which ensures that, for any individual respondent, his cognitive structure has indeed been representatively and adequately explored by the videotaped excerpts. Whilst viewing, the respondent is invited to record on each card his observations and/or comments on the episodes to which he is being exposed. In order to facilitate recall, each card is identified in four ways. Firstly, a letter code which corresponds to a column on the grid form (see Figure 10.1.), secondly by a photograph of the teacher who appeared in that episode, thirdly by a key word linked to the topic being 'taught' and finally by the respondent's own comments. Prior to



commencing grid administration the respondent is informed that all cards must be destroyed before he leaves the room in order that his personal comments on the videotaped teaching episodes, and on the teachers named by him, will remain confidential.

The second stage of the administration process can be undertaken once the card pack is completed. The respondent is asked to write his personalised TARGET reference number (which was drawn at random from a 'pool') in the appropriate space on top of his grid form. Figure 10.1. is a reduced version of the grid form used. The completion now continues in the traditional manner for repertory grids using rating scales described in detail by Bannister and Mair (1968). The triads identified in each horizontal line, by means of shaded cells, are chosen on the basis of a pilot study which indicated that the distribution chosen facilitated the elicitation of independent constructs.

The administration session may be expected to take between $l_2^{\frac{1}{2}}$ and $2\frac{1}{2}$ hours.

THE ANALYSIS

The grid analysis package developed by Patrick Slater (1972) is used for the initial analysis. From this programme, principal components may be identified and a test applied to determine the number of components significant (at the 0.05 level) for each individual respondent. The greater the number of significant components the more likely a respondent is to be able to appraise a teaching act, (be it his own or observed) on its merits rather than against some rigid and restrictive set of criteria.

The TARGET research team investigated the relationships between the constructs, elements and these components and have been able to develop a computer programme, written by Reid, which acknowledges the mathematical relationships, operates on them, having selected the appropriate data from the initial Slater analysis, and prints a selection of teaching profiles. It is these profiles alone which are returned to the respondents as feedback.

THE PRINTOUT-FEEDBACK TO THE RESPONDENT

The reader will have realised that the analysis represents a sophisticated mathematical treatment of the grid, which itself Clearly there is a pressing need to is a powerful instrument. present the output from the operation in a manner which does not bewilder the respondent who will for the most part, be a practising teacher who may have little knowledge of, or a desire to become involved in, the interpretation of such a numerical The TARGET printout, therefore, consists of a number analysis. of simple profiles presented in a way found to be both attractive and easily interpreted by respondents. At level one there are Of these, the first consists of a five pages of feedback. statement designed to assist the respondent when interpreting the profile, and the last a list of all the constructs used by the respondent in completing the grid. It will be recalled (Figure 10.1.) that each line of the grid requires the respondent to write in his own words why two of the three cells (identified by shading) are similar and yet, by the same token, different Reference to Figure 10.2. will reveal how these from the third. constructs were written by one respondent, for this is a repro-The list of constructs on the fifth duction of an actual grid. page of the printout will, therefore, be a complete record of all the words used by the respondent when he originally completed the grid.

с. Г

FIGURE 10.2.

«	(TARGET OO5139														Γ	है	
۶.	Effective	Inelizciare	A	B	c	D	E	F	G	H		L	ĸ	PAIR 5	SINGLE		For Office Use
5	4	I	2	2	1	2	5	ī	Ī	3	2	1	4	Computere & Professional	Badly Organised	1	1
5	3	5	1	3	2	Ī	5	ŧ	1	5	4	1	5	Lack of american	throwed with Audience	2	
1	4	5	5	I	1	ī	1	5	5	5	ī	1	1	Relizze upon notes	Notes not needed	3	16
4	3	Ī	5	1	Γ	÷	3	ī	ī	2	5	5	5	Tao serious	Unclear	+	- E
I	Ī	5	3	4	3	1	1	5	t	5	2	5	2	Poor Planning	Good Planning	5	1 a
Ŧ	5	I	1	3	1	1	5	1	1	+	4	t	2	Good Modelling	and Modelling	6	E .
5	1	T	1	3	3	2	5	3	ł	2	5	5	1	Learning by Experience	Teacher Isolation	7	Pre-
5	5	4	1	2	I	1	5	3	2	ī	3	1	5	Stylish	Dead	8	Creat
I	4	5	5	2	5	3	1	3	4	ī	2	5	Ī	sumbling	Theogenful	9	4
3	2	2	4	1	5	5	3	3	4	3	5	5	I	States Obvious	Patronazing	10	Came.
5	5	2	3	3	I	1	5	i	1	5	t	1	2	Presentation	No Insentation	II	Pavle
5	F	4	i	1	I	3	4	2	1	5	4	1	3	Demonstrations	No Cemandrati me	12	- UNDER
I	I	5	3	3	2	3	I	5	5	2	Ι	5	3	Lack of Programsion	Evidence of Pressoin	13	219 11
+	4	5	3	1	I	I	3	5	I	5	3	2	2	Rushed Concepts	Leiturely Pace	14	Ce.
																	1-10Tel
																	turna F
																	ci. Lea
Ш																	Proje
																	Terge
Ш																	RN 10
\square																	22
(h)						С	14										
					PHYSICS												14
	Γ.			L	rniolos												13

It is to the middle three pages that attention should now be directed. Each of these consists of a 'profile' which itself is sub-divided into two, firstly the "weighted profile" and, secondly, the "construct profile" (See figures 10.3., 10.4., 10.5.)

FIGURES 10.3., 10.4., 10.5. ABOUT HERE

The weighted profile is a bar chart designed to give an immediate visual impression of the characteristics present or absent in the pedagogic style represented by that profile.

FIGURE 10.3.



•

FIGURE 10.4.



FIGURE 10.5.



There will be a number of bars either to the right or to the left of the central vertical line. The number of bars found to be significant in the analysis is clearly indicated. A respondent with more than four significant bars (some respondents have up to eleven) has an ability to appraise a teaching act, be it personal or otherwise, in an open and unbiased manner. However, less than three significant bars on his profile indicates a distinctly lower than average ability to recognise and appreciate differing pedagogic practices for their intrinsic worth.

Visual examinations of the three weighted profiles provides immediate feedback in terms the similarities and differences between the respondent's perception of his own teaching style

and his perceptions of effective and ineffective teaching styles respectively. These weighted profiles provide a visual and immediate comparison of pedagogic styles as perceived by the respondent. There is a need, however, to provide a descriptive analysis of the weighted profile in order to assist the respondent when interpreting the nature of any visual mismatch between profiles. The 'construct profile' is intended to do just this, at two levels.

Firstly, it enables the respondent to interpret the meaning of each bar on the weighted profile by reading his own words, the construct profile being made exclusively from the words he used in completing the original grid. No-one else can interpret his profile as only the respondent truly knows what he meant by the words used. Secondly, by reading downwards, as a list, he is able to formulate an accurate picture of the teaching styles described by the weighted profile.

Directing attention in particular to Figure 10.3. it can be seen that the respondent in this example construes in four dimensions (four bars). This represents his view of effective teaching of his discipline, Physics in this case. It can be seen that each of the four bars represents "qualities" of teaching in descending order of importance: in fact about 27% of "effective teaching" is perceived as arising from component one; 20% from component 2; 10% from component 3 and 8% from component 4, with the balance supplied by a range of other components not found to be significant.

For each bar the four constructs printed under the construct profile can be used by the respondent, and only by him, to determine a meaning for the first component. Similarly, the other bars are described by the four constructs printed alongside the appropriate bar. In every case the sequence of the four constructs, reading from the top, indicates the order in which they define the component in question. Thus, in Figure 10.3., the first component is best described by "Allows student interaction" (whatever the respondent understood that to mean), next by "Inculcates curiosity," and so on.

Experience has shown that, although respondents are able to interpret both the weighted and construct profiles with case and without formal consultancy, many (about 60%) have sacrificed their anonymity by approaching the TARGET project staff and asking for the optional consultancy service offered, and referred to earlier in this paper.

The security of TARGET is highly prized by the authors as a safeguard for potential respondents and the system devised has proved most acceptable. It should be made explicit that at no time would, or indeed could, data be made available to 'administrators' in a way which enabled respondents to be identified. Any further development of the TARGET project will continue to maintain this policy.

In conclusion, comments upon reliability and validity of grid methodology would seem to be appropriate. Reference to the

standard work on repertory grid usage, namely Bannister & Mair (1968) will indicate the special consideration that has to be given to these concepts when applied to repgrids. Tests are built onto the analysis to ensure internal reliability, and respondents for whom these tests indicate suspect responses are informed of the finding but are not provided with a set of profiles. In the region of 2% of all grids completed have to be rejected either because of a failure on the test, or due to the element sample proving unrepresentative for that respondent. Test-retest reliability is poor (0.2 and not significant) when feedback is provided but is significant at the 0.1 level when feedback is withheld, the exercise being replicated with common elements and common constructs four months later. Thus the evidence shows that exposure to the feedback does alter the respondent's perception, both of himself and effective teaching, although in common with the finding of Fransella (1970; it is the mode of definition of components which changes more readily than the number of components found significant.

Face validity is high, respondents readily agreeing that the system works. The research findings suggest this view but, nevertheless, in common with other researchers using repgrid techniques, it must be said that reliability and validity are difficult to quantify in any meaningful way. Thus, for the critic, who adopts a stance specifying the need for a polished and convincing argument based on traditional lines, there is still no satisfactory answer, save the only true test of validity, namely, inviting him to become a respondent conducting

personalised phenomological appraisal. Once experienced, the technique generates an appeal which comes from the explicit revelation of facts implicitly acceptable, but otherwise hidden from the external world of the observer.

In conclusion, such evidence that is available in these early stages of TARGET operation tends to support an initial claim made for the system, namely, that it is seen by teachers to offer a useful service. If this is the case in the longer term, it might be speculated that those concerned with the improvement of teaching effectiveness in the FE/HE sector will find in TARGET a useful tool which, at the very least, will generate conversations about effective teaching, an important first step in improving teaching climate.

10.3. It is not the author's intention to repeat the results in this chapter that have been discussed in the previous two chapters, however it is considered important to highlight the fact that many observations are in opposition to what fexpected by both the hypotheses and popular opinion. Such an example is hypothesis 3 where popular opinion considered, as did the author, that students would see overcoming the hurdle of an examination more important than reading physics for its own sake. There is clear evidence to show the opposite view is true.

One does not require the expertise of a curriculum developer to see that such a mismatch of perceptions would have major implications for the way a teacher might order and present the

material in the syllabus.

Similarly, there is evidence to suggest that pedagogic criteria of teaching effectiveness do feature strongly in the students' repertoire of criteria for assessing teacher effectiveness, and yet our educational system makes no provision, at the time of writing, for teachers and students to explore each other's differing perceptions of what role each should play in developing an effective learning environment. These are but two of the interesting and surprising results. In conclusion one might summarise the results simply by saying that popular opinion amongst teachers and students of physics does not generally reflect the true relationships.

- 10.4.1. Given the comments of the previous paragraph what lessons are to be learnt from the findings and what new questions have emerged which may be worthy of subsequent research?
- 10.4.2. The research was not designed to be, and has not evolved into, a prescriptive scenario. The author can only maintain his position in a descriptive arena and offer the results to those interested in reading them as a means of highlighting the conflicts of perception which have been identified. Many questions need to be answered but the author believes his results don't specify the answers, but show the need to ask questions about areas normally ignored. The questions can only be framed by the practitioner supplied with the information about his teaching arena unique to him.

It is with some regret that the author has not been able to provide more convincing evidence to suggest that mixed ability. common attitude learning groups demand, in order to learn more efficiently, a common pedagogic experience. His intuitive belief that this is so remains unshaken and the evidence which has become available, whilst not statistically significant still identifies this as an area where research might profitably be continued. The author, in conjunction with the approval of his institution of employment have appointed a research assistant specifically mandated to develop the TARGET project over the 1977 - 1980 period, with particular attention being directed towards this attitudinal question. The author will be directing some of his attention towards an investigation of the applicability of the results contained therein to the secondary school system and across discipline boundaries.

10.5. It would be true to report that the thesis contains interpretations of the data which raise many additional questions, many of which merit extensive further research three of which are being actually pursued by research effort on the part of the author and others, namely:

i) The attitudinal question of paragraph (10.4.2.)

ii) The applicability of conclusions to secondary schools(10.4.4.)

iii) The methodological development (10.5.)

The 'off-spin' benefit of the research has already been identified as centred on the methodology. At the time of writing, the basic methodology, developed to make explicit

pedagogic styles, is being used operationally within the TARGET project, of which the author was a director, and for which a research assistant has been appointed, and a research grant sought.

An identical methodology embellished with new videotapes of managers and salesmen has been successfully used by a company of management consultants with clients including a major European Hotor Hanufacturer.

The operational uses of the methodology as described earlier in this paragraph accepts the inherent weaknesses whilst research is being conducted by the author in conjunction with computer programming experts to overcome these difficulties (as discussed in earlier chapters). This activity has resulted in the development of a programme which, if found to be satisfactory after piloting, will enable the analysis used in the research to be conducted without many of the concerns expressed in the discussion chapter dealing with the grid analysis.

10.6. In concluding this thesis the author is painfully aware of how little progress has been made into an extremely complex area, more questions have been identified as worthy of investigation than have been answered and the author is forced to conclude this thesis by saying that far from finishing a research project he finds himself standing on the threshold of an exciting and relatively unexplored area.

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND BEPERTORY GRID STUDY

APPENDIX



APPENDIX CONTENTS

PACE NO.

Appendix 2.

APPENDIX 1

A review of attitude inventories

Appendix 1 considers a detailed review of existing inventories. An effective attitude inventory should exhibit some fundamental characteristics. Firstly, it should reflect some underlying theoretical construct. Secondly, where distinct variables can be identified from amongst its elements, these must be kept distinct from one another and not added together, (nor indeed operated upon by some other unacceptable process), to produce some meaningless total score, and thirdly, if it is used as an instrument to measure outcomes from a 'process' upon which some 'treatment' has been applied, then there must be some defensible connection with the treatment being studied and the instrument. An investigation of the psychological and educational literature on attitudes over a period in excess of half a century reveals instruments which, although designed with the objective of measuring an attitude, appear to exhibit some weaknesses. The majority of these defects result from a failure to ensure that the attitude scale does not exhibit any of the types of defect described earlier.

A lack of appreciation of what measurement means can lead to these defects. An example follows to illustrate how some theoretical defects might arise. Suppose one wishes to measure the physical property of some object, a bottle of wine, for example. We assign numbers to some attribute of the object in accordance with some set of rules. Strictly speaking we do not measure the bottle of wine, we measure some well defined and conceptually clear aspect of it. We may, for example, measure its mass, its volume and its opacity. If a researcher

hypothesises that the rate of increase in blood alcohol content is related to the opacity of the wine, the onus is on him to specify exactly what he means by these terms and how he proposes to go about measuring them. The numbers that have been recorded, corresponding to the measurements of mass, volume and opacity may be operated upon in a variety of ways. The outcomes of such operations may be useful or useless and meaningful or meaningless. Clearly, it is not possible for a measurement to be both meaningless and useful, but the other combinations are possible. One can, for instance, divide the mass of the wine by its volume to compute its density, a meaningful and useful value for both the winemaker and the consumer interested in having a means of knowing (The specific gravity can easily be found the alcohol content. from the data available and this in turn is related to alcohol Consider now the number obtained by adding together content). Such an operation yields absolutely the opacity and the volume. no information, the number is uninterpretable, meaningless and To conclude the example a meaningful but useless useless! Such a concept is MV (Mass x concept needs to be identified. Volume), which will generate a number which will become larger as It can be represented graphically, the quantity of wine increases. and is clearly meaningful but as it has no practical application or To return to attitude measurement value it is therefore useless. the warning pointed up by the example is to avoid combining together different measurements unless there is a compelling theoretical or practical reason for doing so.

An instrument developed by Harrison (1971) exemplifies a kind of defect which could be described as a lack of theoretical construct. The instrument sought to assess the attitudes of students, in a college of technology, to their course. As with the work of Genn (1970), it is clearly possible to identify and define theoretical constructs such as 'satisfaction with course' and develop suitable measuring instruments for such constructs. Harrison, however, omitted to do this. Consider some examples of his elements:

(a) 'Travelling to this college is a difficult business.'

(b) 'No fees should be charged for my course here.'

(c) 'My lecturers tell me what to learn and no mean than that.' Does agreement with (a) reflect satisfaction (It's hard to get there, but worth the effort) or dissatisfaction? Does agreement with (b) reflect attitudes to college, or the respondent's economic condition? Does agreement with (c) reflect satisfaction (I'm a dependent type and I like to be told what to do) or dissatisfaction (the course is far too narrow)? The items quoted are not isolated examples and many others could be cited.

If the user of Harrison's instrument tabulated and intercorrelated the scores, as might be done by opinion survey agencies, the items may be considered appropriate, but as soon as item scores are summed the absence of any theoretical rationale renders attitude scale theory entirely inapplicable and statistical techniques (e.g. split half reliability testing) completely irrelevant.

It is possible to question the inventory of Shah (196. r his 'Attitude to Science Course at Training College.' As · Harrison, a both Shah does not base his elements on a single construct. cases (Harrison and Shah) a high score would indicate a good match between the respondents' opinions and the scale constructors' model of what constitutes a 'good' attitude. By the same token, a low score would indicate a mismatch of perception between the scale constructor and the respondent, but an intermediate score would be meaningless and useless. These test designers should not have applied summated rating techniques, indeed any test/inventory designer should ensure that all the elements do have something in . common before the responses to separate items may be added together to produce a meaningful total score.

Defects associated with the confusion of variables are less elementary than defects associated with summated rating techniques. A review of the scales published reveals that generally some attempt is made to ensure the items possess something in common. Consider the scale constructed by Selmes (1971) and used by Wilmut (1971 and 1973). Some of the scale items are:

- (a) 'The cold dispassionate scientist is a mythical animal.'
- (b) 'Science is a fixed and clearly defined body of knowledge.'
- (c) 'There is no such thing as unprejudiced observation; every act of observation we make is biased.'

(d) 'Science is boring.'

The scale purports to measure 'attitude towards science and scientists.' This assumes that the attitude is a single, unidimensional trait. It can be argued that this assumption is

unwarranted and false, and that the single total score yielded by the scale represents a confused mixture of separate variables.

Wilmut (1971) lists twenty components of scientific attitudes identified by Diederich (1967); scepticism, faith in ability to solve problems, desire for experimental verification, precision, humility, aversion to superstition, and so on. Such a list merely serves to identify attributes that scientists have, or are presumed to have. They are <u>apparently</u> distinct attributes (just as mass, volume and opacity are distinct attributes of bottles of wine). They are not <u>necessarily</u> unrelated attributes; people's scepticism, for example, might be correlated with their aversion to superstition, (just as bottles of wine of large volume tend to be heavier than those of less volume)..

However, many of the attributes listed are likely to be psychologically distinct; humility and desire for experimental verification for example, are probably unrelated. If this argument is accepted, then it is unwise to assume that there is a single, unidimensional trait called 'the scientific attitude' which can somehow be measured by adding up scepticism plus faith plus precision plus humility. This is, however, the argument that Wilmut uses. Following his presentation of Diederich's list, he asserts: 'Thus a scale which seeks to determine the attitudes of sixth form students towards science and scientists would need to contain items which are relevant to each of these attributes.' (Wilmut, 1971).

Wilmut develops a scale based not on Diederich's list but on an amalgam of a shorter list of attitudes identified by Aiken and Aiken (1969), in their review of the literature on attitude to science, and a scale already devised by Selmes (1971). This scale contains items of four types:

- (a) items which refer to scientists and their work;
- (b) items which refer to science in general terms;
- (c) items which stated an opinion about science as a method of investigation;
- (d) items which are either emotive in form or thought to have a latent relationship with other items.

The first three types of items correspond to the categories in Aiken and Aiken's review namely: (a) attitudes towards scientists, (b) attitudes towards science and (c) understanding of the scientific method. The theoretical construct underlying the fourth is more difficult to discern:

The feature of both Selmes' and Wilmuts' Scale which makes them unacceptable to the author is that all the items are treated as if they were contributing to a unidimensional trait, i.e. a student's attitude to science and scientists' is the sum of his responses to all the items in the instrument. Once again the same argument is used, namely that the way to measure an attitude is to identify attributes and then add them all together. The implicit assumption is that attitudes to science, attitudes to scientists and understanding scientific method are all just aspects of a single variable. This is of course simply not the case.

There are some strategies which could be adopted to minimize the For instance the data would be more risk of defects of this kind. valid if at the original questionnaire design stage Thurstone type In this technique, judges are presented with a scales are used. single scale description, together with all the items supposedly belonging on that scale. Thurstone type procedures are then applied (even though the scales might ultimately be scored using The judges rate each item, often on an Likert type procedures). ll point scale, on a continuum running from an extremely favourable attitude to an extremely unfavourable attitude. (The rating refers to the strength of the attitude expressed by the item, and not the judges' personal opinions). When this procedure is carried out, ambiguous or irrelevant items tend to be given either a wide range of ratings or a fairly neutral median rating. Such items can be The same procedure is then used for each of the other eliminated. Items which survive can then be randomly assorted within scales. a single instrument; the separate scales are of course scored Gardner (1972) has used this technique in separately. constructing scales measuring various aspects of attitudes to Both the sorting procedure and the rating procedure physics. require an explicit statement of the underlying construct that is A test of conceptual clarity is the inventory being measured. designer's ability to communiate his constructs to his panel of judges. and his ability to write items which reflect those constructs. Should he fail in this task the resulting range of 'scores' assigned to each element by the panel of judges indicates the unsatisfactory nature of the instrument.

A second strategy might be to use factor analysis which can then provide a further basis for allocating items to scales, for verifying the uniqueness of the various constructs, or for reducing the number of scales when some of them are shown to be redundant.

If empirical evidence were available to show that attitudes to science, attitudes to scientists and understanding are all strongly intercorrelated, then addition to form a single unidimensional scale would be justified. In the absence of such evidence, the variables should be measured separately and not added together. Two strands of evidence are available which suggest that these variables are not strongly intercorrelated, thus supporting the contention that they should be kept apart. The first strand is derived from the work of Jackson (1968) who has reviewed the literature on the relationship between the achievement of elementary school pupils and their attitudes to school. 'Commonsense' might suggest that achievement ought to be correlated with attitude. Surprisingly, the correlations usually turn out to be exceptionally low. Thus Wilmut's assumption that poor understanding might be expected to lead to emotionally unfavourable attitudes is questionable.

The second strand of evidence comes from an investigation (Gardner, 1972) in which Australian high school physics students were measured on four attitude variables; attitude towards discovery learning, attitudes towards the openness of physics (Whether the student saw physics as flexible and dynamic, or static and unchanging), attitude towards scientists (eccentric or normal) and enjoyment of physics. Each of these items' attitude was
measured, by a separate ten-item Likert type scale; the items were randomly assorted within the 40 item instrument. The corrected split-half reliability coefficients for the four scales lay in the range .75 to .89; the six intercorrelations between the scales lay Thus, while the various attitudes were between .24 and .35. correlated with one another (e.g. pupils who enjoy physics tend to regard scientists as more normal), the fact that the intercorrelations are very much smaller than the reliabilities indicates that the variables are quite distinct. Factor analysis of the item responses provides further evidence of the uniqueness of the four Here, then, is clear evidence that attitudes towards scales. scientists and attitudes towards the study of a scientific subject are not the same thing.

The Selmes scale is not the only one to attempt such a reduction. For example, in an instrument entitled Attitude Toward Science (Shah, 1962), some of the items reflect interest in science ('I like science as a hobby') while others reflect the respondent's attitudes to the significance of science in society ('I feel that science has grown up with the primary purpose of serving humanity'). The research evidence shows that these are distinct variables which ought to be measured separately and not mixed together into a single score. Later research using other scales (Nuttall, 1971; Ormerod, 1971 and 1973) shows that these variables are correlated, but are certainly not identical.

A further example is provided by the work of Gaster (1972), who conducted an investigation of sixth-form science, mathematics and

technology courses and administered a number of inventories, including a Science Motivation scale. The scale consisted of The first section presented a list of three distinct sections. 26 occupations of which 14 were applied scientific/technical; respondents were asked to tick all that they felt would be The second section consisted of 12 multiple-choice interesting. cognitive items requiring students to demonstrate understanding of aspects of science in daily life, e.g. 'Which contributes most of the efficiency of a thermos flask? Glass/silvering/vacuum.' The third section (hobbies and influences) contained eight assorted items asking about the respondent's interests, his father's occupation and other matters. Scoring procedures were not explicitly described, but apparently the single Science Motivation score was obtained by adding the number of scientific occupations ticked, the number of "correct" cognitive items and the number of science-based responses in the 'hobbies and influences' section. The rationale for such a procedure is unclear. The researcher was interested in measuring creativity as a criterion variable; the Science Motivation scale, it is claimed, measures 'dominant motivation factors which might have influenced the subjects' performance in the creativity aspect of the battery.' Simply combining cognitive and affective items and assuming that this combination measures motivation represents dubious psychometric technique.

A final illustration here is provided by the work of Bollen (1972) who devised an instrument to measure teachers' attitudes to the 'philosophy' of the Nuffield primary science course. In many

respects this is a good scale; the items are well written and there has been a clear attempt to relate their content to an explicit rationale.

The rationale consists of a number of value judgements held by developers of the Nuffield course. The key phrases in these value judgements are:

- (a) The teacher does not need to be a subject specialist.
- (b) He should not follow a set syllabus.
- (c) The children should be allowed to work in groups.
- (d) Investigation should be by practical means wherever possible.
- (e) The object should be to promote an attitude to open ended enquiry.

(f) The teacher should have faith (in the approach). This collection forms a 'philosophy,' but it does not reflect a single theoretical construct. There are at least two dimensions underlying this collection; belief in child-centredness, and belief in subject specialism. Statements (b) to (e) could probably all be regarded as aspects of the child-centred vs. the teacher-dominated dimension. However, the first value judgement reflected in the item 'A teacher must have a good scientific background' hardly belongs on the same dimension. A teacher with a good Nuffield attitude is supposed to disagree with the item. Yet it seems perfectly rational for a teacher to believe in a child-centred science curriculum and still believe that a good science background is helpful for teaching it. In other words, it is possible to believe that subject specialists make better teachers, or that child centred

education is valuable, or both, or neither. It follows that teachers' views about subject specialism and their views about childcentred education ought to be measured using separate scales; they are not reducible to a single score.

The third, and final, class of defect described earlier relates to the need for an instrument to exhibit some connection with the treatment being applied if the outcome is hypothesised as being related to the treatment. If this study were to make prescriptive comments indicating what teachers should <u>do</u> to change their students' attitudes, then the measure of attitude should exhibit some defensible relationship with that which the teacher is advised to 'do.' (i.e. the treatment). This kind of defect is less crucial in this research which does not conform to the 'experimental prescriptive' type of project, and yet the possibility of this kind of defect needs to be considered in evaluating existing instruments for possible use by the author.

Much research in the social and medical sciences conforms to the model the author has called 'experimental prescriptive,' which may be simply defined as apply treatment X and measure outcome Y.

In any experimental situation, an unlimited number of treatments could be applied and an unlimited number of outcomes could be studied. How, then, does the researcher decide what to do? Frequently, but not always, he is guided by some sort of theory. For example, if treatment X is add fluoride to town water supply and outcome Y is incidence of tooth decay, the two apparently

unrelated entities are connected by a theory which tells us how fluoride salts are incorporated into tooth enamel. If a researcher comes up with the outlandish proposal to study the effects of fluoridation upon the incidence of ingrown toe-nails, the onus is upon him to justify and explain why he thinks there might be a connection between the two.

Unfortunately, many educational experiments belong to the fluoridation/ingrown toe-nails category. It is often less obvious that they do, because the investigator can frequently mask the lack of connection between treatment and outcome by the use of plausible language. For example, in the study by Wilmut (1971 and 1973), the treatment was 'Huffield A-Level physical science' project work, and the outcome was attitudes to science. This is not only plausible, but eminently reasonable until one looks at the instrument used to measure the outcome.

It is important to note that a Likert-type scale is nothing more than the sum of its parts; unless the experimental treatment affects students' responses to a large number of the individual items, it will have no effect upon their attitude score. If the scale contains a large number of items which in no way relate to the experimental treatment, then it is highly unlikely that significant treatment effects will be found.

To persue this kind of defect a little further, we must look at five more items from Wilmut's scale.

(a) 'It is not possible to study human beings scientifically.'

(b) 'Women scientists are less feminine than other women.'

- (c) 'The basic moral rule of a scientific society is simple; mutual respect, intellectual honesty, and good will.'
- (d) 'The most important advances in science are made by a few outstanding men.'

(e) 'Scientists have proved that God does not exist.'

Since the research was specifically concerned with the effects of physical science project work, it is fair to examine each item in the criterion instrument carefully and ask, what is the connection between project work and responses to this item? Are the treatment and the measured outcome manifestly related? Are the objectives of project work - which Wilmut discusses fully in an early chapter of his 1971 work-related to students' responses to these items? Does project work in physical science claim to alter students' views about the applicability of science to the study of human behaviour, about the femininity of women scientists, or about the fundamental morality underlying a scientific society? Is there a theory tying all these elements together? One suspects 'that the answers to these questions are all 'no.'' Since the scale is composed of many more such items, it is hardly surprising that no significant differences were found between the experimental and the control Partly this is a result of the confounding of variables groups. within the scale, but not entirely: even if the criterion instrument were divided into separate, well-defined scales, the question of the relationship between the treatment and outcome would still remain.

In the light of this discussion, it is perhaps pertinent to note that scale validity is not a property inherent in a scale; it depends very much upon the purpose for which the scale is used. This is considered further in the section dealing with the validity of the author's scale. A spatial relations test may be valid for measuring the effects of play with blocks and construction kits but quite invalid for measuring the effects of an audio-visual language course. Dr. A may construct an attitude scale which he claims is valid but this does not mean that Dr. B should use it in studying treatment effects unless the theoretical construct being measured by the scale can be shown to be related in some way to the treatment.

The previous paragraphs may be considered to be excessively critical. It was not the author's intention to give the impression that the area of attitude measurement has not been adequately considered elsewhere, on the contrary, attitude scaling is an area of psychological research which has been extensively studied. However, before adopting an instrument already in existence for a new research project, there is a need to carefully screen the available inventories. As in this case many often have to be rejected. However, some instruments pass the screening and emerge as possible scales for adoption which can then be considered in relation to the specific research needs. The author found eight such scales.

(a) Wilson-Patterson Attitude Inventory (WPAI)

(b) Minnesota Student Attitude Inventory

(c) Pupil attitude measure. (Studies in inter-action ... analysis. Flanders 1960).

- (d) Attitudes of Sixth-forms to School (Josephs 1973)
- (e) Science Attitude Questionnaire (Skurnick & Jeffs 1971)
- (f) Physics attitude index (PAI) (Gardner 1974)
- (g) Physics classroom index (PCI) (Gardner 1974)
- (h) Personal Preference index (PPI) (Gardner 1974)

Each of the eight potentially useful measures identified in the previous paragraph were then considered in relation to their possible adoption, in whole or part, for the research.

In the six years since its first publication, Wilson-Patterson Attitude Inventory (WPAI) has been widely used in research projects around the world and in consequence a large amount of information on its practical application has been assembled. Certain rather dated items in the questionnaire have been amended and others have been revised to facilitate understanding or interpretation. In addition to general conservatism, a number of other attitude scales have been identified by factor analysis. Two are othogonal principal component factors; conservatism-liberalism and realismidealism. The others are oblique primary factors including: militarism-punitiveness, ethnocentrism-intolerance and religionpuritanism.

The Conservatism Scale remains incorporated in the inventory but there is now a rationale for scoring other attitude scales from which a much more detailed picture of the attitudes and belief patterns of an individual or a group may be obtained.

In terms of attitude dimensions this is possibly the best measure available, however, this research does not demand the identification of these factors.

Thus notwithstanding the quality of the inventory it clearly was not suitable for adoption by the author for this research.

A review of the 89 elements of the Minnesota Student Attitude inventory together with comments of the many researchers who have used this measure provides evidence to support the value of this There are some disadvantages not inventory as a research tool. the least of which is the time it takes to answer the 89 elements! The measure is strongly 'American' in style and requires 'anglicising' before application in Great Britain. The inventory is capable of giving an indication of attitude to a particular teacher as well as attitude to school, this is potentially useful as the implication is that those elements relating to each factor could This 'editing' of a measure together be isolated and used alone. with the need to anglicise the questions would necessitate pretesting and validating the measure, and in consequence exhibited little potential in terms of resource saving over the design of a new instrument.

Flanders, in common with the author, was particularly interested in the effects of the teacher on attitude formation. He constructed a 110 element attitude scale from which he measured student attitudes to school, subjects and teacher which he called the Pupil Attitude Measure.

In common with the MAI this instrument required anglicising and was time consuming to complete. Certain elements (particularly from the attitude to science section) exhibited high face validity for this research and were identified as potentially useful elements in the construction of a new instrument should the review of existing instruments fail to identify a measure suitable for direct adoption by the author.

Josephs who, with Smithers at Bradford, worked on Syllabus-bound and Syllabus-free student orientations, developed in his M.Sc. dissertation a very useful short (30 element) attitude measure for sixth-form pupils. (Attitudes of 6th. forms to school). A few, about 20%, of the elements are non-specific to the 6th. form and could therefore be used in a new questionnaire. This measure is particularly valuable in that it demonstrates that attitude measures do not have to be lengthy to be valid, reliable and useful.

Skurnik and Jeffs Science Attitude Questionnaire is also short. The factors which can be identified are:

- (i) Science interest
- (ii) Social implications of Science
- (iii) Learning activities. Although these relate to Science in general most would appear to be applicable specifically to Physics. The section deals with the students' preference for techniques such as reading, being lectured at, doing practical work, etc.

(iv) Science Teachers. Although this section deals adequately with an attempt to measure the pupils attitude to his

teacher it does not represent any substantial improvement over the sections of the general attitude inventories reviewed which purport to measure attitude to teachers. It has exhibited a high reliability coefficient of 0.81. The weakness in relation to this research is in the desire to quantify a student's attitude to <u>his</u> teacher as opposed to his perception of an 'ideal' hypothetical teacher.

(v) Attitude to School. All three items used to measure this factor also appear in the other inventories considered.

All three of Cardner's instruments, the Physics Attitude Index, Physics Classroom Index and Personal Preference Index (Gardner 1974) can be considered together. These instruments are excellent examples of scale development, but nevertheless quite inappropriate for direct 'adoption' for two reasons:

- (a) The instruments have been devised for use with school children and in consequence are inappropriate for undergraduate students and adults.
- (b) The instruments were devised for use in the Australian culture and no evidence exists to suggest that a direct transfer to the British culture would prove acceptable.

APPENDIX 2

The Attitude Inventory

.

			· · ·	
YOU FOR AGREEING TO ASSIST BY REPLYING TO THIS QUESTIONNAIRE.		1/KSS/AV/A.	The best way to learn about Physics is to listen to locturer, who knows a great deal about the subject, and who is a good teacher, and summarize the facto and laws he tells you.	
B READ THE INSTRUCTIONS NOY.		AA+	5 4 3 2 1	
UCTIONS.	• .	1/253/17/4.	By Physica lecturers frequently encourage no to do extra work in Physica which I do alwaya or alacut alwaya.	
This is not a test or exam, there are no 'correct' answers so please rark the response which applies most readily to you.		AE+ 1/741/24/AV	54321 I would aniow having a vacation job in a Physics	
Please write your name below in the space indicated. No one that you know (including your teacher/lecturer and other		DC+	Inboratory. 5 4 7 2 1	
students) will see your reply.			I like getting as much fun as I can out of life, even if it means neglecting my private studies in Physica	
(iii) Read each of the following statements then reply by placing a line through the number (5, 4, 3, 2 or 1) which most			5 4 3 2 1	
readily applies to you. Please reply to all the statements.		1/101/39/Att.	advice for real life situations as well as the forund academic content.	
The response is	•	1/PC1/KSS.	"All work and no play" sytly describes my Physics	
		DD	5 4 3 2 1	
5 If you strongly agree with the statement.			/A. We spend a lot of time in Physics leasons going ove examplers, I enjoy this.	
1 If you mildly agree with the statement.		AC+	54321	
3 If you are not sure if you agree or not.			Good performance on tests is really recognized in my Phymics group, I like this. 5 4 3 2 1	
If you mildly disagree with the statement.	÷	1/X55/9.	It is reasonable to EAK students to learn laws, but they should not be told the laws before they do	
If you strongly disagree with the statement.	•	- BC+	experiments, they should try to discover the laws for themselves first. 5 4 3 2 1	
ŢŢ	•	1/TA1/18. CE+	I like the way Physics challenges us to find out why things happen. 5 4 3 2 1	
	÷.	1/845/6.	A Physics experiment is best when the instructions are all laid out in the toxtbook.	
		Bi-	54321	
Score:		1/PC1/34. Da-	In Physics, we seem to dither about for a long time without learning very much. 5 4 3 2 1	
	YOU FOR ACREEING TO ASSIST BY REPLYING TO THIS QUENTIONNAIRE. <u>A RAD NHE HISTRUCTIONS NOW</u> . <u>METHODS</u> . This is <u>not</u> a test or exam, there are no 'correct' answers so please rark the response which applies nost readily to you. Please write your name below in the space indicated. No one that you know (including your teacher/lecturer and other students) will see your reply. Read each of the following statements then reply by placing a line through the number (5, 4, 5, 2 or 1) which most readily applies to you. Please reply to <u>all</u> the statements. The response is If you strongly agree with the statement. If you mildly agree with the statement. If you mildly disagree with the statement. If you mildly disagree with the statement. If you strongly disagree with the statement.	You FOR AGREEING TO ASSIST BY REPLYING TO THIS QUESTIONNAIRS. <u>A RAD END HISTRUCTIONS HOW</u> <u>UCTIONS.</u> This is not a test of exam, there are no 'correct' answers so please mark the response which applies nost readily to you. Please write your name below in the space indicated. No one that you know (including your teacher/lecturer and other students) will see your reply. Read each of the following statements then reply by placing a line through the number (5, 4, 5, 2 or 1) which most readily applies to you. Please reply to <u>all</u> the statements. The response is If you strongly agree with the statement. If you mildly egree with the statement. If you mildly disagree with the statement. If you mildly disagree with the statement. If you strongly disagree with the statement.	YOU FOR AGREEING TO ASSIST BY REPLYING TO THIS QUENTIONNAIRE. A* YARAD THIS HISTRUCTIONS NOW. V/XS/AV/A. NATHORNAL V/XS/AV/A. This is not a test or exan, there are no 'correct' answers No. So please rark the response which applies nost readily to you. V/XS/AV/A. Please write your name below in the space indicated. No one No. students) will see your reply. Need each of the following statements then reply by placing No. a line through the number (5, 4, 3, 2 or 1) which most V/rei/Ys/AV. Please reply to all the statements. No. The response is No. If you strongly agree with the statement. V/rei//A. If you are not sure if you agree or not. D* If you strongly disagree with the statement. No. If you strongly disagree with the statement. D* If you strongly disagree with the statement. No. Score: No.	

.*

•.

i

.

335

	1/721/26. CA-	Learning about the sort of thing that anyone can see happening makes Physics toring for re. 5 4 3 2 1	
,	1/241/33. ED+	In Physics there are any problems for which physicists at present do not even know the procedure for obtaining the answers. 5 4 3 2 1	
			· •
	1/271/50/71	. The heat Physics teacher would be the one who tells you as many facts about Physics as peosible, rather than concentrating on explaining principles.	•• -
	Al:	5 2 3 2 1	
	1/XSS.	I like spending my tire thinking about and discussing to mplet problems in Physics. 5 4 3 2 1	
•		5152.	
• .	1/PA1/5. CC+	Physics is worth doing whother you want to be a scienti- ot or net. 5 4 3 2 1	
-		The second second state to be able to first over	
•	1/PA1/36. BC+	widely accepted ideas and think without restriction. 5 4 3 2 1	
	•		
	1/FA1/2. CH+	I cajoy reading books about Physics. 5 4 3 2 1	
336	1/P11/17. HJ-	I don't see why a Physicist would want to spend most of his life doing experiments and analysing the results. 5 4 3 2 1	· ·
	1/PA1/16. PI:+	A Physics experiment is good when students are given ap- paratum and told to solve a problem without instructions of how to do it. That' a real experiment. 5 4 3 2 1	
	1/221/19/2. Až-	When you learn a new law in Physics it is very important to thoroughly learn the statement of it so that you know it word perfectly, rather than concentrating on the 'meaning'. 5 4 3 2 1	
	1/501/20		
	DF+	discussions that have coon sparked off by seasthing the teacher has introduced into the lesson. 5 4 3 2 1	
	1/PC1/1. CI+	I like potting difficult goals for myself in Physics. 5 4 3 2 1	
		· · · · ·	• •

•

-

...* · . ··

•

• . · ·

• ..

PEDAGOGIC STYLES IN PHYSICS TEACHING: AN ATTITUDE SCALING AND BEPEBTOBY GRID STUDY

BIBLIOGRAPHY

ITEMS MARKED WITH AN ASTERISK ARE OF INTEREST, AND RELEVANT, BUT ARE NOT SPECIFICALLY CITED WITHIN THE THESIS.

*ADAMS, G.S. (1973) "<u>Measurement and Evaluation in Education</u>, <u>Psychology and Guidance</u>." New York: Holt, Rinehart and Winston.

*ADCOCK, C.J. (1959) "Fundamentals of Psychology." Metheun.

*AICKEN, F. (1975) "Can Television Break the Monotony of Science Education?" Visual Education.

AIKEN, L.R.J. & AIKEN, D.R. (1969) "Recent research on attitudes concerning science." Science Educ., 53, 295-305.

*AMERICAN INSTITUTE OF PHYSICS (1960) "Physics in your high school." McGraw-Hill, New York.

AMIDON, E. & FLANDERS, N.A. (1963) "<u>The Role of the Teacher in</u> <u>the Classroom</u>." Paul S. Amidon & Associates, Minneapolis.

*ANDERSON, G.J. & WARLBERG, H.J. "Classroom climate and group learning." <u>International Journal of the Educational</u> <u>Sciences 2</u>, 1968, 176-180.

*ANONYMOUS. "Still Useful Despite Disillusionment."

Education 14, Feb. 1975.

APPLEBEE, A.N. (1976) "The Development of children's responses to repertory grids." <u>Brit. Journal of Soc. and Child</u> Psychology. Feb 1976. Vol 15(1), 101-102.

*ARDREY, R. (1966) "The Territorial Imperative." Collins.

ARBIB, P. & HANSCOMB, J.R. (1972) "<u>An Audio-tutorial Approach</u> to <u>Physics</u>." Sydney, University of New South Wales.

ARMIDON - see AMIDON

*ASHTON, B.G. & MEREDITH, H.M. (1969) "The attitudes of sixthformers to science and scientists." <u>S.S.R</u>. 51, 174.

AYLGREN, A. & WALBERG, H.J. (1973) "Changing attitudes towards science among adolescents." <u>Nature 245</u>, 187-190.

BAILEY, G.D. (1977) "Self-made classroom observation instruments:

An aid to self-assessment. Educational Technology.

BALES, R.E. (1950) "<u>Interaction Process Analysis</u>." Addison Wesley, Reading, Pa.

BANNISTER, D. (1962) "The nature and measurement of schizophrenic thought disorder." <u>Journal Ment. Sc. 108</u>, 825-842.

BANNISTER, D. & FRANSELLA, F. (1977) "<u>A manual of the repertory</u> grid technique." Academic Press, London.

BANNISTER, D. & MAIR, M. (1968) <u>The evaluation of personal</u> <u>constructs</u>. Academic Press, London & N.Y.

*BEARD, R. (1970) "<u>Teaching and learning in higher education</u>." Penguin.

*BENNETT, D.J. & J.D. (1971) "Making the Scene" in "<u>School and</u> Society." R. & K. P.

BERNARD, C. (1865) Introduction of L'Etude de la Medecine Experimentale. Paris 1865 p.148.

BILES, BILES, KEEN & HOPWOOD (1975 oral presentation) (1978 published) "Traditional methods of teacher appraisal; Target and Idea." <u>Proceedings of the</u> <u>Third International Conference: Improving University</u> <u>Teaching</u>, 4, paper D468. University of Maryland U.S.A.

BILLING, D.E. & FURNISS, B.S. (1973) "<u>Aims, methods and assessment</u> in advanced science education." London Heyden, 169

BINNER, P. (1958) "Permeability and Complexity: Two dimensions of Cognitive Structure and Their Relationship to Behaviour." Unpublished Ph.D. Thesis. University of Colorado.

*BLOOM, B.S. "The Toxonomy of Educational Objectives, Handbook 1. Cognitive Domain." Longmans 1956.

BOLLEN, F.A. (1972) "Attitude assessment in science teaching." School Science Review 53, 217-35.

*BONEAU, C.A. (1960) "The Effects of violations of assumptions underlying the t-test." <u>Psycho. Bulletin</u> 57, 49-64.

*BOOTH, N. (1975) "The impact of science teaching projects on secondary education." Trends in Education.

*BOSWORTH, D.P. (1971) "Investigation of the attitude towards science by children taught using a programmed approach." Edul. Dev. 11 no. 2, 16-20.

*BOX S. & FORD J. (1967) "Sociological approach to measurement."

Sociology 1 No. 3.

*BOUD, D.J. (1973) "The laboratory aims questionnaire - a new method for course improvement?" <u>Higher Education 2 (1)</u>, 81 - 94.

1. 1 .

*BOULIND (1957) Paper presented at conference. <u>Science Masters</u> <u>Association</u>.

*BRADLEY J. & HUTCHINGS, D. (1973) "Concepts of science and scientists as factors affecting subject choice in secondary schools. <u>School Science Review 55</u>, 8-15.

*BRANDWEIN, P.F. (1951) "Selection and training of future scientists." <u>Science Educ</u>. 35, 251-253.

*BRIDGHAM, R.G. & WELCH, W.W. (1969) "Physics enrolments and grading practices." <u>Journal of Research in Sc</u>. Teaching 6, 44-46.

*BROOKS, M. & HOUNSHELL, P.B. (1975) "A study of science achievement." <u>Journal of Research in Sc. Teaching</u>, vol 12, no 2, 175-181.

BROWN, S.A. & DAVIS, T.N. (1973) "The Development of an attitude to science scale for 12 - 14 year olds." Scottish <u>Ed. Studies</u>, vol 5, 85-94.

*BRUNER, J.S. (1966) "<u>Toward a Theory of Instruction</u>." Harvard University Press, U.S.A.

*BURMESTER, M.A. (1952) "Behaviour involved in the critical aspects of scientific thinking." <u>Science Education</u>, 36, 259-363.

*BUTCHER, H.J. (1969) "An investigation of the 'swing from science.'" <u>Rsch. in Ed. 1</u>, 38-57.

*BUTZOW, J. & PARE, R. (1972) "Physical Science: a multi-media facilitated course." <u>Journal of College Sc. Teaching 2</u>, no 1, 29-32.

*CAMPBELL, D.T. (1947) "The generality of social attitudes." Unpublished Ph.D. Thesis, University of Calif, U.S.A.

*CAMPBELL, D.T. (1963) "Social attitudes and other acquired behavioural dispositions. In S.KOCH (ed). <u>Psychology; a study of science vol. 6</u>, New York, McGraw Hill.

*CARPENTER, R. (1963) "A reading method and an activity method in elementary science instruction - comparison." Science Education <u>44</u>, 256-258.

CASHDAN, A. & PHILPS, J. (1979) "Teaching styles in nursery education." <u>Open University SSRC Project</u> (current).

*CATTELL, R.B., EBER, H.W. & TATSUOKA, M.M. (1970) "<u>Handbook for</u> <u>the sixteen personality factor questionnaire</u>" (16PF) Institute for Personality and Ability Testing U.S.A. (Printed and distributed in G.B. by NFER, Slough).

*CHAMPAGNE, A.B. & KLOPFER, L.E. (1974) "Formative Evaluation in Science Curriculum Development." <u>Journal of Research</u> in Science Teaching vol 11, no. 3, 185-203.

CHEIN, I. (1948) "Behaviour theory and the behaviour of attitudes." <u>Psychological Review 55</u>, 175-188.

*CHILD, D. (1970) "<u>Essentials of Factor Analysis</u>." Holt, Reinhart & Winston.

*CHILD, D. & SMITHERS, A. (1973) "An attempted validation of the Joyce-Hudson scale of convergence and divergence." British Journal of Educ. Psych. 43 (1) 57-61.

*COCHRAN, W.G. & COX, G.M. (1950) "Experimental Designs." New York, Wiley.

*COHEN, A.R. (1964) "<u>Attitude change and social influence</u>." Basic Books Inc. New York, Chapter 8, 121-128.

*CONNOR, D.V. "A study of problem-solving in physics." <u>Australian Journal of Higher Ed</u>. 1967, 3(1), 56-66.

COOK, S.W. & SELLTIZ, C. (1964) "A multiple indicator approach to attitude measurement." <u>Psych. Bulletin 62</u>, 36-55.

C . . .

*COOK, T.D. (1970) "<u>Changing attitudes to the teaching of history</u> <u>in schools</u>." Unpublished M. Litt dissertation. University of Lancaster.

*COOLEY, W.N. & REED, H.B. (1961) "The measurement of science interest: an operational and multi-dimensional approach" <u>Science Educ</u>., 45, 320-326. COON, L.H. (1969) "Attitude: a rating scale for calculus." <u>The Journal of Educ. Research</u>, 63(4), 173-177.

*COTTRELL, T. (1972) "Subject barriers broken down at Stirling." <u>The Times Higher Education Supplement.</u>

*COX, D.L., JOHNSON, P.E. & CURRAN, T.E. (1970) "<u>Perceptual</u> <u>structure of scientific knowledge</u>." University of Minnesota, School of Education, Minneapolis, Minnesota.

*CRABTREE, J.F. (1967) "A study of the relationship between score, time, I.Q., and reading level for fourth grade students using programmed science materials." Science Ed. 51, 298-304.

CROMWELL & CADWELL (1962) Quoted in Fox. <u>The Educational Process</u> <u>in Action</u>. Holt, Rinehart & Winston Inc.

CRONBACH, L.J. & MEEHL, P.E. (1955) "Construct validity in psychological tests." <u>Psych. Bulletin 55</u>, 281.

CRONBACH, L.J. & SNOW, R.E. (1977) "Aptitudes and Instructional Methods." John Wiley N.Y. Halsted Press.

*CROPLEY, A.J. (1960) "Creativity and intelligence." Br. Journal of Ed. Psych. 36, 259-266.

*DEDIJER, S. "Models of Science for Science Policy." <u>The</u> <u>Advancement of Science</u>, 1968, 24(122), 498-506.

DIEDERICH, P.B. (1967) "Components of the scientific attitude." Sc. Teacher 34, 23-4.

DOOB, L.W. (1947) "The behaviour of attitudes." <u>Psychological</u> <u>Review 54</u>, 135-56.

*DREVER, J. (1952) A Dictionary of Psychology. Penguin.

*DREW, D.E. (1974) The College Science Improvement Program: an empirical evaluation. <u>Research in Higher Ed.</u>, 2(1), 25-43.

*DWYER, F.M. (1975) "The Relationship Among Theory, Research and Practice." Audio-Visual Instruction.

*EAGLY, A. & MANIS, M. (1966) "Evaluation of message and communication as a function of involvements." <u>Journal of Personality and Social Psychology</u>, 3, 483-85.

*EBEL, R.L. (1938) "What is the Scientific Attitude?" <u>Science Educ. 22</u>, 75-81.

EDWARDS, A.L. (1957) <u>Attitude Scale Construction</u>, New York: Appleton Century Crofts. EDWARDS, A.L. (1957) <u>Techniques of Attitude Scale Construction</u>. New York: Appleton Century Crofts.

ELLIOTT, W.E. (1971) "Perceptions of high school physics and physics teachers." <u>The Physics Teacher 9</u>, 33-36.

*ELTON, L.R.B. (1968) "The assessment of students - a new approach." Universities Quarterly 22(3), 291-301.

*ELTON, L.R.B. (1969) "The making of physicists." <u>Physics</u> <u>Education</u> 4(4), 236-244.

*ELTON, L.R.B., HILLS, P.J. & O'CONNELL, S. (1970) <u>Self-teaching</u> <u>Situations in a University Physics Course</u>. Institute for Educational Technology, University of Surrey (unpublished), 23.

ENGLISH, H.B. & ENGLISH, A.C. (1958) "<u>A comprehensive dictionary</u> of psychological and psychoanalytic terms." Longmans.

*ENTWISTLE, N.J., NISBET, J., ENTWISTLE, D. & COWELL, M.D. (1971)
"The academic performance of students - 1,
prediction from scales of motivation and study
methods." <u>Brit. Journal of Educ. Psych</u>. 41, 258-267.

ERAUT, M. (1975) "Promoting innovation in teaching and learning problems, processes and institutional mechanisms." Higher Education 4, 1, 13-26.

*ETZIONI, A. (1964) "Modern Organisations." Prentice Hall 1964.

*EVANS, K.M. (1965) "<u>Attitudes and Interests in Education</u>." R & K Paul.

*EYSENCK, H.J. & EYSENCK, S.B.G. (1964) <u>Manual of the Eysenck</u> Personality Inventory. London U.P.

*FACHNIE, J.D. & SCHILLACE, R. (1973) "Attitudinal effects as a function of undergraduate statistics course design: personalized (self-directed) versus traditional. Higher Education 2(1) 73-80.

FENKER, R.M. (1975) "The organization of conceptual materials: a methodology for measuring ideal and actual cognitive structures." <u>Instructional Science 4</u> (1975) 33-57.

FESTINGER, L. (1957) <u>A Theory of Cognitive Dissonance</u>. Evanston, Ill. Row Peterson.

*FIELD, T.W. & CROPLEY, A.J. (1969) "Cognitive Style and Science Achievement." Journal of Research in Sc. Teaching 6, 2-10.

FINCH, C.R. (1969) "The Development of an Instrument to Measure Student Attitude Toward Individualized and Laboratory Instruction." <u>Vocational - Educational Research Report</u>. Research Series no. 2. Pennsylvania State University.

FISHER (1965) Statistical tables - reprinted (1972). See FISHER (1972).

FISHER (1972) Statistical Tables for Education. 0.U.P.

FJELD, S.P. & LANDFIELD, A.W. (1961) "Personal Construct System in Psychotherapy." Raud McNally, Chicago.

*FJELD, S.P. & LANDFIELD, A.W. (1961) "Personal construct consistency." <u>Psych. Rep.</u> 8, 27.

FLANDERS, N.A. (1960) "Teacher Influence, Pupils' Attitude and Achievement." <u>U.S. Office of Education Co-operative</u> <u>Research Project no. 397</u>, University of Hinnesota.

FOX, D.J. (1969) <u>The Research Process in Action</u>. Holt, Rinehart & Winston, Inc.

FRANSELLA, F. (1970) "Measurement of conceptual change accompanying weight loss." Journal of Psychosomatic <u>Research</u>, 14, 3, 347 - 51.

FRANSELLA, F. (1972) Reported in BANNISTER & FRANSELLA (1977)

GARDNER, P.L. (1972) "<u>Attitudes to Physics</u>." Unpublished Ph.D. Thesis, Monash University.

GARDNER, P.L. (1973) "Changes in attitudes of PSSC physics students: a second look." <u>Australian Science</u> <u>Teachers' Journal 19</u> (1), 71-78.

GARDNER, P.L. (1973) "Attitudes to Physics: Personal and Environmental influences." Journal of research in Sc. teaching.

GARDNER, P.L. (1974) "The Physics Attitude Inventory, PAI." <u>The Australian Science Teachers' Journal</u> vol 21(1) 75-85.

GARDNER, P.L. (1974) "Deterioration of high school students' attitudes to physics." <u>Nature 250</u>, 465-6.

GARDNER, P.L. (1974) "Pupil personality, teacher behaviour and attitudes to a physics course." In MUSGRAVE, P.W. (Ed.) <u>"Contemporary Studies in the Curriculum.</u>" Argus & Robertson (in press).

贤

CARDNER, P.L. (1974) "Attitude Measurement; a critique of some recent research." Ed. Research vol 17 no. 2.

GARDNER, P.L. (1974) "Research on teacher effects: critique of a traditional paradigm." Srit. Journal Ed. Psych.

GARDNER, P.L. (1974) "Sex differences in achievement, attitudes, and personality of science students: a review." <u>Paper presented at the Australian Science Ed. Research</u> Assoc. Conference, May 1974.

GARDNER, P.L. (1975) "Projection effects and the needs - press model." <u>Journal of Social Psychology 95</u>, 91 - 97.

GARDNER, P.L. (1975) "Scales and Statistics." <u>Review of Educ</u>. <u>Research vol 45 no. 1</u>, 43-57.

GARDNER, P.L. (1975) "Attitudes of PSSC physics students: relationships with personality and teacher behaviour." <u>The Australian Science Teachers Journal vol 21(1</u>), 75-85.

*GARRETT, H.E. (1958) "Statistics in Psychology and Education." Longmans, London.

*GASKELL, D.C. (1973) "Levels of conceptual difficulty." School Sc. Review 55, 374-8.

GASTER, E.J. (1972) "<u>The measurement of some aspects of</u> <u>integrated science, mathematics and technological</u> <u>courses at 6th. form level</u>." Unpublished M. Phil. Thesis, University of Nottingham.

GLUCKSBERG, S. (1976) "Evaluation of semantic competence. How do adults and children understand the words 'same' and 'different'?" <u>Bulletin de Psychologie 289-294.</u>

GELLER, M.S. (1963) "Pupil's attitudes towards programmed learning." <u>Science Teacher</u>. Washington D.C. Vol 41.

GENN, J.M. (1970) "Chemistry scholars in school and university contexts: a study of transition from school to university."

"Scholars in Context." Wiley, Sydney.

GILBERT, T. (1978) "<u>Human Competence: Engineering Worthy</u> <u>Performance</u>." McGraw-Hill Book Co. N.Y. (In print).

*GOFFMAN, E. "The neglected situation in school and society." R & K. P. 1971.

GUILFORD, J.P. (1965) "<u>Fundamental statistics in Psych. and</u> <u>Education</u>." McGraw Hill Book Co.

*GUTHRIE, E.R. (1952) "<u>The Psychology of Learning</u>." New York: Harper.

*HALL, E.T. (1959) "<u>The silent language</u>." Faucett World Library. 149-157. *HALLWORTH, H.J. & BREBNER, A. (1965) "<u>A system of computer</u> programmes for use in psychology and education." London: Brit. Psych. Soc.

HANDLEY, H.M. & BLEDSOE, J.C. (1967/8) "The personality profiles of influential science teachers, regular science teachers and science research students." <u>Journal of</u> <u>Research in Science Teaching 5</u>, 95 - 103.

HARRISON, I.V. (1971) "<u>Development of an attitude test for ONC</u> <u>students in sciences</u>." Unpublished 14 Sc. Thesis, Bradford University.

HAYS, W. (1973) "<u>Statistics for the Social Sciences"</u> Second ed. N.Y. Holt, Rinehart & Winston.

*HEDGES, W.H. & MACDOUGALL, M.A. (1964) "Teaching fourth grade science by a means of programmed science materials with laboratory experiences." <u>Sc. Education 48</u>, 64-76; also Sc. Education 49, 348-358.

HEIDER, F. (1946) "Attitudes and cognitive organisation." Journal of Psychology, vol 21, 107-112.

*HESKETH, R. (1971) "Pitfalls of modern physics." <u>T.H.E.S. No. 7</u>, 26 Nov. '71, 5.

HESS, E.H. (1965) "Attitude and pupil size." <u>Scientific American</u>, 212, 46-54.

HESS, E.H. (1959) Quoted in HESS, E.H. (1965) "Level of cognitive awareness; its measurement and relation to

<u>behaviour</u>." Unpublished Ph.D. Thesis, University of Colorado.

*HEYWOOD, J. (1971) "A report on student wastage." University Quarterly 25(2), 189-237.

*H.E.L.P. (1970-75) "Higher education learning project in physics." Newsletters. <u>Nuffield Foundation</u>, c/o Birmingham University.

*HILL, W.F. (1964) "Learning: <u>a survey of psychological</u> interpretations." Methuen.

*HIRST, P. (1971) "E281 Personality, growth and learning."

0.U.P.

*HOCKEY, S.W. (1972) Arts and science, Nature 235 (5335), 233-4.

HORNSBY-SMITH, M.P. (1973) "Styles of teaching and their

influence upon the interest of students in science." <u>Durham Research Review 7(31</u>), 807-815, and Research into tertiary science education, papers presented at the conference on tertiary science education at Thames Polytechnic December 1972, London Society for Research into Higher Education (in press).

HORNSBY-SMITH, M.P. (1973) "Styles of teaching and their influence upon the interest of students in science." <u>Durham Research Review</u>, 807-815.

*HOROWITZ, E.L. & HOROWITZ, R.E. (1938) "Development of social attitudes in children." Sociometry 1, 301-38.

*HUDSON, L. (1966) "Contrary Imaginations." Metheun, London.

*HUDSON, L. (1970) "Frames of Mind." Penguin.

*HULL, C.L. (1943) "<u>Principles of behaviour</u>." New York: Appleton-Century Crofts.

*HUTCHINGS, D. (1971) "First degree courses in physics." <u>Physics Bulletin 22</u>, 342 - 345.

ISAACSON, G.S. & LANDFIELD, A.W. (1965) "Meaningfulness of personal versus common constructs." J. Indus Psychol. 21, 160.

JACKSON, P.W. (1968) "Life in Classrooms." N.Y. Holt, Rinehart & Winston. *JOHNSON, P.E. (1964) "Associative meaning of concepts in physics." Journal of Ed. Psychology 55, 84-88.

*JOHNSON, P.E. (1969) "On the communication of concepts in science." <u>Journal of Ed. Psychology 60</u>, 32-40.

JOSEPHS, A.P. (1973) "<u>Syllabus-bound and syllabus-free</u> orientations in the 6th. form." Unpublished M.Sc. dissertation. University of Bradford.

JOSEPHS, A.P. & SMITHERS, A.G. (1975) "Personality Characteristics of syllabus-bound and syllabus-free sixth-formers." <u>Br. Journal Educ. Psychology 45</u>, 29-38.

KAPLAN, A. (1964) <u>The conduct of enquiry; Methodology for</u> <u>behavioural science</u>. San Francisco: Chandler (1964).

*KEEN, T.R. (1975) <u>Assessment in H.E</u>. A paper presented to a day conference entitled 'Assessment at Plymouth Polytechnic' held in tape format in Plymouth Polytechnic Learning Resources Centre.

*KEEN, T.R. (1975) <u>The historical development of medical education</u> <u>in Great Britain</u>. Paper presented at the Devon General Hospital Postgraduate Medical Centre.

KEEN, T.R. (1976) <u>Repertory Grids used to facilitate the</u> <u>improvement of medical training</u>. Paper presented at a day conference of nurse teachers. Devon General Hospital Postgraduate Medical Centre.

KEEN, T.R. (1976) <u>Small group interaction, the problems of</u> <u>implementation</u>. Paper presented at a day conference at Plymouth Polytechnic Learning Resources Centre.

*KEEN, T.R. (1976) The Dip. H.E. Is there a need for it?

Plymouth Polytechnic for academic staff circulation.

KEEN, T.R. (1977) Guided Learning. <u>Association of Programmed</u> <u>Learning and Educational Technology Journal</u> February 1977, vol 2, no 1, 26-32.

KEEN, T.R. & HOPWOOD, W.H. (1977) Target - an alternative to Idea. <u>SCEDSIP</u> 1977

KEEN, T.R. (1978) Repertory grid techniques for teaching

17 8

appraisal. S.R.H.E. Conference, December 1976.

Published in the proceedings of that conference in a book entitled "Staff Development in Higher Education."

KEEN, T.R. (1978) Cognitivist systems of teaching appraisal.

A paper presented to the Second Internation Congress on Personal Construct Theory in July 1977. Abstracted

in the book of abstracts from that conference. Personal Construct Theory <u>Academic Press</u> 1978.

- KEEN, T.R. & HOPWOOD, W.H. (1978) Target: A New Approach to the appraisal of Teaching. <u>Association of Programmed</u> <u>Learning and Educational Technology Journal</u>, August 1978, vol 15, no 3, 187-195.
- KEEN, T.R. & HOPWOOD, W.H. (1978) What kind of teacher do you think you are? <u>Impetus</u> Co-ordinating Committee for the training of University Teachers, March 1978, vol 8, 20.
- *KEEN, T.R. (1978) Developing students' learning skills. <u>Teaching</u> <u>News</u>, University of Birmingham, no 5, 19, Dec. 1978.
- *KEEN, T.R. (1979) <u>Resource Centres in Secondary Education</u>. Paper presented to CET Conference, June 1979.
- KEEN, T.R. (1979) <u>Target</u>: Two years on! Paper presented at the <u>Third International Conference on Personal Construct</u> <u>Theory</u> - Utrecht, July 1979.
- KELLY, G.A. (1955) <u>The Psychology of Personal Constructs</u>, vol I & II. Norton, New York.

*KELMAN, H.C. (1961) "Three processes of social influence." <u>Public Opinion Quarterly</u> 25, 57-78. Priceton, U.S.A.
*KELMAN, H.C. & HOVLAND, C.I. (1953) "'Reinstatement' of the communicator in Delayed Measurement of Opinion Change." Journal of Abnormal and Social Psychology. 48, 327-35.

- *KENSHOLD, G.E. (1969) "An experiment in undergraduate teaching using audiovisual aids." <u>Physics Education</u> 1969, 4(3), 157-160.
- *KENT-SMITH, K. (1972) "<u>A study of the sex differences in attitudes</u> <u>to physics</u>." Unpublished M.Ed. research report. Monash University of Australia.
- *KEOHANE, K. (1975) "Review of Nuffield Physics." <u>Education</u>, (7.2.75.)

*KEOHANE, K. (1975) "Nuffield ten years on - and two to go." Printed in supplement to Education, 7.2.75.

*KERR, J.F. (1968) "Changing the Curriculum." U.L.P.

*KERR, J.F. & NEAL, G. (1966) "Understanding in junior science." Educational Review 18, 167-176.

*KLUPFER, L.E. (1970) "Student behaviour and Science content categories and sub-categories for a Science program." <u>L.R.D.C. Working Paper 54</u>, Learning Research and Development Centre, University of Pittsburgh.

*KLOPFER, L.E. (1971) "Individualized Science: Relevance for the 1970s." <u>Science Ed. 55</u>, 441-448.

*KLOPFER, L.E. (1971) "Evaluation of Learning in Science." In Bloom B.S., Hastings J.T. and Madans G.F. (Eds) "<u>Handbook of Formative and Summative evaluation of</u> student learning." McGraw-Hill, New York.

*KRATHWOHL, D.R. (Ed) (1964) "<u>The Toxonomy of Education</u> <u>Objectives</u>." Handbook II. Affective Domain. Longmans.

*LACEY (1970) "Hightown Grammar." <u>Hanchester Univ. Press</u>.

LANDFIELD (1967) "<u>Personal Construct Systems in Psychotherapy</u>" Rand McNally, Chicago.

*LAVIN, D.E. (1965) <u>The Prediction of academic performance</u>. N.Y. Russell Sage Foundation.

LAUGHTON, W.H. & WILKINSON, W.J. (1968) "Pupil attitudes to Science Teaching." Educ. in Science 36, 31-33.

LAUGHTON, W.H. & WILKINSON, W.J. (1970) "The science opinion poll - a study of attitudes for use in G.C.E. Science Courses." <u>Educ. in Science</u>, vol 8, no 37, 25-7.

LAUGHTON W.H. & WILKINSON, W.G. (1973) "Study of some pupil, home and school variables associated with attitudes to science." Durham Research Review no. 30, 748-51.

*LEIDERHAN, P.H. & SHAPIRO, D. (Eds) (1964) "<u>Psychobiological</u> <u>approaches to social behaviour</u>." Stanford U.P.

*LEWIS (1967) <u>Statistical Methods in Education</u>. London: University of London Press.

*LOEWENFIELD, I.E. (1958) "Mechanisms of reflect dilation of the pupil: historical review and experimental analysis." Documenta Opthalmologica 12, 185-448.

*LOVELL, K. (1961) "Educational Psychology and children." University of London Press.

*LUEPTOW, L.B. (1968) "Need for achievement and occupational preferences; some operations with value-orientations as intervening variables in need-goal relationships." <u>Sociometry 31(3)</u>, 304-312.

MCGUIRE, W.J. (1969) "The Nature of Attitudes and attitude change in Lindzey & Arouson (Eds). <u>The Handbook of</u> <u>Soc. Psych.</u>, 2nd. ed. Addison-Wesley vol 3, 142-44, 147-53.

MACKAY, L.D. (1971) "Changes in Victorian physics students during two years of physics study." <u>The Australian Physicist</u>. July, 103-109.

*MACKAY, L.D. (1971) "Changes on affective domain objectives during two years of physics study in Victorian Schools." <u>Australian Science Teachers' Journal 17(1)</u>, 67-71.

*MAYER, R. & GREENO, J.G. (1972) "Structural differences between hearing outcomes produced by different instructional methods." Journal of Ed, Psych. 63, 165-173.

*MEISTER, M. (1966) "Strengthening science education for youth and industry: factors in the early motivation of scientists." Sc. Educ. 50, 409-410.

MEYER, G.R. (1959) "<u>An enquiry into factors accompanying</u> <u>secondary school pupils interest in science</u>." Unpublished Ph.D. Thesis, University of London.

MEYER, G.R. (1961) "Factors accompanying the scientific interest of a selected group of English secondary pupils." <u>Australian Journal of Educ. 5</u>, 24-40, 105-115.

MEYER, G.R. (1963) "Factors related to scientific attitudes with secondary schools of an Australian City." <u>Australian Journal of Educ. 7</u>, 21-40.

MEYER, G.R. (1966) "<u>Pupils reaction to trial education of</u> <u>Nuffield '0' level science materials.</u>" Unpublished thesis, Macquarie University, New South Wales, Australia.

MEYER, G.R. (1969) "<u>A test of interests</u>." Jaceranda Press, Milton, Queensland.

MEYER, G.R. (1970) "Reaction of pupils to Nuffield Science Teaching Project Trial Materials in England at the '0' level of G.C.E." <u>Journal of Research in Science</u> <u>Teaching, 7</u>, 283-302.

*MILLER, N.E. (1941) "<u>Social Learning and Imitation</u>." New Haven: Yale U.P.

MOLLET, D. (1977) "Students' evaluation of teaching methods and changes in teaching methods." <u>Assessment in H.E. Ed</u>. Services Unit, Univ. of Bath, vol 2, no 2, 135-155.

*MONTGOMERY, S.R. "Changes in the teaching of the thermodynamics." Physics Education 1968; 3(4), 202-204.

*MURRAY, H.A. (1938) "Explorations in Personality." O.U.P. New York. *HUSGROVE, F. & BATCOCK, A. (1969) "Aspects of the swing from science." <u>Brit. Journal of Ed. Psych</u>. 1969, 39, 320-325.

N.F.E.R. "Science Attitude Questionnaire." Slough N.F.E.R.

*NEAVE, G. (1973) "The 'swing to science.'" Universities Quarterly 27(4), 420-430.

*NEWCOMB, T.M. (1956) "The prediction of inter-personal attraction." <u>American Psychologist</u>, vol 11, 575-86.

*NICODEMUS, R.B. (1974/5) "Validating a questionnaire on aims and objectives in Sc. Teaching." <u>Journal of App. Ed.</u> <u>Studies</u>, vol 3, no 2.

*NOVAK, J. (1970) "Relevant research on audio-tutorial methods." <u>School Science & Mathematics 9</u>, 777-784.

*NOVAK, J.D. (1966) "The role of concepts in science teaching." H. J. Klausmeier & C.W. Harris (Eds), <u>Analyses of</u> <u>Concept Learning</u>, 239-254, New York: Academic Press.

NUTTALL, D. (1970) "<u>Administrators' manual for pupil opinion</u> poll 104." N.F.E.R. Tests division.

NUTTALL, D.L. (1971) "Science attitude questionnaire administrators' manual," London: N.F.E.R. & Schools Council.

*OBOUN, E.S. (1956) "An analysis and check list on the problem solving objectives." <u>Science Ed. 40</u>, 388-392.

OPPENHEIM, A.N. (1966) "Questionnaire design and attitude measurement." Henemann.

ORMEROD, M.B. (1971) "The social implications factor in attitudes to science." Brit. Journal Ed. Psych., 41, 3 & 335-338.

ORMEROD, M.B. (1973) "Social and subject factors in attitudes to science." <u>School Sc. Review 54</u>, 645-60.

OSGOOD, C.E. & TANNENBAUM, P.H. (1955) "The principle of congruity in the prediction of attitude change." Psychological Review, vol 62, 42-55.

OSGOOD (1963) "Studies on the generality of affective meaning systems." <u>Am. Psychol</u>. 17, 1, 10.

*PARE, R.R. & BUTZOW, J.W. (1975) "The relationships among independence of work habits, attitude and achievement in an audio-tutorial physical science course." <u>Journal of Sc. Teaching</u>, vol 12, no 1, 1-3.

*PERKES, V.A. (1975) "Relationships between a teacher's background and sensed adequacy to teach elementary science. <u>Journal of Research in Sc. Teaching</u>, vol 12, no 1, 85-88.

PERRODIN, A.F. (1966) "Children's attitudes towards elementary school science. <u>Science Educ. 50</u>, 214-218, Tampa, Florida.

*PIAGET (1955) Quoted in Peel, E.A. "<u>The Pupils Thinking.</u>" Oldboune Press.

POPE, M. (1978) "<u>Constructive Alternatives in Education</u>." Unpublished Ph.D. Thesis, Brunel University.

POWER, C.N. (1976) "Competing paradigms in science education research." <u>Journal of Research in Sc. Teaching</u>, vol 13, no 6, 579-587.

*PUMFREY, P.D. (1962) "Junior school science." Education Review 14, 142-151.

REED, H.B. Jr. (1961) "Teacher variables of warmth, demand and utilization of intrinsic motivation related to science interests: a study illustrating several potentials of variance - covariance. <u>Journal of</u> <u>Exp. Education 29</u>, 205-229.

REID, F. (1976) "A note on the measures of similarity and scaling metrics for users of the repertory grid." <u>Centre for</u> <u>the Study of Human Learning</u>. Brunel University Technical report no 45.

REID, F. (1977) "A preliminary evaluation of the advanced learning and reading course." <u>Assessment in Higher</u> <u>Education</u>, 2, 1, 5-30.

REITZ, J. (1973) "The flight from Science reconsidered: career choice of science and engineering in the 1950's and 1960's." <u>Sc. Educ.</u> 57, 121-134.

*RICHMOND, K. (1971) "<u>The school curriculum</u>." Routledge & Keegan Paul.

*ROBBINS REPORT (1957), 189-192, 127-8.

*ROE, A. (1963) "Psychological approaches to creativity in science." Quoted in M.A. Coler (ed) "Essays on creativity in the sciences." New York University Press.

ROKEACH, M. (1968) "<u>Beliefs, Attitudes and Values</u>." Jossey Bass Inc., San Fransisco.

*RODZAK, T. (1969) "<u>The making of a counterculture</u>." Doubleday, New York.

ROTHMAN, A.I., WELCH, W.W. & WALBERG, H.J. (1969) "Physics teacher characteristics and student learning." <u>Journal of Rsch. in Sc. Teaching 6</u>, 59-63. *RUSSELL, I.L. (1969) "Notivation for school achievement: measurement and validation." Journal of Ed. Research, 62(6), 263-266.

S.C.E.D.S.I.P. (1974) "<u>Standing Conference of Educational</u> Development Services in Polytechnics."

SALMON, P. (1976) "Grid measures with child subjects": in "Explorations of Intrapersonal Space."

(ed) P. Slater vol 1, Wiley, London.

SAYER, S. (1977) "Effectiveness and Efficiency" Letter to T.H.E.S.

*SAYLES, P.C. (1972) "<u>Syllabus-bound and svllabus-free</u> <u>orientations, achievement motivation study habits</u> <u>and related personality characteristics of students</u> <u>via technological university</u>." Unpublished M.Sc. dissertation. University of Bradford.

*SCHOOLS COUNCIL (1965) "Science for the young school leaver." Working Paper no. 1. Schools Council.

SCOTT, W.A. (1968) "Attitude measurement." In G. Lindzey & E. Arouson (Eds) <u>Handbook of Social Psychology</u>, vol 2. Addison Wesley. SCRIVEN, M. (1966) "The contribution of the philosophy of the social sciences to educational development. In G. Barrett (ed) <u>Philosophy and educational</u> <u>development</u>. New York Houghton Mifflin.

*SELKIRK, F. (1973) "Subject choice at G.C.E. Advanced level." Universities Quarterly 27(1), 74-89.

*SELLTIZ, C.H., EDRICH, H. & COOK, S.W. (1965) "Ratings of favourableness of statements about a social group as an indicator of attitude towards the group." <u>Journal of Personality and Soc. Psych. 2</u>, 408-18.

SELMES, C. (1969) "The attitudes of 12/13 year old pupils to science and scientists." <u>S.S.R</u>. 174,51.

SHAH, S.A. (1962) "<u>A study of the attitudes of students towards</u> the science course at a certain training college in <u>Trinidad</u>." Unpublished N.A. Thesis, University of London.

SHAVELSON, R.J. (1970) "Some aspects of the relationship between content structure and cognitive structure in physics instruction." Unpublished Doctoral Dissertation, Stanford University, California.

SHAVELSON, R.J. (1972) "Some aspects of the correspondence between content structure and cognitive structure in physics instruction." <u>Journal of Educ. Psych</u>. 63, 225-234.

SHAVELSON, R.J. (1973) "Learning from physics instruction." Journal of Research in Sc. Teaching, 10, 101-111.

SHAVELSON, R.J. (1974) "Methods for examining representations of a subject-matter structure in a student's memory." Journal of Res. in Sc. Teaching vol 11, no 3, 231-249.

*SHERIF, M. & HOVLAND, C.I. (1961) <u>Social Judgement</u>. New Haven, Yale, U.P.

SHERIF, C.W., SHERIF, M. & NEBERGALL, R.E. (1965) <u>"Attitude and Attitude change</u>." Saunders, Calif, U.S.A.

SHUBSACHS (1975) "To repeat or not to repeat? Are frequently used constructs more important to the subject? A study of the effect of allowing repetition of constructs in a modified Kelly repertory test. Brit. Journal of Medical Psychology, 48, 31-37.

*SINCH, A.J. (1969) "Interests, values and personality traits of students specialising in different fields of study in university." <u>Brit. Journal of Ed. Psych</u>. 39, 90-91.

SKURNIK, L.S., JEFFS, P.M., KAWA, T. (1971) "Science attitude questionnaire." N.F.E.R. Slough.

SLATER, P. (1972) "The measurement of Intrapersonal Space," vols 1
 & 2. John Wiley and Sons.

SMETHURST, S. (1978) "<u>Combined Studies Degrees</u>." Unpublished M.Sc. Thesis, C.N.A.A.

*SNOW, A. & COHEN, L.K. (1968) "Student attitudes toward the sciences and the humanities." Journal of Ed. Res., 61(10), 456-461.

*SPENCE, K.W. (1956) "Behaviour Theory and Conditioning."

New Haven: Yale Unilever Press.

*SPLAINE, J. (1975) "Research for the Practitioner."

Audio Visual Instruction.

*STEPHENS, M.D. & RODERICK, G.W. (1973) "Changing attitudes to education in England & Wales, 1833-1902: the governmental reports with particular reference to science and technical studies." <u>Annals of Science 30</u>, 149-64.

*STERN, G.C. (1958) "People in context." Wiley, New York.

*STEVENS, S.S. (1946) "On the theory of scales of measurement." Science, 103, 677-680.

*STRAUSS, A.C. (1964) "Psychiatric Ideologies." Free Press 1964.

*THOMAS, A. (1966) "An experiment in programming junior science." Prog. Learning 3, 163 - 170.

THOMAS, A. (1970) "A study in the evaluation of some junior school science programs." <u>Programmed Learning</u>.

THOMAS, L.F. (1976) "The Kelly grid in education." Wiley (in press 4.1.76.)

THOMAS, L.F. (1976) Interactive computer programmes for eliciting personal models of the World. "FOCUS & PEGASUS" Centre for the Study of Human Learning. Henley.

THOMAS, L.F. & HARRI-AUGSTEIN, E.S. (1976) "Further development of techniques for studying and influencing reading as a learning skill." <u>The self organized Learner and the</u> <u>Printed Word</u>. Final Report, SSRC Project Centre for the Study of Human Learning, Brunel University.

THORDIKE, R. & HAGEN, E. (1955) reprinted (1969) "Measurement and Evaluation in Psychology and Education." Wiley International.

*TROHANIS, P. & MEIDENBAUER, N. (1975) "An introduction to media outreach services." <u>Audio Visual Instruction</u>.

TUCKMAN, B.W. (1976) "The Tuckman teacher feedback form TTFF." <u>Journal of Ed. Measurement</u>, vol 13, no 3, Rutgens University.

*U.S.S.P. The Unified Science Study Programme at the Massachusetts Institute of Technology. Research team: Prof. Dr. Jerrold Zacharias, Institute of Physics Profession and director of Educational Research Centre at the Massachusetts Institute of Technology.

U.T.M.U. (1976) "Evaluating teaching in higher education." U.T.M.U., London Univ.

•

*VREELAND, R.S. & BIDWELL, C.E. (1966) "Classifying university departments: an approach to the analysis of their effects upon undergraduates' values and attitudes." <u>Soc. of Ed</u>. 39 (3) 237-254.

WAITE, G.A. (1961) "<u>Study of attitudes</u>." Unpublished M.A. (Edn) Thesis, Birmingham Univ. WALBERG, H.J. (1967) "Dimensions of scientific interest in boys and girls studying physics." <u>Sc. Educ., Tampa</u>, Florida, vol 51, no 2, 111-116.

WEICK, K. (1968) "Systematic observational methods in G. Lindzey & E. Aronson (eds) <u>Handbook of Social Psychology</u>, vol 2. Addison-Wesley (Mass.)

*WELCH, W.A. (1968) "Curriculum decisions - how can evaluation assist science teachers?" <u>The Sc. Teacher</u>, 35, 8, 22-25.

WICK, J.W. & YAGER, R.F. (1966) "Some aspects of the student's attitude in science course." <u>School Sc. and</u> <u>Mathematics</u>, 66, 269-273.

*WILKINSON, W.J. (1967) "Disappearing B.Sc's." <u>Letter to "The</u> <u>T.E.S.</u>" Sept 15, 1967.

*WILKINSON, W.J. (1972) "Factors affecting choice of subjects in the sixth form." Durham Research Review, 6, 623-626.

C

*WILLIAMS, W.F. "A follow-up investigation on A level physics candidates." Physics Educ. 1969, 4(5), 305-308.

WILMUT, J. (1971) "The effect of project work on pupils' attitudes to science." Unpublished M.Ed. Thesis, Univ. of Bath.

- WILMUT, J. (1973) "The effect of project work on the attitudes to science held by sixth-form pupils." <u>Educ. Research 15</u>, 128-33.
- *WYNN, D.C. & BLEDSOE, J.C. (1967) "Factors related to gain and loss of scientific interest during high school." <u>Sc. Educ., 51</u>, 67-74.
- YOUNG, M.F. (1971) "<u>Knowledge and Control.</u>" London: Collier & Macmillan.

*ZAJONC, R.B. (1960) "Balance, congruity and dissonance." <u>Public Opinion Quarterly</u>, vol 242, 280-96. Priceton, U.S.A.

*ZIMBARDO, P.G. (1960) "Involvement and communication discrepancy as determiants of opinion conformity."

> Journal of abnormal and social psychology, vol 60, 86-94.



Υ

•

.

ERRATA

Please note the following

amendments to the Thesis

....

 $C^{(1)}(t)$

p.51.	Para. 3.3. line 17. For "reject" read "reflect"
p.64.	Para. 4.3. line 7. Insert "the" between "in" and
	"appendix" and delete "l" after "appendix"
p.65.	Para. 4.4.2. (b). line 4. For "alternated" read
	"associated"
p.71.	Para. 4.4.5. line l. Insert "The" before "appendix"
	and delete "2" after "appendix"
p.76.	Para. 5.2.1. Delete from "appendix 2 " line l,
	to " paragraph 5.3.)." line 5.
p.87.	Para. 5.4.4. line 17. For "leads" read "lends"
p.89.	Para. 5.4.5. line 4. For "11" read "10"
p.96.	Para. 5.5.1. For "X" read " " (Chi).
p.97.	Para. 5.5.2. Delete last two lines after "Figure 5/6
	about here."
p.106.	Para. 6.5. Delete last line starting "Appendix"
p.110.	Para. 6.6.2. line 23. Delete remainder of line after
, . .	" paragraph 7.5."
p.115.	Fig. 6/2. Cross in lower right segment should be
	. labelled "b" and the cross on left hand end of
	sporty component should be labelled "e"
p.128.	Line 20. Insert "and Mair" after "Bannister"
p.136.	Add to end of page "(a new package is currently being
	produced by the author together with R. Bell
	(University of Western Australia) and is due for
·	publication in November 1979)."
p.138.	Para. 6.9.2. Line 14. Delete all line after "Prefan"
	and all of line 15.

3<u>7</u>8

p.179. Para. 7.3.2. line 4. For "different" read "difference" p.181. Fig. 7/1. Insert asterisk in second box from bottom

on left hand side starting "Invite

participation"

p.194. Formula near top of page - label as 8/1.

p.210. Para. 8.8.4. line 4. Delete "however" after "(3:)"

The following three lines should then be amended to read "There is a significant relationship between their construction of the first principal component and so the substansive hypothesis is not supported."

-p.214. Page 214 is totally replaced by the new version attached.

---- at the end of errata.

p.224 & p.225. Renumber these pages

224 becomes 225

225 becomes 224

p.227. Para. 8.11. line 17. Delete "intrinsic motivation" p.228. Fig. 8/20. Insert "SIG" in cell "Factor C/Physics

Teachers"

t

p.240. Formula at foot of page should read ...

- M2 М

p.252. Para. 8.15.1. line 6. Slater (1972) not (1977)
p.257. Line 1. For Slater (1977) read Slater (1972)
p.258. Line 8. For Slater (1977) read Slater (1972)