THE REPERTORY GRID : A Critical Appraisal

by

Dudley Mansbridge Yorke BSc MEd AdvDipEd

VOL. I

Thesis submitted to the
University of Nottingham
for the degree of
Doctor of Philosophy

October 1983

THE REPERTORY GRID: A Critical Appraisal

The attempt to clarify one's own practice belongs to the very meaning of science; indeed, that is what method in science is all about.

A. Giorgi
1976: 288.

Contents

Volume 1

Volume 2

Abstract

The Repertory Grid: A Critical Appraisal

George Kelly claims a scientific status for Personal Construct Theory, but I argue that it is more appropriately characterised among the interpretive human sciences. Examination of the theory from the point of view of the grid user discloses a number of weaknesses, the most directly relevant being Kelly's assumption of the dichotomous nature of constructs. Even when this assumption is weakened by allowing grading between oppositional poles, the grid matrix retains a positivism that appears at variance with the main thrust of Kelly's theorising.

The central chapters appraise technical aspects of grid methodology, dealing sequentially with elements, constructs, bipolarity, the completion of a grid matrix, analysis, and the stability of grid data. Analysis of underlying assumptions, reflection upon the 'grid literature', and some empirical studies indicate that grid methodology is often flawed in both conceptualisation and practice. While some improvements may be made regarding technique, element X construct interactions radically undermine the grid as a research instrument, as does Kelly's later claim for the importance of events.

I further argue that short verbal labels are inadequate to bear the load of meaning that respondents wish to convey, and that grid

methodology excludes the richness of figurative language: developments based upon fuzzy set theory are unlikely to improve matters. If, as I suggest, communication of meaning is a prime requirement of construct theory research, then alternative approaches to the elicitation of constructs are necessary.

I conclude by sketching a possible response to the criticisms that have been advanced, and argue for a 'personal construct hermeneutics' in which theory and method are brought into a closer alignment. Indications are given of how this might be operationalised in terms of 'accounts methodology' and of some of the implications for the conduct of research in the human sciences.

Acknowledgments

This appraisal of the repertory grid has depended to a substantial extent upon the help of many people during the last seven years, and it is fitting that I fully and warmly acknowledge their contributions here. I wish to thank the following.

George Brown for his warm tutorial encouragement and support (especially at those moments of difficulty and doubt); for being prepared to allow me sufficient space for my meandering progress to run its course; and for persisting in the belief that the outcome would, at some indeterminate point in the future, be worthwhile.

Mick Youngman for devising computer programs in response to my needs and for patiently initiating me into the intricacies of multivariate analysis despite all the signs of being faced with a slow learner.

Other staff of the School of Education, Nottingham University, who have gone out of their way to give me help when it was needed.

All those with whom I have discussed and debated Personal Construct Theory and other issues: their conceptualisations - often radically different from mine - have played a significant part in the development of the thesis advanced here.

The many students who have helped me by willingly contributing data via repertory grids and other instruments, in the knowledge that, despite feedback sessions, the greater benefits from the research would be more likely to be gained by others.

Ann Bromiley for her care and commitment in turning the bulk of the final draft - a nastily mangled palimpsest - into a readable text.

Anita Mahomed for similarly coping with Chapters 13 and 14, and the Appendices.

Ian Reid for his help with the design and production of a number of the Figures.

Manchester Polytechnic for giving my research a generous measure of time and financial support over many years.

1 Introduction

...in order that [a man] may have any success in learning he must be penetrated with a sense of the unsatisfactoriness of his present condition of knowledge.

C.S. Peirce
Collected Papers, 5.583.

1.1 GENESIS

In my beginning is my end.

The opening line of T.S. Eliot's 'East Coker', in capturing the ambiguity of 'end', is simple and yet profound. We can foretell but a minute fraction of what will unfold in our lives, yet we often insist on planning with the utmost attention to detail, using as a foundation the current and imperfect state of our knowledge. Such a picture might constitute a vignette of much research conducted within the realm of the human sciences,[1] in which a particular theoretical stance is adopted, empirical work is conducted, and the findings are interpreted in the light of the grounding theory. Events which were not anticipated at the outset become 'error noise' to be ignored or to be explained away by further ad hoc theorising. This positivistic approach to research lacks a genuine openness to aspects of 'end' as yet unrevealed and, I shall argue, is inappropriate to research conducted within the framework of Personal Construct Theory.

The problem of 'end' has pervaded the research described in the following pages. My intention was to use the repertory grid merely as a tool, but the difficulties I came across in using it led me ever deeper into an investigation first of the grid and its theoretical substrate and, second, of how theory and method might evolve to the benefit of future research. In Radnitzky's (1968) terms, this has meant a shift from a predominantly T-type (technical) piece of research mainly concerned with solving a relatively small problem to a Q-type (question raising) study which problematises that which is often taken for granted.[2]

My original end was to investigate what kind of staff development programme might be most beneficial to science schoolteachers newly-

embarked on their careers in secondary schools. The first, probationary year in teaching is a critical time in that it involves a process of socialisation more far-reaching than is likely within short periods of teaching practice, and at the same time the probationer is under scrutiny regarding his or her acceptability as a member of the teaching profession. The Manchester Local Education Authority has given consistent support to in-service education in a variety of ways: the focus of my interest in 1978 was an LEA-sponsored induction course for science teachers which drew upon a range of expertise within schools, institutions of higher education and the authority's own advisers and inspectors. The questions uppermost in my mind were the extent to which the course programme matched the needs identified by the course participants, and whether the course content might be changed in order to accommodate any unmet needs.

George Kelly's psychology of personal constructs seemed to offer an approach to exploring the way in which these probationary science teachers construed the task of teaching without the imposition of my own categorisations and prejudices upon them[3]. Moreover, the repertory grid seemed an admirable instrument for systematising in a rigorous way the collection of the teachers' constructs.

Fourteen teachers agreed to participate in a study of their construing of first year science teaching during their probationary year (1978-1979). Nine came from Manchester and five from a nearby authority which did not run a comparable induction programme: there seemed to be potential advantages in contrasting the two groups even though nothing could be done to control the variables involved. Repertory grids were administered as near to the beginning and end of the academic year as was possible, the elements being fifteen previously piloted aspects of science teaching and the constructs being elicited on each occasion by

Kelly's method of triads. After each grid elicitation feedback was provided based on INGRID analyses (Slater, n.d.), and this provided a measure of cross-checking on both computer output and the inferences which I drew. Many of these teachers found the repertory grid procedure interesting in itself, and some commented that it made them think in new ways about what they were doing as science teachers.

However, during the year in which this study took place it became apparent to me that the repertory grid, as a research instrument, was more problematic than the literature had led me to believe. Whilst a grid could be subjected to sophisticated statistical analyses, it seemed to be giving a very limited depiction of the respondents' construing. Further, a scrutiny of the elicitation process revealed that there were a number of serious flaws in the technique that either had been glossed over in the literature or had simply not been recognised. Given that I was becoming increasingly uneasy about the quality of the information I had acquired by means of the repertory grid (even though this was supported by recordings of conversations I had had with respondents), it seemed that there was insufficient warrant for drawing the 'staff development' conclusions I had originally hoped to make. Were I to write up the research as it stood (and thereby strongly imply its substantiveness), I would be in a morally untenable position.

1.2 TOWARDS A CRITIQUE OF METHOD

The alternative option (which rapidly became Hobson's choice) was to invert my original intentions. Instead of using the repertory grid to investigate the personal constructs of science teachers, I could use the understandings gained from my work with them as the basis of a critical appraisal of grid methodology. Returning for a moment to

T.S. Eliot, the 'obvious' original end to which I had devoted the initial phase of my research had become subverted, confronted and overwhelmed by a latent end which had only become manifest after a couple of years of research and reflection.

By now highly critical of my own work I began to turn my attention to that of others, which I had hitherto accepted rather unquestioningly. Had they found similar difficulties, and - if so - had these been marginalised or overlooked? Fundamental questions arose. What <u>does</u> a grid actually elicit, and what does it fail to elicit? How valid is grid methodology, and could any general estimate of validity be made despite the assertion that each grid is inevitably unique? To what extent does the repertory grid cohere with its parent theory - and how does the theory itself stand up to scrutiny in the light of some of the problems thrown up by my inquiry into methodology?

Questions such as these have a range of implications that is too vast to be tackled within the scope of the present work. I have chosen to make the repertory grid the initial focus of convenience of this study: this has entailed a limited critique of Kellian theory, but a full appraisal has not been attempted. Similarly, matters which bear upon grid methodology (such as memory and linguistics) are discussed more briefly than their potential importance warrants. As my account develops, it will become increasingly obvious that its focus shifts from the plane of the grid itself to that of a different methodology which is nevertheless consistent with what I take to be the central tenets of Personal Construct Theory.

1.3 <u>ARCHITECTONIC</u>

The transformation of this study (from one in which a particular method

was chosen to collect evidence to one in which the evidence collected is used to support a critique of the generative method) has been rather like taking one turn around a Möbius strip. But even this simile implies an ordering of thought far greater than occurred as the study unfolded. In exploring the repertory grid I worked - often haphazardly - from mainstream considerations to tributaries and then back to the mainstream again. It is only in the writing of this account that the full shape of my argument has begun to emerge with a measure of coherence from a series of rudimentary sketches ranging from the distant peaks of philsophical abstraction to local artefacts of empirical investigation.

I have chosen an approach to presentation which I hope the reader will find both logical and helpful. Rather than provide a review of the literature at the outset (and this would have been lengthy indeed), it seemed preferable to subdivide this appraisal into a sequence of thematic chapters, each with its own evidence, review of literature and discussion, my own empirical work being drawn upon where appropriate. As I worked on this study the empirical work diminished in importance when set against the theoretical argument I was developing: rather than overburden the text with a plethora of empirical detail, I have collected the main data from this side of the study into a series of appendices for the reader who wishes to examine it in greater depth.

There is a price to be paid for this mode of presentation, but it seems to be small: there is some duplication of material between chapters in the interests of both the continuity of argument and cross-referencing. I have nevertheless endeavoured to keep this to a minimum.

The text has been written with three types of reader in mind. First,

some readers will be interested in appraising (and perhaps improving) the quality of the grids which they are using. These aspects of grid method are fundamentally technical, and the relevant material is concentrated in Chapters 5 to 10. A summary of the main methodological findings is given in Section 12.2[4].

In view of the comparative lack of criticism of Kelly's theory in the literature, some readers may be interested in the discussion presented in Chapters 2 to 4 and Chapter 11, which provides the framework for the appraisal of method mentioned in the preceding paragraph.

Finally, those readers who share my unease regarding theory and method (or who come to share it) may also be prepared to accept the radical aspects of my critique which emerge at certain points in the text (for instance, in Section 8.4). These confront the "established orthodoxy" of the repertory grid and deem as totally inadequate the mere tinkering with the details of a methodology that stands in fundamental contradiction to its parent theory.

Taken as a whole, Chapters 2 to 12 can be read as a sustained antithesis to the "received view" of construct theory and grid method. Chapter 13 is an attempt to go beyond the contradictions in theory and method and to sketch a new synthesis suggestive of a direction in which research within the Kellian tradition might begin to move.

1.4 INFRASTRUCTURE

Kelly is rather coy[5] about the writers who influenced his philosophical and methodological position, save for John Dewey (perhaps the most acclaimed figure in American philosophical circles, though the mantle of greatness may turn out to have been the original property of

C.S. Peirce). In Chapter 2 I make a limited attempt to track Kelly's thought back through the American tradition of pragmatism and to indicate where the Kantian metaphysics of personal construing constitutes a departure from the realist assumptions of much pragmatist thinking. The metaphor of 'man the scientist' is discussed, and this leads into a consideration of the 'goodness of fit' between Personal Construct Theory and modern developments in the philosophy of science. A comparison between Kellian and Lakatosian theorising emphasises that, in much construct theory research, there is a gap between super-ordinate theory and practical methodology. I propose that the way in which Kelly appears to operationalise his theory (and hence the theoretical substrate itself) suggests a more appropriate home within the domain of the interpretive human sciences. This proposition indicates that I intend to 'read' Kelly in a critical, and at times unorthodox, way: the philosophical and methodological implications are left in abeyance until Chapter 13.

In Chapter 3 the focus is the assumptive structure of the theory itself. The user of repertory grids necessarily employs Personal Construct Theory from a cognitive stance, and my appraisal emphasises the cognitive aspects of the theory in so far as they are relevant to grid use. A number of counter-arguments are developed against Kelly's assumptions, the most important for the grid user being the challenge made to the Dichotomy Corollary. At this point I develop a rather lengthy argument that, in essence, dichotomous construing is but a subset of the ways in which construing may take place. If this argument is valid, it exposes the repertory grid as imposing unwarranted restrictions upon the ways in which information may be gathered from respondents.

A fair proportion of Chapter 4 is given to a discussion of the inability of the repertory grid to reveal much about the structure of a construct system (or, put another way, to capture the intention of the Organization Corollary). This is not a fault to be laid at the door of the repertory grid, for structural considerations largely lie outside its terms of reference: however, the discussion is necessary because of the number of studies that have made structural claims on the basis of repertory grid data. More pertinent to research grounded in construct theory is the linguistic impoverishment 'imposed' by the repertory grid. I press the argument that, if communication of meaning is the prime consideration, then the short verbal tags typical of grid method are inadequate. (Linguistic aspects of grid methodology are treated more fully in Chapter 11.) There is a further, and more fundamental, point made in this chapter: the use of such labels to define the ends of scales gives the grid a positivistic orientation that renders it incompatible with the major tenets of Kellian theory.

Chapters 5 to 10 embrace a sequence of technical matters relating to grid use. Chapter 5 is given to a consideration of the choice of elements to be used in a grid, a matter to which little attention has been given in the literature. I argue that this is a critical issue for the grid user and that it is important that elements be carefully selected, bearing in mind the context of the research and their adequacy as a sample from the domain being studied. Other issues discussed include the decision regarding the elicitation or provision of elements, and problems relating to the elements' meaningfulness, salience and stability.

Constructs are the theme of Chapter 6. I begin by examining Kelly's assumptions regarding constructs before turning to the practicalities

of eliciting constructs from respondents. As with elements, whether to elicit or to supply constructs is an issue for the grid user. After a fairly lengthy discussion of the matters involved (which includes that of the relationship between personal relevance and extremity of rating) I come down on the side of elicitation. Kelly pragmatically identifies a number of types of construct, and other workers have attempted similar classifications in the light of particular problems: such evidence as I have found suggests that research claiming to support some of Kelly's typal distinctions is inadequately grounded, and that the ad hoc classification of constructs is of limited value. At the end of the chapter I suggest that the existential status of many constructs is ambiguous and that, without considerable care on the part of the researcher in specifying the context of the research and the nature of the elements, a grid will contain a pot-pourri of constructs whose ontological significance is indeterminate.

Kelly argues strongly in favour of dichotomous construing (combining both opposition and relevance), and equally strongly against the notion of 'concept' enshrined in classical logic. In Chapter 3 I make a challenge to Kelly's position at the level of theory, and Chapter 7 offers an elaboration from the standpoint of practical utility. The evidence adduced by Kelly in support of his claim is shown to be meagre, and I advance a number of objections by considering how some ostensibly bipolar constructs in both repertory grid and semantic differential research have been, or might be, used.

The cells of a grid matrix may be filled in a number of ways - by dichotomous allocation of elements to construct poles, by ranking, or by rating: each method has its particular advantages and disadvantages.

Whilst all three methods are considered in Chapter 8, I give the bulk of my attention to rating since this seems to be the most widely used in research. The evidence I present suggests that, far from being a simple linear psychological continuum, a rating scale (whether in a repertory grid or in some other instrument) can subsume complex relationships between its metrical, semantic and evaluative aspects. Further, and given the existence of rating scales, it seems likely that there is a variety of ways in which respondents actually use rating scales - a matter that seems to have been overlooked in previous research. As with the semantic differential there seems to be a strong possibility of interaction effects in a grid, particularly between elements and constructs: it is in respect of this issue that the radical aspect of my critique begins to manifest itself with some sharpness. If element X construct interaction is a major feature of repertory grids, then the whole concept of the repertory grid is drastically undercut.

Given a completed grid matrix, and ignoring the doubts about the validity of rating scales, there arises the problem of analysing it in such a way as to preserve the maximum ammunt of information in the minimum of space - the 'minimax problem', as Kelly put it. A number of analytical routines are available for grid data, but only Slater's INGRID principal components analysis and Thomas and Shaw's FOCUS two-way cluster analysis are widely accessible. Chapter 9 surveys the advantages and disadvantages of these two approaches (concentrating on the analysis of single grids), and of an alternative approach combining one-way cluster analysis and profiling. All of the routines discussed are in difficulty when a grid has blank cells (or when it contains 'not applicable' responses), and I argue that a modified version of the 'cluster and profile' method is best able to cope with the problems

involved. Whichever analytical approach is adopted, the researcher is faced with the interpretation of the output: I suggest that this is not always a straightforward matter for an individual grid, and that the aggregation of grids is likely to highlight commonality to the almost complete exclusion of individual differences. Towards the end of the chapter I discuss some unresolved issues in analysis, and then conclude by arguing that if my radical scepticism regarding the grid is accepted, then an entirely different approach to analysis is required.

The last of the technical chapters, Chapter 10, deals with the notions of stability and change. Whilst Kellian theory sees virtue (and not vice) in change, I argue that, if a grid is to have other than a transient meaning, there must be a core of stability around which change may take place. A review of the evidence, coupled with some empirical investigation on my part, suggests that the concept of stability is underlain by a·complexity that is ill-recognised in the literature. A number of factors are tentatively put forward as likely influences on stability, but these await a more rigorous investigation than has been possible here. As with analytical routines, missing data makes it difficult to complete meaningful stability coefficients based on correlational procedures, and I suggest that stability is best calculated on the basis of the summation of 'per cell' changes compared with the maximum change possible in the original grid.

The problem of language, adumbrated in Chapter 4 and at other points in the succeeding chapters, is addressed directly in Chapter 11. Taking a deterministic perspective, I argue that construing is profoundly influenced by language norms but that the coexistence of an idiolectic overlay makes normative interpretation of a respondent's

construing a questionable procedure. Further, natural language contains far more communicative potential than can be actualised through the short construct labels typical of grid use. Taking as axiomatic the Kellian researcher's desire to maximise an understanding of a respondent's construing, I suggest that an approach to research which acknowledges speech act theory is more appropriate than grid methods, whilst fuzzy set theory is leading to a methodological cul-de-sac. Speech act theory offers a prospect of congruence between method of inquiry and interpretation, although the latter ultimately requires the broader perspective of hermeneutics.

In Chapter 12 I discuss the validity of repertory grid methodology and summarise the main technical points that have arisen in preceding chapters. I suggest that many of the problems of the repertory grid threaten its validity as a research instrument, but that the threats I have identified extend to other instruments as well. At the end of the chapter I return to the philosophical conflict between theory and method identified in Chapter 4. Holding the discussion back to this point allows me to use it as a lead into Chapter 13 by taking up Pepper's (1942) contextualist root metaphor and by developing the theme of Kelly the historian in contrast to that of Kelly the scientist.

Chapter 13 is frankly speculative, but necessary in the light of the criticism spread throughout the preceding chapters. Despite its length it is no more than an outline sketch for a possible reorientation of theory and method which retains the standpoint of existential phenomenology whose fundamental coherence with Kelly's intentions I believe has not been compromised by my critique. This reorientation allies Personal Construct Theory to Heideggerian ontology and emphasises the radical historicity of 'the events as they appear' (<u>contra</u>

Husserlian phenomenology). It is also hermeneutic in Dilthey's sense and, if a few liberties are taken with the so-called 'hermeneutic circle', it seems possible to draw into the sketch some of the exploratory aspects of pragmatist thinking and thereby to combine the retrospection of <u>verstehen</u> (and its inward spiral towards understanding) with the outward spiral of exploration.

In this sketch the natural and applied sciences, with their experimental and observational methods, are treated as a subset of human endeavours[6]. I argue that, although experimentation has a part to play in a humanistic psychology, the understanding of another's construct system is based upon a hermeneutic process which draws upon a quasi-legal approach to 'procedural evidence'. An important component of this process is the appreciation of the other's intentions. These constitute the link between construing and purposive behaviour (action)- a link that is sparsely represented in the research literature (apart from clinical studies), perhaps because it is unavailable to experimental method.

The outline of what I have chosen to call 'personal construct hermeneutics' has drawn on theoretical perspectives in which understanding, interpretation and explanation are foregrounded, notably from analytical history and that branch of literary criticism in which authorial intentions are allowed as legitimate. Whilst both respondent and researcher are seen primarily as interpreters of their invironments, the demands of research make the latter's approach a far from naive process which calls upon, and ultimately goes beyond, phenomenology and phenomenography.

It seemed to me that I would be ducking the issue of research practice

were I to leave the argument at the level of abstraction, and the second phase of the chapter attempts to indicate how what is largely an 'accounts' based methodology might be operationalised. Following a consideration of some of the problems and practicalities regarding the elicitation of accounts, I propose a systematic approach which has as its focus the respondent's construing of salient events and the implicative relationships surrounding and connecting them. Not being satisfied with testimony alone, I indicate where cross-validation might be undertaken, and this leads to an examination of some of the problems of interpretation.

The methodology which I propose is not without its broader implications for research, and I conclude my account by considering some of these. Perhaps the most critical to methodological individualism is the issue of generalisation, and I argue for an abductive approach (using C.S. Peirce's terminology) in contrast to the actuarial-inductive approach typical of research conducted under the rubric of sampling statistics. Abductive generalisation emphasises the building of theory, rather than its confirmation: given the fragmented state of theory in the human sciences, this would seem to be a not unreasonable aim.

1.5 ON CRITICISM

The conceptual and methodological research described in the following pages is partial in at least two senses. It is partial in that it is based firmly on a cognitive perspective despite Kelly's claim for the unification of the cognitive, conative and affective within his theory. It is also partial in the sense that I have emphasised the historical side of Kellian theory and practice in contrast to the scientific. The reader must judge whether the case I have attempted to make is sufficiently convincing to merit further exploration, despite

its acknowledged incompleteness.

At this juncture I must anticipate two criticisms.

Since I have centred this inquiry upon the repertory grid it might be felt that I have indulged myself in a lengthy, and at times violent, attack on a straw man. Many users of grids might wish to contend that grid data are only part of the information which they collect from their respondents. Two points need to be made here. First, the literature contains a considerable number of studies in which reference is only made to grid content, and hence the reader may be justified in inferring that these reflect a mono-methodic approach. Second, in many cases where information derived from interviews is incorporated into the research report it is strongly implicit that this is a subsidiary aspect of the methodology which is used to support the (primary) grid findings[7]. In either case the researcher is viewing the respondent from what is tantamount to a fixed point: what is seen may be an illusion determined by the standpoint, rather like the Pozzo ceiling of the church of S. Ignatio, Rome[8]. If, as I claim, the grid is seriously flawed as a research instrument it will be insufficient for a critic to offer the defence that the repertory grid is but part of the Kellian's methodological equipment: what the grid user must do is to give the main points of my argument against its validity a convincing rebuttal.

The second potential criticism relates to the empirical work which is largely documented within the appendices. The investigations which I have undertaken have not been tightly controlled psychological experiments. A number of problems of fieldwork in social research (such as differential co-operation and 'mortality') have not been

overcome, although I have minimised their effects as far as has been feasible. It might have been possible to have been more rigorous in this respect, but this would either have required more time than was available to me or, crucially, would have curtailed the conceptual span of this study. My choice has been for broad brush-strokes on a large canvas and against the detail of a miniature: the empirical work, therefore, has been used to aid my conceptualisations and does not pretend to offer a filigree of finely-wrought conclusions. (There is something of a methodological parallel with Bartlett's, 1932, work on remembering, though at a less exalted level).

What I have written will probably be contentious to many working in the field of personal construct psychology, and perhaps particularly so to grid users. In developing my arguments I have found it necessary to criticise the work of others - at times quite trenchantly. Nevertheless, the writings I oppose have been a vital stimulus to my thinking, and I owe their authors a considerable debt. I hope that, in turn, my writing will be treated likewise by others. Controversy is essential to progress in research and is surely consistent with constructive alternativism.

All that I have read of Kelly leads me to picture a man with a mind open to alternative possibilities and to the challenge and counter-challenge of argument. No doubt he would have vigorously taken issue with many of the points I make in this text, and he would probably have insisted that I have badly misconstrued aspects of his theory. He would, I am sure, have recognised that my spending a great deal of time in criticising both method and theory is an indication that I take his formulations very seriously: despite his likely disagreement with my

analysis and conclusions my guess is that, all things considered, he would have approved the intention - if not the outcome.

2 A scientific theory ?

Construct theory and the philosophy of science

Behaviour is indeed a question posed
in such a way as to commit man
to the role and obligations of an
experimenter.

G.A. Kelly
1970 b: 260-261.

2.1 INTRODUCTION

George Kelly is at pains to make the point that Personal Construct Psychology is a scientific psychology concerned with people's life-experiments in the interests of prediction and control. This chapter explores the pragmatist background of his thinking and indicates where the notion of personal construing departs from the realist assumptions of the pragmatist tradition.

The guiding metaphor of 'man the scientist' is discussed, and this leads into the largest section of the chapter which attempts to look at Kelly's theory in the light of developments in the philosophy of science. During this discussion a number of problems emerge whose solutions may lie beyond the scientific realm within which Kelly construed his own theory.

2.2 THE PRAGMATIST BACKGROUND

Kelly's Psychology of Personal Constructs owes much to 'America's natural philosophy'[1], the tradition of pragmatism. Kelly specifically acknowledges his indebtedness to the philosophy of Dewey[2], but it is posible to detect in his work echoes of ideas from other pragmatist philosophers (notably C.S. Peirce and William James), though the routes by which this thinking has entered Kelly's writing are unclear.

The pragmatist tradition deriving from Peirce's pioneering work is realist and empiricist in character: Peirce made explicit his rejection of ontological metaphysics as a 'meaningless gibberish' of verbal circularity or absurdity, and asserted that what remained was a series of problems capable of being investigated by the observational methods of the true sciences[3]. Truth, for the pragmatist (James's 'satisfaction' criterion notwithstanding[4]), was a reality which could

only be approached asymptotically through a series of approximations and could never be reached in practice. But while truth might rest on an ultimate consensus reached as a result of progressively refined experimentation, it would not necessarily have any practical utility, as Peirce was aware[5]. In practice, people accept some things as "true"[6] (the provisional nature of this being often unacknowledged) in order to explore other things: in terms of Neurath's metaphor[7], we replace the rotting timbers of our ship, one by one, as it sails along, trusting the while in the structural integrity of the remainder. Kelly charts a similar course. Though he makes a claim for the reality of the universe[8], in practice he is more concerned with the psychological processes of existence and development than in the quest for absolute truths. For Kelly truth is a relativistic construct; not a stationary achievement, but something provisional and dynamic[9].

Why should one wish to explore the reality of the world? Both Peirce and Dewey stressed the importance of a sense of doubt, uncertainty or perplexity as a stimulus which people sought to eliminate through solving the manifested problem. Such a position is fundamentally homeostatic and would appear to be the philosophical ground-bass for cognitive dissonance theory[10]. The cognitive aspect of the problem-solving activity set in motion by a sense of uncertainty was clearly outlined by Dewey (1916a) in his description of the stages of what he termed 'reflective experience'. These stages are worth recording here because of the implicit parallel with Kelly's guiding metaphor of 'man the scientist'.

Dewey's five stages of reflective experience are

(i) perplexity, confusion, doubt, due to the fact that one is implicated in an incomplete situation whose full character is not yet determined;

(ii) a conjectural anticipation - a tentative interpretation of the given elements, attributing to them a tendency to effect certain consequences;

(iii) a careful survey (examination, inspection, exploration, analysis) of all attainable consideration which will define and clarify the problem in hand;

(iv) a consequent elaboration of the tentative hypothesis to make it more precise and more consistent, because squaring with a wider range of facts;

(v) taking one stand upon the projected hypothesis as a plan of action which is applied to the existing state of affairs; doing something overtly to bring about the anticipated result, and thereby testing the hypothesis.' [11]

Dewey went on to observe that the extent and accuracy of stages (iii) and (iv) marked off a distinctively reflective experience from trial and error, though he recognised that, in practice, one could never get away completely from trial and error since no hypothesis could ever be constructed which would take into account all the impinging factors[12].

Kelly's metaphor of 'man the scientist' catches the empiricism present in pragmatist thinking. Following Dewey, he makes the central assumption that people are inherently active[13] and seeking to explore their environment in order to predict and control it. Acting in the manner of scientists (albeit loosely, as will be argued later), their behaviour is purposive, an on-going experiment[14], involving the testing of hypotheses implicit in the ways in which they construe the world.

The notion of personal construing is crucial to Kelly's formulation of Personal Construct Theory: he sees people as examining the world through 'transparent patterns or templets' (constructs) which they create for themselves and which they attempt to match to the reality

of the world[15]. In Kellian terms, this is a dynamic activity in which people are continually seeking to improve this 'match' by adjusting their systems of constructs in the light of corroboration or refutation of their hypotheses[16]. While much of this experimentation necessarily involves sense-data, it is important to note that Kelly sees construct systems as involving hierarchical abstractions which may well be removed a considerable distance from the sense-data to which they relate. It is thus possible to conduct 'thought experiments' - and hence influence construct systems - without making direct reference to sense-data: in this, Kelly is close to Dewey's (1933) later description of the phases of reflective thought[11]. It is evident that Kelly's theoretical position (his own transparent templets) is tinted by a Kantian metaphysics which pulls it away from the empiricism predominant in pragmatist philosophy, though not to the point of total separation.

Kelly recognises that construct systems (or, to be more accurate, parts of construct systems) can be communicated and hence shared[17]. In this he is in harmony with the pragmatist view of knowledge as public and general; pragmatism as a <u>social</u> theory of truth, reality, knowledge and meaning[18]. However, Kelly's emphasis on the personal, individual nature of construing[19] (some of which may never be verbalised) has led him towards a position of methodological individualism which is at variance with mainstream pragmatist assumptions and which - as Holland (1970) suggests - has much in common with existentialism[20]. This can be seen, for instance, in the opening pages of 'The Psychology of Personal Constructs' where Kelly is at pains to draw attention to the autonomy of all people within democratic systems[21], contrasting this with the asymmetry of power in systems - religious or otherwise - of

rulers and ruled. This - almost Nietzschian - view of self-responsibility runs like a leitmotiv through the psychology he has constructed.

2.3 THE BOUNDARIES OF KELLIAN THEORY

It must be stressed here that Kelly does not presume to offer any kind of universal theory. He makes it clear that his theory has a restricted field of application, limited to the field of human personality and having a specific focus on interpersonal relationships. He also indicates that his theory represents an interim position which might become extended or contracted in the course of time. Since 1955, when the theory was formally published, there has been a steady growth of interest in his work, and the range of convenience of the theory has been widened to subsume a diversity of human activity far beyond that discussed by Kelly in either his original work or his later writings. As a result, it is necessary for those interested in Personal Construct Theory to consider whether widening the aperture of the theory has restricted its depth of focus and led to greater conceptual and empirical fuzziness[22].

The arguments presented in the following pages offer partial answers to the questions of conceptualisation and empirical investigation, the greater emphasis being given to the latter in that this study centres on repertory grid methodology - an offshoot of Kellian theory that seems to flourish even where it is severed from its parent plant. However, an appraisal of methodology cannot ignore theory, and, although a full critique of Kelly's position is not attempted, the theory is examined in Chapter 3 largely from the perspective of its implications for the repertory grid.

In taking the view that people are active explorers of their environment, Kelly indicates that he was, in part, reacting against what he termed 'push' theories based on stimuli and 'pull' theories based on inner needs[23]. He sees both of these types of theory as failing to take account of the contribution that people (qua persons) make to the determining of their own behaviour. The general tenor of Kelly's writing suggests that he set up his own theory in contradistinction to others (in particular, behaviourism). To do this, however, is to treat the respective theories as comparable in terms of both level of conceptualisation and practical utility, and hence to claim a polarisation which, from another perspective, might be less easily justified.

An alternative construal might see Kelly's theory and, for example, behaviourism as complementary. Throughout their lives people are bombarded with sought and unsought stimuli in immense profusion and variety. Kelly's theory (in its concern with purposiveness, anticipation, prediction and control) emphasises the sought at the expense of the unsought, whereas in a behaviouristic psychology the reverse is the case. The position sketched in Chapter 13 is located in the complementarist's perspective.

It may be that Kelly's opposition to behaviourism accounts for his giving relatively little space in his writing to unintended happenings impinging on the person, for Dewey (1916a) had earlier made a point of drawing attention to both the active and the passive components of experience, seeing activity as relating to 'trying' (and made explicit in the term 'experiment') and passivity as 'undergoing'[24]. Or, as Rychlak (1973) observes (in the context of existential analysis, and

using a terminology closer to Kelly's), 'activity' could be seen as the self's conceptualisation of its experience, and 'passivity' as the experience's impact on the self: one can learn either way[25]. The point is important because of its implications for Kelly's Fundamental Postulate, which is discussed in Chapter 3.

The point is also important because of its implications for the development of construct systems themselves. Kelly makes it clear that, for his purposes, the physiological aspects of psychology are largely taken for granted. The focus of convenience of his theory is interpersonal relations as seen through human explorative and interpretive activity. But in order to interpret one has to have a framework for interpreting, and there is a serious gap in Personal Construct Theory in that it has very little to say about the establishment of an interpretive framework (i.e. a construct system), despite offering rather more about how an already existing system may be modified in the light of experience[26]. Had Kelly addressed himself to the problems of relating Personal Construct Theory to developmental psychology (thus going beyond the boundaries of his clinical experience with adults), he would have been faced more directly with the questions of, first, the origin of construct systems[27] and, second, their subsequent elaboration through interactions between individuals and between individual and environment. In this wider frame of reference it would be expected that more attention would have been given to the internal dialogues between perception and conception ('construction' in Kelly's terms) and thought and word (in Vygotsky's, 1962*, sense), the act of construing itself developing the construct system. Such an approach would not deny the essential activity inherent in human existence and would subsume more clearly both the responsive and purposive aspects of human functioning[28].

Criticism has been levelled against Kelly on the grounds that he gives inadequate consideration to the power of social forces to influence construct systems[29], and it may be as a result of this that recent writing (for example, Stringer and Bannister, 1979; Bannister and Fransella, 1980) has begun to foreground the more broadly social aspects of the theory which Kelly, with his clinical interests, does not develop. For the present purpose it is sufficient to indicate that the 'transparent patterns or templets' through which people see the world are <u>not</u> wholly self-constructed (as Kelly strongly implies), but are to a very considerable extent determined by the cultural environment[30]. The crucial factor here is language. The importance of language to construing, and to repertory grid methodology in particular, is discussed in Chapter 11.

2.4 <u>THE METAPHOR OF 'THE SCIENTIST'</u>

Kelly is well aware of the limitations of his metaphor and makes it clear that he is referring to the scientist-like aspects of all people in their attempts to explore their environment with the aim of predicting and controlling it. The danger with metaphor is that it can be pressed too far. It is doubtful whether people are as explorative as Kelly suggests, and it can be argued that construct systems provide the framework for a set of rules for behaviour as well a basis for exploration[31]. Maier's problem-solving experiments (such as the 'hatrack problem': Maier, 1945) and Hudson's (1966) work on convergent and divergent thinking are two examples which demonstrate what might be termed 'inertia in construing', the former showing the inertia to persist even when it was apparent that the constructs being used were failing to lead to an adequate solution.

In addition, Luchins's (1942) work on the <u>Einstellung</u> effect showed

how, once a particular way of solving the 'water jug' problem had been worked out, alternative - and quicker - solutions were not seen.

Peterson and Beach (1967) examined the cognate metaphor of the 'intuitive statistician' who seeks to behave optimally. They pointed out that it is very difficult to sample events multidimensionally, and that the likely outcomes would be a failure to extract all the information that was available, resulting in a conservative approach to behaviour. The complexity of human experience would seem to take the 'intuitive statistician' far from the simple, unidimensional tasks that Attneave (1959) found to be undertaken with remarkable degrees of accuracy[32]. Postman and Tolman (1959) argued that people were forced to adopt a flexible strategy of subjective probability regarding cues enabling inferences to be made about objects capable of varying multidimensionally.

Taylor and Fiske (1978), in criticising H.H. Kelley's (1971) notion of the naive scientist, take a more jaundiced view of human behaviour and suggest that, instead of using scientist-like processes, 'many perceivers seek a single, sufficient and salient explanation for behavior, often the first satisfactory one that comes along'[33]. They suggest (reminiscent of Newcomb's, 1931, 'logical error') that positive instances weigh heavier than negative instances in the subjective estimation of correlations.

The view of Taylor and Fiske receives some support from a study by Wason (1960) who presented psychology students with the set of integers 2, 4, 6 and asked them to identify the underlying rule. There was a strong tendency for respondents to use strategies which were confirmatory rather than refutational, and some of the

respondents therefore found considerable difficulty in ascertaining the rule which was simply 'three integers in ascending order'. More recently, Snyder and Swann (1978) found a similar confirmatory bias when students had to outline a strategy to gain evidence as to whether a "target" person who had been labelled extrovert or introvert actually was so. They found that, even when the probability of confirmation was as low as seven in thirty, confirmatory strategies were typically employed[34].

The evidence presented here suggests that it is doubtful whether people typically follow to the full the stages of reflective thinking outlined by Dewey (see page 22 above), since it seems that stages (iii) and (iv), relating to problem clarification and the refining of hypotheses, tend to be underused and may even be absent. This effectively short-circuits Dewey's schema and would seem likely to lead to the testing - in stage (v) - of a very loose, perhaps implicit hypothesis derived from stage (ii). Invalidation would then allow the loose hypothesis to be maintained with little or no change whilst other aspects of the problem were explored. Alternatively, and as Dewey himself suggests, the effect is to reduce the problem-solving process to trial and error, the trials being ill thought-out in the first place.

Kelly recognised that a good scientist would wish to put a theory to the test and would be prepared to change it in the light of invalidation. Presumably a very good scientist would wish to put the theory to the most carefully thought-out and stringent test possible. On the evidence it seems plausible to conclude that, in general, people are not very good scientists, being given to inadequate testing of their hypotheses and to a reluctance to change

them even when they are disconfirmed. One might further conclude that there is a tendency for construct systems to be possessed of a psychological inertia which manifests itself as a resistance to change.

The point of this discussion has not been to invalidate Kelly's metaphor, but rather to point out that more may be read into it than can be justified in practice. Everyday human behaviour is partial: if this be construed as a failing, perhaps it is as well to remember that scientists are human beings too.

2.5 PERSONAL CONSTRUCT THEORY AND THE PHILOSOPHY OF SCIENCE

At various points in his writing Kelly strongly suggests that his theoretical formulation is scientific[35], and the reflexivity embedded in the metaphor of 'man the scientist' extends the scientific attitude across the breadth of human action. His point of view implicitly invites the reader to construe the theory in scientific terms, and in this section I examine construct theory in the light of recent developments in the philosophy of science.

It may be objected that it is unfair to appraise a theory published in 1955 against the tenets of more recent philosophising. However, Kelly made only minor modifications to the theory before his death in 1966 and, despite some subsequent attempts to add statements to the theory, it still remains substantially unaltered. The theory continues to attract new followers, further suggesting that an examination be conducted in the light of current thinking rather than be grounded within its historical horizon. This course of action is potentially to Kelly's advantage, given the shift of emphasis in the philosophy of science from the structure of scientific

theory to the life-world within which the scientist acts.

For the moment, though, I wish to focus briefly on the structure of theory. Nagel (1961) draws attention to the painstaking care with which a scientific theory needs to be stated and to the necessity for precision in the articulation of the interrelationship between its component parts[36], and Popper (1972a) points to the importance of consistency, arguing that a self-contradictory system allows any conclusion to be derived from it[37]. Tested against the cutting edge of this view of scientific theorising, Personal Construct Theory can only be construed as inadequate. The formal content of the theory, expressed through its Fundamental Postulate and eleven Corollaries, is imprecise in the articulation of its components and their interrelationships. There is sufficient slack in the theoretical framework to absorb the shock of challenges from empirical investigation - not that all of its statements are open to empirical appraisal, for the empiricist would find considerable difficulty in testing (for instance) the tautologous Commonality Corollary, and also the Choice and Dichotomy Corollaries which can always be defended by appealing to the existence of initial conditions of which the researcher was unaware.

Kelly emphasises in his writing that a good theory 'provides an explicit framework within which certain deductions may be made and future events anticipated'[38] and, more loosely, 'a general framework within which certain facts may be held in place, pending one's induction of some specific principle among them'[39]. Either way, the emphasis is towards the future, towards prediction and the testing of hypotheses. In Kelly's eyes, the nature of a scientific theory coheres with his metaphor of people as scientists who test

hypotheses: it entails experimentation - though, as was pointed out earlier[40], the experiment may be conducted in the mind as well as through observable actions.

Kelly's view of scientific theorising is ambiguous, and it would be unfair to attack the theory for its lack of scientific rigour when he may be intending to offer only a climate (or Weltanschauung[41]) within which theorising can take place, rather than a 'completely logic-tight' system[42]. However, such an open-textured approach to theorising undermines his assertion of the virtue of testable hypotheses (cf. Popper, 1972a, above) when he writes:

> ' A criterion of a good psychological theory is its production of hypotheses which are testable. In contrast to other construction systems, any scientific theory should enable one to make predictions so precise that they are immediately subject to incontrovertible verification. This means that the hypotheses which are deduced from the theory should be brittle enough to be shattered whenever the facts they lead one to anticipate fail to materialize.
>
> The theory itself need not be so fragile as its offspring hypotheses. If it is a comprehensive theory it is likely to possess some degree of elasticity even though the hypotheses deduced from it are brittle. Rarely does a scientific theory wholly stand or fall on the outcome of a single crucial experiment.'[43]

Here Kelly fails to disentangle the development of a network of theoretical statements from the practicalities of testing hypotheses. There is always a disjunction between statements in the language of theory and those in the language of observation; as Nagel (1961) points out, the co-ordination between observations and theory is 'comparatively loose and imprecise'[44], involving further theoretical assumptions which are very often left unarticulated. In other words, hypotheses are more difficult to shatter than Kelly would lead one to believe - but such is a problem for all scientific inquiry.

Kelly places great value on the fertility of a theory[45] (its capacity to subsume creative ideation), thus implicitly denying his earlier claim that a good theory offers an explicit framework for deduction. His later writing emphasises the polarisation when he contrasts the narrowness and rigidity of 'accumulative fragmentalism', in which nuggets of truth are collected piece by piece, with his own position of 'constructive alternativism'[46] - a kind of liberal pluralism in respect of the individual's supposed freedom to construe the universe in his or her own way. Kelly does not say that accumulative fragmentalism is wrong; indeed, he points out that it is not inconsistent with constructive alternativism. The implication, however, is that it is trivial and of short range compared with the broad sweep of constructions of the world.

But his colourful presentation again over-polarises the issues. Construct systems evolve with time as the individual reinterprets the life-world. For the most part, this evolution is likely to be slow, and to take place in peripheral rather than core constructs. Occasionally the construer may, like Saul, undergo a quantum shift to a new world view - but this can hardly be more than infrequent without implying mental instability. I would argue that most change is more akin to accumulative fragmentalism, in that people test their construction-hypotheses in limited ways (Kelly himself says that people try them out 'in test-tube proportions'[47]): what emerges is not a closer approximation to some absolute truth but - perhaps - a more coherent world-view.

Whilst a strong case can be made for the fertility of a theory, the operationalisation of construct theory has been accompanied by considerable variation in interpretation and hypothesis generation.

As cases in point, the very flexibility of the theory has permitted workers to claim on its behalf the validity of various indexes of cognitive complexity[48] and structure[49] which appear to bear little relationship to each other within their respective domains.

The critical issue for construct theory is the conceptual gap between the formal content of the Fundamental Postulate and Corollaries on one hand, and testable hypotheses on the other. What is missing is a set of related subordinate theoretical statements capable of being formulated as testable hypotheses: D.E. Bolton (1978) half makes the point when he observes that, whilst construct theory is capable of subsuming much of empirical psychology, there is a difficulty when one tries to tie it (as a metatheory) to experimental procedures[50].

Kelly is well aware of the gap[51], but defends his position by saying, in effect, that it is not his problem: operationalisation is a primary concern for the experimenter, but only a secondary concern for the theorist[52]. What this does is to thrust the burden of middle-level theorising on to the experimenter since, by his own admission, his theoretical framework does not entail particular deductive statements[53]. Bannister and Mair (1968) admit that the high level of abstraction of the theory will be considered a weak point by some, but point to the advantages of a theory which deals with the way that people organise their experience and anticipate events, rather than with the minutiae of existence[54]. To judge from the literature, researchers basing their work on construct theory have responded to its general conceptual attractiveness without paying much attention to its inadequacies regarding middle-level theorising.

To be fair to Kelly, the criticism advanced in the preceding

paragraphs is a little too sweeping. Kelly acts as his own middle-level theorist and experimenter in the domain of interest dealing with dimensions of transition. In Chapters 9 and 10 of 'The Psychology of Personal Constructs' he not only locates his theory at the focus of its convenience, clinical psychotherapy, but also ' gives the clinician a set of professional constructs within which he can subsume the personal constructs of his clients'[55] - to all intents and purposes, a middle-level theory. It is at this point in his book that Kelly's writing begins to communicate to me some of the implications of his theoretical position and to bridge the gap between theory and method that, for me, yawned wide in the preceding 451 pages.

Kelly's descriptions of constructs such as 'threat', 'anxiety' and 'guilt' are elaborations of the basic theory, but are limited to transition within a psychotherapeutic setting. They provide theoretical statements at a level low enough for correspondence rules to be articulated linking them to statements in the observational language - that is, the client's behaviour. These relationships are not free from ambiguity, and Kelly's explications of clinical practice in the second of his two volumes indicate that he has had to incorporate concepts from outside construct theory in order to maximise its usefulness. Indeed, Kelly openly admits to an eclectic subsumption of psychotherapeutic procedures within the framework of his theory[56], observing that the appropriate procedures vary from case to case and that 'a procedure which is used predominantly with one client may never be used in precisely the same way with another'[57]. In other words (and in tune with the Individuality Corollary) each client is a unique 'event' - a point to which I shall return later.

Mackay (1975) sharply criticisés therapy grounded in construct theory as being supported by an assortment of plagiarised techniques which are used in an ad hoc manner[58]. In so far as techniques are not given a rationale within the theoretical framework, the criticism is fair. However, if psychotherapy is construed in terms of experiment and observation one can expect, following Nagel's point noted earlier, that it may well involve theoretical assumptions and techniques from outside the articulated theoretical framework. Theory in the human sciences generally is subject to a multiplicity of interrelating perspectives, and is unlikely to be adequate to its purposes if it is developed deductively from a limited number of basic postulates[59]. Or, turning the point round, any item of data from the human sciences is likely to be interpretable from a - probably indefinite - set of theoretical positions: this issue of 'overdetermination' of data by theory will recur at various points, and especially in Chapter 13.

Although, as will be noted in Chapter 3, criticisms have been levelled against the detail of Kelly's theory, these have been sporadic and fragmented and have dealt with specific aspects of the theory rather than with the theory as a whole. The evidence of the literature suggests that these have been insufficient to stimulate an appraisal of the strengths and weaknesses of stating a theory at a high level of abstraction. It must be said straightaway that there is a strength in centring theory on personal construing since it reflects the now widely acknowledged 'theory-ladenness' of observation and interpretation[60]. It is when interpretations diverge under Kelly's licence that problems appear.

The literature to be cited in the subsequent chapters provides plenty of evidence for divergence as far as grid-based studies are concerned,

perhaps the weightiest being in the (already identified) area of cognitive complexity where, in the name of Kellian theory, a wide range of indexes is manifested - and few high intercorrelations appear to obtain. There is little guidance from the theory as to which of these indexes is most compatible with it - making the assumption that the notion of cognitive complexity is itself compatible (a matter which is open to debate).

In the light of my reading of the literature pertaining to construct theory and its application I am tempted to suggest that methodological developments (and, by inference, middle-level theory) have tended to develop radially from the hub of Kelly's original formulation with little attempt having been made to link the radii through a web of interconnections (that is, where the theoretical acknowledgement to Kelly has been more than a cursory nod towards the virtue of respondents 'telling it as it is').[61] Hence there has been a failure to establish the meshing of theoretical statements that Nagel sees as an essential part of scientific theorising.

To make this point is not to imply that a psychological theory which claims to be formulated scientifically can be stated in terms as precise as those used in, say, physics: all I am arguing is that there is a need for cross-links to be established in order that the coherence of the theory may be made more explicit. With people as the 'objects' of such a theory it seems unrealistic for this articulation to be feasible at other than a fairly broad level.

Kelly's theory, then, has to be appraised at two levels: at that of psychotherapy, and within the broader context claimed for it by its growing band of adherents. Drawing parallels with the work of

Lakatos (1970) on 'research programmes', and for the moment considering the theory from the point of view of psychotherapy, the Fundamental Postulate and Corollaries can be construed as components of the 'hard core' (or 'negative heuristic') to which the modus tollens of classical logic is not applied. Kelly, it will be remembered, asks his readers to accept the Fundamental Postulate (and by implication the Corollaries) as the basis upon which the theory is constructed.

The hard core has to be oriented towards the actual testing of theory, and Lakatos gives the term 'positive heuristic' to the (often loosely articulated) set of framing devices which suggest the paths to be pursued by the researcher. It is only where Kelly explicates the dimensions of diagnosis and transition[62] that there is a reasonably clear indication of the psychotherapeutic direction which he wishes his theory to take: here, perhaps, is best captured the idea of the positive heuristic. The third element of Lakatos's characterisation, the 'protective belt', refers to the body of theories which are the specific instantiations of the general research programme. In Kellian terms these can be construed as the theoretical statements made by the therapist about the client, which are tested in the light of the latter's acts and actions[63].

Outside the domain of psychotherapy - and Kelly makes it clear that he does not draw the same boundaries as other psychotherapists - the Lakatosian parallel is of value in that it highlights the inadequate development of positive heuristics to cover, for example, the social aspects of construing in the life-world. This weakness renders research beyond the confines of the psychotherapist's room vulnerable

to being criticised as ad hoc and inadequately grounded in theory. So, whilst Kelly's claim for his theory - that it is in keeping with 'the model of the modern philosophy of science'[64] - would seem to be broadly justified in the context of psychotherapy[65], the extension of the theory into other milieux has yet to attain that status.

But the Lakatosian simile must not be pressed too far. Laudan (1977), for one, has pointed to the difficulties associated with Lakatos's assumptions regarding the inviolability of the hard core and the cumulative entailment of successive theories subsumed by the research programme[66]. Laudan puts forward the idea of 'research traditions' which offer a "broad steer" for theory and method, and which overcome the rigidities in Lakatos's position[67]. Laudan observes that, although some theories may be acceptable at the empirical level, they may nevertheless fall into question at the level of conceptualisation; that is, in respect of the way they fit into broader frameworks of assumptions[68]. Empiricism alone, according to Laudan, is insufficient to account for movements in the conceptualisation of theory. He might have gone on to draw attention to the often implicit - and sometimes well concealed - assumption of absolute truth in empiricist methodology[69] which believes that conjecture and refutation will result in successive approximations to, or increasing correspondence with, "reality" (rather in the manner of those mathematical problems which have snails climbing up the slippery walls of wells)[70].

Laudan's theorising is broadly in harmony with three points I wish to make. First, although a number of studies involving repertory grids can be said to have 'worked' (in that broad support has been given to the predictions made) the relationship between empirical findings and theoretical position is often far from clear. One has

only to examine the various 'grid measures' employed by researchers to recognise the point being made: the difficulty lies in trying to assess whether theory and method are coherent or, if the empirical and the theoretical appear discordant, whether this is due to a disjunction between theory and method or to flaws within the theory itself. In Chapter 3 a number of challenges will be made to basic tenets of construct theory which have serious implications for methodology and hence for empirical findings. Laudan's suppression of the negative heuristic's immune defence system legitimates the questioning of basic theory and its evolution in the light of conceptually-grounded criticism.

Second, there is no need for successive refinements of theory to subsume progressively greater amounts of empirical and logical content. As a theory develops, some initial assumptions may be eliminated as new assumptions are found to be more coherent with the theoretical structure: the theory, while remaining within the ambit of its research tradition, grows at one boundary as it decays at another (the processes sustaining lenticular clouds above and to the lee of hills offer a useful physical analogy). In making this switch of emphasis the notion of truth has moved away from 'correspondence' to 'coherence', the latter position being more congenial to the theoretician in the human sciences given the complexity embedded in the simplest piece of research involving people. The 'coherence' theorist seeks the most internally consistent set of statements that is possible against a background of knowledge that is currently sparse - and likely to remain so for the foreseeable future[71]. Although Kelly wants it both ways, his hankerings for a correspondence theory of truth (in the pragmatist sense of Peirce) are ultimately subordinated to a coherence theory based on the validation of constructs by reference to others in the same person's system[72].

Laudan makes it clear that the relationship between a research tradition and theories is not one of entailment in either direction:

> 'A research tradition, at best, specifies a general ontology for nature, and a general method for solving natural problems within a natural domain. A theory, on the other hand, articulates a very specific ontology and a number of specific and testable laws about nature'[73].

Kelly would probably have been happy with Laudan's distinction[74] although many working in the human sciences, whilst accepting the general thrust of Laudan's point, might jib at the precision implicit in his elaboration of 'theory'. Laudan emphasises the looseness of fit between research tradition and theory:

> '... there are a number of mutually inconsistent theories which can claim allegiance to the same research tradition, and there are a number of different research traditions which can, in principle, provide the presuppositional base for any given theory'[75].

This leads me to the third point I wish to make. If a researcher is working within the ambit of a specific theory, it is likely that the work will be informed by a particular research tradition and that the outcomes will be construed in terms of the chosen framework. But Laudan's perspective allows the researcher to track back along other paths to different research traditions which offer the prespect of illuminating the work being undertaken. Atkin's (1981) idea of overlapping cover sets at different levels of abstraction neatly captures the point[76], and indicates that it would be a false polarisation to suggest that interpretations of data need be mutually exclusive according to the particular theoretical tributary being followed. Kelly, with his emphasis on dichotomous construing at all levels of cognitive activity, runs the grave risk of elevating difference above commonality[77].

If, following Laudan, one can tap into a range of research traditions,

the question of scientific prediction becomes more problematic. The physical scientist may not experience any difficulty, in that there may be a tightly intermeshed set of theoretical statements which lead to predictions subsequently confirmed by experiment. There may be no need to raise the eyes beyond what is happening on the laboratory bench: such might characterise a Kuhnian paradigm of 'normal science' in action. A 'scientific revolution' can only come about when the theories in use are realigned against other conceptual frameworks[78].

Successful prediction in science depends upon a determinate structure of theory. In the human sciences theory is much less tightly articulated and, taking into account the fact that the subjects of research are people who construe and react to situations in their own individual ways, prediction is an altogether more hazardous undertaking. Kelly emphasises prediction and control in his writing[79], but most of this is related to the somewhat <u>ad hoc</u> predictive activities of people in the life-world and is not of any great help to the researcher concerned with the framing and testing of formal research hypotheses. In the context of theorising in psychology, the Kellian standpoint on prediction can only be described bluntly as inadequate and unconvincing.

But not all scientific endeavour is directed towards control in a practical sense. The observational sciences, of which astronomy is perhaps the best example[80], stress the building of theories through which the universe may be better understood and explained - for instance, the 'big-bang' theory of the origin of the universe. The opportunity to test hypotheses may present itself (as witness Eddington's test of Einstein's prediction that light rays are bent by gravitational fields), but this involves an opportunistic recognition that circumstances are

favourable rather than a conscious manipulation of those variables believed to have an influence on the situation being studied.

Moreover, prediction is not an essential aspect of scientific inquiry. Sophisticated analytical techniques are employed to determine the origin of oil slicks found on the surface of the sea. The analysis will identify the type of oil, enabling those with a wider perspective to set that information against the knowledge of the vessels recently in the area. The purpose of the investigation is the attribution of origin and the explanation of how the slick came about - no doubt with a view to the apportionment of blame and the institution of legal proceedings.

The point of the two preceding paragraphs is to begin to shift the discussion of construct theory from the original ground chosen by Kelly to a ground which takes greater cognisance of the explanatory function of science. Work conducted within the framework of Kelly's theory has acknowledged the explanatory function of empirical inquiry in addition to the predictive, and Adams-Webber (1979) speculates that the growth of constructive alternativism derives from its offer of 'a viable approach to psychologists who are interested in making sense of human experience as well as anticipating behaviour'[81]. To emphasise 'making sense' is to accentuate the commonality of interest of science and history at the expense of their differences.

When one examines Kelly's descriptions of how he operates as a psychotherapist one is struck by the way in which he seeks to explore why clients are as they are, in order that they might 'conduct experiments' through their subsequent behaviour[82]. In other words,

by seeking to uncover the antecedents and current existential position of individuals Kelly is acting more in the manner of a historian than a scientist[83].

Thus there seems to be a major contradiction between what Kelly claims for his theory in the name of science and the way in which he operationalises his theory. The elevation of eclecticism to a virtue[84] undercuts his attempt to construct a scientific psychology grounded in humanism, and reveals the epistemological flabbiness of his thinking. Yet the pluralism of constructive alternativism could be turned to better account were Kellians to recognise the strength of subsuming his theory under the philosophical banner of the interpretive and explanatory human sciences. This is too large a challenge to be taken up here: signposts towards this goal appear at various places in the following chapters, and an attempt is made in Chapter 13 to re-read Kelly from a human sciences perspective and to sketch out some of the methodological implications of this "paradigm shift".

2.6 SUMMARY

In this chapter the pragmatist substrate of construct theory was examined, the stress on personal construing being shown to be a deviation from mainstream pragmatist thought. The metaphor of the scientist was shown to be less persuasive than is implied in Kelly's writing, in that research indicates that people tend to show a bias towards confirmation rather than refutation - in other words, they tend not to test their theories in the scientific sense.

The case was argued that Kellian theory has much in common with the

comparatively flexible philosophical position of Laudan's 'research traditions'. The main difficulty with the theory was shown to be associated with its high level of abstraction and the failure of workers within the Kellian tradition to produce a coherent middle-level theory which could be translated unambiguously into a programme of hypothesis-testing.

Constructive alternativism is both epistemologically and methodologically pluralistic, allowing data to be overdetermined by theory - a position shared by the human sciences, notably history. There is a strong filament of historical method running through the operationalisation of construct theory, implying that a historical perspective might suggest ways of resolving the contradiction between a Kellian world-view oriented towards prediction and a methodology strongly grounded in understanding, interpretation and explanation.

3 Construct theory

A cognitive perspective

[Constructive alternativism] does remind us that all our present perceptions are open to question and reconsideration...

G.A. Kelly
1970 a: 1.

3.1 INTRODUCTION

Before embarking on an apparaisal of Personal Construct Theory it is necessary to indicate that this appraisal will concentrate upon those aspects of the theory which have the most direct implications for repertory grid methodology. Apart from the Fundamental Postulate, these are taken to include the Construction, Dichotomy, Range and Organization Corollaries. This is not to deny that the other Corollaries have implications for grid use; however, their implications lie at a remove from grid methodology *per se*, and they will hence be considered more briefly.

Though admitting that his theory was built upon an intellectual model, Kelly himself rejects the subdivision of psychology into cognition, conation and affection, since he sees all three as falling within the range of convenience of his theory[1]. As far as repertory grids are concerned, the manifestation of construct systems through them is largely limited to what can be communicated through language[2]. And whilst that language may express conative and affective aspects of a person's being, these are necessarily filtered through cognition into the overwhelming majority of grids. For the purposes of the present study, therefore, Personal Construct Theory will be examined through templets that are coloured with cognition.

3.2 THE FUNDAMENTAL POSTULATE

> 'A person's processes are psychologically channelized by the ways in which he anticipates events'[3].

If Kelly had been asked to indicate which single word of his Fundamental Postulate was the most important to him, it is likely that he would have chosen 'anticipates'. He saw anticipation as the driving force

('the push and pull') of personal construct psychology and, in using this word, he built into his theory its predictive and motivational features. The metaphor of the scientist is hardened here when he writes 'Like the prototype of the scientist that he is, man seeks prediction'. This is the nub of the Fundamental Postulate; but is Kelly's assertion wholly sufficient?

Kelly pointed out that a postulate was an assumption, taken to be true for the purposes of developing a theory: to question the postulate would mean the use, explicitly or implicitly, of other postulates. And though, in presenting his theoretical structure, Kelly asks the reader to accept his postulate as an ad interim statement of truth, there is no reason why it should not be held up to scrutiny from an alternative perspective - indeed, one can imagine the philosophy of constructive alternativism welcoming it.

There are two ways of examining Kelly's Fundamental Postulate: first, to consider it within its own terms and, second, to look at it from an external perspective. Examination from an internal perspective suggests that the linear way in which the postulate is stated is an inadequate representation of what Kelly tries to convey in the bulk of his writing. In his elaboration of the Fundamental Postulate he makes it clear that it is the ways in which a person anticipates events that determine behaviour, appropriate psychological channels being established in the light of anticipation. The linearity of Kelly's statement seems to conceal a circularity, for how can a person anticipate events without in some way having already established psychological channels towards them?

The circularity is not necessarily vicious, but the imprecision in Kelly's writing does not help the reader to avoid viciousness. A reading which avoids the trap might focus on the ambiguity inherent in Kelly's terms and his failure to give adequate weight to past experience. For the present, attention will be given to terminological ambiguity: experience will be considered shortly when the Fundamental Postulate is examined from an external perspective.

Take, for example, 'processes'[4]. This subsumes a range of levels of meaning, from which I shall select but two - cognition and purposive behaviour. Kelly asks the reader to construe the behaving organism as a process which operates through a network of pathways (a construct system). In the Organization Corollary he makes it plain that he views this network hierarchically: some constructs are superordinate to others. Kelly would probably have argued that all processes were interlinked in the construct system, and it would be difficult to challenge such a view. I would wish to argue, however, that treating 'processes' in this undifferentiated way is fundamentally unhelpful.

Purposive behaviour is a function of prior cognition. Construed in psychological terms, cognition is superordinate to purposive behaviour both structurally and temporally, even though new cognitions depend upon the outcomes of behaviour. This circularity can be represented diagrammatically[5] as in Figure 3.1.

Events have meaning only in so far as they can be construed, thus requiring the person to attempt to map them on to the existing construct system. A close matching at this point is unlikely to make much difference to the construct system and the event is assimilated: one might say that the event is a further corroboration of the existing

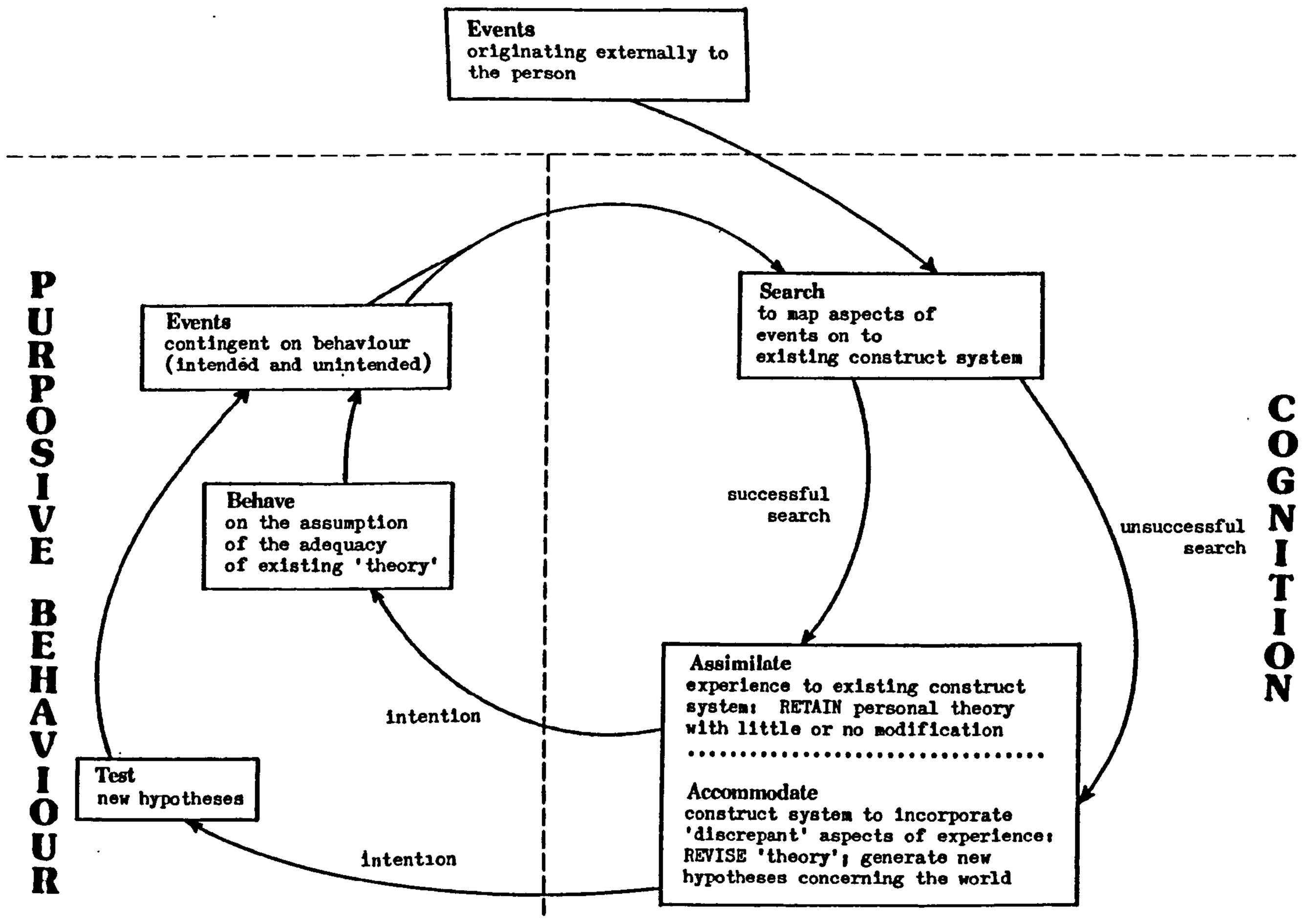


Figure 3.1 A schematic representation of the relationship between cognition and purposive behaviour.

inductive generalisation(s) already made, and the implicit personal theory can be maintained. Further behaviour is undertaken on the assumption of the truth of this implicit personal theory (which is treated as a rule, or law). However, when the existing construct system is inadequate to subsume the event, the search strategy fails, and the person is forced to adapt the construct system to accommodate the discrepant aspects of experience. The personal theory has to be revised, implying the development of new hypotheses (though the rigour with which this test is conducted is open to some doubt, bearing in mind the earlier discussion of the metaphor of 'the scientist').

It is worth pointing out that Figure 3.1 represents the growth of construct systems since the number of events construed inevitably increases with time. The 'assimilation cycle' comes close to Kelly's description of 'definition' (of which more later, when the Choice Corollary is discussed), whereas the 'accommodation cycle' relates more closely to the extension of construct systems[6].

Had Kelly been more explicit about action in his theoretical formulation, perhaps acknowledging Dewey's (1916a) idea of the 'end-in-view' as a hypothesis or plan guiding activity[7], his theory might have gained in clarity. Schütz (1967*) makes the distinction between levels of 'process' when he suggests that action is always carried out in accordance with a 'plan more or less implicitly perceived'[8]. The plan is seen by Schütz as an intuitive advance picturing which is quite vague when compared with the real action that finally occurs, and it can be subdivided into a series of intermediate goals, the action towards each of which can be pictured in advance in the same way. Schütz's plan clearly implies a prior psychological channelling towards the projected event, whether or not the person is capable of articulating it.

So far the argument presented has been concerned with purposive, intended behaviour, reflecting Kelly's emphasis on anticipation. But Figure 3.1 acknowledges that the world can 'happen to' the person (Dewey's recognition that people 'undergo' as well as 'try' was noted on page 25). To be fair to Kelly, it must be recorded that he discusses experience in terms of undergoing[9], making the point that people can only make sense of events if they construe them: this is consistent with the schema of Figure 3.1. Yet if a person is to make sense of the world by construing it, it may be that he or she is seeking an explanation for the features of the world being as they are, rather than using the constructions in an anticipatory way. There is clearly a potential for anticipation in any explanation, since it is implicitly a theory, but it does not seem necessary to cast the Fundamental Postulate solely in terms of anticipation: I would wish to argue that the role of 'undergoing' experience should be given greater prominence when the sedimentation of a construct system is considered.

Kierkegaard once observed that we live life forwards, but understand backwards, implying a model of cognition and action that contains both anticipatory and explanatory features. It can be argued against Kelly (as Pepper, 1942, argues against pragmatists in general[10]) that he gives insufficient weight to the explanatory function of the scientist, and hence generally to the role of explanation in human experience.

The following vignette, drawn from personal experience, illustrates the point I am trying to make here.

I am riding my motorcycle, in the small hours of a wet November

morning, along a familiar road through a wood. The journey is routine. Rounding a sharp left-hand bend the motorcycle slides from under me, and I suddenly find myself bumping along, face-down, among the fallen leaves on the muddy verge. Uninjured, I pick up my stalled machine in order to restart it, only to find that something is preventing the engine from turning over. I push the motorcycle to the top of the rise and begin to coast down the slight hill on the other side. I let in the clutch, and to my surprise the engine fires and continues to run, albeit making ominous clattering noises. Nevertheless, it gets me home. Subsequent examination of the engine reveals a number of teeth broken from cogwheels in the gearbox.

There are two points to be made here. First, I learn that wet leaves on roads are a hazard to motorcycles: I learn this not as a result of any anticipation (clearly such was lacking, or at least unavailable, in my construct system), but because, from my point of view, 'the world happened to me'. My processes have, as a result, been psychologically 'channelized' - and that channelization now influences the way in which I anticipate further motorcycle riding. In other words, Kelly's Fundamental Postulate is reversed: my anticipations have been psychologically channelized by the ways in which I have construed events[11].

The second point relates to my curiosity regarding why an engine which has apparently suffered a seizure can suddenly return to life. I am puzzled (Peirce and Dewey might have said I was in a state of doubt) but the puzzlement is removed once the crankcase is split open and the internal damage is revealed. I have an answer to the implicit question 'why did it happen?': the explanation that broken cog-teeth temporarily jammed the mechanism seems wholly reasonable. This does not add any

markedly new construct to my system, for I am already aware that loose bits can jam intricate machinery: all that is added is another instance falling into an already defined category. Kelly might well have argued that the event would have helped me to define further my construct system and to anticipate future events (in that I would be better able to hypothesise the cause of jammed machinery), but in this case the argument reverts to that made in the previous paragraph: the construing of events influences my anticipations. For my part, the emphasis is on the defusing of what can be construed as a situation of cognitive dissonance, the function of science being explanatory rather than predictive in this case.

It may be felt that it is inadequate to use a relatively simple explanation of a physical event as the starting-point for an argument for the role of explanation in the development of construct systems. However, in this wider human context, questions dealing with why people fall in love, commit murder, and suchlike are typically answered by sentences beginning with 'because'. In a court of law a lawyer might claim that a killer was not responsible for the action on the grounds of diminished responsibility; or a person might claim to have fallen in love with another because of the latter's similarity to a parent. Such explanations may be grossly simplistic[12], yet may satisfy the questioner: whatever their level of sophistication, they are constructs of events in the world, whose potential for prediction may never be realised.

Explanations are qualitatively different from hypothetico-deductive procedures in that they seek to identify the causal influences on an event that has already taken place, instead of testing the effect of selected variables believed to have a causal influence on a situation. Explanations are retrospective, whilst experimentation is prospective:

both fall within the realm of the scientist, and require painstaking investigation and the careful weighing of evidence. I labour the point because Kelly's view of the scientist seems unwarrantedly restricted in its emphasis on experimentation, prediction and control, and because the explanatory function of science receives little attention in his writing. In later writing Kelly (1970b) gives more weight to events[13]. I shall argue in Chapter 13 that the explanation of events is a major concern for the personal construct theorist and researcher, and that the scientist and the historian are a lot closer than Kelly may have believed.

3.3 THE COROLLARIES

Kelly develops his Fundamental Postulate by stating eleven propositions which are partly derived from it and are partly further elaborations of his theoretical position. He admits that his theory is not logic-tight, but he sees greater importance in a theory's capacity to be provocative and to stimulate fertility of thought. The consequence of this - as he is well aware - is that everyone starting from the Fundamental Postulate will not necessarily make the same deductions from it[14]; this presents difficulties for those who wish to test the theory. Kelly's theory is perhaps best construed as a metatheory[15], a theory about theory making, which requires the user to provide theoretical content at an appropriate level for empirical testing: the problems presented by this were discussed in the preceding chapter.

It is not, however, proposed here to attempt to assess the relationship of the eleven Corollaries to the Fundamental Postulate, nor to offer a different set of derivations. Rather, in the pages that follow, an attempt is made to appraise the Corollaries as they stand from the

points of view of their internal coherence and of their standing in the light of cognate research.

3.3.1 The Construction Corollary

> 'A person anticipates events by construing their replications'[16].

In his discussion of this Corollary Kelly construes 'construing' as 'placing an interpretation', a construct system being a framework for interpretation. Within this framework that which is being construed (in Kelly's terminology, the 'element') takes on meaning. He points out that it is the person, not the element, who is responsible for this structure.

However, Kelly is not entirely consistent in his use of the term 'construct'. In addition to its discriminative meaning[17], constructs are described as 'being imposed upon events'[18], as 'controls'(defining and limiting the options open to the person)[19], and finally as 'pairs of hypotheses' which may be applied to any new element which the person wishes to construe[20]. But a construct cannot in itself constitute a control or a hypothesis, for these require the intersection of a construct with an element before the construct's potential for action can become meaningful. To criticise Kelly's own example of construct-as-hypothesis[21], the construct 'black-white' can only become part of a hypothesis when brought into contact with an element. Offered a clenched fist containing a chess piece (and initially construing this as non-violent in intent), I can hypothesise that the piece is black (or white) on such grounds as are available to me: the hypothesis depends upon both element (chess piece) and construct (black-white). On this 'reading', it is clear that T. Mischel (1964) is technically incorrect in his discussion of whether constructs can or cannot be falsified[22].

Mischel takes the view that constructs are rules for making decisions about action[23], and, as with hypotheses, this view is open to the objection that it does not take into consideration the element(s) to which the rule is potentially applicable. Kelly would probably have argued that to claim a construct to be a rule is to indulge in oversimplification. He makes the point that one should avoid the pre-emptiveness of construing elements as nothing but elements, and constructs as nothing but constructs, since this closes off options that might be explored with benefit. Why not, then, construe constructs as <u>elements</u>, abstracting from them those features to be taken as determinants of behaviour? The point can be illustrated in the following way.

I have a view of Mediterranean resorts which subsumes the following construct poles: sunniness, warmth of sea, good night entertainment, interesting food, and so on. However, I am also aware of reports of pollution in the Mediterranean Sea. Wanting a holiday in which I can enjoy sea-bathing, I construe the constructs in terms of a superordinate construct

'constructs which I will treat as determining my choice' - 'constructs which I will not treat as determining my choice'

under which I assign to the left hand pole the negatively evaluated aspect of pollution, and to the right hand pole the other (positively evaluated) features of a Mediterranean holiday. The original constructs are not the rules: the rules arise from construing the original constructs in the light of a situation in which I need to act.

Mischel is on stronger ground when he points out that constructs contain both descriptive and evaluative components[24], the distinction being implicit in the illustration presented above. Harré and Secord (1972)

present a threefold dissection of constructs in their analysis of the description of a person as a coward: the description

(i) states what the person is like (descriptive/interpretative);

(ii) indicates the sort of behaviour the person is likely to show (anticipatory); and

(iii) is a negative evaluation of the person[25].

The construct could be said to 'resonate' at different levels in the system in that it interacts with a different set of constructs in respect of (i), (ii) and (iii). Whilst the sets for (i) and (ii) might be quite similar, that for (iii)would seem to include a range of culturally-embedded constructs and might imply operation at a higher level of abstraction.

Description (ii) can clearly be regarded as predictive in interest, and the prediction can only be appraised if the person is placed in an appropriately testing situation and due consideration is given to the person's behaviour in the light of the prevailing circumstances. This appraisal, as Kelly suggests, is likely to involve constructs at a different level from that at which the prediction is made[26].

The role of constructs in interpretation has yet to be considered. In Figure 3.1, following Kelly, the construct system was presented as a framework on to which incoming events were mapped, with varying degrees of success. Much of that framework may lie beneath consciousness, leaving the person to produce some form of <u>post hoc</u> rationale as a plausible reconstruction of the interpretation already made. Hirsch (1967) suggests, in the context of literary criticism, that an initial guess is made as to the 'genre' of the meaning-complex being confronted, this being done on a gestalt-like sense of the whole.

This genre then acts as a heuristic framework for the further investigation and interpretation of the event[27]. It is the initial guess that is of main interest here, for it is an implicit hypothesis that the event being construed is of a particular type. This hypothesis can only be grounded in prior experience, that is, in the construct system already established (whether that be substantially available to consciousness or not). The guess, as a provisional identification of type, will tend to divert processes into the assimilation cycle until such time as the subsequently-unearthed evidence is discrepant - in which case a different genre may be substituted, or the existing construct system may be revised.

Hirsch's view is hermeneutic, and it seems capable of being extended into the broad area of cognitive psychology. Interpretation begins with an act of categorisation, the category system being, in Bateson's (1973*) terms, 'comparatively hard-programmed' to enable human functioning to be economical through the establishment of habits[28]. As Bruner et al (1956) observe, categorisation

(i) provides a means by which objects may be identified;

(ii) allows the ordering and relating of classes of events;

(iii) reduces the complexity of the environment, and hence the necessity for constant learning; and

(iv) provides directions for human instrumental activity[29].

But categorisation is not a static, one-and-for-all activity. Memory is active and constructive, as Bartlett (1932) showed. Whilst people may store genres or standard episodes ('scripts' as Schank and Abelson, 1977, term them), these are open to revision in the light of experience since no event can be identical to any previous event or 'ideal type': deviation from the norm is the norm, and modification of the norm is the consequence.

Some of the disagreement over the status of constructs may have its origin in a semantic ambiguity. Though Kelly writes of 'construction systems', the literature generally seems not to make a distinction between 'construction systems' and 'construct systems'. At the level of system there may well be little pragmatic difference, but when events are being considered the distinction may be of greater importance for the theorist. If a <u>construct</u> is a dimension of discrimination, it can be argued that a construction relates to the application of the construct to an event. A construct may, in Tulving's (1972) terms, be laid down in semantic memory, whereas a construction (an event-having-been-construed) would be stored in episodic memory. A system will contain both constructs abstracted from events and constructions of the actual events themselves. One can then talk of invalid constructions without calling the construct into question[30].

In this discussion of the Construction Corollary there has been a crucial omission, paralleling an omission in Kelly's writing: that is the importance of <u>language</u> in construing. The question of language is also a critical one for repertory grid methodology and I defer the main discussion of this issue to Chapter 11, since I wish to treat it at greater length than would be appropriate here.

3.3.2 <u>The Dichotomy and Range Corollaries</u>

> 'A person's construction system is composed of a finite number of dichotomous constructs'[31].
>
> 'A construct is convenient for the anticipation of a finite range of events only'[32].

One word of the Dichotomy Corollary has stimulated an immense, and still unresolved, debate: 'dichotomous'. Had Kelly omitted this word from the

Corollary it is quite likely that the Corollary would have been accepted without much dispute. The debate regarding the dichotomous nature of constructs is of great importance for repertory grid methodology, much of which is explicitly based upon the assumption of dichotomy.

The Dichotomy Corollary is based upon a contrast theory of meaning coupled with the notion of relevance[33]. For Kelly, a construct is a discrimination which involves both similarity and contrast: this requires a minimum of three elements in order that two may be seen to be similar in respect of some relevant subjective criterion and to stand in contrast with the third. The criterion may be, for example, an attribute or a function.

Kelly's triadic method of elicitation produces both explicit negations and oppositions. Resnick and Landfield (1961) draw a different distinction between 'logical' and 'peculiar' constructs, but their categorisation obscures the subdivisions within the 'logical' category which have implications for both the completion and analysis of a grid matrix[34]. These different types of distinction are shown in Figure 3.2 [35].

Traditional logic operates on the principle of set inclusion: elements are assigned to a set if they possess the relevant property X[36], and are excluded from it if they do not. The constructs in traditional logic are strict negations, 'X - not X'.

Kelly attacks this method of approach on the ground that it lumps together everything that is 'not X', irrespective of the elements' degree of relevance to the situation being construed: one would not wish, for

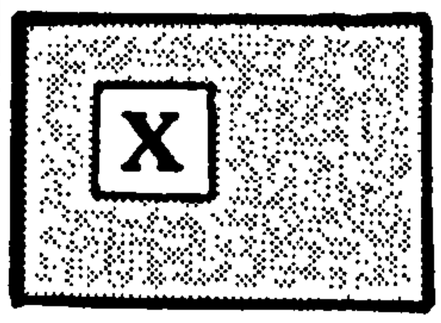
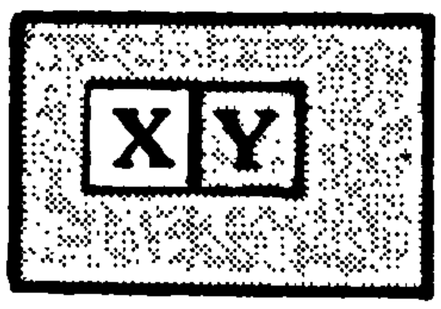
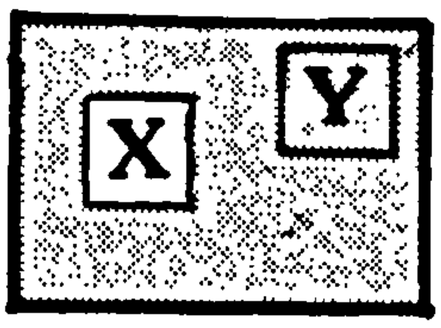
TYPE OF CONSTRUCT (Resnick and Landfield)	DIAGRAMMATIC REPRESENTATION	NATURE OF BIPOLARITY	EXAMPLE
Logical		Negation	happy - not happy
		Negation/ opposition	happy - unhappy
		Opposition	happy - sad
Peculiar		Non-contiguous opposition	happy - businesslike

Figure 3.2 Negation, opposition and non-contiguous opposition.
The shaded area in the diagrams represents 'not X'.

example, to construe a rainbow in terms of 'trustworthy - untrustworthy'. He replaces the negation of the 'X - not X' construct by the opposition of the 'X - Y' type of construct, in which X and Y represent contrasting properties within a superordinate framework to which the distinction between X and Y is relevant[37]. The Kellian 'X - Y' construct thus excludes as outside its 'range of convenience' anything to which the distinction could not be applied - hence the Range Corollary.

Kelly pays little attention to non-contiguous opposition, making the assumption that the oppositional poles remain subordinated to a framing concept and that the respondent does not give, for X and Y, the emergent poles of two different dimensions. My experience of administering grids shows that non-contiguous oppositions can occur, particularly if the administrator remains detached from the elicitation procedure[38]. The variation in types of construct has considerable implications for the ways in which grid matrices are completed and analysed, as will be discussed in Chapters 8 and 9.

Setting aside the problems introduced by non-contiguous opposition, Kelly's claim that his approach to constructs departs from that of classical logic must be examined. An analysis of an example which Kelly himself gives[39] suggests that he may be closer to formal logic than he recognises.

Four elements are presented: A and B are men, C is a woman, and O is the time of day. Kelly suggests that, faced with these four elements, an aspect of A, B and C is abstracted (e.g. sex), which is not applicable to O: in other words, Kelly implicitly sets up the construct 'have sex - do not have sex'. He then takes A, B and C (subsumed under the 'have

sex' pole) and differentiates them in terms of 'male - female' (or, to a first approximation, 'male - not male' or 'female - not female': in terms of logic, it makes no matter which). The correspondence of Kelly's explication with the canons of classical logic is highlighted in Figure 3.3 .

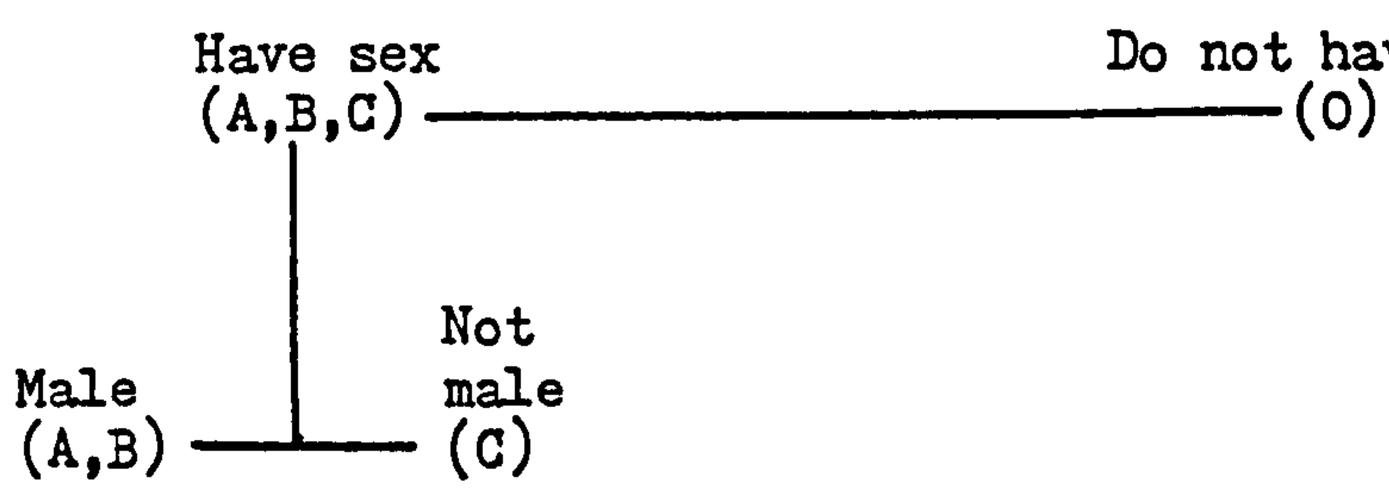


Figure 3.3 An example of Kellian construing construed in terms of classical logic: cf. Kelly (1955:60).

It will be noticed that, in my interpretation of Kelly's example, I have injected a change: I have replaced 'sex' by the bipolar construct 'have sex - do not have sex'. Kelly introduces 'sex' without any indication of what its contrast might be, and in doing so glosses over the problem of opposition being applied to noun forms. His discussion of proper names as constructs implicitly indicates the uneasiness of his position, in that he only refers to noun contrast in a negating 'X - not X' sense rather than in an oppositional 'X - Y' sense[40]. Elaborating the point, what is the opposite of, say, 'table'? The logician would answer 'not table', whereas a Kellian reply would perhaps be 'chair, within the range of things called furniture'. To this, the questioner might ask for the opposite of 'furniture', and I suspect that sooner or later the Kellian respondent would be forced to use a 'not X ' negation.

Ogden (1967*) indicated the difficulties associated with noun

opposition[41], and it would seem that similar difficulties obtain with verbs. However, where linguistic qualifiers (adjectives and adverbs) are concerned the problems seem markedly reduced, though not completely eliminated. It may be that repertory grid methodology, which tends to concentrate on the antonymy of adjectives or adjectival phrases[42], has masked the difficulties associated with the Kellian perspective on contrast[43].

Whilst it is clear that words can have a number of semantic opposites depending on the context (for instance, 'woman' may be oppositional to 'man' or 'girl')[44], and whilst naming a functional opposite may be helpful to researchers and clinicians[45], the use of bipolarity does not entail that all construing is dichotomous[46]. The difficulties associated with such an assumption suggest that a search beyond the confines of dichotomy might prove fruitful.

Miller and Johnson-Laird (1976) advance the notion of 'contrastive sets' of mutually exclusive terms existing within the same semantic field[47]. Thus, in the realm of metals, one can identify gold, silver, copper, aluminium, iron, and so on. For many practical purposes the criterion of mutual exclusivity may be necessary even when the boundaries are fuzzy - for instance, in the realm of colour terms. Returning to the discussion of Figure 3.3, the category-division 'male-female' is simplistic: hermaphrodites require a third category of 'both' under the 'have sex' pole.

Kelly himself implicitly uses the idea of contrastive sets from time to time. In delimiting his psychology of personal constructs he contrasts psychology with both sociology and physiology[48]; and later, when

considering contrasts to red hair, he refers to the non-redness of the (categorically-equivalent terms) white, yellow, brown and black. In the latter case he ducks the issue when he observes that 'Our language gives no special word for this nonredness, but we have little difficulty in knowing what the contrast to red hair actually is[49]. Of course there is no difficulty: the contrast is any or all of the shades Kelly mentions and - to judge from observations of the hairdresser's art - not a few others besides!

The contrastive set gets the construer off the hook of having to reconcile dichotomous construing in circumstances where a number of semantic items inhabit the same realm of meaning. Reverting to the previous concrete example of the metals, how, under a dichotomous construing model, do I come to construe a particular sample as being of copper? I could make a series of pairwise comparisons in my mind until I arrive at a plausible identification. The sequence might go 'Gold or silver? Not silver. Gold or copper? Not gold. Copper or iron? Not iron....' and so on until it emerges that copper is the most likely option. Alternatively, I might work through a series of questions of the form 'Gold or not gold?' until I come up with an answer that satisfies me[50]. Neither approach would seem parsimonious.

Work on recognition of category membership suggests a strategy of matching the stimulus to reference exemplars or ideal types stored in memory. Collins and Quillian (1969) propose a network model of memory in which instances are subsumed under hierarchically-ordered sets of increasing abstraction, a typical example being the subsumption of 'canary' and 'ostrich' under 'bird', which is in turn subsumed under 'animal'. Each level in the hierarchy is associated with particular

properties; thus an animal has skin, eats, breathes, and so on, whilst a bird is additionally specified in terms of the possession of wings and feathers. Collins and Quillian recognise this as an over-simplification when they note that people seem to store certain properties at more than one level, and in a later paper they indicate that concepts may be stored in terms of encyclopaedia-like knowledge instead of limited dictionary-like information[51]. Reaction-time studies suggest that, in general, recognition of an instance as a member of a category takes longer, the greater the number of levels between instance and category. The model is not entirely satisfactory, since there are exceptions: 'mammal' is, in reaction-time terms, more distant from 'dog' than is 'animal': Rips et al (1973) observe that familiarity of the concept, and commonsense usage might have a determining influence on the results[52]. They go on to suggest a feature-comparison model of memory in which each instance presents both defining and functional features, the latter being common to all members of the category whilst the former allow discrimination between instances[53]. As Collins and Loftus (1975) point out, the model presents problems regarding the extent to which defining features are recognised as such: and one might ask how many slow worms are killed, for instance, on the assumption that they are snakes?[54]

The trouble with natural concepts is that most of them are ill-defined and have no simple set of criterial features, leaving them open to confusion when salient but non-characteristic features are present[55]. Rosch (1977), having worked extensively in the field of category recognition, concludes that focal examples ('prototypes') of categories are developed in the mind and act as reference standards against which stimuli may be judged[56]. The work of Posner and Keele (1968) and

and Reed (1972) on pattern recognition suggests that prototypes can be mentally constructed ('abducted', in Peirce's terms[57]) even though the basic pattern is not actually presented, or even does not exist[58]. Rosch points to the advantages to the individual of being able to discriminate finely and to integrate information into cognitively manageable proportions and suggests that the most cognitively economical strategy to resolve the conflict between the demands of differentiation and integration is to encode a category in terms of an image of an average category member. Rosch's 'prototypes' can be seen as combining aspects of both the network and feature-comparison models whilst overcoming at least some of their deficiencies. In her experiments Rosch has demonstrated that 'good' instances of category membership are recognised more quickly and more accurately than 'poor' instances, and Shepp (1978) notes that 'sameness' is processed too fast for multi-dimensional consideration[59]. These findings suggest that the initial recognition strategy is holistic (involving simple wholes, templates or configurations in pattern-matching: Garner, 1978[60]), which is followed, should the matching process fail, by comparisons of component attributes[61]. In other words the pattern-matching embedded in the gestalt notion of the ideal type would appear to have a critical part to play in cognitive funtioning - even though the process is likely to be much less sharply defined in the world beyond the psychological laboratory, where abstractions and highly complex stimuli have to be processed.

This has been a lengthy digression from the problem of construing a sample of metal as copper. In Rosch's terms, 'copper' becomes a category subsuming a disparate range of exemplars of copper metal, such as polished copper, a freshly exposed fracture surface of copper, green-tarnished copper and black oxidised copper. The first three

might match the 'prototype' of copper quite well, but the last is likely to be ambiguous: for example, when cleaning the blackened interior of a central heating pump, I was surprised to find the metal to be copper. Then there is the problem of brasses and bronzes, which are less good members of the category 'copper'.....

To return from the problems of cognitive metallurgy to that of whether one construes dichotomously, it is necessary to refer again to the notion of the 'contrastive set'. The argument presented above points fairly strongly to the conclusion that, for physical instances at least, construing in terms of a contrast theory of meaning does not require the restriction to dichotomy and that contrast in terms of multiple oppositions (some of which may be 'submerged'; that is, unavailable to awareness) is a plausible extension to Kelly's position. It may be that, in construing a sample of copper, I restrict the range of multiple oppositions to those metals of a yellow/pink hue - in which case I may merely be using an intermediate multiple opposition involving colour as a filtering criterion, setting yellow-pink against light grey and dark grey.

To restrict the argument to physical examples is not enough, for much construing takes place in terms of abstractions. Is there, then, a case to be made for extending the notion of contrastive sets to encompass abstractions? Reflection suggests that this might be justified, though the argument needs the support of formalised research if it is to be considered more than speculation.

Take, for instance, a literary critic reading a book. He or she might construe it as reflecting a Marxist (or liberal, or conservative, or whatever) standpoint on the part of the author. How would such a

construal take place? There might be clues in the language (such as the use of words like 'struggle' or 'bourgeois') and in the subject matter (which might highlight exploitation and oppression under capitalistic circumstances). To return to Hirsch (1967), the 'genre' of what is printed would be recognised through a series of cues, some of which may lie outside the text itself (the rhetoric of the dust-jacket) and many of which will certainly lie within it. The initial presumption as to genre may or may not be a good one (as Kelly recognises in his explication of constructs), and may need to be revised as more of the text is read. The categories of judgment in this example are likely to be fuzzy since boundaries between rival political philosophies are typically imprecise.

The use of cue validity as a probabilistic concept (Rosch and Mervis, 1975; Rosch, 1977) would seem helpful here. The validity of a given cue X as a predictor of membership of category Y increases with the frequency with which, in general, X is associated with Y and decreases with the extent to which X is associated with categories other than Y^{62}. On this probabilistic model the cue validity of a category would depend on the summation of the cue validities of the attributes used in judgment which, in turn, may be influenced by expectancy and context.

Work by Tversky and Gati (1978) shows that judgments are influenced by context. As an example, when the pair of countries Portugal and Spain was provided, and the respondents were asked which one of France, Argentina and Brazil was most like the pair, the respective percentages of choice were 45, 41 and 14. On introducing Belgium in place of Argentina, the percentages were altered to France 18, Belgium 14 and Brazil 68. The change in context of judgment exercised a marked shift on the (implicit) weightings in Tversky and Gati's set theoretic

equation for similarity.

Findings such as this - and the authors present a number of others - are consonant with those from experiments in recognition which show the power of a contextual expectancy set: one need look no further for examples than Bruner and Postman's (1949) aberrant playing cards; Abercrombie's (1960) X-ray photographs; Anderson et al's (1977) ambiguous prose passages, and

TAE CAT

which probably caused no problems as far as recognition is concerned[63].

Tversky and Gati's (1978) features-based set theoretic model of similarity/difference assesses the similarity of a pair of instances A and B in terms of three variables: the features common to A and B, those possessed by A but not by B, and those possessed by B but not by A. Each variable is weighted for salience in judgment and similarity is assessed by subtracting the weighted last two variables from the weighted first. This model considers the effect on similarity of the context, the direction of comparison, and whether the task is formulated in terms of similarity or difference. It also makes the assumption that people attend more to shared features in judgments of similarity than in judgments of difference, that they attend more to the subject than the referent of the comparison, and that features having a classifactory significance are more salient[64].

The work of Rosch and of Tversky and Gati emphasises judgments of similarity in respect of both concrete and abstract referents. If recognition tasks are processed too fast for multidimensional consideration[59], then the suggestion of construing by a holistic, positive pattern-matching strategy must be seriously considered.

If people store in their memories 'prototypes' of concrete objects, and schemata or 'scripts' regarding human actions (and - perhaps - complex abstractions), construing may take place by selection of the most subjectively probable from a range of possible referents. The cases of the slow worm, the marsupial mouse and the helium atom[54] indicate that instances can exist on the borderlines of categorial distinctions (at the subjective level, if not with respect to more objective criteria).

I have to agree with Kelly that, in construing an element as 'X', I am implicitly denying its membership of all the possible 'not X' categories within the relevant field. This is a weak and pragmatically unhelpful form of bipolarity. To articulate a specific contrast in the Kellian fashion may be satisfactory for some contrasts (say, 'beautiful' versus 'ugly') but not for others ('iron' as opposed to other metals; colour terms; and so on). The articulation of a specific contrast may help in the communication of meaning but, as I argue later[65], this is likely to be of value in respect of a particular element rather than a range of elements.

So far in this section I have argued that Kelly's rejection of the 'concept' is unjustified, as is his assumption that constructs are pragmatically dichotomous. Bannister and Fransella (1980) suggest that Kelly is not asserting that constructs <u>are</u> bipolar but, rather, that it might be more useful to think about them <u>as if</u> they were bipolar[66]. Maybe: Kelly admits to having been influenced by the propositional 'As if' philosophy of Vaihinger (1924*)[67]. However, a reading of Kelly reveals that he presses very hard the case for dichotomy in construing and in action - too hard, in my opinion, for

although many constructs can be viewed in dichotomous terms (particularly linguistic qualifiers) some involve multiple oppositions and others, like 'sex' and 'statistics', seem to have no meaningful opposition at all.

The strength of Kelly's claim for dichotomy in construing may partly derive from his experience as a clinician. To focus on people as elements, and to ask his clients to construe them in terms of similarity and contrast is likely to produce a vast majority of adjectival or quasi-adjectival constructs regarding personal qualities and capable of being used in a dichotomous way[68]. Kelly makes the assumption that all the elements fall within the range of convenience of his constructs, and implicitly assumes that 'unticked' elements in his repertory grid protocol are appropriately located under the contrast pole[69]. In such circumstances it is possible that problems of dichotomous construing, such as membership of <u>both</u> category poles would not emerge with sufficient acuteness for him to feel the point. For Kelly, an object cannot be both black and white - yet he must surely have seen a chessboard, black-suited playing cards, and heads of 'greying' hair.

Why go to such lengths to challenge Kelly's notion of dichotomous construing? My reasons lie in the importance of communication of meaning in research and the limiting effect of the dichotomy assumption on methodology.

If one thinks of sentences, it is a truism that they develop their meanings cumulatively, only closing their signification with the last word. Each word is understood against the context of those preceding

it, its meaning being sealed only by the retroactive effect of those that follow[70]. Consider the following sentence;

The tall man stood and looked at the grey mouse.

'Tall' only becomes meaningful when its referent 'man' is read, and 'tall man' implies (within the limits of fuzziness: Lakoff, 1972) the contrast with shortness, but at this point in the sentence, what is the contrast of 'man'? With no guidance from context (and no guidance is subsequently available) it is most likely to be drawn from - or to consist of - the contrastive set of human beings, a contrastive set in which the categories may overlap since there is no evidence on which to include or exclude certain members. This contrastive set may in these circumstances include 'child' as well as 'boy' and 'girl', for example. In the Kellian sense of 'submergence', all of this contrastive set may never be evoked, in which case the argument for prototypic pattern-matching is strengthened. The rest of the sentence may be treated similarly, noting that 'grey' (as a colour term) is likely to be a member of a set of colour terms that are appropriate to mice.

The oppositional forms of the sentence might take any paradigmatically substitutive route through the basic syntagmatic chain, using the contrasting terms shown below: admittedly, some of the routes make more sense than others.

The	tall	man	stood	and	looked	at	the	grey	mouse
	short	woman	sat		shot	behind		brown	cat
		child	walked		cursed	in front of		white	dog
		boy	ran		swore	away from		black	rabbit
		girl	jumped		shouted				
		baby							

The point being made rather laboriously here is that whilst a Kellian dichotomous contrast theory of meaning may be logically valid, the complexity of meaning of even a simple sentence implies that cognitive operations upon such lines would not be parsimonious. If the contrastive sets and dichotomies <u>are</u> submerged as the sentence is construed (even though they may be identified upon reflection), in practical terms the construer is tending to operate prototypically.

If this argument is valid, it calls into question Kelly's assumption that the dichotomous approach 'comes near the way people actually think'[71]. It also raises serious questions about the capacity of a research instrument that restricts itself to dichotomous constructs to elicit the richness and complexity of human thinking.

3.3.3 <u>The Organization Corollary</u>

> 'Each person characteristically evolves, for his convenience in anticipating events, a construction system embracing ordinal relationships between constructs'[72].

On logical grounds it is, strictly speaking, unjustifiable to discuss the Organizational Corollary at some length within a context of repertory grid methodology since the repertory grid does not address the question of structural relationships between constructs (save in the sense that it can highlight cophenetic relationships). Any claims regarding structure are necessarily inferential. Yet over the years claims have been made, both explicitly and implicitly, that the repertory grid reveals aspects of structure: for instance, Fransella (1972) writes that all existing forms of repertory grid

> '.... are such that statistical tests of significance can be applied to the rankings, sortings or comparisons that each individual has made. This implies the basic assumption that the statistical relationship reflects the psychological relationship between two constructs'[73].

But a statistical relationship tells nothing of the implicative relationship that may obtain between two constructs: to imply otherwise is to confuse correlation with causation. A study reported by Landfield (1982) illustrates the point that near-zero correlations can reflect a uni-directional implicative relationship, and he indicates the need for new methodologies in the approach to superordinate-subordinate relationships.

With regard to the Organization Corollary, Kelly is generally construed as saying that construct systems are hierarchically organised, but it may be that he was rather less pre-emptive in his thinking. A hierarchy suggests metaphors such as 'pyramid', with superordinate constructs at the top and the remaining constructs subsumed at progressively lower levels. Yet Kelly contrasts the notion of 'pyramiding' of ideas with the penetration of ideas with insights - an ambiguous proposition which he does not develop. Stated as a contrast to pyramiding, it suggests a more recursive model in which superordinate constructs subsume subordinates themselves. Such a model would be characterised metaphorically by the flexible woven metal cross-ply tyres of the American astronauts' moon buggy or the geodetic airframe of the Wellington bomber of the Second World War. It is outside the scope of this work to explore this issue further; I mention it merely to draw attention to the possibility of structural models other than the hierarchical.

It is not surprising that structure is typically construed in hierarchical terms, for Kelly writes that '... man systematizes his constructs by concretely arranging them in hierarchies and by abstracting them further'[74]. The key word here is 'systematizes', and Kelly shows that the ordinal relationships may be developed either by elaborating the meanings of the two poles of a construct taken separately or by subsuming the whole of a construct under a superordinate.

In the quotation presented in the preceding paragraph Kelly is not describing the totality of a construct system's formation. Arrangements and abstraction are only part of the process, and the risk in overemphasising them is that it might distract attention from the refining of comparatively gross distinctions; and in the Choice Corollary Kelly makes it clear that improving the definition of a construct system is one of the options available. Kelly's position is not inconsistent with developmental theorists such as Piaget, Werner and E.J. Gibson who see systems being developed in two directions from what are initially rather crude discriminations - 'upward' by increasing abstraction (i.e. integration) and 'downward' by increasing differentiation[75]. And increasing differentiation may provide a basis for the generation of new, or the redevelopment of old, superordinate constructs. The more cognitively complex the person is, presumably the greater the integration and differentiation in the construct system, as Zimring (1971) recognised. The difficulty is to find an instrument that measures both adequately, and this (together with some rather shakily-grounded conceptualisations) seems to have resulted in a literature on cognitive complexity that can be charitably described as highly differentiated[76].

In Kelly's view it is inappropriate to use the canons of classical logic

to construe the organisation of construct systems, since these would produce nested sets of categories and classes which would approach immutability and rigidity. In contrast, if the system is construed in psychological terms it is much more open to shifts in relationships between superordinate and subordinate constructs. The context in which the system is being used may well overturn a hierarchy, as Kelly made clear.

Relatively little empirical work has been done truly focusing on the hierarchical aspects of construct systems, though rather more is claimed in the literature. For the moment it is sufficient to note Hinkle's (1965) limited attempt to probe the nature of structure through the use of a grid (the 'implications grid' or 'impgrid') which elicited implicative relationships between constructs. Ten Kate (1981) has attempted to extend the discussion of Hinkle's theoretical position, but his analysis is flawed first by the nature of the questions he asks regarding the superordinate-subordinate relationship, and second, by his limitation of discussion to constructs of the 'X - not X' type[77]. Given the structural complexity implicit in both Hinkle's and ten Kate's positions, it would seem that empirical investigation of structure would be a brave and hazardous undertaking.

It is likely that the structural relationships between constructs are more complex than the scanty literature on the subject tends to suggest. Hinkle's own work, and that of others, suggests that there may be circularities and intransitivities[78] inconsistent with the rather static notion of hierarchy that is often an implicit assumption (despite Kelly's own recognition of the vulnerability of hierarchies[79]).

The question of structure is a crucial issue for adherents of Personal Construct Theory. The relationship of repertory and implication grids to the Organization Corollary is discussed further in Chapter 4, and some leads forward for the researcher are proposed in Chapter 13.

This consideration of the Construction, Dichotomy, Range and Organization Corollaries covers those parts of Kelly's theory which impinge most directly on research involving repertory grid technique. In the interests of completeness the remaining Corollaries are considered briefly below, in the recognition that a fuller appraisal lies beyond the boundaries of the current work.

3.3.4 The Individuality and Experience Corollaries

> 'Persons differ from each other in their constructions of events' [80].

> 'A person's construction system varies as he successively construes the replications of events' [81].

Kelly is wont to point out that the psychology of individual differences often turns out to be the psychology of group similarities. In the Individuality Corollary he draws attention to the impossibility of any two people experiencing events identically, in that they have both different standpoints and different construct systems with which to cope with their experiences. That is not to deny the similarities in human experiencing and behaviour for, as Bakan (1956) writes, 'after all, we are all pretty much alike'[82]. What might have been added here is that much of the similarity is due to the influence of the cultural environment, especially implemented through the medium of language. Kelly indicates his awareness of cultural relativism when he writes of the possibility of people living in 'altogether

different subjective worlds'[83], but his own interests seem to have led him away from pursuing the matter.

The Experience Corollary emphasises the dynamic aspect of construct systems: people reconstrue in the light of a succession of events unfolding with time[84]. The changes are not necessarily for the good: they are pragmatic and productive of interim positions open to further change in the light of experience. To construe the replications of events does not, in itself, imply the necessity of changing, for the replications may be consonant with the construct system already erected. Kelly observes that continual change would be chaotic, and that parts of construct systems can be expected to remain stable whilst others vary. Stability and change coexist: excess of either is pathological. For the user of personal construct theory, and <u>a fortiori</u> the repertory grid, it is important to know where stability exists and where change is taking place: stability and change are discussed in Chapter 10.

3.3.5 <u>The Fragmentation and Modulation Corollaries</u>

> 'A person may successively employ a variety of construction subsystems which are inferentially incompatible with each other'. [85]

> 'The variation in a person's construction system is limited by the permeability of the constructs within whose ranges of convenience the variants lie'.[86]

In these two Corollaries Kelly develops his ideas regarding change, The Fragmentation Corollary indicates that a subsystem may vary in such a way that the relationship between the later state and an earlier state may not be directly inferable. What may happen is that the earlier state contributes to a change in its superordinate

relationships and the revised superordinates then permit a revision of the subordinates: a dialectic between the levels is involved. (One can imagine, loosely, the parallel with a branch bank manager's initial refusal to accede to a customer's request for a loan but, after consulting head office, a change in his or her position: the latter position is not directly inferable from the former). Thus the states of the subordinate system may be only indirectly related, whilst consistency lies in the superordinate levels above those at which the change is made manifest, and depends upon the 'permeability' of the superordinate system (i.e. its capacity to assimilate change at subordinate levels).[87]

Bannister and Mair (1968) give the Fragmentation Corollary a somewhat different reading from Kelly's explication which (as is evident from the above) focuses on change in subordinate systems, and hence suggests development. Bannister and Mair concentrate upon selection from a range of possibilities when they discuss fragmentation in terms of behaviour which, though it is at first glance inconsistent, is nevertheless consistent within a superordinate setting. They see fragmentation when 'A parent may kiss and hug a child at one moment, smack him a little later and shortly afterwards ignore him when he insists on showing off ...', all of which are consistent when the parent's superordinate constructs regarding child training are considered.[88] This example does not entail change on the part of the parent: he or she may merely be selecting the behaviour construed to be appropriate. Given that Kelly links the Fragmentation and Modulation Corollaries, and that he explicates the latter with reference to the former, it would appear that the interpretation offered by Bannister and Mair fails to capture Kelly's original intention.

In discussing the Fragmentation Corollary, an implicit consideration has been given to the Modulation Corollary. The degree of permeability in superordinate constructs is a determinant of the capacity of the construct system to change. Permeable constructs can cope with elements and subordinate constructs not yet construed within their framework: taking the example in the paragraph above, the espoused theory of child training could be considered permeable if it allowed other training behaviour (such as 'rewarding with tokens') to be subsumed.

But permeability has systemic implications, which go beyond individual constructs. Weick (1968) suggested that a highly interconnected cognitive system would be resistant to an inconsistent element, but if the element 'penetrated' it, the system would be more vulnerable than its more sparsely-connected counterpart which could incorporate inconsistent information by fewer systemic changes[89]. Crockett and Meisel (1974) subsequently found that the amount of structural change triggered by disconfirming information varied according to the'degree of interconnectedness' of the system and to the closeness of the discrepant information to the person's most central (highly interconnected) construct. These findings are summarised in Figure 3.4.

Crockett and Meisel's work, which depended on Hinkle's (1965) implication grid method for its identification of centrality, is one of the few studies to tackle empirically the questions of structure and change, and their findings give support to the model of cognition and purposive behaviour summarised in Figure 3.1. These findings lead me to offer the thought that the notion of <u>systemic</u> permeability/ impermeability may be rather more important than Kelly's explication of Modulation Corollary seems to allow.

Degree of 'Centrality' of disconfirmation	Extent of change	
	(i) Highly connected system	(ii) Sparsely connected system
Invalidating most central construct.	Large change	Little change
Inferentially invalidating most central construct.	Moderate change	Moderate change
Peripheral: respondents merely told they were relatively inaccurate in their construing.	Slight change	Moderate change

Figure 3.4 A summary of Crockett and Meisel's (1974) results regarding interconnectedness and systemic permeability.

3.3.6 The Commonality and Sociality Corollaries.

> 'To the extent that one person employs a construction of experience which is similar to that employed by another, his processes are psychologically similar to those of the other person'[90].
>
> 'To the extent that one person construes the construction processes of another, he may play a role in a social process involving the other person.'[91]

It is in these Corollaries that Kelly addresses himself to the social aspects of personal construing. In the Individuality Corollary he makes the point that every individual has a unique standpoint from which to construe events. The Commonality Corollary presents the other side of the coin in that it focuses upon the fact that, despite idiosyncratic differences, there nevertheless exist broad similarities in the construing of experience. It was noted earlier (page 79-80) that whilst Kelly was aware of cultural relativism, his interest in developing a psychological theory may have led him away from the social dimension of construing. After all, the Commonality Corollary leads towards a psychology of group similarities - and Kelly chose to attack this in setting up his individualistic psychology.

A problem with the Commonality Corollary is the level at which commonality is to be recognised. Kelly might well have said that the degree of 'overlap' between two systems would be a suitable criterion, but a criterion of this sort does not help when one is faced with the practical problems of research. Two people might, let us say, support the Labour Party with sufficient fervour to become activists, yet one of them may see this support in terms of working class solidarity whilst the other may see the party as the best bet for the attainment of a socialist utopia. Pressing the

implications of this potential schism might reveal the first person as conservative and the second as radical in their construings of possible futures. The critic would no doubt make the point that it would besimplistic in the extreme to accept party affiliation as an index of political support without exploring the foundations upon which that support is grounded: one can but agree. However, the suspicion remains that conceptually similar examples probably exist at a range of hierarchical levels in construct systems and may be altogether more difficult to tease out. So, from a research point of view, at what point is it acceptable to claim commonality of construing? It is very doubtful that a single, unequivocal answer can be given, but the question indicates the need for the researcher to be alert to the danger of being satisfied by apparent manifestations of commonality.

Given that the Commonality Corollary and the Individuality Corollary together provide a basis for exploring similarity and contrast in individuals, the Sociality Corollary provides the ground on which similarity and contrast can be seen in action - or, at least, it is Kelly's statement of the conditions for meaningful social action.

Holland (1970) is particularly critical of this Corollary[92]. First, he argues that the statement is tautologous in that 'taking the attitude of the other' is itself the definition of a role-based social process. Kelly would no doubt have countered Holland's criticism by saying that construing the construction processes of another was rather different from what appears to be a psychologistic empathy in that the construer has to subsume the relevant parts of the other's system under his or her own. He might also have said that this subsumption

was a necessary condition, and not a definition, of a role-based social process[93].

Holland's second criticism is that in a number of social processes it is not necessary for a person to construe the construct system of another: for example, the confidence trickster in respect of the victim, and some doctors in respect of their patients. But here he seems to have overlooked Kelly's emphasis on the constructive nature of the social process. One might well criticise Kelly for not exploring social processes other than the constructive, but within Kelly's own conceptualisation (however limited it is) Holland's critique is wrongly directed.

The construing of the construction processes of another was foreshadowed by Dilthey who observed that we understand others by transference from our own inner life, making the assumption that they have a broadly similar mental structure to ours. This does not entail an assumption of psychological isomorphism, for it can be argued that understanding is based on analogising and extrapolating from personal experience (or episodic memory: Tulving, 1972). Thus if I have suffered the death of a close relative I can 'understand' another person's cognate loss even though the details are very different. I am then in a position to offer support in a much more deep sense that if I had only read about and 'intellectualised' how people feel after a bereavement: on this argument semantic memory would appear less satisfactory than episodic memory as regards the playing of a constructive part in a social process.

In this type of approach 'understanding' and social process might be sketched in the following way. On hearing of another's existential circumstances, a person searches his or her store of cognate

experiences and abstracts such superordinate constructs as seem appropriate. These are then communicated to the other who construes the superordinates in the light of his or her (subordinate) existential position. In practice this may well be a more complex hermeneutic dialogue through all the channels available to the participants. As the dialogue deepens understanding, so the actors are enabled to involve themselves in a constructive social process.

The tone of the whole of Kelly's explication of the Sociality Corollary is consensual - he is interested in people getting along with each other. But outside the sheltered environment of the clinician's room not all dialogue is harmonious and constructive[94]. In their explication of the Corollary Bannister and Fransella intimate the possibility of conflict in social processes when they write about confusing and defeating others[95]. It may be that conflict for Kelly is the submerged contrast pole of consensus, since conflict emerges obliquely in some of his definitions of constructs relating to transition (in particular, those of 'threat', 'fear', 'guilt' and 'hostility') if they are read with social processes in mind.

Kelly sees the Sociality Corollary as the take-off point for a social psychology. A more apt metaphor might be "in the hangar, under construction" since he does not elaborate the 'group' aspects of a social psychology. But this would require additional theoretical formulation, and there is little evidence yet that this is being pursued by personal construct theorists.

3.3.7 The Choice Corollary

> 'A person chooses for himself that alternative in a dichotomized construct through which he anticipates

> the greater possibility for the elaboration of his system'.[96]

In the Sociality Corollary Kelly partially develops the relation of action to personal construct systems, but the relationship is strengthened in the Choice Corollary. In this Corollary, too, there is the most explicit recognition of people as being active and exploratory within their environment. However, as Tyler (1981) notes, it is not easy to accept that people are always sufficiently energetic to follow the implications of the Corollary[97] which, at the root level, offers the person Hobson's choice regarding elaboration.

The underlying construct of the Choice Corollary is something like

'anticipates less possibility of elaboration - anticipates greater possibility of elaboration'

and, in Kelly's formulation, the option is limited to the right hand pole[98]. Once this pole has been 'chosen', it is fair to consider whether that choice relates to the extension of the construct system beyond its previous boundaries, or to increasing the degree of definition within its existing scope. The argument presented above can be illustrated with reference to the following hypothetical situation.

Suppose I have had a hard day at work followed by a good evening meal, and it is now seven o'clock. I could go to sleep early, go to the bridge club with my regular partner, or go to a dance club in the hope of meeting up with a new partner (among a range of other possible actions). Going to the bridge club may offer me the opportunity to increase definition in that I might learn more accurately when to use certain techniques - say, a Vienna Coup and

squeeze. Going to a dance club might have considerable potential for extension, but it is difficult to see how going to sleep would increase either definition or extension of my construct system.

It might be felt that such an exemplification is unfair to Kelly's intentions, but he implicitly acknowledges the point being made when, in discussing experience, he gives the example of the veteran school administrator who had 'one year of experience - repeated thirteen times'[90]. According to the letter of Kelly's writing this cannot be strictly true, but the spirit suggests that the administrator did not seek greatly to elaborate his construct system as far as school administration was concerned. Bannister and Fransella (1980) remark that people can over-define to the point of boredom[100], but they seem to be confusing repetition of experience with the development of an increasingly defined (refined?) system: the two are not identical. I can go to an air display merely to enjoy the sight and sound of aircraft, but I can also go to develop my discrimination of types and sub-variants. The first, though anticipatory, offers far less for intentional elaboration than the second.

Bannister and Fransella are themselves unhappy with the Choice Corollary, seeing it to be in need of both definition and extension itself. There is no indication as to why the individual should choose definition in preference to extension (or vice versa); all that is suggested is that the person will try to move away from confusion and towards understanding. There seems, as Holland (1970) suggests, no way in which the Corollary can be tested empirically[101], and Bannister and Fransella (1980) seem overoptimistic when they suggest that the Corollary is testable provided enough is known about an individual's system in order

for predictions to be made in terms of that system[102]; such a proviso is of a truly enormous magnitude.

It will not have escaped notice that the 'evening out' example used a few paragraphs ago did not involve a dichotomous choice. If one does not accept the Dichotomy Corollary as it stands, the implications reverberate through other aspects of the theory, and particularly loudly in respect of the Choice Corollary. There is no particular reason to suppose that decisions about action are taken on a dichotomous basis (though any multiplex of possibilities can be presented in terms of a binary algorithm). The 'evening out' decision related to a series of categories (as often does judicial sentencing), but decisions are not necessarily 'digital' in character. It is quite possible to construe some decisions in 'analog' terms - for instance, choosing one's speed in driving along a motorway. Action does not necessarily imply only opting for one construct pole or the other.

Returning to Tyler's (1981) point, on the whole it seems improbable that people constantly act to maximise the elaboration of their construct systems: there might be anarchy if they did! Critiques of the notion of personal freedom from a variety of standpoints are united in their recognition of the power of social forces to limit elaboration[103], in some cases coming close to adapting the structuralists' claim for language to say that "our culture behaves us". Kelly's theoretical framework, with its emphasis on individuality, betrays its country of origin - in which the espousal of individual freedom is encapsulated in the idea of the 'American Dream'. And yet, as Ward (1974) points out, there is a paradox - if not a

contradiction - in a society which enshrines individualism as an ideal whilst at the same time socialising individuals to accept as their own the demands of the society in which they live, and yet from which they believe themselves to be independent.

3.4 SUMMARY

It must be reiterated that this review of Personal Construct Theory is but partial. It is partial in the sense that it has not sought to submit the theory to searching scrutiny through a variety of philosophical templets. It is partial, too, in the sense that its broad perspective is that of the user of repertory grids, with the inevitable bias towards cognition that is entailed.

Whilst the fundamental Weltanschauung of Kelly's theoretical position has much to offer the researcher in the human sciences (and this theme is taken up later, in Chapter 13), this review has been critical of a number of aspects of Kelly's specific formulations (not least because there seems a tendency amongst adherents of the theory to accept its propositions uncritically). The main points of criticism are summarised below.

1. The Fundamental Postulate is inadequate in that it gives insufficient weight to the effect of adventitious experience upon psychological 'channelizing'.
2. The socially determined aspects of construing are consistently underestimated.
3. Construing is not necessarily limited to a dichotomous operation: both (classical) concepts and multiple oppositions may well coexist with dichotomous constructs in a construct system.

4. Construct systems may contain features (such as vicious circles) which are not easily subsumed under the notion of hierarchy.
5. The relationship between construing and action is insufficiently developed.

Of these points, (3) has immediate implications for the user of repertory grids, whilst (1) (2) and (4) - as amplified in the review - refer to problems which may not surface during the administration of a repertory grid, but which may colour the interpretation of any grid data collected. Where research relating grid data to action is envisaged, (5) assumes importance.

In the following chapters the focus is narrowed more specifically to repertory grid methodology and, in that context, to some of the practical strengths and weaknesses of Kelly's Dewey-eyed methodological individualism.

4 Theory and method

In essential tension ?

What we observe is not nature itself,
but nature exposed to our method
of questioning.

W. Heisenberg
1962*: 58

4.1 INTRODUCTION

The main content of this chapter is an initial consideration of the links between repertory grid methodology and Kelly's theoretical formulation, and the view is advanced that grid and theory are reciprocally related - but only to a limited extent. In this outline sketch attention is given to three general issues which bear upon the use of grids: the nature of the cognitive task, language and meaning (which is treated more fully in Chapter 11), and the relationship between researcher, respondent and task.

The question of the structure of construct systems looms disproportionately large in the middle of the chapter. The length of this discussion - which argues that repertory grids elicit little, if anything, about structure - is justified on the grounds that a persistent confusion exists in the literature regarding the grid's capacity to capture the intentions of the Organization Corollary.

In considering the relationship between theory and grid some of the implications of the challenges made to theory in Chapter 3 are explored, but the broader purpose of this chapter is to set the scene for the more detailed appraisal of repertory grid methodology in Chapters 5 to 11 - and for the alternative position advanced in Chapter 13.

4.2 REPERTORY TEST AND REPERTORY GRID

In presenting his original Repertory Test[1], Kelly refers to it as being methodologically an application of concept-formation test procedures. However, instead of sorting into categories items such as the blocks of the Vygotsky test, the person sorts people - people with whom he or she has a role relationship. Kelly suggests that this direct approach through the construing of relations and

acquaintances is very appropriate to the investigation of problems presumed to be of an interpersonal nature.

In the Repertory Test Kelly asks the respondent to supply names to fit a number of role titles. He then presents triads of these names (chosen to suit the purposes of his investigation) and requests the respondent to construe the triads in terms of the similarity of two of them and the contrast of the third, hence producing a dichotomous construct. This could be seen as a disadvantage, in that the procedure relates only the triad of elements to the construct and does not seek to generalise the construct to the remaining elements. A particular verbal label might be elicited in response to other triads, but the picture of relationships between elements and constructs as revealed by this procedure seems very 'patchy', to judge by Kelly's example of Mildred Beal's repertory test[2].

The grid form of the test[3] allows each element to be located dichotomously on each construct (or, possibly, to be judged to be outside a construct's range of convenience), thus creating a complete element X construct data matrix which is open to mathematical analysis. Kelly himself developed a form of nonparametric factor analysis to help reveal the relationships between the constructs, and between the elements, in the grid.

Kelly's grid test has been adapted in the course of time to cater for the demands of research in a wide range of fields. Elements and constructs have been elicited from the individual in a variety of ways, or have been provided by the researchers. The original dichotomous allocation has been largely superseded by other ways of relating elements to constructs, notably by ranking and rating procedures, and

increasingly sophisticated methods of analysis have been developed to cope with the information captured in the grid matrix. What remains common to all forms of repertory grid is the interpretation of the full range of elements (stimuli) in terms of the full range of constructs (dimensions) and the presentation of the data in the form of a complete element X construct matrix.

Much has been claimed for this methodology: for instance, Thomas (1978) writes that the repertory grid 'is an obvious choice for exploring [personal construct] systems'[4], and Salmon (1978) goes rather further when she claims that

> '... the whole range of systematic enquiries that Kelly called the repertory grid technique can reveal the structure underlying major aspects of intuitive personal understandings'[5].

Others are less optimistic about the potential of repertory grid methods. D.E. Bolton (1978), for example, points to the grid's inability to tap the historical development of constructs and to investigate both memory-processes and the relationship between speech and action[6]. Watts and Pope (1982) are much more cautious about claims such as Salmon's when they indicate that the grid is too often seen as the <u>only</u> methodology compatible with Kellian theory, and that the numbers in the matrix are too easily congealed into absolute statements about people[7]. The arguments presented in the following pages will be found to be more in accord with the sceptic than with the believer.

In repertory grid methodology a number of issues appear to have been taken for granted. Little attention has been given, for example, to the nature of the task of completing a repertory grid, to the

implications of language, and to the effects of the interaction between researcher and respondent during grid administration: these matters are taken up in the succeeding sections.

4.3 REPRODUCTION OR RECONSTRUCTION?

As far as cognition is concerned, a definitive answer can probably never be given, but it is worth exploring the issue to see what the researcher might be collecting through the use of the grid.

If the elements are concrete objects placed in front of the respondent and he or she were required to construe selected triads in the standard Kellian manner, it is probable that the constructs would reflect discriminations previously made (explicitly or implicitly), rather than be new constructs made up on the spot: in effect, following the assimilative cycle set out in Figure 3.1. The elicitation, as far as the grid is concerned, would seem to be limited by the language already acquired. In these circumstances, construing is likely to depend very heavily on constructs already stored in the system[8]. Research using the grid in this way has been conducted in the area of marketing (e.g. Frost and Braine, 1967) in which the objects of interest have been placed before the respondents. It can be argued that the use of photographs of people as elements comes close to this, but the constructs elicited in this case are likely to be mediations of 'social' constructs projected on to the photographs: photographs, in themselves, do not have personality characteristics.

In mentioning photographs, the problem begins to surface. Suppose, instead, that the respondent is asked to construe people from a list of named acquaintances. He or she cannot construe them 'in the flesh', but must rely upon aggregated conceptions of them as people developed

through interaction and observation. These people are reconstructions, probably based on the existing construct system, but subject to the distorting effects of selective sedimentation in memory[9], constructive processes in memory[10], selective retrieval, recency of meeting, the context in which the grid is elicited, and so on. They may, in some cases, be very good reconstructions, but they are nevertheless reconstructions in the mind and inaccessible to scrutiny by the researcher. The same is true of other elements drawn from episodic memory (Tulving, 1972), such as critical incidents.

The point is made not in criticism of the grid, for such validity as this argument has extends to other elicitation procedures and instruments. It is made in order to emphasise that the grid task is, in the majority of cases, likely to be reconstructive at more than one level, despite the use of the existing system as a framework for the reconstructions.

As I argue in Chapter 6, it seems rather unlikely that developments in the construct system (other than minor) take place during grid administration.

Is the situation any different when the elements are presumed to be drawn from semantic memory? Take, for instance, the general teaching situations which were the elements I used in my work with the science teachers (Appendix 1). Here the teachers were asked to construe elements such as 'teacher exposition' and 'pupil practical exercises' with the help of a short sentence amplifying each in respect of what I, as researcher, sought to investigate within the framework of my study. In these circumstances what the respondents did is more uncertain. They may have located practical work in terms of a

specific (perhaps untypical) instance, a generalisation derived from a number of specific instances (like a 'prototype'), or a more abstract characterisation derived from the theoretical side of teacher education: the categories are unlikely to have been so clear-cut in practice[11]. They would all, however, be reconstructions of one sort or another. There is a further possibility in that, beginning with one of the generalisations, the respondent may have constructed a hypothetical 'scenario' of practical work in operation. To construe this would be to construe a construction, a fabrication, rather than to reconstruct 'relatively straightforwardly' from memory[12]. These are difficult matters for the researcher to probe.

In most cases the researcher is collecting a picture of the past as seen through the eyes of the present. The further the elements are from direct experience both temporally[13] and through the level of abstraction, the more problematic becomes the nature of the sorting task of the grid, and the greater the risk of the researcher collecting a picture incongruent with the intentions of his or her work - and perhaps of being enchanted by the distant view.

4.4 LANGUAGE AND MEANING

Kelly makes a most curious statement when he observes that the researcher's understanding of the respondent's verbal labels is less crucial in the grid form of the test than in the original Role Construct Repertory Test. He goes on to remark:

> 'The use of the grid permits the psychologist to make a rather extensive analysis of the protocol without once looking at the terms which the client has employed. It is this advantage that argues strongly for the use of the grid...' [14]

A little later, when writing about his homespun version of factor analysis, the point is amplified:

> '.... we have suggested a type of analysis.... which transcends both the subject's verbiage and the particular individuals who make up his personal-social milieu. The test protocol can be meaningfully analyzed even if one clips off the names from the top of the test blank and the constructs from the side.' [15]

To take Kelly's statements at face value would be to do him an injustice, given his interest in what his respondents chose to tell him in their own words. What is meant is that the researcher can identify statistical associations between elements and between constructs without having to make inferences from the language used by the respondent: the mathematical potential of the data matrix makes this possible.

Yet Kelly's own language hints at a bias towards the (objective?) mathematical and away from the (subjective?) linguistic. The connotations of 'verbiage' are somewhat slighting, and to say that a grid can be 'meaningfully analyzed' without reference to its linguistic frame is to make a claim for the meaningfulness of a structure devoid of content - if such a thing exists. Here, and elsewhere in the exposition of his psychology, Kelly betrays a fascination with the binary mathematical to the extent that one suspects it on occasion of determining the nature of the psychology.

Statistical associations, irrespective of the formulae used to compute them, are very crude indexes of the relationships of meaning between rows and between columns of the grid matrix. Such associations may be adventitiously cophenetic, in that the sample of elements and/or constructs may be too small to permit differentiation: Kelly himself recognises that the patterns of interrelationships can be idiosyncratic manifestations of the particular elements and constructs in the grid.[16]

There are many possible implicative relationships between constructs (Hinkle, 1965; ten Kate, 1981) and these may be embedded in measures of statistical association. As Landfield (1982) shows, the relationship between implication and association is complex. It is simply not good enough to process a grid matrix through a program which computes associations and to make claims for relationships thus 'revealed'[17]. A much more penetrating analysis is necessary to substantiate any inference beyond the reporting of statistical association. That analysis must make use of language.

It is in the realm of language that repertory grids exhibit substantial weakness[18]. In the previous section of this chapter, attention was drawn to the need for a clear understanding between researcher and respondent regarding the basis on which elements are being construed: such understanding is inevitably based in language. Elicited constructs (and 'supplied' constructs, for that matter) very often take the form of adjectives or short adjectival phrases, and this presents difficulty for the intersubjectivity of understanding. There is an indefinite set of connotative meanings for each such verbal label, and, without further elaboration, the researcher has no guidance as to which the respondent does and does not intend. Further, the researcher's (personal) indefinite set acts as a filter for interpretation.

It is almost a commonplace for users of repertory grids to discuss elicitation in terms of a structured conversation: if practice is consonant with rhetoric, and appropriate records are made (perhaps using a cassette recorder) much amplifying detail can be obtained and used in conjunction with the linguistically-restricted grid data:

this would be consistent with Ziff's (1972) point that the continued modulation of meaning is characteristic of all discourse in natural language[19]. Given the complexity of the task of grid administration, and its lengthiness, it is difficult for the researcher at this stage systematically to explore structural relationships, though some initial intimations might be obtained. It would seem that the identification of structural relationships embedded in grid data depends upon a subsequent retrospective evaluation, unless the elicitation procedure can incorporate the systematic exploration of them. The practical difficulties of so doing are considerable: the work of Eden and his colleagues, though not aimed at eliciting complete grid matrices[20], suggests that a computer program may well be the most efficient administrator in this respect.

A further problem, which also has linguistic implications, arises when a construct is applied to elements beyond those (typically three) in respect of which the construct is elicited. The assumption is made that the construct can be applied consistently to all the elements being subsumed. In other words, the construct is being used invariantly (like the physicalist scale of length) and hence in a manner strongly reminiscent of Hempel's (1949*) positivistic covering laws. This point is discussed in more detail in Chapters 8 and 13: it is sufficient for the present to adumbrate the opposite view - that a construct is likely to shift in meaning as different elements are successively considered.

Language and meaning are critical issues in repertory grid methodology, and yet they have been very little explored to date. The points made in the preceding paragraphs are in the manner of abstracts of the explorations undertaken in Chapter 11, in which I elaborate upon some

of the communication problems associated with grid methodology.

4.5 INTERACTION BETWEEN RESEARCHER AND RESPONDENT

The advantages of treating grid administration as a conversation were pressed in the previous section of this chapter. There is another side to this particular coin: the disadvantage that, the more the grid data is augmented through conversation, the more open is the procedure to 'contamination' from the researcher. The dilemma is captured by Rosenthal (1969) when be observes that the interaction of experimenter and subject is a major source of knowledge in the behavioural sciences[21], yet is open to the profound influence of nonverbal cues such as nods, smiles and glares - and a whole host of direct and indirect experimenter expectancy effects. Orne (1962, 1973) indicates the importance of understanding how the respondent is construing the situation with which he or she is faced, since this can be radically different from what the researcher thinks is taking place. Bannister and Fransella (1980) indicate clearly the reflexivity of construct theory when they remark that the respondent, like the researcher, is construing the interaction[22]: the closeness with Orne's point here is marked.

There is, inevitably, an asymmetric power relationship whenever a researcher asks a respondent to complete a grid, though this is true of any circumstance in which the researcher solicits the assistance of a respondent. Weick (1966) observes that a respondent makes two decisions in respect of participating: the first is whether to participate, and the second relates to the level of commitment to be given [23]. Although some respondents become very interested and involved in repertory grid work[11], this is not always the case.

Murphy (1978) for instance found difficulty in using grids with organisational consultants, Armstrong and Eden (1979) obtained a similar response from Local Authority professional valuers, and Scott (1962) avoided potential problems of this sort when he rejected Kelly's original Role Construct Repertory Test as too cumbersome for a non-captive audience. It has to be pointed out that a lot of grid-based research has involved students whose participation is rewarded by some form of course credit. One might surmise that grid methods will be maximally acceptable where respondents see some kind of 'payoff' for themselves, provided that the cost-benefit equation is not adversely affected by the particular procedures chosen by the researcher.

To complete a grid is an act of self-disclosure. Cozby (1973), reviewing the literature, indicates a number of points which may have a bearing on grid administration though the evidence is rather equivocal. Situational variables, including the relationship between researcher and respondent, consistently influence what is disclosed. Positive encouragement, and disclosure on the part of the researcher both appear to encourage the respondent to disclose more. There is conflicting evidence regarding the effects of race, acculturation[24], and the sex of researcher and researched - and a positive correlation between disclosure and extraversion is reported. As far as repertory grid research is concerned, Reid (1976) remarks on differences in 'openness' in his interviewees[25], and Lipschitz (1972) observed that, when a 'familiar female experimenter' was replaced by an unfamiliar male, less 'private' information was given[26]. My own experience in eliciting constructs has been mixed in this respect. Some of my respondents have been very frank - even indiscreet - in their

comments on their colleagues and on the systems operating within their schools, trusting my commitment to confidentiality and finding it helpful to discuss issues that concerned them as probationary teacher, with someone who would not, at some later date, don the mantle of institutional judge. Others have been less open in this respect, limiting their responses to the specific task in hand. Having read Labov's (1972*) account of interviews with alienated black children[24] (and the asymmetries involved) it seemed important to make some attempt to find out how my respondents had perceived my role as researcher. At the outset, two of the fourteen probationary science teachers had seen me as something akin to an inspector, but after the first grid administration and feedback session none of the respondents appear to have construed me in terms other than those of a researcher with no axe to grind as far as their careers were concerned.

The completion of a grid, particularly where this involves the elicitation of elements and/or constructs, can be a demanding process in terms of both the nature of the task and the time taken for its completion. This leads to pressure on researcher (as I have experienced) and respondent, since both are likely to be aware of constraints such as the time available and the other's patience, and there may be temptations to take short cuts in the interests of completing the grid. The respondent may work superficially and carelessly, whilst the researcher may 'lead' the respondent towards the 'provision' of information that otherwise might not have emerged. There is a danger that 'getting the grid completed' could become a more dominant interest than meaningful communication about the problems towards which the research is directed.

The user of grids, like any other researcher in the human sciences,

is not immune from the conflicts between social process and 'objectivity'. If meaningful data are to be gathered, this can only be done at the risk of the researcher's influence on that which he or she seeks to investigate: the human sciences' analogue of Heisenberg's uncertainty principle in physics. Fiske (1971), writing about 'personology' (and whose comments are broadly generalisable within the field of the human sciences), suggests that one way to resolve the problem is to recognise that the data collected are inseparable from both the observer and the methods used in its collection. As with a standard laboratory experiment in physics, some attempt can be made to assess the magnitude of the 'errors' involved. This problem of 'contamination' raises a number of theoretical and methodological issues which it would be inappropriate to pursue at this point, and discussion of them is deferred to Chapter 13.

4.6 THE REPERTORY GRID AND COGNITIVE STRUCTURE

It will be recalled that earlier in this chapter Kelly was taken to task for overvaluing the mathematical relationships that could be derived from a grid matrix at the expense of understanding the embedded relationships through the medium of language. Some users of repertory grids[27] have taken the further step of assuming that statistical analysis of the grid matrix can reveal aspects of cognitive structure. This assumption can only be true in the weak sense of indicating "what goes together" and structural claims beyond this are at best inferential - hence the title of this section is not a little ironic. Repertory grids alone can provide no indication of a construct's degree of superordinacy or subordinacy, or whether it is a core or peripheral part of the construct system, and so on. Hinkle (1965) recognised this and devised 'implications grids' and 'resistance-to-change grids'

in order to explore issues relating to the hierarchical status of constructs and their psychological salience.

As was noted in the previous chapter, a consideration of structural aspects of construct systems is - strictly speaking - outside the remit of a study of repertory grid methodology. However, in this and the next section I explore some of the inadequacies of repertory grids as far as structure is concerned, and also Hinkle's work on structure, since these are essential precursors to the arguments I present in Chapter 13.

Some of the blame in respect of the overvaluation of statistical association can be laid at Kelly's own door. In Chapter 6 of 'The Psychology of Personal Constructs', which is entitled 'The mathematical structure of psychological space', he is concerned to explore the factor analysis of grids in order to point up the relationships between elements and between constructs. Although Kelly does not claim in his book that the <u>psychological</u> structure can be revealed by factor analysis, it might well be inferred that this is what is in his mind when a statement such as the following is met:

> 'Thus we may have a mathematical basis for expressing and measuring the perceptual relationships between the events which are uniquely interwoven in any person's psychological space' [28].

The assumption in factor analytic procedures (whether based upon correlation coefficients or upon nonparametric matching scores) is that all the rows (or columns) of the grid matrix are of equivalent status. In most uses of the grid this is more likely to be true of the elements than of the constructs, and the following discussion is directed towards the latter rather than the former.

Factor analysis, then, would treat the following (hypothetical) constructs relevant to life-situations as equivalent:

(a) 'vital to my well-being - unimportant to my well-being'

(b) 'encourages tidiness - does not encourage tidiness'

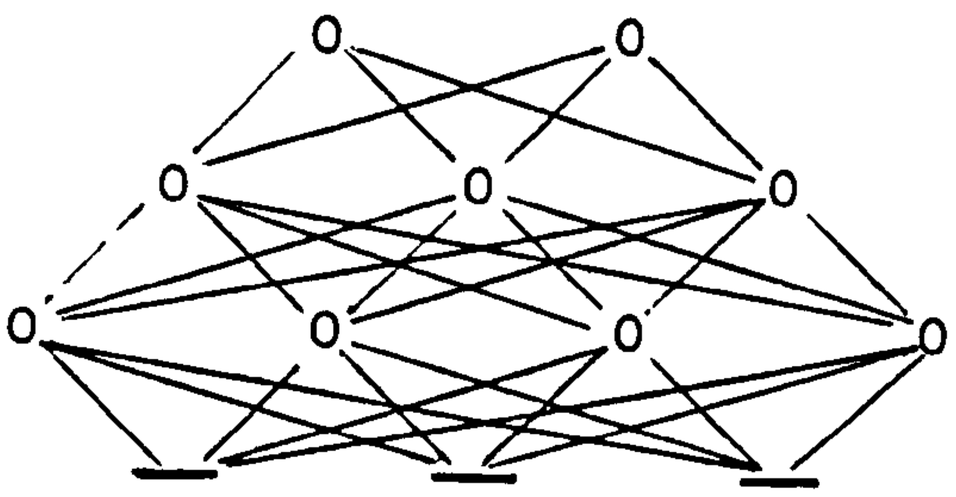
(c) 'social aspect of living - individual aspect of living'

despite the fact that it could be argued that (a) was of more central importance in most, if not all, contexts.

Suppose, now, that (b) and (c) correlate + 0.9 whilst each individually correlates zero with (a). Factor analysis would produce a first factor accounting for the bulk of the variance and relating the encouragement of tidiness to the social aspect of living. The psychologically more central construct (a) would, in effect, constitute a much smaller second factor. The example is patently artificial, but if a grid were to contain a number of constructs that could be subsumed under 'good-bad' (i.e. that are evaluatively 'loaded'), these might intercorrelate to produce a very heavy first factor which would 'background' other dimensions[29] (the weight of the 'evaluation' factor consistently found in semantic differential research comes to mind in this respect). The problem is that the researcher, without further information from outside the grid, has no way of interpreting the outcome of the factor analysis. The corollary to this argument is that it is invalid to presume that factors derived from an analytical routine represent superordinate constructs: to do so is grossly to confuse statistical association with psychological structure.

Considering the issues at a more abstract level, the congruence between Kelly's Organization Corollary and the notion of integrative

complexity[30] proposed by Schroder et al. (1967) seems quite close. The latter, as difficult to validate as Kelly's corollary, nevertheless contributes as a heuristic to the argument. Schroder et al. see the integratively complex person as possessing four levels of cognitive organisation rising from basic dimensions for discrimination via rules for combining dimension scale values and rules for comparing results to - at the highest level of abstraction - a structure for generating complex relationships[31]. Inverting their diagram of these relationships allows it to be aligned with Kelly's notion of hierarchy, given the addition of elements (implicit in the Schroder et al model), as shown in Figure 4.1.



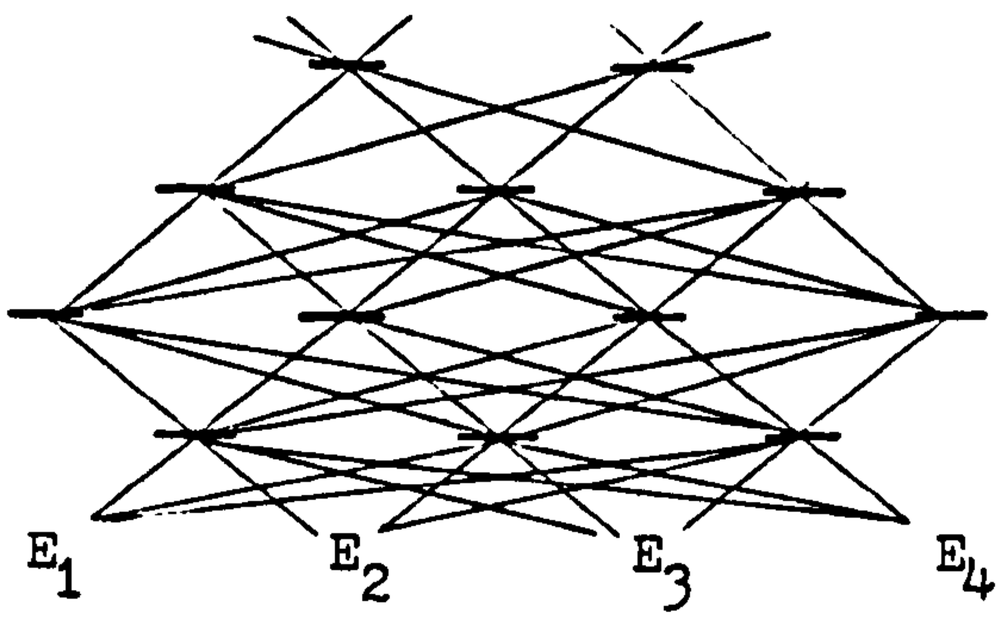
Structure for generating complex relationships

Rules for comparing results

Rules for combining dimension scale values

Dimensions

(a) Psychological structure of the integratively complex person (after Schroder et al, 1967:22).



etc. etc.

3rd order constructs

2nd order constructs

1st order constructs

Elements.

(b) An illustration of the Kellian notion of hierarchy in construct systems

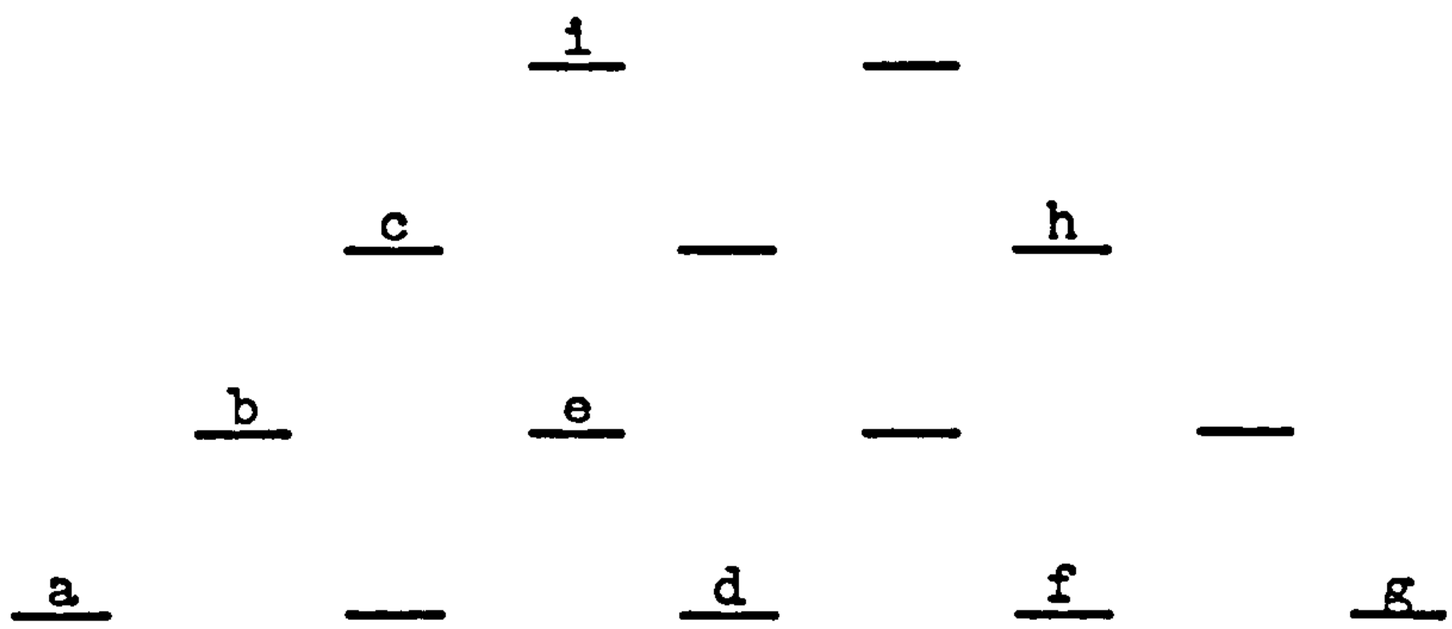
Figure 4.1. The relation between integrative complexity and Kelly's Organization Corollary.

Figure 4.1 gives a rather rigid, mechanistic impression and does not attempt to portray the inconsistencies apparent between subsystems or the general flexibility that Kelly describes. Kelly sees

construct systems as very much open to change, perhaps even reversing superordinate-subordinate relationships from time to time. It may also be that the relationships between constructs are not rigidly hierarchical, allowing direct links between some subordinate constructs and some superordinate constructs, or the location of constructs at mezzanine levels in the structure: there seems no particular a priori reason why one hierarchical model should be preferred to another.

Hierarchies are probably far more complex than can be represented in a simple two-dimensional model. Kreitler and Kreitler (1976), in considering the nature of responses, construe orienting, defensive, adaptive and conditioned responses as not only very complex behaviours involving high neuronal centres, but also as differing from each other in so many respects that it is doubtful whether they all belong to the same class or represent the same level of behaviour[32]. One might conclude that, if behaviours are complex both in themselves and in the ways in which they are related, any model that can be drawn on paper is likely to be the grossest of oversimplifications. These criticisms notwithstanding, the Kelly - Schroder et al heuristic has value in that it can show why factor analysis or any other clustering algorithm is inadequate to reveal hierarchical aspects of construct systems.

With reference to this model, an elicitation procedure might result in the production of a number of constructs (a) to (i) at different levels in the person's system. Without showing the elements and all the interconnections of Figure 4.1 (b), the situation might be represented as



Yet, when analysis is undertaken, the implicit structural assumption is equivalence of level:

a b c d e f g h i

overlooking the fact that the constructs stand in different implicative relationships with each other - which may be more important to an understanding of the system than the various statistical associations thrown up by the analysis[33]. For example, what could be inferred from a correlation coefficient of + 0.5 between constructs (i) and (f)? And what might have been the effects on analysis had the "missing" constructs been elicited?

4.7 HINKLE'S WORK ON STRUCTURE

Hinkle's (1965) work falls outside the purview of 'standard' repertory grid procedures in that he is interested in the superordinate-subordinate relationships between constructs, rather than in the associative relationships capable of being derived from the repertory grid matrix. Whilst his study is open to criticism[34], his methods highlight matters unapproachable by way of the repertory grid: a brief summary is offered here[35] since his work is of intrinsic interest and, more importantly, some features are drawn upon in the construction of the arguments presented in Chapter 13.

In his research Hinkle elicited ten constructs from each of his 28

respondents, using the triadic 'difference' method[36]. He then asked, in respect of the first of these, which pole the person preferred and followed this by asking for the reason for the preference. This led to (what Hinkle interpreted as) a superordinate construct which was then treated in the same way, the procedure being repeated until the respondent ran out of superordinates: the process was then begun afresh with another of the original constructs. Ten superordinate constructs were elicited by this technique.

The constructs elicited (now 20 in number) were taken one at a time, the respondent being asked on which of the remaining 19 he or she would need to change preference if forced to change on the selected one. This gave Hinkle a full matrix of implicative relationships in which asymmetry of implication could be accommodated. In general, he found that the superordinate constructs had more implications for change than did the subordinates, supporting the suggestion that superordinates really did exist at a higher level in the system.

Hinkle also devised the 'resistance-to-change' grid. In this type of grid the respondent was first asked to indicate the preferred pole for each construct, after which each construct was paired once with each of the remaining constructs[37]: the respondent was then asked on which of the two preferred poles he or she would be more prepared to give way if forced to choose. It was open to the respondent to indicate that it was impossible to make a choice, or that change on one logically entailed the necessity of changing on the other. This procedure enabled the rank order of 'resistance-to-change' to be computed for the 20 constructs. Hinkle found that the superordinate constructs were, in general, more resistant to change and that the rank orders of constructs, using the two separate criteria of number

of implications and resistance-to-change, were highly significantly correlated.

In the light of criticisms of his methods it is clear that Hinkle's work must be adjudged as a pioneering study opening up new terrain for others, rather than as providing any definitive statement on structure in construct systems. The 'resistance-to-change' grid, for instance, is not as simple to interpret as might seem the case at first sight. First, there is no guarantee that 'laddered' constructs are superordinate to 'unladdered' constructs and, second, the nature of some constructs may make them more resistant to change than their apparent degree of superordinacy might suggest. Having used 'resistance-to-change' grids with science teachers, there were occasions when they wished to use constructs like 'compulsory-voluntary' and 'in school activity-home activity' which are, in a sense, 'situational' constructs. These proved to be very highly resistant to change, yet it is plausible to infer that this resistance derives from the broad context in which the teachers were operating (reflecting curricular decisions taken at a level beyond their power to exert much, if any, influence) and not from the ways in which they construed science teaching <u>per se</u>.

Resistance-to-change and implications grids do, however, offer some insights into the structure of construct systems. It is not surprising that Hinkle's findings from the two were significantly correlated, for they seem to overlap considerably regarding what they were tapping. In the absence of further evidence it must be presumed that such grids collect data about 'ought' aspects of structures, rather than those aspects reflected in some sort of action. The resistance-to-change grid asks the respondent to indicate preferences

and then to give evidence regarding the strength of those preferences, thus establishing the order of importance of constructs on a personal basis. The implications grids, in contrast, seem less personally involving once the initial elicitation has been completed, despite Hinkle's use of the 'if you were to be changed...' type of questioning: the task can be approached in a more detached manner than can that of the resistance-to-change grid.

It may be that these two types of grid are, somewhat loosely, investigating the salience of the respondents' constructs. Underlying aspects of structure would appear to be being tapped (particularly in the implications grid), but the implications of Hinkle's findings for hierarchy in structure are rather less clear-cut than might be assumed. In particular, the 'laddering' procedure does not necessarily elicit superordinate constructs sequentially, to judge from the implications grid findings: the relationship between laddering and structure would appear to be indeterminate. The example from Hinkle's work which is quoted in detail in Bannister and Mair (1968) serves as a case in point[38].

In his thesis Hinkle did not indicate the subordinate origins of particular superordinate constructs, which makes it difficult to produce an accurate appraisal of his work. However, it seems very likely that the construct

(i) 'face problems - escape from problems'

was laddered to produce, in ascending order of superordinacy,

(ii) 'solve problems - can't'

(iii) 'achievements - few achievements'

(iv) 'boost ego - tear it down'.

One would expect that, if laddering does elicit progressively more superordinate constructs, the implicative relationships would be undirectional from the relatively superordinate to the relatively subordinate. However, although Hinkle's results show a perfect implicative transivity between (i), (ii) and (iii) in the expected direction, (iv) is shown to be implicatively reciprocal in respect of each of the other three - in other words, its hierarchical status is ambiguous. Other examples can be found, passim, in Hinkle's results.

Without access to Hinkle's subject, it is impossible to do other than speculate, but it seems at least plausible that some sort of loop or circularity is present in the example, linking the development and effects of a boosted ego to the capacity to face and solve problems. Presenting only the left hand poles for simplicity, this situation would appear as in Figure 4.2.

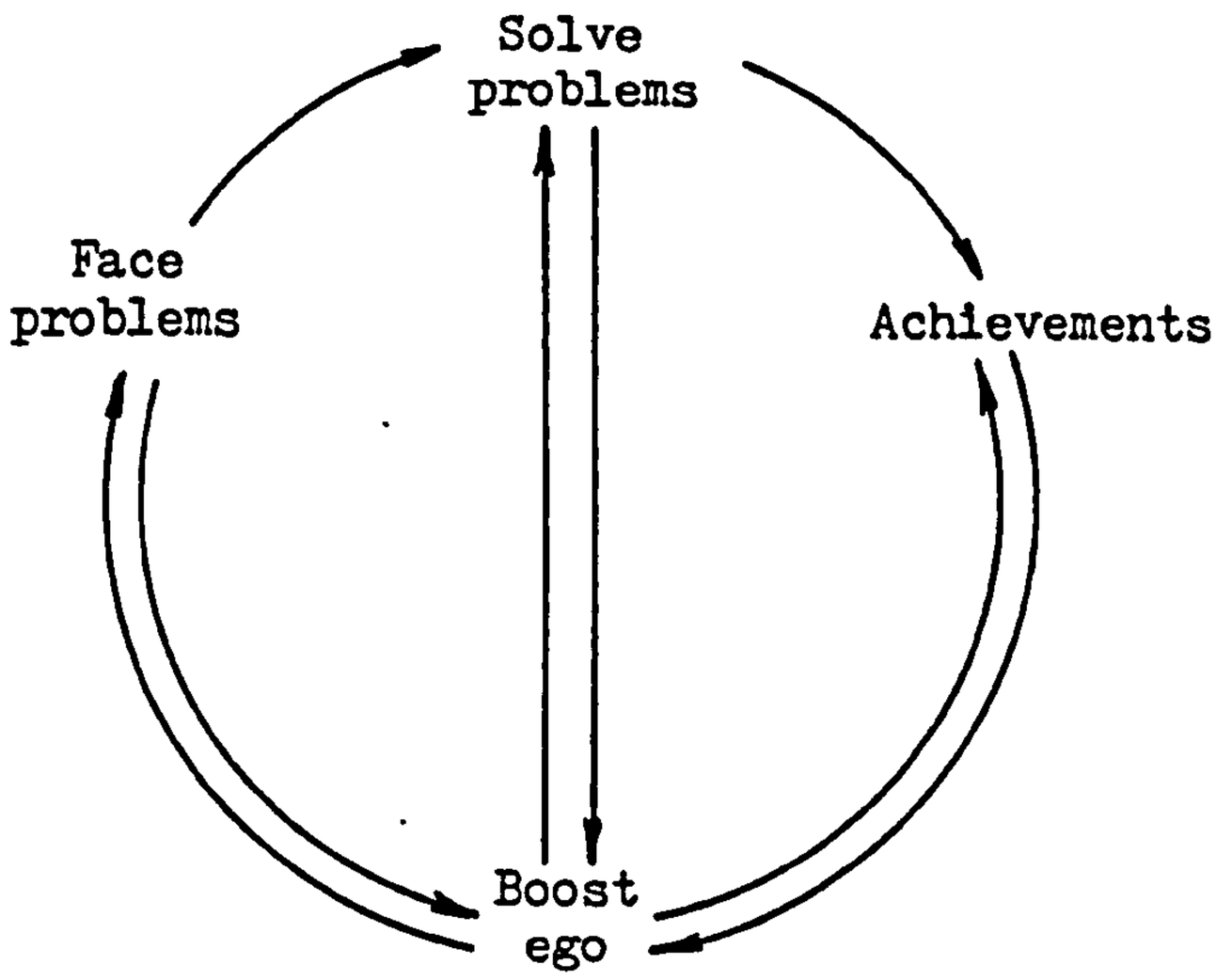


Figure 4.2. A possible circularity in Hinkle's findings.

There is further reason (apparently overlooked by commentators on Hinkle's work) to doubt that the constructs elicited by laddering

stand in a strict hierarchical relationship. Hinkle remarks that on occasion a specific construct label was given at different levels in the (presumed) hierarchy[39], the construct 'happy-unhappy' being most frequently noted in this respect. Operating from an assumption of hierarchy he is somewhat puzzled by this, and suggests weakly that the meaning of the labels may have varied according to their position in the hierarchy.

It could be the case that respondents 'loop' between superordinate and subordinate constructs[40], and (if the implications of 'reciprocal implication'[41] between constructs are explored) between constructs at the same hierarchical level, in either case the constructs being linked by overlaps in their respective meaning-complexes.

If looping does take place it could be reflected in 'vicious circles', as represented in the hypothetical example of Figure 4.3 (which is presented in unipolar terms in the interests of clarity).

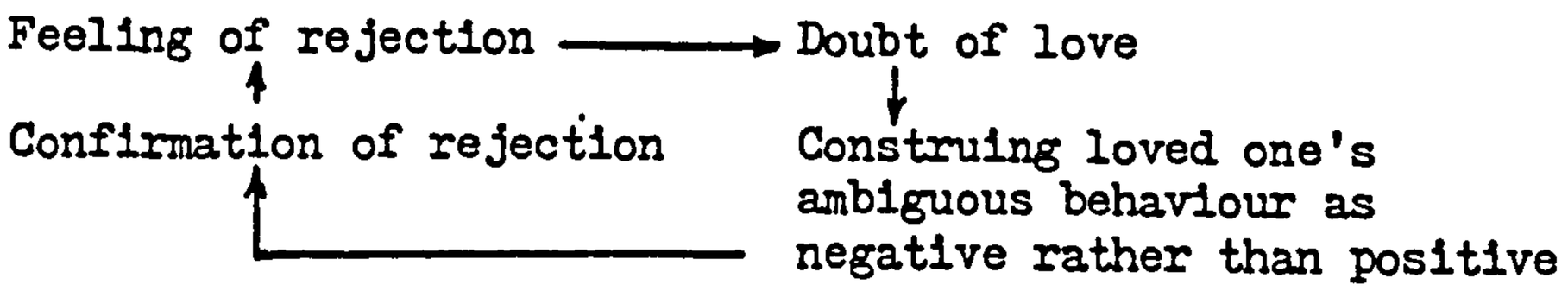


Figure 4.3. An illustration of a vicious circle of construing

Examples of vicious circles appear in Kelly's own writing, in Diesing's (1972) analysis of Baruch's (1952) detailed case study of 'One little boy', and by Eden et al (1979)[42].

This consideration of work on structure leads me to speculate on the relationships between dynamics of process and cognitive structure, and whether these are confounded in Hinkle's technique. What can be

concluded from this limited appraisal is that the relation between laddering and implication is likely to be more complex than seems generally to be accepted in the literature. It may be that Hinkle's technique, which probes the relationships between constructs in a context of abstraction rather than of practical living, is in need of development: some moves in this direction are discussed in Chapter 13. Further research in this area is difficult, but is necessary if the ambiguity of findings derived from laddering procedures is to be reduced.

4.8 THE REPERTORY GRID AND PERSONAL CONSTRUCT THEORY

Bannister and Mair (1968) define grid method in the following terms:

> 'A grid may be defined as any form of sorting task which allows for the assessment of relationships between constructs and which yields these primary data in matrix form' [43].

The definition is broad, since it opens out beyond the repertory grid to include Hinkle's implications and resistance to change grids, as well as other possibilities yet to be constructed. Although the definition as stated is not specifically linked to Kelly's theory save, perhaps, in respect of the word 'constructs') the explication provided makes it clear that the authors have that theory in mind. But a definition as broad as this runs the risk of being construed as vague, and as allowing a range of superordinate theoretical formulations. Thus it is not surprising to find Chetwynd (n.d.) and Slater (1977) pointing out that there is no necessity to relate grid methodology in general to Personal Construct Theory, but Chetwynd would seem to be going too far when she claims that grid technique is relatively free from theoretical assumptions and constraints[44]: grid-based research will have its underlying theoretical statements even if these remain

unrecognised by the researcher. It would appear reasonable to conclude that the relationship between grid method and Kellian theory is both partial and reciprocal: the theory does not necessarily imply the use of grid method, and grid method is not necessarily subsumed under the theory.

Bannister and Mair (1968) state that every aspect of grid method is closely related to an equivalent assumption in construct theory, supporting their claim by referring to the assumptions of construct bipolarity, finite ranges of convenience for constructs, and so on[45]. The coherence of Kelly's formulation and grid method (thus broadly defined) is high, though the repertory grid - as a subset of grid method - does not require the full range of theoretical assumptions: for instance, it depends upon the Dichotomy and Range Corollaries, but is indifferent to the assumptions specific to the Organization Corollary. None of the other statements of the formal content of the theory exert any influence on the grid, though many of them could be (but do not necessarily have to be) investigated by grid methods.

In his elaboration of the Fundamental Postulate, Kelly makes it clear that the focus of interest of his theory is the individual person rather than groups of people[46]. The process of data-gathering, by grid or other means, is idiographic[47] and consistent with the Individuality Corollary, though Allport (1962) observes that it is not entirely 'free' since the respondent is obliged to fit his or her construing to the particular elicitation protocol devised by the researcher[48]. In arguing that the repertory grid requires the application of a construct to a range of elements in the fashion of a

covering law, I am claiming that the grid smuggles in a nomothetic positivism that Kelly's theoretical position would appear to reject[49]. There are further nomothetic implications in the selection by Kelly of role titles, and the situation becomes more complicated in this respect when the researcher supplies elements and/or constructs. If, as Kelly stresses in his later writing, <u>events</u> are important in the Psychology of Personal Constructs[50] it would appear that the repertory grid (as it has been used to date) may be a less appropriate instrument than many researchers currently believe.

The challenge to the repertory grid is increased if construing is seen as not necessarily dichotomous - even the limited extension to bipolarity threatens the grid, as is argued in Chapter 8. Where elements are people, as in Kelly's original Role Construct Repertory Test, the dichotomous constructs may apply to all - but other types of element increase the risk that some will fall outside the range of convenience of some constructs. This presents no problems for the grid matrix, since blank cells are permissible: the problems come when statistical analysis cannot accommodate 'gaps' in the data, as is shown in Chapter 9.

It would be improper to attack the repertory grid for its failure to offer evidence on the structural aspects of construct systems (and hence to fail to relate to the Organization and Fragmentation Corollaries) since it is clearly an inappropriate instrument for such research, though not all researchers have appreciated the point. It may be that developments of Hinkle's (1965) original implications and resistance-to-change grids will be able to chart a way forward in this difficult area of research.

I argued in Chapter 2 that part of the problem of the relationship between theory and method lies in the high level of abstraction of the former. The breadth of view of the theory encompasses far more of the panorama of methodology than the narrow enclosure of the repertory grid - and it is interesting to note that, following the chapter in which Kelly describes it, there is no further mention of the repertory grid in Kelly's (1955) text[51].

The difficulties in which grid-based research finds itself spring in part from the failure of researchers to bridge the gap between Kelly's theory and method. Too often the assumptive structure is taken as the whole theory, and all that is deemed necessary is to administer a grid on the naive assumption that it is <u>the</u> instrument appropriate to the theory[52]. A substantial proportion of the blame for this state of affairs, this failure to recognise and tackle the problem of conceptual disjunction, must lie with the authors of standard texts on construct theory and method - including Kelly himself.

4.9 <u>SUMMARY</u>

Three main points were made in this chapter.

1. The latent positivism of the grid, which seems at odds with Kelly's general philosophical stance, widens the gap between theory and method that was identified in Chapter 2. That gap is opened still further if the challenges to construct theory made in Chapter 3 are taken into account.
2. The linguistic limitations of the repertory grid make it an inadequate medium for the communication of meaning. The researcher will need to look beyond the grid content and analysis in order to obtain a rich understanding of what the respondent wishes to convey.

3. The repertory grid does not provide more than scant and oblique evidence regarding the structure of construct systems, for there is no connection between grid methodology and the Organization and Fragmentation Corollaries. The researcher interested in structure may find Hinkle's work on implications grids a useful point of departure.

Some writers have made considerable claims for the repertory grid. The succeeding seven chapters constitute an analysis of the degree of congruence between that rhetoric and reason.

5 Elements

For the elemental creatures go
About my table to and fro,
That hurry from unmeasured mind

W.B. Yeats
To Ireland in the coming times

5.1 INTRODUCTION

Given the variety of grids being used by researchers it is surprising that relatively little has been written on the nature of elements and their role in grid methodology[1]. The elements chosen (whether by researchers or respondent) give an indication of the realm to be explored: that is, they define - if only loosely - the boundaries within which the respondent is expected to construe.

The central importance of elements in grid methodology requires that they be considered from the points of view of

(i) the context of the research;

(ii) their adequacy as a sample;

(iii) whether they should be elicited or provided; and

(iv) their meaningfulness, salience and stability.

These aspects are interrelated, but it is convenient to treat them separately from the point of view of the arguments being put forward.

5.2 CONTEXT

In his development of the Repertory Test, Kelly makes it clear that its context is the personal-social behaviour of those who came to him for therapy[2], and this led him to construct lists of role titles (which varied slightly according to the circumstances), against which his clients could identify people who were salient in their own lives. The 'personalised' list of elements is then used as the starting-point for the elicitation of personal constructs, usually by the method of triads[3].

The examples given by Kelly, however, indicate that the grid form of his Role Construct Repertory Test tends to elicit personal attributes (both concrete and inferential) rather than the elements' behaviour, though there are behavioural inferences to be drawn[4]. In other

words, there is some disjunction between what the grid elicits and the context in which it is set. This is not a problem for Kelly as a clinical psychotherapist, since he is interested in using the grid as a way of revealing, inter alia, 'the pathways along which the client is free to move'[5]. The grid, then, is seen as a technique for exploring the client's construct system, and as an interim step towards helping the client towards a revision of his or her personal-social behaviour.

Kelly's description of the grid gives very little emphasis to the contextual background against which his clients are to construe their chosen elements, thus rendering questionable the inferences that can be drawn from the data matrix[6]. To use the repertory grid as a research instrument in its own right implies the need for a closer connection between grid and context, and for the latter to be made more explicit than has often been the case in repertory grid research.

Various writers on grid methodology emphasise the importance of having a clear definition of the context of the grid, Mair (1967), for instance, pointing out that failure to do so may result in the respondent flicking from one context to another as he or she works through the grid[7] - a danger which Nash (1973) failed to guard against in his studies of teachers' perceptions of their pupils[8]. I am well aware that had someone asked me - as a beginning teacher - to construe science teaching, he or she would have got a markedly different response from me in the safety of the staffroom than when under threat of losing control in the laboratory!

There are practical limitations to the tightness with which contexts can be defined. In the research which I conducted with science

teachers (see Appendix 1), the sample was limited to teachers in their probationary year and sought their construing of the teaching of science to first-year pupils only: in my experience, science teaching has different problems and emphases according to the age of the pupils. But even specifying the context as tightly as this (which was as tightly as was feasible), any nomothetic analysis of the data gathered would be open to criticism on grounds such as variation in catchment area, school ethos, curriculum being followed, facilities available, and so on[9]. The analyst seeking inductive generalisation in such circumstances is faced with the problem of attempting a justification on the grounds of adequacy of 'overlap' between the instances studied.

Grid contexts, as well as being defined by the area the researcher chooses, are also defined by the specific elements being used. If Nash (1973) can be criticised for imprecision in his specification of the setting against which the pupils were to be construed by their teachers, he can point to the fact that his elements were at least 'homogeneous' (i.e. all of the same kind) in that they were all pupils. 'Heterogeneous' elements are a weakness in grid design: Bannister and Mair (1968) note a tendency in grid based research to use an 'undifferentiated' - the inference is that they mean 'heterogeneous' - list of elements originating from the respondent and they remark that 'this achieves neither specific relevance to problem areas nor general representativeness'[10]. Their warning seems to have gone unheeded by some researchers, to judge from the studies discussed below.

In her doctoral research Pope (1977) asked student teachers for 'a list of things which come to your mind when you think about teaching'[11]. This very broad interpretation of context elicited lists of elements

such as shown in Figure 5.1.

STUDENT No.7	STUDENT No.8	STUDENT No.19
Learning	Children	To broaden a child's knowledge
Teaching	Books	Widen child's interests
Communication	Chalk	Build a bridge between home and
Schools	Headmaster	school
Pupils	Classroom	Work in an atmosphere of fun
Teachers	School	Try to get on with the rest of
LEAs	Ability	staff
Books	Board	Relationship with the head
Classes	Pens	Plenty of space
Assessment	Exams	Trying to adapt method to suit
Arts	Worry	child
Sciences	Sense of	Be where the child is at
Social training	achievement	Happy relationship with children
Academic training	Film	Understanding each child
Interest	Attitudes	Making allowances for individual
Boredom		problems.
Discovery		

Figure 5.1. Examples of elements elicited by Pope (1977)[12].

The examples show that Students No.7 and 8 have construed 'thing' in terms of concepts and objects, whereas No.19 has seen the task largely in terms of aims. The situation is analogous to Nash's failure to specify the context in his research. But the methodological problems are deeper than this. If the aim of the research had been to investigate associations of the word 'teaching', no doubt the lists of elements would have been of considerable interest. However, Pope's aim was to monitor the viewpoints of students before, during, and shortly after a major teaching practice, and this required the elicitation of constructs (beginning in Kelly's triadic fashion, but becoming more flexible in approach once each element had appeared in at least one triad). Some of Pope's respondents remarked on the difficulties of using heterogeneous elements in the elicitation procedure, and this may account for the small number of constructs in some of the grids and also the very general constructs that were often produced[13].

It would appear that Pope obtained a diverse list of elements at the subsequent cost of limited construing - limited both in number and in insight into the problem area being investigated. To be fair, there is evidence that the feedback conversations regarding the grid analysis gave rise to a fuller exploration of 'teaching' than the grid on its own,[14] but this patches up weakness in her grid method rather than justifies it.

Keen (1979), in a doctoral study of physics teaching, employed heterogeneous elements in his grids despite the fact that all the elements to be construed were people. He originally intended that his respondents should construe, in addition to 'myself as a teacher of physics', a number of physics teachers known to them. Some of his respondents were first year undergraduates, and their lack of experience of an adequate number of teachers fitting the specification induced Keen to incorporate short videorecordings of seven physics teachers (employing a range of pedagogic styles) into the set of elements to be construed. Thus his respondents were faced with a mixture of teachers who were personally memorable, together with unfamiliar teachers mediated by videorecording. How they responded to the task of construing heterogeneous triads, or triads composed solely of 'video-elements', is not clear[15].

This mixed-element technique is developed further in the TARGET[16] project developed at Plymouth Polytechnic by Hopwood and Keen (1978), and my personal experience of responding to a TARGET grid provides me with an existential basis for arguing against the mixing of personally relevant elements with video-elements.

In the TARGET procedure there are three personally relevant elements

(self, effective and ineffective teachers) and eleven video-elements (short extracts from microteaching sessions): the respondent has the opportunity to make limited notes about each video-element. In construing the triads, which are selected in a predetermined pattern, I found difficulty in providing constructs when I was faced with the video-elements (who were unfamiliar to me and mainly using a lecturing approach in their microteaching environment). I found myself making discriminations on superficial criteria with regard to both the physical appearance of the video-elements and the pedagogic styles manifested. A broader range of constructs regarding teaching - and, as a 'staff developer', I can claim a measure of experience in construing teaching - was not encouraged to emerge. Whilst I am well aware of the danger of generalising from an isolated experience, I am led to suggest that the use of video-elements in a manner akin to the TARGET procedure may well fail to penetrate to the heart of the area of investigation. One might contrast the TARGET approach with an alternative, apparently little used, in which individual teachers are faced with excerpts of their own teaching[17], or of teaching by themselves and others in a teacher education group, in which one might expect a much richer understanding of the relation between the actions of teaching and the underlying context of aims, resources, constraints, and suchlike to be obtained.

My critique of the TARGET procedure leads to a second problem, hinted at in the opening of this chapter: that is, whether the grid methodology matches the intentions of the study being undertaken. In the TARGET approach the emphasis appears to be on teachers' characteristics rather than on teaching (the personally relevant elements setting the orientation) and hence one would expect that constructs would focus upon matters such as personal appearance and style rather than upon the

nature of teaching itself[18]. A focus on <u>teaching</u> would seem to require a grid in which the elements were teaching <u>situations</u> presented in such a manner as to maximise construing in terms of pedagogy and to minimise "teachers' characteristics" save for those seen as salient to pedagogic practice[19]. The alternative presented in the preceding paragraph would make it possible to go a long way towards 'partialling out' superficial characteristics, particularly if the respondents were asked to construe, for example, in terms of underlying aims and ideologies.

Although the validity of the grids used in the work of Perrott et al (1976) and Hopwood and Keen (1978) is compromised by the relationship between the elements and the intentions of the researchers, it can be argued - though, in my view, not persuasively - that construing teachers is closely related to construing teaching. The link between elements and intentions is much more tenuous in a study reported by Kevill et al (1982). In this research the authors aimed to assess the effectiveness of the Diploma in Literacy Development at Middlesex Polytechnic through the use of repertory grids during, and at the end of, the course. Whilst doubt might be entertained whether the time of administering the first grid is appropriate for what is intended to be a pre-course/post-course comparison, a more substantial doubt arises on examination of the way in which the grids were set up.

The course members were asked to list 'all the elements which regularly made up their teaching job'[20], and two examples of the lists are given in Figure 5.2.

Elements of a primary school teacher	Elements of a secondary school teacher
Teaching	Teaching
Organising library	Requisition
Socialising	Testing and timetables
Discussing	Liaising with staff
Duties	Discipline
NUT representative	Visits
Out of class activities	Preparing school leavers
Relations with parents	Form teacher.

Figure 5.2. Lists of elements provided by two respondents (from Kevill et al, 1982: 47).

To judge from the examples provided, and from the authors' failure to indicate that they asked their respondents to construe the elements from the point of view of literacy development, it is difficult to see how the grids - and hence the change measured by the grids - can be used in the context of the evaluation of the Diploma. To be fair to the authors, it must be made clear that a substantial contribution was made to their study by feedback interviews in which the grid analyses were explored further, but this does not absolve them from the criticism that their study might have been far richer had the respondents been asked to list, as elements, aspects of teaching (in the sense of extended professional role) in which the language development of children was, or might be, involved.

In all the examples discussed so far, the assumption has been made that the individual elements have been unconfounded in the respondents' minds. This assumption may not always be valid. Canter et al (1974), using photographs of people against different backgrounds, found that the background characteristics influenced the ratings that were given. The potential problem is easily seen in respect of visual elements, but may be more difficult to perceive where the elements are reconstructions in the respondents' minds. And what the respondents

reconstruct may not be what the researcher has in mind.

There is a further contextual problem to be mentioned: that of the self. It is now a commonplace that each self in its time plays many parts, assuming many role relationships (often in parallel), switching from one to another as the occasion demands - a self-referential version of the problem of contextual shift noted earlier in conjunction with Nash's (1973) work. There is a range of past, present and possible future selves: which one(s) does the element 'self' actually tap in a repertory grid? Some researchers are aware of the need to specify the 'self' they are interested in, or to specify a range of selves (for instance: 'self as I am now', 'self as I was', and 'self as I'd like to be'), in which case the context of elicitation is on rather firmer ground. Failure to recognise this potential problem, or to anticipate it in designing a grid, could result in findings that are ambiguous to interpret - particularly where the relationships between selves, or between self and others, are a focus of the research.

Shaw and Thomas (1978) claim that, in grid work, elements of all kinds are suitable[21]. The arguments developed in this section suggest that, construed naively, this claim could result in findings whose validity is seriously compromised. There is no substitute for rigorous thought at the outset of designing a grid.

5.3 SAMPLING

Much of the discussion in the preceding section has also been implicitly concerned with the question of sampling, for any question about the validity of grid elements implies a question about the sampling of elements from their realm membership. Kelly assumes that the elements provided by his clients in response to his list of role titles would be

representative of their realms of social interaction[22]. He bases this assumption upon the cumulation of his experience, but it is possible that salient individuals might fall outside this normative framework - for instance, policemen and (in the 19 - element grid) priests. In the United Kingdom, social workers and probation officers might be salient, yet not be elicited by Kelly's role titles. It might be useful - if the researcher intends to provide role titles - to allow the respondent to identify salient persons beyond those prompted, provided that they fall within the intentions of the grid.

Kelly recognises that different sets of events might lead to differences in construct interrelationships[23]. From the context of his writing it is probably fair to say that his attention is not directed to the sampling of elements from a realm membership, but the point nevertheless has some force for the grid user. Take, for example, the TARGET grid and its requirement that the respondent specify one effective and one ineffective teacher. It is not difficult to imagine a respondent wishing to specify more than one effective teacher - for instance, one who gains success through a traditional, authoritarian and didactic mode of teaching, and another who is successful by way of using a pupil-centred, inquiry-based approach. Whichever one is chosen to fit the TARGET protocol will have an effect on the construing and on the construct interrelationships, both of which would influence the 'self' and 'effective teacher' profiles which are output by the TARGET analytical routine: this output provides the basis for the respondent to reconstrue his or her teaching. One might surmise, in the light of the argument just presented, that the TARGET procedure would be enhanced by the inclusion of a greater number of designated-effective and designated-ineffective teachers at the expense of the video-elements.

The choice of effective and ineffective teachers as elements does draw attention to contrast in the grid. Both Pope and Keen (1981) and Easterby-Smith (1981) argue for the inclusion of 'good' and 'bad' elements among the sample[24], though, as the latter observes, a strongly-made contrast might inhibit element elicitation. Drawing on his experience with managers, Easterby-Smith remarks that a number of them find it difficult to name someone they dislike, and he suggests that the category 'a colleague you dislike' might profitably be amended to 'someone you like less'[25].

It may be the case that element-specifications are irrelevant to some respondents. Easterby-Smith gives the example of the element 'chairing meetings', which may be highly relevant to a works manager but irrelevant to a graduate trainee[26]. Similarly, I found, whilst working with probationary science teachers, that very few of them made any use of the provided elements 'outside visits' and 'science club' in the context of their teaching of science to first year pupils. There is a temptation, when providing elements on the basis of both experience and pilot work (as was the case here) to assume their general applicability to the sample being studied. In this case the pedagogic relevance of each element was tested by the use of the provided construct 'used a lot in my teaching - not used in my teaching'. To use a construct of this type allows the 'irrelevant' element(s) to be eliminated at the outset, or, less satisfactorily, during analysis [27].

The representativeness of sampling of elements is a problem for grid users which, in most cases, is not susceptible of resolution. In a few cases the sample is the population (for instance, all the pupils in a particular class[28]), but in the majority of circumstances it is impossible to define the population of elements objectively and hence be

sure of extracting a random sample[29]. Here the best that can be done is to make a reasoned case for the appropriateness of the chosen elements to the problem being investigated.

Despite the force of this argument, Keen (1979) and Hopwood and Keen (1978) claim to be able to determine when the sampling of elements has been unsatisfactory, the criterion being the dispersion of the elements in the component space computed by principal components analysis[30]; this criterion has been used to reject grids. But if one supposes that the - largely normative - sample of elements <u>is</u> a representative sample in the majority of cases, one must question why it appears unsatisfactory for a minority. The answer, one might speculate, could lie in the constructs elicited and the way in which the elements were located on them: the problem does not necessarily lie in the columns of the grid, but might reside in the rows, or in the interaction between rows and columns. Although the sampling may be suspect, it would seem to be invalid to reject grids on the grounds of their failure to fit the normative model that these writers' decision-making processes imply.

5.4 ELICITED OR SUPPLIED ELEMENTS?

Should the grid user elicit or provide elements? The question can only be answered in terms of the purposes of the research. For Kelly the answer is a midway position in which he goes as far as to specify role titles but requires his clients to give personal meaning to these by 'attaching' acquaintances to each role title on the list. Fundamentally, the elements must be adjudged to reflect Kelly's view of the compass of elements necessary to elicit a series of constructs sufficiently wide-ranging to provide an adequate understanding of his clients' construing of social relationships[31].

Easterby-Smith (1981) conveniently summarises the ways in which a set of elements can be put together[32].

(i) Supply elements to the respondent.

(ii) Provide role or situation descriptions.

(iii) Define a 'pool' (e.g. name five subordinates).

(iv) Elicit through discussion of the topic of interest, prompting if necessary. Jointly draw up list of elements.

Supplied elements - (i), and to some extent (ii) above - reflect the researcher's view of the problem. Even where the researcher has developed the list of elements through experience and/or discussion with a pilot group of people similar to the intended respondents, there is no guarantee that they will reflect the problem as any particular individual may see it, nor will they necessarily form a representative or sufficient sample (as was noted in the preceding section).

To supply elements requires the respondent to construe their meanings before engaging in the formal operation of completing a grid. The level of construing required will reflect the nature of the elements. Olson (1981) provided sentences relating to science teaching[33], implicitly recognising that a short verbal 'tag' would be insufficient to ensure that researcher and respondent were sharing a substantial measure of understanding. In my work with science teachers I adopted a slightly different approach, choosing to give short verbal labels (for convenience of reference) but supporting these with amplifying sentences. 'Assessment', for example, was amplified in such a way as to indicate that this referred to ongoing, day-to-day assessment of pupils rather than end-of-term or end-of-year testing; formative, rather than summative assessment. This proved in practice to have

been an important precaution, since some of my respondents were disposed to construe assessment in terms of terminal examinations or tests, rather than in terms of a more informal approach to the monitoring of pupil progress.

This difference in construing highlights the question raised at the start of this section: whose perception is, or should form, the focus of the research?[34] To draw attention to the problems associated with the provision of elements is not to imply that there is no place for supplied elements in repertory grid methodology. Given an adequate degree of intersubjective agreement regarding the elements, they may prove to be a superior sample than those elicited directly from respondents. And where the focus of interest lies in the respondent's construct system alone (the implicit espoused theory rather than the theory-in-use), rather than in the relationship between elements and constructs, the provision of elements[35] may be satisfactory.

To elicit elements, however, seems more in keeping with the Kellian philosophy of asking how the respondent construes his or her world - or, at least, that part of it upon which attention is being focused. Whilst it is highly likely that the elicited elements will be relevant to the respondent, the procedure is open to distortion in a number of ways: lapse of memory, selective attention, or fixation on a particular subset of the context[36], for example. In some circumstances it could be fruitful to analyse the elements elicited against those which might have been expected, but most grid users would probably wish to demonstrate that their procedures result in a representative sample of elements for the ensuing sorting procedure. Reid (1976), for example, draws attention to the possibility of obtaining a skewed sample of elements and makes the further point that this is likely to bias the

elicitation of constructs[37].

As Pope (1977) observes, interviewing respondents within the framework of a repertory grid requires considerable skill on the part of the researcher[38], who has to be sensitive to, and empathic with, them whilst at the same time controlling - in a flexible way - the course of the proceedings. Despite her awareness of interviewing skills, she records that some of the students who completed grids found difficulty in providing a satisfactory list of elements: as was noted earlier[39], the elements produced were heterogeneous and the problem may only have surfaced when the students were faced with the triads.

The nature of interpersonal elicitation of elements opens the procedure to the possibility of interviewer bias, though my experience suggests that this is more likely to be a problem in respect of constructs, whose elicitation is often encompassed within an extended conversation. Shaw and Gaines (1979) remark that interviewer influence is eliminated when the respondent interacts with a computer program such as PEGASUS[40]. In this program the respondent is asked to define the field of interest and to identify appropriate elements according to a standard routine. The full list of elements can be built up as the grid is being completed: it is not essential that the respondent specify all the elements at the outset. Although it is expected that the respondent will produce a set of homogeneous elements, there is no way in which heterogeneity or bias can be eliminated from the list of elements elicited.

The iterative addition of elements to an original list is also a feature of Green and Tull's (1978) 'consumption grid' used in marketing research. They begin by presenting a set of initial stimuli (such as 'toast pop-up') and asking their respondents to list all the occasions when each item

would be considered appropriate for consumption (in effect, the constructs in a grid matrix). The matrix is expanded by asking, in respect of each occasion, for the names of other appropriate products. The full matrix - which can by now be very large - can be completed dichotomously, each 'appropriate' element X occasion intersection being scored 1, and each 'inappropriate' intersection being scored zero. As far as the elements in such a grid are concerned, there would seem to be considerable scope for each respondent to build on the 'core' of elements provided by the researchers.

This blending of a 'core' of researcher-generated elements and a set of elements provided by the respondent may be an acceptable compromise between the requirements (in so far as they may be met) of sampling rigour and those of idiography. It is, however, not necessarily the case that the parties to the construction of a grid share a common framework of meaning - particularly at the outset - and it will be important to try to develop a shared understanding as the grid protocol unfolds, or else any conclusions that might be drawn from the subsequent analysis would be distinctly suspect. There is, of course, a potential problem regarding shared meaning when the respondent nominates elements to match a list of role titles, or provides from experience examples of critical incidents matched to given general situations. Some of the ambiguity, if such exists, may well be resolved in the light of the ways in which elements are located on constructs.

5.5 MEANINGFULNESS, SALIENCE AND STABILITY OF ELEMENTS

In the earlier part of this chapter reference was made, en passant, to the degree of meaningfulness of the elements to the respondent, particularly in those cases where the elements are provided by the researcher. The question of meaningfulness of elements (as opposed to that of

constructs) has received little attention in the literature, yet embedded in it are issues which the grid user might need to consider before embarking on research.

The identification of specific people to fit a list of role titles comes closer to relating the grid to everyday life than the construing of the role titles themselves[41]. Easterby-Smith (1981) argues that results from the latter may not mean much[42], though this can be countered to some extent by suggesting that such a grid might be very revealing of prejudice and stereotyped thinking. The point made here calls to mind Reid and Holley's (1972) work with potential university entrants who were asked to construe a number of universities from what appears to have been a standpoint of relative ignorance.

Even when specific people or incidents are provided in response to a researcher's rubric, little or no evidence is gathered regarding their psychological salience - yet salience could influence the results gathered from the grid. No doubt some would wish to argue that the less salient or less known elements will be revealed as such by the ratings given to them on constructs and by the subsequent grid analysis. However, this is by no means a safe position to take, for such elements may be assigned, on the grounds of lack of definition, to either the mid-point or the 'not X' end of a rating scale: the researcher may have a difficult time trying to infer the meanings of ratings.

It may be that part of the problem lies in a confusion between 'salience' and 'meaningfulness': if an element is salient it must entail meaningfulness, but the reverse implication is invalid. A clergyman is meaningful to me, but is not salient in my life. But to identify a list of salient elements may not be enough, since salience may vary

with context[43]: here I revert to the theme of an earlier section of this chapter.

Problems of meaningfulness and salience are likely to be more acute when the researcher provides the elements. Earlier in this chapter reference was made to the attempts of Olson and myself to define elements in such a way that there would be a substantial degree of shared understanding between researcher and respondent regarding meaning, but it is acknowledged that no similar statement may be made regarding salience. In each of these two studies the researchers were interested in the ways in which their respondents construed a series of given situations within the context of science teaching - in other words, the location of the elements in their context was of major importance. But this may be of negligible importance in other circumstances, such as in the Bannister-Fransella (1966) grid test of schizophrenic thought disorder in which the elements are photographs of minimal personal importance[44] in order to allow the respondents to 'project' aspects of their construct systems in a way that is relatively uncluttered by the intrusion of interpersonal matters.

Ravenette (1975), however, points to the difficulty that may be experienced when a respondent tries to construe 'unknowns' - a problem I found when trying to make notes on TARGET video-elements, when my attention was (paradoxically) diverted by the boring style of presentation of two of the teachers from the task of appraising critically what they were actually doing. The pressure of time in the TARGET procedure - viewing, and making notes on, eleven 'clips' of teaching in approximately half an hour - precluded the possibility of going back to the elements to answer, in a more detailed way, the question of why I had evaluated these

particular teachers as boring: in these instances, the potential salience of these elements had not been fully realised. To make the underlying point more generally: elements of low salience may fail to tap constructs of interest to the researcher[45] and hence be a weakness in the grid method being used.

Meaningfulness and salience are in practice not easy to disentangle in grid methodology, and both are likely to be partial determinants of the stability of ratings of elements. This raises the question of the stability of the element itself.

Stated in grid terms, Kelly's notion of 'loose construing' seems to be expressed in terms of the shift of an element upon a construct, and he gives the example of 'breakfast' being loosely construed as a meal eaten over a wide range of time in the morning[46]. The example is trivial, the point less so. Collett (1979) suggests considering the loosening of the element, arguing that if 'mother' invalidates predictions, one could dislodge the element within the system[47]. Given that the distinction between elements and constructs is one of convenience rather than of principle, it is difficult to see what Collett is adding here: to dislodge an element within a system is to reconstrue it, presumably assimilatively.

There is a further point to be made. Element instability can be construed in terms of the relocation of an element or a construct: I may construe a person as 'distant' on one occasion and as 'friendly' as I get to know him, shifting the element's position on the construct 'distant-friendly'. But many elements are complex - people and situations, to give two examples. What is construed as instability may result from differential focusing on the element's components on two

occasions, and not a bodily shift of the element in toto. In the example above, I might simply construe two different samples from the person's overall behaviour, focusing on the 'distant' and 'friendly' aspects on different occasions[48].

It is probably true to say that all elements are vague, with a sphere of meaning rather than single unambiguous point of meaning. The larger the sphere of meaning is (i.e. the less well defined the element), the greater is the likelihood of subsequent reconstrual in either of the senses noted above, to the detriment of stability in the grid[49]. Given Kelly's view of people as 'a form of motion'[50], change can be expected over time: the difficulty lies in identifying where the cause of the change might lie. The finding of low stability in the grid - and this can be assessed on a 'per element' basis[51] - might well lead the researcher to explore the respondent's construing of the elements (inter alia) and also the fuzziness of their definition[52].

The more abstract the element is, the more likely it is to be fuzzy. Carter et al (1968) provide some limited evidence to this effect. They asked 135 female teachers on a summer school conference to rate, on a semantic differential, sets of humans (e.g. the football coach), 'inanimate objects' (e.g. 'the new math') and 'relational objects' (e.g. relationships between teachers and administration). Ease of rating was found to relate to the type of 'object' being rated, humans proving to be the easiest and 'relational objects' the most difficult. One of the difficulties with this study, however, is the confounding of element fuzziness with the appropriateness of semantic differential scales (which is likely to vary with the type of 'object' being construed)[53].

On the whole, the literature suggests that the construing of people who are personally well known tends to be more stable than the construing of other elements, despite Mair's (1967) claim that such people are nevertheless complex. This probably reflects not only the relatively low ambiguity of the elements concerned, but also the use of constructs that are fairly well ingrained in the culture. On similar grounds, very familiar objects construed in terms of commonplace physical constructs (such as 'large-small') could be expected to provide stability values as high as the +0.92 found by Bannister and Mair (1968)[54]. These authors suggest that similar stability values might well be obtained if people were to be rated on the same set of physical constructs, but that, where 'psychological'(and therefore more abstract) constructs are used, the stability value could be expected to fall. It is not surprising to find the stability of rating of photographs to be somewhat lower than for physical objects since there is, again, the likelihood of a confounding of element and construct fuzziness: one might suspect the first of being responsible for the larger proportion of the instability on the grounds of the culturally-ingrained nature of the constructs being used.

Whilst general issues of stability are discussed in Chapter 10 - they are not the focus of this section - some further exploration of element stability is necessary, drawing upon empirical work I have undertaken, and which is more fully reported in Appendix 10.

I first began to ponder on the notion of element fuzziness when I noticed that the stabilities of the grids produced by the science teachers - who were using relatively abstract supplied elements (e.g. 'pupil practical exercises') - were distinctly lower than those of grids completed by a group of teachers on an in-service course who were using

self-selected pupils as elements. Assuming a degree of homogeneity among all these teachers, it was nevertheless unclear whether the difference in stability might be more attributable to differences between the two types of element being used, or to the difference between supplying and eliciting elements.

I investigated the matter further by asking each of different groups of teachers to complete test/retest grids in which the elements consisted of one of the following: (a) pupils; (b) specific teaching situations from their own experience; and (c) general descriptions of aspects of teaching. In each case, the elements were elicited from the respondents. The hypothesis was that the grid stabilities would, in descending order of magnitude, follow the order (a), (b), (c) on the following grounds. In (a) the elements would be sharply etched in the mind, and the resulting constructs could be expected to be familiar characteristics of pupils; in (b) the elements might be sharp but the constructs might be broad and vague after the manner of those produced by Pope's respondents; and in (c) both elements and constructs might be vague.

Up to twelve constructs were elicited for each grid using a dyadic rather than a triadic approach (see Section 6.5.2). Each element was located on a seven-point rating ladder (cf. Yorke 1978), a 'not applicable' category being available. One week later the same teachers were asked to repeat the rating procedure on a set of blank forms of the rating ladders, the elicited constructs having been added, one per sheet. 20 grids were completed by 17 respondents, since three respondents found time to complete a second grid (using a different type of element) on the first occasion,

Analysis of the stability indexes showed that they were significantly higher for pupils than for either of the 'teaching situations' grids ($p < .05$), but that no significant difference could be detected between the 'specific' and the 'general' teaching situations grids[55]. This latter finding may have been influenced by the fact that there were fewer pairs of grids completed in these two categories than for pupils, reflecting the happenstance of differential absence of teachers on the second occasion. Although it is possible to point to weaknesses in the design of this investigation, the findings lend some support to the view that the nature of the elements exerts a considerable influence upon the stability of the grid.

5.6 SUMMARY

The main points raised and discussed in this chapter lead to the suggestion that it is necessary, in grid-based research, to give serious consideration to the following:

(a) the relationship of the elements to the aims of the research;

(b) the extent to which the chosen elements are a representative and adequate sample;

(c) the justification for supplying or eliciting elements (or for opting for a mixture of both); and

(d) the potential meaningfulness, salience and stability of the elements.

The arguments presented in the chapter suggest that elements of all kinds may not be as suitable in grid methodology as some writers claim - and it may be necessary for considerable control to be exerted over the imagination in the development of novel approaches to grid methodology. If it has done nothing else, this chapter has drawn attention to the

need to subject proposed elements to a careful scrutiny before embarking on the use of a grid.

Although the following quotation is taken out of its proper context, it seems to sum up the burden of this chapter quite well: a clinician, Kelly says,

> '... should be careful about assuming that a test which is concerned with the perception of ambiguous forms or the sorting of objects will reveal all that he needs to know...'[56]

6 Constructs

Without Contraries is no progression.
Attraction and Repulsion, Reason and
Energy, Love and Hate, are necessary
to Human existence.

W. Blake
The Argument
from The Marriage of Heaven and Hell.

6.1 INTRODUCTION

This chapter takes a horizontal look at the repertory grid, focusing upon the constructs. Rather more has been written on constructs than on elements, necessitating a longer discussion than in the preceding chapter. Kelly makes explicit his assumptions regarding elicitation, and these are appraised before moving on to consider the methods by which constructs can be elicited. The question of whether the researcher should elicit or supply constructs is then discussed and this leads to the final section of the chapter, in which typological aspects of constructs are considered. No overall summary is offered to this lengthy chapter on the grounds that each subsection is fairly well self-contained: where appropriate a summary is given for a particular subsection.

6.2 ASSUMPTIONS UNDERLYING THE ELICITATION OF CONSTRUCTS

Kelly makes a number of assumptions in respect of his Repertory Test and of the grid form[1]. Since the first three of those relating to the former can be subsumed under the latter they are considered together.

6.2.1 Permeability

The constructs used are assumed to be permeable; that is, they must be open to the inclusion of new elements. If a construct is elicited by the 'triadic' method (see page 172), it is possible that the construct might apply only to the triad of elements which 'triggered' it (or to only a subsample of the list of elements in the grid): in such circumstances the construct would be restricted in permeability to an extent consonant with its range of convenience. Two men, for example, might be contrasted in terms of an interest in home decorating: the construct 'interested in car maintenance - interested in home decorating'

might well be inapplicable to other people presented as elements in the grid. Such a construct might be applicable to situations not yet confronted, but in the circumstances of the grid it is likely to be of little value. The question then is whether the grid methodology should determine the constructs to be included, or whether the constructs should be 'accepted' by the researcher even at the expense of gaining an incomplete grid matrix[2]. The implications of this question are too far-reaching for them to be pursued here, but there are implicit explorations of this issue at a number of points in succeeding chapters: a fuller consideration is offered in Chapter 13.

In Kelly's repertory grid the explicit assumption is made that all the elements fall within the range of convenience of the constructs, and this assumption is made in the majority of repertory grid studies in the literature. The mathematical analysis of the grid demands it, as do the more sophisticated computer packages developed for grid analysis, which are discussed in Chapter 9. The problem of the non-applicability of a construct to an element is recognised by Kelly, but he fights shy of the technical problems this entails[3] and, in the dichotomous allocation procedure that he favoured, he tacitly let the 'awkward' elements be subsumed under the implicit pole[4] of the construct. More recent work, involving rating scales instead of dichotomous allocation, has been unable to get over the difficulty: the tendency is to use - and even recommend - the mid-point of rating scales as a kind of 'dustbin' for those elements lying unconformably[5] upon the constructs' range of convenience, a blind eye being turned to the conceptual implications. The literature suggests that the problem is at a minimum when people are being rated on logically-opposite constructs of personality, and at a maximum when 'peculiar' constructs[6] and heterogeneous elements are being

rated on inappropriate dimensions - a problem particularly acute in semantic differential research[7].

6.2.2 The pre-existence of elicited constructs

Kelly makes the assumption that pre-existing constructs are elicited. Whilst he points out that sorting tasks are frequently associated with concept formation, his opinion is that it is relatively unlikely that new constructs are concocted in response to whatever elicitation procedure is being adopted. This opinion is based on the focus of convenience of Personal Construct Theory (interpersonal relationships) and, even though this may be true for the construing of familiar people, it is open to debate whether this assumption is generally valid - in fact, some grid users might wish (like Tully, 1976, in respect of social workers) to use the technique for heightening awareness and therefore seek that their respondents construe in ways other than those they habitually use[8]: the boundary between construct reproduction and construct formation would appear to be indeterminate, as was argued in Chapter 4.

Three categories of construing need to be considered, the second of which is often subsumed under either or both of the other two according to the perspective taken. The categories are, in respect of an element:

(i) using a set of existing constructs found to be appropriate;

(ii) using a set of existing constructs, finding them to be inappropriate, and thus switching to other pre-existing constructs; and

(iii) producing and using new constructs.

Category (i) is consistent with a successful pattern matching strategy, in the sense of finding a 'fit' with a Roschian 'prototype', and would seem to be applicable to both concrete, physicalistic constructs (e.g. extroversion construed as subsuming a range of 'prototypic' behaviours such as animation, initiative, style of dress and suchlike), to the consistency of data to abstract theory, and to consistency between theoretical formulations. In respect of the last two examples presented here, the matching process goes further than the work on object recognition would strictly warrant.

In category (ii) 'matchings' such as those exemplified in the preceding paragraph are unsuccessful, and the person has to search around in his or her system for other constructs that allow matchings other than those originally tried. Behaviourism, for instance, was based on a set of mechanistic constructs, but the 'match' of people with this framework proved inadequate and required the importation of some aspects of a Kantian metaphysics in neobehaviourism's admission of intervening processes. The Kantian constructs were not new: they had been in existence for a century and a half and were available to those who wished to make use of them. At the level of personal construing, an initial assumption based on, say, style of dress might be a construal of self-indulgence and amorality which might need to be tempered in the light of knowledge of efforts spent supporting charity work. In this category, as category (i) the constructs used are those already present in the system, and construing in both these senses is fundamentally assimilative (see Figure 3.1).

Category (iii) breaks new ground. Here existing constructs prove inadequate for the load they are required to bear, and new constructs have to be developed to take the strain. At the macro-level, this

would correspond to 'quantum jumps' in theorising - for instance, the contributions of Marx, Darwin, Freud and Einstein. At the personal level this is the accommodative, developmental aspect of construct systems. People learn new discriminations in the light of experience, often attaching the culturally-embedded verbal labels (children often invent their own labels until they find they have to adapt to the demands of society: 'goggy goggy goggy' becomes something more widely communicable in the course of time, even though the meaning expressed may be fairly stable). I learn, for instance that doing 'classical' experiments in social science is often scientistic rather than scientific, and I have to find new constructs such as verstehen and hermeneutics in order to escape from the crevasses of a pseudo-science. I emphasise here that these are my own personal constructions which may or may not find acceptance in the outside world: the transitions in my thinking do not have to be 'right'. I also acknowledge that verstehen and hermeneutics have also been in existence for a long time - however, they were constructs outside my ken until comparatively recently. I have grafted these on to my personal construct system, thus enlarging it: it remains to be seen whether experience validates or repudiates these for me in such research as I might in future undertake.

What is the relevance of this lengthy digression to the practicalities of construing in repertory grids, and in particular to Kelly's assumption of the elicitation of pre-existing constructs? Essentially it forms a partial challenge to Kelly's assumption, and therefore to what it is legitimate to infer from the grid itself.

Take the widely used 'triadic' method of elicitation in which the respondent is asked to discriminate two elements from a third.

Category (i) construing is unproblematic: the instances are discriminable on the basis of readily-available constructs at the appropriate level of abstraction (at the levels of 'gestalt', subsumed characteristics, and so on). Even if, as M.T. Taylor (1976) suggests, people do not construe in terms of triads there would appear to be little cognitive strain associated with the elicitation of pre-existing constructs.

Category (ii) construing provides the researcher with a problem. The respondent, presented with the three elements separately, might well be able to offer a number of constructs in respect of each. However, the presenting of a triad might make the original discriminations inoperable, and it is easy to see that this could be blatant with heterogeneous elements. There are also more subtle potential manifestations in that the researcher could collect a discrimination concocted by a respondent on the spot (in both senses) which, whilst drawing upon existing constructs in his or her repertoire, nevertheless uses them for the first time in respect of one or more of the elements. 'I hadn't thought of it that way before', reflects comments which both Olson (1980a) and I have heard during elicitation and sounds a warning bell that the grid might be eliciting novel rather than pre-existing construct uses.

The distinction is made here between novel construct uses and novel constructs. As far as the latter are concerned, it seems unlikely that the repertory grid will induce new ways of construing (unless the researcher introduces them to the respondent). Reflection upon category (iii) construing leads me to speculate that this takes place when an inconsistency is noticed between two or more subsystems of a construct system, and something has to be done to eliminate what Dewey referred to as 'perplexity, confusion, doubt'[9]. In construct theory terms this

can only be done by 'taking the relevant parts of the system to pieces' and reassembling them in a different way, allowing a new superordinate structure to emerge - which may, of course, have further subordinate implications. This is far more than giving a new name to an old discrimination, since it involves major reconstruction of a section of a system. The repertory grid, with its reductionist emphasis on analysis rather than synthesis, would seem particularly ill-suited to the induction of category (iii) construing: hence it is reasonable to conclude that category (iii) construing is unlikely to be represented in repertory grid data.

I have focused in this discussion on the triadic elicitation method since it is widely used in repertory grid methodology. The other methods discussed later in this chapter are likely to be differentially susceptible to the problems outlined above, but it would take me too far from the course of my appraisal of Kellian assumptions to consider them exhaustively here.

So how does Kelly's assumption regarding the elicitation of pre-existing constructs stand? The argument presented above suggests that, at a general level, it is a tenable assumption, but that it may fail in respect of the practicalities of the grid respondent's use of constructs. If the research is to probe the way the respondent uses constructs, rather than what constructs the respondent might call upon, then, as an instrument, the repertory grid is on uncertain ground[10].

6.2.3. The functional communicability of constructs

Here Kelly makes the assumption that the words used are adequate for the researcher to understand what the respondent means: it will be

remembered that the words themselves are not the constructs (though the common use of 'construct' as a synonym for the bipolar verbal label makes this easy to overlook); they are symbolic of the discrimination which the respondent is making. Put another way, the respondent encodes discrimination; the researcher decodes it and relates it to his or her own construct system. There is a potential for meaning to slip sideways at both the encoding and decoding stages because verbal labels in natural language often have multiple meanings and always a range of connotations. Whilst the primary meaning of a respondent's verbal label may be unambiguous, the connotative colouring may not be available to the researcher who has his or her own connotative concatenation in respect of the same symbol[11].

There are occasions when communication through the use of verbal labels can go remarkably wrong. I can recall the example of a science teacher who gave the label 'encourage interest' to two types of teaching activity, offering 'drollness' as the contrast. This seemed to be a very 'peculiar' construct so I enquired further, finding that by 'drollness' she meant 'making science seem like so many mechanical exercises' (my initial supposition having been that she meant something like 'drily humorous'). The importance of conversation for the elucidation of meaning is well exemplified here[12], but it is a glaring instance of an apparently puzzling construct: how many commonplace 'obvious' constructs pass as unproblematic when no adequate sharing of meaning has taken place?

Kelly recognises that, in his repertory <u>test</u>, communication of meaning through the elicitation is a critical assumption - indeed, he goes so far as to state that it is the most precarious of the assumptions he makes[13]. However, where the <u>grid</u> form is concerned, Kelly seems to

see this as being of lesser importance on the grounds that the - presumably mathematical-analyses would allow the researcher greater access to meaning[14]. Some criticism was levelled at Kelly's position in Chapter 4, but the assumption of communicability needs further exploration than was undertaken at that point.

I have already referred to the failure of repertory grids to provide data from which conclusions can legitimately be made about hierarchical structure, but it remains to be considered whether measures of association derived from the grid matrix can aid the researcher's attempts to extract meaning.

It is often assumed that a perfect association between constructs implies their functional equivalence[15]. This may be so (for instance, with synonymous or near-synonymous constructs) but is not necessarily so. The smaller the number of elements and the more restricted the score-range of the construct scales, the greater is the chance of adventitious isomorphism of distributions of ratings given on two constructs[16]. Further, isomorphic rating distributions may reflect the influence of an underlying construct (which may not have entered the grid)[17]: one might find a person allocating a set of elements identically on the constructs 'happy-sad' and 'extrovert-introvert', but the claim of functional equivalence would be bold, even if the range of consideration were limited to the particular set of elements concerned.

In practice isomorphism of rating seems to be relatively rare between constructs, though it is common to find some high indexes of association between constructs in a grid. Where this is the case, it

may be legitimate to infer that the constructs overlap in meaning to a substantial extent, but the inference cannot be taken for granted: cross-checking with the respondent would seem to be advisable. However, where a construct is not closely associated (in statistical terms) with others in the grid, the crutch of connotation would appear to be out of reach.

Looking at the matter from a linguistic standpoint, the verbal labels provided by a respondent signify (at least) the specific distinction he or she is consciously making, yet those selfsame labels can signify other possible distinctions: to use the examples of a couple of paragraphs ago, what do the discriminations 'happy-sad' and 'extrovert-introvert' mean when a respondent offers them? They are by no means unproblematic. These signifiers are overdetermined[18], and the researcher's task is to try to elucidate as far as possible which of the range of possible 'signifieds' is intended.

Olson (1979a) makes the point that brief verbal labels on their own are inadequate to define the meaning of complex ideas. In grid work, constructs are often highly elliptical[19]. Where the researcher listens to the conversation that is taking place with the respondent during elicitation, much of the ellipsis can be rounded out: to use the conversation merely as a means of getting a set of grid dimensions is to ignore a wealth of meaning.

6.2.4 Representative sorts

Kelly's assumption is that the task of triadic sorting requires the respondent to make discriminations representative of those made in real life. The more heterogeneous the set of elements being used, the less likely this assumption would appear to hold. Even a set of

homogeneous elements (for instance, people) may give rise to triads which do not reflect the discriminations the respondent would make in real life, though they may elicit distinctions potentially available to the respondent. There is embedded in this assumption an 'is/ought' problem for the researcher, to which further attention is given later in this chapter.

Setting this problem aside for the time being it would seem that, given a grid whose context is meaningful to the respondent in which the elements are representative and homogeneous, it is not unreasonable to suppose that Kelly's focus on role-relationships can be widened out to encompass a far wider range of discrimination tasks.

6.2.5 The manageability of grid completion[20]

Kelly's original procedure provides the respondents with a detailed set of instructions to enable them to complete the supplied grid blank, there being no indication that Kelly intended to, or did, converse with them during the completion of the grid. Kelly assumes that respondents can cope with the 'pencil and paper' demands of grid completion.

Experience with rating forms of the grid indicate that, if respondents are left to themselves to complete grids (as in group administration), a number of errors can creep into the procedure: these are detailed in Chapter 8, which deals with matrix completion. The use of rating scales is a more complex undertaking than Kellian dichotomous allocation, but the latter would not appear immune from a number of the errors discussed.

Other workers, myself among them (Yorke, 1978) have described approaches in which all the recording is done by the researcher. There are distinct advantages in so proceeding: not only is some of the pressure taken off the respondent (grid completion shifts in emphasis from being a test towards being a dialogue), but also the potential for error is reduced in respect of response and recording.

6.2.6 Construct stability during elicitation

Kelly makes the assumption that constructs remain stable during the elicitation procedure. He might have added that the assumption extends to the location of elements on constructs, since this is strongly implied in grid methodology. Kelly's narrower assumption is that the respondent does not shift ground during triadic elicitation to produce emergent poles from two distinct bipolar constructs, one in respect of the pair construed as similar and the other in respect of the contrasting singleton.

The tenability of Kelly's assumption seems to rest in part upon the precise way in which the triadic elicitation procedure is formulated. Kelly asks the respondent whether two of the elements are 'alike in some important way that distinguishes them from the third...'[21] Following Kelly, many workers[22] have asked their respondents to discriminate within the triads on the criterion of difference. However, this type of rubric seems to undercut the assumption, producing constructs which, whilst having contrasting poles, are not functionally antonymic. If the researcher is merely concerned with the constructs the respondent produces, this may not be a problem; but if the completion of a grid is at issue, the problem begins to loom large. Although grid completion is discussed in Chapter 8 it is appropriate to consider the assumption of construct stability further at this point.

Resnick and Landfield (1961) found that respondents produced both 'peculiar' and 'logical' constructs, the former referring to constructs in which the contrasts appeared idiosyncratic and illogical judging by the criterion of dictionary meaning, whereas the latter reflected normative antonymy. Both types of construct were found to be used consistently, though 'peculiar' constructs were significantly inferior to 'logical' constructs in this respect. Subsequently, Bonarius (1971) made a point of mentioning the non-oppositionality of construct poles[23].

The question of bipolar opposition was pursued by Epting et al (1971) who found that a more marked bipolarity was obtained if respondents were asked to give, as a contrast, the <u>opposite</u> characteristic to that construed as the similarity of the pair, instead of simply allowing the singleton to be 'different' in whatever way the respondent chose. These findings have influenced other users of grids to use the 'opposition', rather than the 'difference' method of triadic elicitation[24].

However, it would be dangerous to infer that the 'opposition' approach is necessarily superior to the 'difference' approach, for the <u>logical</u> opposition may not be the discrimination the respondent would wish to make in real life[25]. The psychological opposite of 'ambitious', 'does not trample on colleagues', may be far more meaningful than the logical opposite 'unambitious' through the implied relationships between ambition and trampling, and between lack of ambition and not trampling. It would seem likely that this particular construct would, in Resnick and Landfield's terms, be classed as 'peculiar' and be the coalescence of two related constructs. At the level of meaning, the above interpretation is at least plausible, but at the level of grid usage

there could be difficulties where the (presumed normative) implicit personality theory relating ambition to trampling does not hold. The more 'peculiar' the construct (and the example is not extreme in this respect) the greater the potential problem for the grid respondent even though the meaningfulness of the original distinction (which, it must be remembered, is likely to have been produced in response to only three elements) may have been adequately communicated.

It would seem that the researcher, in designing the study, has to bear in mind the likely implications of the elicitation technique chosen. There is a choice between on one hand the greater chance of elements all lying within the range of convenience of the 'opposition' generated construct coupled with the construct's potentially lower psychological meaningfulness, and on the other the greater risk of elements not conforming to a 'peculiar' (though psychologically meaningful) construct treated as if it were a linear continuum, and hence of a potential increase in the incidence of 'gaps' in the grid matrix.

The points presented here raise a number of questions regarding the bipolarity of rating scales. The issues are too broad to be considered at this point and have been deferred to the succeeding chapter in which there is space for a fuller treatment.

6.2.7 Other assumptions

Kelly also makes two other assumptions which relate to the role relations at the focus of convenience of his theoretical position[26]. First, he assumes that the person is able to construe to some extent the construct systems of the people serving as grid elements[27] and, second, that the constructs elicited refer to the respondent's own identity and behaviour (construed in terms of Kelly's definition of role).

In view of the fact that grid methodology has diffused the focus of convenience of Personal Construct Theory from its origins in inter-personal relations, these two assumptions now pertain to only a subset of grid methodology, and will not be considered here.

6.3 ADEQUACY OF THE SAMPLE OF CONSTRUCTS

The adequacy of the sample of the elicited constructs is a further assumption in grid methodology to which Kelly gives little attention, yet which is cognate to the representativeness of the sample of elements in the grid. Kelly does quote some preliminary studies conducted by Hunt indicating that, on the Role Construct Repertory Test, forty sorts of twenty figures elicited nearly all the constructs which the respondent was capable of expressing. In this work Hunt found that students produced some 70 per cent of common constructs in responses to two different role title lists balanced as to salient attributes. Kelly reports this work under the heading of 'consistency' (with respect to the Repertory Test, and not the grid form), and the tone of his writing leads the reader to infer that this indicates that this elicitation procedure does adequately sample the respondent's construct system.

Fjeld and Landfield (1961) also investigated the ability of students to repeat their constructs a fortnight after completing an initial Repertory Test, and found correlations near to +0.80 irrespective of changes between test and retest in the figures nominated to fit the role titles. Bannister and Mair (1968) implicitly, and Adams-Webber (1979) more explicitly, seem satisfied by this evidence that, as far as Kelly's Repertory Test is concerned, the triadic elicitation procedure provides the researcher with adequately representative samples of respondents' construct systems[28]. However, the strongest claim for adequacy is

made by Bonarius (1965), who writes that the directions supplied in respect of the test 'are such that the sampling is relevant and representative of the person's construction system'[29].

But consistency should not be mistaken for representativeness. If I go to a number of fairgrounds and find, in the light of my failure to hit the bull's eye in the shooting galleries, that all the gun barrels are slightly bent, am I justified in castigating all the gun manufacturers on their inability to produce a 'straight' gun? Clearly not. However, to challenge the equating of consistency with representativeness it is necessary to do rather more than indulge in light-hearted analogies. Evidence must be produced to show that the triadic approach may not elicit constructs which would generally be agreed to be salient within the context of the problem being studied.

At this point I draw primarily on my experience in using the repertory grid with science teachers. A key factor here is that the teachers were in their probationary year in school, at the end of which their head teachers would judge whether their performance was satisfactory or whether an extended period of probation was necessary. There is plenty of evidence to show that aspects of class control are a major concern of a beginning teacher[30], and the realities of the profession make the ability to control a class a crucial element in the successful completion of a probationary year. In the teaching of science there is an added premium on class control because of the hazards to pupils from chemicals and equipment. Yet of 27 grids collected from fourteen such teachers, only two contained constructs relating to discipline and control (and at that, one construct per grid) despite a mild prompt in the direction of management and control issues. Further, in respect of one of the grids it seemed to me that the 'control' construct

was produced as a result of the noise emanating from an adjoining laboratory in which a colleague appeared to be having difficulty in controlling the class. On readministering the grid to this same teacher in a quiet classroom at the end of the year, no 'control' construct was elicited despite the fact that, immediately prior to the grid administration, he had had to take a pupil down to the deputy head for punishment because of receiving some 'aggro' from him during a lesson. This teacher subsequently remarked in conversation that he did have serious discipline problems and that he was aware that there was some doubt that he would be deemed satisfactory at the end of his probationary year.

It would be going too far to claim these findings as substantive evidence against the ability of triadic methods to elicit a representative sample of constructs. Despite the mild prompt, the teachers appear to have construed the task much more in terms of aims of science teaching, subject content and its implications for the pupils, and other ideological matters such as authoritarianism in the laboratory or classroom. It is likely that, for some of them at least, the issue of control was not problematic or was not construed as lying within the scope of the research. It must remain an open question whether the specification of the context was adequate to my intentions as researcher, or whether the teachers fell into a cognitive 'set' regarding the task. A more conversational elicitation procedure, rather than an approach which deliberately sought to avoid injecting the researcher's constructs (and I had had nine years of experience as a science teacher in schools), might have produced a very different pattern of constructs.

The research reported here does have some support, though it comes

from a study not set within the framework of Kellian theory. Kerry (1980), in a study of primary school teachers, found they tended not to identify pupils' reading competence when they were asked to indicate important factors in classroom functioning, but this emerged generally as important when they were subsequently asked to complete a checklist containing this among a range of suggested factors.

The problem of representativeness in the sampling of constructs is not a simple function of the way in which the respondent construes the elicitation task. The researcher may fail to recognise a critical dimension to a problem and hence fail to orient the elicitation appropriately. Bannister and Bott (1973), for example, report a clinical study of a married couple whose sexual relationship was unsatisfactory. Repertory grids were used in an attempt to understand the problem but failed to detect a crucial connection between the husband's need for control in his job and the state of the couple's sexual relationship. Viewing the problem as limited to sexual matters seems to have blinkered the therapist to other possible constructions pertinent to the problem.

Pressure of time may militate against an adequate sampling of the construct system. Grid methods are often time-consuming, and respondents (for example, teachers in free periods or at the end of the day) may have only a limited time available. Wood and Napthali (1975) limited their elicitation to twelve constructs because this was 'the most which could reasonably be asked for in the circumstances'[31].

Taken together, these pieces of evidence raise a little disquiet that the sampling of construct systems can suffer from biases of which the researcher may be unaware. Users of Kelly's Role Construct Repertory

Test and its grid form may be justified in the claim of sampling representativeness, but such a claim is based on metaphysics and is not open to empirical test. Validation of such sampling is necessarily subjective, but the subjectivity may be reduced if the researcher seeks to cross-validate what appears in the grid with other evidence which might be available (such as might be gained by conversation, formal interviews or behaviour in relevant circumstances). The implication here is that the researcher should be existentially familiar with the problem area being studied in order that he or she may maximise the chances of detecting 'gaps' in the sampling: the risk is an increased possibility that the researcher's constructs will be 'injected' into the respondent's grid, as Shaw and Gaines (1979) point out[32]. But their preferred solution, elicitation by way of a computerised algorithm, gives no greater guarantee that the respondent's sample of constructs will be representative.

6.4 HOW MANY CONSTRUCTS?

The question of the number of constructs to be elicited in a repertory grid is one to which no satisfactory answer can be given, since it is very closely related to the problem of sampling discussed in the preceding subsection. The literature offers rule-of-thumb advice[33]. Hunt's suggestion that 40 sorts will virtually exhaust the respondent's capacity to produce different constructs seems to allow for constructs to be repeated, and so might be regarded as an upper boundary[34]. Bannister and Mair (1968), having in mind the productions of a different construct for each sort, suggest that no set number of sorts need be made, but that a common range is from ten to 25[35]. Pope and Keen (1981) take a similar position, indicating that the limits are likely to depend upon practicalities such as the time available, the

stamina of the parties involved and the size of any computer program which might be used, as well as the exhaustion of the respondent's repertoire of constructs[36].

Bell and Keen (1981) have attempted to provide the researcher with a statistical criterion to assist in the decision of the appropriate time to cease eliciting constructs. They assume that the element intraclass correlation (which is equivalent to the average inter-correlation among elements) is an inverse function of cognitive complexity[37] (differentiation). The more constructs that are elicited, the closer the intraclass correlation approaches to an asymptotic value. So, if the researcher updates this measure after each construct has been elicited and used as a rating dimension, it would appear to be of little advantage to pursue the elicitation once the intraclass correlation appears to have reached its asymptotic value.

At first sight this approach is attractive, but closer inspection casts doubt upon the validity of the argument presented and on the implicit assumptions on which it appears to be predicated. First, Bell and Keen present evidence which shows that the element intraclass correlation is unlikely to follow the hypothesised relationship though it is likely to approach an asymptotic value as the number of constructs is increased. One would expect this correlation to fluctuate un-predictably when the number of constructs is small since it will depend on the similarities and differences in the rating patterns on the constructs so far used: one person may start with two or three quite similar constructs whilst another might differentiate quite sharply

between successive constructs. There seems no adequate ground on which the authors label four categories of relationship: these seem merely to be artefacts of happenstance.

Bell and Keen suggest that to cease elicitation when the element intraclass correlation is rising is almost certain to deprive the researcher of additional worthwhile data. It seems that they base this comment on one of two assumptions: either the fact that the correlation is rising implies that further elicitation will sooner or later cause it to decline to an asymptotic value, or that an increase in correlation reflects increasing differentiation. If the latter is the case, they are surely mistaken in that differentiation is inversely related to the intra-class correlation[38].

One would expect, with their method, that fluctuations in the correlation coefficient would tend to be damped with increasing number of constructs; after all, the third construct is more likely to have a marked effect on the correlation derived from the use of the previous two than is the thirteenth in respect of the previous twelve. So even if a salient, atypical construct is produced late in the elicitation its psychological significance may not be picked up by the statistics.

It is this psychological significance which is the critical issue in elicitations and which a statistical approach overlooks. A respondent may produce a personally important construct at any stage in an elicitation: in 27 grids elicited by me from science teachers the most important construct, as identified in a resistance-to-change grid (Hinkle, 1965), tends to have been elicited later rather than earlier[39]. The results are summarised in Figure 6.1.

	Most important construct elicited in the		
	First third of constructs	Middle third of constructs	Final third of constructs
Frequency	8	12	16

Figure 6.1 Location of most important construct in 27 grids. (The total is greater than 27 because of constructs appearing jointly as 'most important' in some grids).

It is quite possible that a number of these late-appearing important constructs would not have made a significant ripple in the smooth approach of Bell and Keen's element intraclass correlation towards an asymptotic value, for a psychologically significant construct is not necessarily distinguished by an unusual or statistically disruptive pattern of ratings. In other words, Bell and Keen's measure substitutes statistics for meaning: it is an easy but invalid way out of the problem of deciding when to discontinue the elicitation process.

It can further be argued that the more constructs that are elicited, the more meaning is potentially available for the researcher to interpret. Constructs which are related statistically in the grid may give the researcher valuable clues as to which segment of meaning is intended by the respondent. In my own experience of eliciting grids, respondents have usually produced between ten and twenty constructs before 'drying up'. I have become increasingly uneasy about accepting these as limit beyond which the respondents might be unable to go, for informal conversation during and after elicitation has hinted at the availability of other constructs (maybe not explicitly bipolar) which might have emerged in a different interview format and which might

have added richness to meanings already expressed - or possibly even radically new meanings.

Put another way, I have to advance the opinion that the typical repertory grid elicitation procedure may be such a wide-meshed sieve that only the broadest of constructs are entrapped by it. If, as Oswalt (1974) found, people are capable of exhibiting an extensive repertoire of person-oriented constructs in a relatively 'free' elicitation procedure, why is it that the repertory grid typically elicits a much smaller number - of the order of twenty-five?

The ability to construe in a more wide-ranging fashion than the repertory grid seems to allow is shown in other contexts than personal attributes. In Olson's (1979b) transcripts of his science teachers' elaborations of grid constructs there are other constructs (or, at least, single ends of constructs) which do not enter the grid. Rowe's (1978) work on depression exhibits a similar profusion though her own interests have led her away from making use of the grid in other than exploratory contexts.

So how many constructs should the grid administrator elicit? The answer, for me, is as many as possible within the limits of time, patience and stamina noted by Pope and Keen (1981). It is always possible that the next construct will prove to be the vital one! More work needs to be carried out on the question of elicitation since such evidence as is available suggests that formal methods may limit the respondent's production of constructs to an undesirable extent. But if large numbers of constructs are elicited by other approaches, the grid may become unacceptably large. If this were to be the case,

which should be discarded - constructs, or the grid itself? In Chapter 13 I offer the outline of one type of answer to the dilemma.

6.5 METHODS OF ELICITING CONSTRUCTS

Kelly describes a number of methods which can be used for the elicitation of constructs in the Role Construct Repertory Test, though these are reducible to three when variations in element content are eliminated from consideration. These three methods are

(a) the minimum context form, in which the elements are presented in successive triads;

(b) the full context form, in which the respondent is asked to discriminate among all the elements (which are presented together); and

(c) the self-characterisation sketch, in which the respondent is not required to make specific contrasts in the manner of first two above.

Some workers have amended triadic elicitation to what might be termed 'dyadic elicitation' in which pairs of elements are presented, the researcher being asked initially for a similarity or a difference. The free form of the self-characterisation sketch has encouraged a few workers to experiment with elicitation procedures which have more in common with conversational practice than with the more formalised approaches. Finally, Hinkle (1965) extended the range of elicitation procedures to include 'laddering', which is based on the implicative relationships between constructs.

These approaches to elicitation are reviewed briefly below, the main attention being given to triadic elicitation on the grounds of its commonness of use.

6.5.1 Triadic elicitation

Kelly's term 'minimum context form' reflects the way in which he considers constructs to be formed; that is, to note features of similarity and contrast among a number of elements. The minimum number of elements to allow discrimination of both similarity and contrast is three, hence the sorting of triads. Kelly had a clear idea of the dimensions of discrimination he thought likely to be relevant to his clients and chose particular triads in order to focus the client's attention in particular directions. For example, the 'Threat sort', comprising Brother, Ex-pal and Threatening Person, is seen by him as offering the client the opportunity to construe threat: that is, aspects of threat are more likely to surface than anything else when these three elements are presented.

Keen (1979) and Hopwood and Keen (1978) describe grids in which a fixed protocol of sorting is adopted, though the rationales for the successive triads are not presented. Glossop et al (1975), who were studying constructs of value in a wide range of 13-14 year old children, selected triads on the grounds of maximum potential contrast on the two dimensions 'good-bad' and 'valued-not valued'. Collett (1979) suggests that triads of dissimilar elements give rise to more superordinate constructs than triads of similar elements, but gives no theoretical rationale[40]. If Pope's (1977) experience is anything to go by, one tends to get very general constructs from triads of dissimilar elements. Whilst these might be described as superordinate, they might also be described as vague.

The selection of triads by the researcher is an indication of his or her perception of the problem and of the discriminative dimensions believed to be important in the mind of the respondent. It is reasonable

to suppose that the choice of a particular triad is likely to predispose the respondent to use certain dimensions of construing at the expense of others, and Keen's assertion that it does not matter how the triads are selected[41] must be open to some doubt.

Random selection of triads removes the element of bias on the part of the researcher, but does allow the possibility that individual elements may be proportionately over-represented or under-represented in the grid. Stratified random sampling, through the use of balanced incomplete block designs[42], would seem to be the best way of overcoming the problems associated with the sampling of elements - unless, of course, there are strong a priori justifications for selecting specific sorts. Smith and Leach (1972) and the empirical work which I undertook both drew upon such designs. The number of triads required for a full balanced incomplete block design would have made the completion of the grid an unacceptably lengthy task: half-designs were used in which the frequency of occurrence of each element was identical, but in which each element did not equivalently co-occur with other elements.

The sequence of presentation of the triads has attracted some attention. Kelly suggests a list of triads for the Role Construct Repertory Test in which repetition of an element in successive triads is almost completely eliminated. In the Sequential Form two elements are common to any successive pair of triads, and Kelly remarks that this was a more exacting test of the functional adequacy of a person's constructs, but does not elaborate on the point. Keen (1979) for example, followed Kelly's sequential approach quite closely, his grids exhibiting considerable commonality from one sort to the next.

Bender (1974), noting that Warr and Coffman (1970) had unexpectedly failed to detect a difference in rating extremity between elicited and supplied constructs, decided to investigate whether the anomaly might have been attributable to the particular elicitation procedure that had been used. Warr and Coffman had used the Sequential Form of the Repertory Test in the elicitation of constructs, and Bender hypothesised that this method might tend to produce trivial constructs. He reasoned that few 'really meaningful' constructs could be verbalised and that a respondent, having produced one of these in response to a particular triad, might find it impossible to produce another such construct in response to a succeeding triad in which two elements remained the same as before.

Bender individually tested 60 junior social workers, using a grid in which 'self' was an element common to all triads and in which there were seven instances of another element being 'carried over' from one triad to the next and nine instances where this was not the case. He found a tendency for 'sequential' sorts to produce 'unimportant' constructs and for 'non-sequential' sorts to produce 'important' constructs, 'important' and 'unimportant' being defined as those five with the highest and lowest loadings respectively on the first factor of Slater's principal components alalysis. Considering that 600 constructs were elicited, a χ^2 value of 6.54, though statistically significant at the 0.02 level, is not particularly impressive. The criterion of importance is open to challenge, for it could be argued that high loadings of constructs on the first factor are commensurate with high semantic correlation: though labels may vary, the constructs may exhibit substantial functional similarity[44]. In this case, 'important' constructs would not necessarily load highly on the first factor - on the contrary, they might idiosyncratically define

subsidiary factors.

Whilst Bender's research design can be criticised, his basic reasoning is not without its merits. He makes a strong case for not persisting with the 'self' element in all triads on the grounds that it may restrict the elicitation of new constructs to the two remaining elements of the triad, so that even if the focus of interest is the self, the maximally informative procedure may be one which often approaches the self obliquely. It would appear probable that Bender's argument can be generalised to other grid contexts as well.

My own experience in using grids indicates that some triads seem to 'block' rather than facilitate the elicitation of constructs when the task requires a new construct to be produced for each sort. A switch to a completely new triad has often been sufficient to overcome the problem, save where the respondent's flow of constructs has dried up.

Any interaction between researcher and respondent is open to challenge in respect of the bias that may creep into the elicitation procedure. Shaw and Gaines (1979) see a distinct advantage in using a computer program to elicit grids, in that it interacts only in terms of what the task demands and does not interfere with the cognition that the respondent chooses to produce in respect of the triadic elicitation routine. An interviewer thus becomes redundant; but if such a procedure is built into a research design, the researcher will need to gain access to the grid data so remotely collected - the price being the loss of the individual's privacy and the diminution of meaning regarding the terms used in completing the grid. Post hoc interviewing could help to amplify the grid information, but the amplification would not necessarily be the same as that which might have been

collected at the time of elicitation.

The PEGASUS routine provides some instant feedback to the respondent regarding the associations and discriminations being made, thus allowing the respondent to reflect on 'the position so far' as the grid elicitation proceeds. Whilst there are advantages in so doing, the counter-argument can be advanced that such feedback can induce a revision in construing which might otherwise not have taken place. A similar balance of argument exists in respect of interviewing a respondent: one can be 'detached' at the price of minimising communication, or one can be involved in a conversation at the risk of 'contaminating' the data.

Despite Kelly's logical point that a construct relates to the copresence of similarity and contrast, not everyone is satisfied with the notion of triadic elicitation[45]. M.T. Taylor (1976) questions whether this does tap the way people normally think[46], and the argument developed in Chapter 3 suggests that a contrast may be so submerged that it never surfaces or, as another possibility, it may be multi-categorical. If either of these two deductions is correct, it suggests that triadic elicitation may not be the most appropriate methodology in all cases. The demands of logic and of 'psychologic' are not necessarily congruent - an assumption that is a foundation stone in the Psychology of Personal Constructs.

6.5.2 Dyadic elicitation

Fransella and Bannister (1977), following Kelly's view that pre-existing constructs are being elicited, indicate that it is not essential to use triads of elements during elicitation and that pairs may be adequate. In some circumstances dyadic elicitation may be

distinctly superior. Frost and Braine (1967), for instance, found that the construing of tastes was a very difficult task for respondents when triads of taste sensations formed the context of elicitation. Both memory and 'taste-fatigue' were limiting factors for the respondents. Frost and Braine, for these reasons, elicited as many constructs as possible from each dyad.

Fjeld and Landfield (1961) used dyadic elicitation, having found from clinical experience that some respondents were unhappy with the triadic approach, being unable to provide a similarity before stating a difference, yet being able to draw a distinction between two of the three acquaintances presented as elements. When presenting elements in pairs, Fjeld and Landfield asked their respondents to articulate a contrast between the two elements or to indicate the grounds on which the pair were similar (in which case a third element had to be provided who could be described in contrasting terms). Other users of dyadic elicitation include Allison (1972), Ryle and Lunghi (1970) and Smith (1978).

I, too, have used dyadic elicitation in a study of element stability (see Appendix 10), in which grids were completed individually within a group administration. My previous experience with grids elicited individually from group members showed that the triadic approach caused difficulties for some respondents, and that these often did not become (or were not made) apparent until the completed grids were collected. As far as I am able to tell the dyadic approach did not give rise to similar problems.

Easterby-Smith (1981) recognises the difficulty some people have when faced with triads, and implies that dyadic elicitation is particularly

advantageous where the elements are complex, an example being the relationships between people that form the elements in Ryle and Lunghi's (1970) dyad grid. On the other hand, there is a tendency for the construct elicited to embody logical opposites rather than psychological opposites, the comparative virtue of each being an unresolved issue in elicitation technique.

In asserting a contrast between two elements one is presumably linking each element to at least one other stored in the memory and possessing attributes similar to those which give rise to the verbalised contrast. And the assertion of similarity, in Kellian terms, implies the existence in memory of at least one element which does not possess the attribute shared by the presented pair. Dyadic elicitation then, can be construed as a method which implicitly recognises the Kellian assumption of similarity and contrast, though it may well draw (again implicitly) upon a wider context than the three elements in triadic elicitation. Its advantages would seem to be sufficient to warrant its further study as a method of elicitation.

6.5.3 'Full context' elicitation

The 'Full Context Form' of Kelly's Repertory Test involves the discrimination of similarity and contrast through the consideration of all the elements taken together. Kelly gives little detail regarding the use of the Full Context Form once the respondent has given an initial indication of the reason for placing the first pair of similar elements together, but it is possible to infer from his own description, and from a slightly fuller description in Bannister and Mair (1968), that similar elements are collated (as piles of named cards) according to a discrimination made by the respondent[47].

Elements may be removed from piles or transferred to other piles as the basis of the discrimination alters. At each stage the researcher keeps a record of the location of the elements in their groups and of the constructs associated with the piling of the elements.

In that elements can be shifted from pile to pile, Scott's (1962,1963a) approach to cognitive structure shares a likeness with the Full Context Form, though Scott's respondents were asked to sort elements (countries) into sets which were not necessarily mutually exclusive and he did not ask for the basis for the discriminations to be articulated. Zajonc's (1960) approach to cognitive structure has Kellian echoes in that he requires his respondents to articulate differentiations from a full set of elements, but his method demands that the elements be successively partitioned in a manner similar to that of divisive hierarchical cluster analysis. Thus, in contrast to Scott's, a strict set-inclusion principle operates in Zajonc's procedure.[48]

Kelly does not describe a use for the Full Context Form in repertory grid methodology, but it is not difficult to see that the partition of elements into two groups could form the basis of a dichotomously-scored grid. It is probable that the instructions would need to specify the 'opposite', rather than the difference, form of verbal labelling in order to avoid the problem of elements falling outside the range of convenience of the construct dimension being used. Successive sorts, using different constructs, would enable the grid matrix to be built up.

Reid (1976) describes how he used a variant of the full context form,

asking respondents to select two elements to exemplify the pole (sic) of a construct that was considered to be 'most important and most immediately evident'[49]. The location of element on the construct so identified was accomplished by laying out the cards on the table according to a five-point rating scale.

Nash (1976), in contrast, used the full context sorting of elements as a preliminary to a series of dyadic elicitations of pupils' constructs regarding their teachers. He initially asked the pupils to sort into two piles cards bearing the names of their teachers, the criterion being whether the individual pupil 'got on with' the teacher or not. This preliminary sort having been completed, Nash took one card from each set and asked in what ways the teachers behaved differently. To judge from the transcript appended to his paper[50], the subsequent comparisons did not necessarily require that the pairs of elements considered came from the opposite piles - and it is difficult to see what the original dichotomous assignment of elements was intended to achieve, other than perhaps to help the pupils feel at ease at the start of their interviews and to establish a cognitive 'set' towards the elicitation of contrasts. It must be pointed out that Nash was seeking not to collect data in a format suitable for repertory grid analysis, but rather to ascertain the pupils' constructs of teachers and teaching.

Olson (1980b) justifies the use of the Full Context Form on the grounds of the type of elements involved (classroom activities) and because he wanted to avoid the production of superficial constructs. Implicit in Olson's remarks is the assumption that the triadic approach, with which his own is contrasted, might have elicited a number of constructs

with a range of convenience inadequate for analysis - in other words, he feared that he might have obtained very 'gappy' grids.

But Olson's argument for the Full Context Form lacks conviction, particularly when one looks at the constructs he elicited, which are often very general and parallel those which I obtained from my group of science teachers using triadic elicitation with slightly more general elements. Olson's elements are to some extent heterogeneous, focusing sometimes on pupils activity and sometimes on teacher activity, and hence one would tend to expect rather broad and general constructs. Olson's work has considerable strength in that he asked his teachers to <u>elaborate</u> the general constructs, and these elaborations are rich sources of understanding of how his respondents construed teaching activities. The methodological penalty (if such it be) is that the elaborations are so idiosyncratic that they will not fit into a repertory grid (not that Olson intended that they should).

On the whole, I would suggest that the full context approach to elicitation is likely to collect broad, general constructs since these by definition have to apply to <u>all</u> of the elements presented. The triadic and dyadic approaches, whilst being more narrowly focused in context, offer a greater possibility of constructs that are less bland, but at the risk of a number of elements falling outside their ranges of convenience.

The evidence indicates that, to date, little use has been made of the Full Context Form of elicitation in repertory grid work. Though no criticism of the approach has been found in the literature, its relative unpopularity may be connected with the difficulty of providing a linear dimension appropriate to a plethora of elements, each of which

presents its own distinctive features to the respondent's attention.

6.5.4 Laddering

Hinkle's (1965) 'laddering' procedure was discussed in Chapter 4 in connection with the limitations of the repertory grid's capacity to provide evidence regarding the structure of a construct system, and it is therefore treated briefly here.

It will be recalled that, in elaborating the Organization Corollary, Kelly suggests that constructs stand in hierarchical relationships to each other, superordinate constructs being held to subsume subordinates, the extent of subsumption depending on the hierarchical level of the superordinate. Kelly does not indicate whether the hierarchies he envisages are of the strict set-inclusion type or are based on overlapping sets, but he does hint that strict hierarchical organisation might not be the structural model upon which construct systems are based[51].

'Laddering' was Hinkle's method of approaching the question of hierarchy, respondents being asked repeatedly why they preferred one pole of a construct to another, the responses being taken as superordinates. Sooner or later the respondent reached a point in this hierarchical ascent beyond which further progress was impossible, and Hinkle turned his attention to a different (basic) construct derived from triadic elicitation. Hinkle found in pilot work that students could usually produce, from each base construct, between eight and twelve superordinate constructs in this way, though Bannister and Mair (1968) indicate that in Britain considerably fewer superordinates are typically produced in an individual ladder[52].

Although Hinkle states that there was a tendency for ladders to converge on common superordinate constructs, he does not present adequate evidence to justify the hierarchical distinction he makes between superordinate and subordinate constructs: some constructs which were not laddered may have been of high position in the system. Figure 6.2 gives a schematic picture of the situation. If constructs A, J, P, S and T are 'base' constructs the first three may converge on the superordinate F whilst the relationship of S and T to F is not explored: in this hypothetical illustration S, though a 'base' construct, is hierarchically superior to D, K, C, J, Q, T, B, P and A.

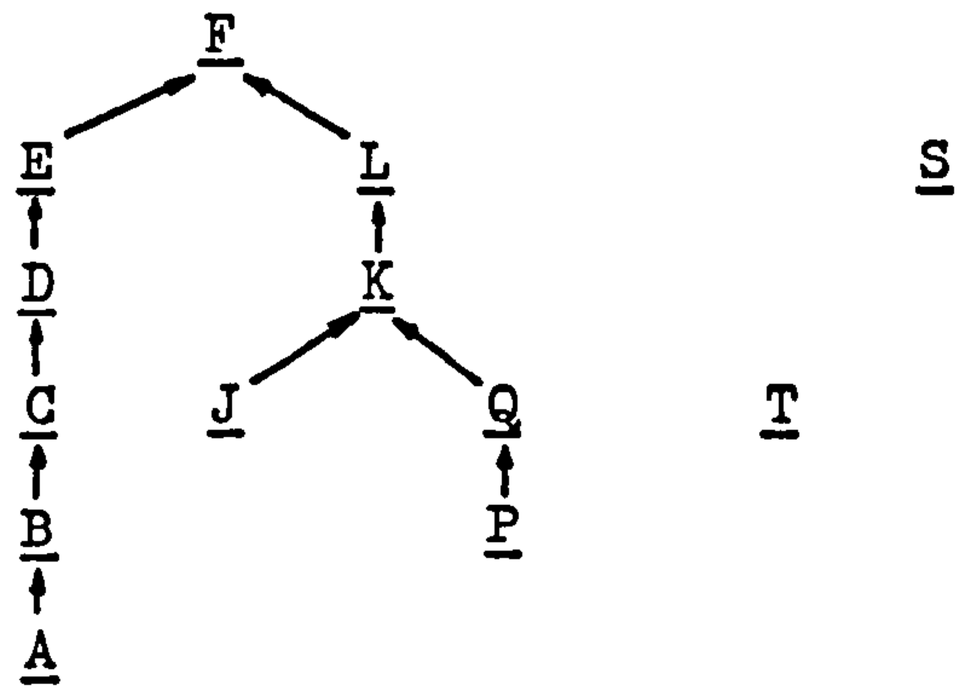


Figure 6.2 An illustration of the weakness of Hinkle's assumptions of superordinacy/subordinacy. A, J and P are 'base' constructs which are laddered to a common superordinate F, whilst S and T are unladdered.

Laddering is a time-consuming process. In fact, Hinkle found it so time-consuming that he was only able to ask his respondents to ladder a few of the ten 'base' constructs which they produced.

Notwithstanding the structural and temporal problems ingrained in Hinkle's work, a number of workers have used laddering in construct elicitation though in some cases (e.g. Landfield, 1971) the laddering has been 'downwards', focusing on subordinate implications through sequential enquiry as to the rationale underpinning the poles of the

construct already verbalised. Honikman (1976), whose interest lay in the construing of living-space, laddered downwards towards the more 'concrete' subordinate implications of constructs such as 'formal-informal'. When the <u>importance</u> of constructs was investigated Honikman found occasionally that a construct low in the hierarchy was seen as important by the respondent, and he quotes the instance of a librarian whose mention of 'books in the living room' came late in the laddering procedure yet was the second in importance among twenty constructs. There are a number of ways of interpreting findings of this sort (for instance, contrasting the statics and dynamics of construct systems or postulating a determining superordinate structure in the unconscious), but I restrict myself here to the simpler, pragmatic point that it appears dangerous to assume a monotonic relationship between importance (or perhaps salience) and hierarchical level as manifested through a laddering procedure.

Hinkle's laddering focused its attention upon the preferred pole of the construct under consideration, asking for its superordinate implications, and Armstrong and Eden (1979) seem to have taken a similar methodological view in their study of the purposes held by local authority valuers. The implicit assumption here is that the superordinate implications of the non-preferred pole follow in parallel. Such a position is in accord with what Hinkle termed a parallel implicative relationship in which, for two constructs A-B and X-Y A implies X and B implies Y.

But, as Hinkle (1965) and ten Kate (1981) are well aware, the implicative relationships between constructs are not necessarily as straightforward as the parallelism described above[53]. To follow through the implications of a preferred pole, therefore, is not

necessarily to follow through the implications of its non-preferred opposite: the implicative pathways might diverge.

There is little direct evidence bearing upon the problem since most laddering seems to have followed Hinkle's original rubric. Wright (1970), however, reports the study of a single case in which the superordinate implicative relationships of both poles were followed and the divergence of implicative pathways became apparent. Landfield's (1971) pyramiding technique, focusing on subordinate implicative relationships, seems likely to produce a similar divergence, to judge by the example given in his text[54]. Evidence of a more indirect nature, yet to the same effect, can be gleaned from Fransella's work with bi-polar implications grids in which the pattern of implicative relationships does not totally reflect parallelism in Hinkle's sense of the term[55].

There can be little doubt that laddering produces constructs which elaborate 'base' constructs both superordinately and subordinately. Whether this can include recursive, looping processes[56] remains a matter of debate, for it is not at all clear whether laddering elicits constructs of sequentially greater superordination or subordination depending upon the way in which the implicative relationships are elicited. Given the present state of the art, it may be safer to use the laddering as a way of eliciting elaborations of a person's constructs - akin to the way in which Ravenette (1977) seeks to explore the meaning of verbal labels - rather than to use it to determine hierarchical relationships between constructs in the absence of a specific context.

6.5.5 Freer forms of elicitation

Though laddered constructs could be incorporated in repertory grids this seems not to have been done, researchers preferring to follow the implicative networks rather than the statistical associations that can be derived from the repertory grid. The freer forms of elicitation are likewise typically directed towards the constructs used by the respondent together with their implications (though these tend to be less systematically explored than is the case with Hinkle's approaches), and it is rare for the constructs so produced to be used in repertory grids.

The work published by Nash between 1973 and 1978 indicates a marked shift in attitude to the repertory grid as an instrument of research. In 'Classrooms Observed' (1973) the grid featured prominently in his investigations of the ways in which pupils construed their teachers. The study reported in 1976 saw him move away from the repertory grid and use (following an initial full context sort) a dyadic elicitation procedure focusing on contrast. In response to a paper by Clift et al (1978), Nash (1978) indicates that he finds standard elicitation procedures eliminating some perhaps meaningful constructs and he has moved to a still freer form of elicitation in which he asks teachers to talk about pupils as if they were discussing them with a colleague. The discussions are tape-recorded and Nash extracts from the transcripts a number of constructs which are taken as central to the teachers' perceptions of their pupils. He finds this to be quicker and less threatening to his respondents and argues that, since it is less formal (and apparently less powerful) than other approaches, it can be practised with respondents such as education officials who might otherwise be unwilling to co-operate. How this information is used is not reported:

presumably at the time of publication of this short response Nash was referring to work still in progress.

Nash's description of what might be termed a 'classroom character sketch' is reminiscent of Kelly's use of self-characterisation as a method of eliciting constructs. In this approach Kelly asks the respondent to write a character sketch of him- or herself from the standpoint of an intimate and sympathetic friend. In this approach the constructs elicited are anchored to the construction of what one might call, after the existentialists, 'self-in-the-world' (or, at least, that part of the world of immediate interest). As Fransella and Bannister (1977) point out, such an approach is by no means as structured and 'tidy' as formal elicitation procedures[57]; there seems no reason in principle why constructs elicited in this way could not be incorporated into a grid format[58] though, with self at the focus of elicitation, there could be limitations on the ranges of convenience of the constructs used. Nash's interviews would seem less restricting in this respect.

Fransella and Bannister (1977) note that self-characterisation has been little explored, but suggest that it tends to seek out superordinate, rather than subordinate, constructs when used in the self-oriented sense of Kelly's formulation[59]. It remains an open question whether the superordinate emphasis holds when the respondent freely describes circumstances exterior to the self. Bannister and Fransella (1980) make the point more strongly when they claim that self-characterisation elicits more superordinate constructs than triadic elicitation[60]. However, little evidence seems to be available to support either the stronger or the weaker claim. In Kelly's own account he indicates that he is interested in something more than superficial appearances

when he asks for a self-characterisation, but this does not necessarily imply the production of superordinate constructs: in the case study of 'Ronald Barrett' it would appear that some superordinate constructs were elicited since the implicative evidence is present in the sketch produced by the respondent[61].

Freer forms of elicitation are not limited to self-characterisation and its analogues. Ravenette (1980) makes use of a 'portrait gallery' in which he presents a child with two schematic faces (one 'happy' and one 'sad') and asks the child to distinguish between them and to say three things about each[62]. C.P. Hargreaves (1979) provides respondents with wedge-shaped cards bearing the names of their acquaintances and asks that they be laid out meaningfully on a table: he reports that sometimes the construct emerges only after the layout has been completed. Honess (1978) used essays about liked and disliked same-sex peers to provide him with twelve frequently used unipolar constructs to which children had to supply the opposite verbal labels. Karst and Groutt (1977), like Nash, preferred to avoid the formal interview appearing as a psychological test, but were unwilling to adopt an unstructured approach. Their solution was a compromise - to use triads but, having presented a triad, to ask respondents to compare and contrast the three elements in as many ways as possible. 'Questions were inserted into the proceedings in order to keep the interaction lively and to allow for the unveiling of other significant constructs or for the elaborations of those previously articulated to be revealed.

Perhaps the most 'element-free' elicitation was used by Bonarius (1971) who asked 125 students to provide two constructs by writing down what they felt were the 'two most important dimensions for describing people'[63].

These constructs were later used in an experiment on the extremity of rating, but his so-called 'Free Formation Procedure' has not been extended to grid method proper.

Keen and Bell (1981), echoing M.T. Taylor's (1976) doubts regarding the capacity of standard grid administration to reflect the ways in which people think, suggest that the triadic technique may be unusual and unnatural in this respect. They take the view that reflective thinking and conversing is much less systematic than is consonant with triadic elicitation and in consequence have developed a computer algorithm, DYAD, which elicits both elements and constructs in an interpenetrating sequence. A computer is critical to the success of the elicitation of a full grid of ratings because of the difficulty a human interviewer would find in keeping track of all the partly-recorded relationships already elicited and of the remaining questions essential to the completion of the grid matrix. I have not worked through a DYAD routine, but it seems likely (to judge from the example appended to Keen and Bell's paper) that it is open to the same sort of reservations as, for example, Shaw's (1980a) PEGASUS program in respect of bias, linguistic impoverishment, and tedium during completion.

6.5.6 Methods of elicitation: a concluding perspective

Triadic, dyadic and full context elicitation rely explicitly on similarity and contrast, whereas many of the freer forms of elicitation tend to leave similarity and contrast at an implicit level. Laddering, which can be seen as complementary to the elicitation of 'base' constructs still rests upon the assumption of contrast. The question that has to be asked here (and that is implicit in the whole of this appraisal of the repertory grid) is whether the emphasis on similarity and contrast is sufficient for the researcher to gain an adequate

understanding of the respondent's world.

Kreitler and Kreitler (1976) identify four categories of meaning-relationship of which only the first two seem related to the completion of repertory grids. The four categories are

(i) attributive (of qualities to the referent);

(ii) comparative (similarity, dissimilarity, complementariness, relational);

(iii) exemplifying-illustrative (of instances, situations, scene), and

(iv) metaphoric-symbolic (interpretative, metaphor, symbol)[64].

Kreitler and Kreitler suggest that categories (i) and (ii) are mainly appropriate for conveying interpersonally shared lexical meanings and that categories (iii) and (iv) are more appropriate for meanings in the realm of the personal-subjective and symbolic. As the authors point out, the latter pair are more open to error in encoding.

Such repertory grid data as I have seen are strikingly lacking in category (iii) and (iv) meaning relationships when set against other information collected from respondents: the work of Olson (1980 a,b) who collected constructs explicitly in repertory grids and more implicitly when his respondents elaborated the rather general constructs elicited by his grid protocol, points up the distinction very clearly. The conclusion I draw from Olson's work and the range of repertory grid studies I have also studied is that the formalised elicitation techniques can at best give but a partial indication of the respondent's construing and that the richness of understanding is likely to accrue from freer approaches which allow the respondent to convey meaning through exemplification and metaphoric allusion.

This standpoint is not without its risks of selectivity and distortion: a rigorous approach to research into personal construing requires more than testimony, as will be argued more fully in Chapter 13. Any grid user worthy of the title of researcher or clinician is aware of the importance of the embellishments to the grid data that conversation (that freest of forms of interactive elicitation) can provide, and much interpretation of grid data takes such information into account. It is in these aspects of elicitation that formalisation - exemplified par excellence by computer algorithms - is at its weakest.

Many studies reported in the literature do little other than report lists of elements, of constructs and of their statistical relationships and couple these to interpretations which lack the vibrancy of the existential experience of the respondents. Leading a workshop at the 1982 British Conference on Personal Construct Psychology, Peter Stringer drew the conclusion that perhaps the users of repertory grids did not spend enough time talking to their respondents; to which I would add that maybe they do not always spend enough time listening to the wealth of information that respondents are often willing to give.

6.6 ELICITED VERSUS SUPPLIED CONSTRUCTS

6.6.1 Introduction

The question of whether the researcher should elicit constructs from respondents, or supply constructs to them, has received a considerable amount of attention. Bannister (1973) dismisses the issue as a pseudo-problem arguing (like Hinkle, 1965) that, whoever provides the verbal label, the respondent necessarily provides the underlying construct from his or her repertoire. Bannister further argues that if 'the experimenter tries to supply verbal labels which are in the

native tongue of the subject and which relate to constructs likely to be important to the subject ... there will be no difference in the subject's ability to use "supplied" as contrasted with "elicited" constructs'[65].

Bannister's first point is well taken, but the second begs a number of questions. Despite acknowledging that the second point depends on the adequacy of the experimenter's understanding of the subject, he gives no indication regarding either how this might be achieved or how the degree of understanding might be assessed. One is left with the feeling that Bannister would call upon a normative framework derived from experience in order to (implicitly) settle the issue. His position also skates over the question of whose constructions of the problem are being researched.

The burden of these comments on Bannister's position is that the question of 'elicited versus supplied constructs' is by no means a pseudo-problem but is a problem that must be addressed at the outset of research involving repertory grids.

Bonarius (1965), reviewing studies pertaining to the Individuality Corollary, concluded that individuals preferred to express themselves in their descriptions of people by using their own verbal labels, rather than by employing labels supplied by others[66]. Later Adams-Webber (1970) came to a broadly similar conclusion in his review of elicited and provided constructs, though he added that carefully selected lists of adjectives might be used in some circumstances as effectively as the verbal labels provided by respondents[67]. More recently, work by Metcalfe (1974) and Kuusinen and Nystedt (1975) has been interpreted as showing the desirability of allowing the

respondent to use his or her own verbal labels though, since both studies investigated aspects of cognitive complexity rather than construing as such, the inference is not entirely clear. As far as personal choice is concerned, Burgoyne (1981) observes that the respondent's own constructs are superior to those provided from outside.

6.6.2 Extremity of rating

Much has been made in the literature of the link between extremity of rating and personal relevance or importance, though it is not always clear whether the relationship would be better characterised in terms of personal meaningfulness to the respondent.

In an early study Cromwell and Caldwell (1962) asked 44 students to provide six constructs in response to a shortened form of Kelly's Repertory Test, and then to rate acquaintances on the six constructs that they themselves had provided and also on six constructs from other students. The respondents were divided into two groups, the order of presentation of elicited and provided constructs being reversed in the second group. As the authors had predicted, acquaintances were rated more extremely on the personally-elicited constructs than on the supplied constructs, and this is interpreted as support for the hypothesis that 'own' constructs are used more decisively than supplied constructs in judging the characteristics of people. Isaacson and Landfield (1965) found a similar tendency towards greater extremity of rating 'self' when respondents' use of personally-elicited constructs was compared with their use of Butler-Haigh Q-sort statements. And when clients were asked to rate themselves on constructs which they had provided and on constructs provided by their therapists, Landfield (1965) found the ratings on the former

to be significantly more extreme than those on the latter. Landfield interpreted these findings as demonstrating the greater meaningfulness of personal constructs, but his question to the respondents was couched in terms of usefulness to describe people, and one would expect a present or ideal self to be construed more clearly (and therefore show a tendency to greater rating extremity) than a comparatively unfamiliar therapist, particularly when the respondent's own verbal labels formed the scales for judgment. It is not surprising to find that no significant difference in extremity was found between the two 'selves' and the therapist when the therapist's constructs were used.

Other work in the area of meaningfulness and extremity tends to indicate that greater extremity (Bonarius, 1971; Stringer, 1972) and greater differentiation (Delia et al, 1971) are obtained with the respondent's own constructs than when the constructs are supplied[68]. These findings would appear to be in harmony with Koltuv's (1962) finding that variance of scoring tended to be higher on traits felt by the respondent to be personally relevant.

Bender (1969) used a slightly different approach to the problem. He elicited constructs from 52 university students using triads drawn from a list of eighteen people (nine 'who influence you most' and nine nominated against role titles). A resistance-to-change grid (Hinkle, 1965) was used to establish the order of importance of the elicited constructs. Following the rating of the elements on the constructs (a scale from +25 through zero to -25 was used), Bender found that the range of rating was significantly higher for the five most important constructs than for the five least important constructs, and that a similar tendency existed in respect of the distance of the

rating of 'self' from the centre of the scale.

An alternative to the resistance-to-change grid (which is a time-consuming task) is the direct rating or ranking of constructs against a criterion of meaningfulness or importance, though little use appears to have been made of it in practice. Adams-Webber and Benjafield (1973) asked 30 undergraduates to rank twelve supplied constructs in order of their perceived usefulness for describing people and found that the higher the perceived usefulness of a dimension, the more extremely respondents rated themselves and others upon it. Wood and Napthali (1975) used a similar ranking approach in respect of their study of teachers' perception of pupils, the criterion in this case being the usefulness of the construct for a teacher taking over a new class.

Tajfel and Wilkes (1964) found evidence to suggest that salient constructs (adjudged so on the criterion of rating extremity) tended to be those produced early in an elicitation routine and also frequently repeated, though their failure to offer any convincing explanation of the origins of salience attracted criticism from Eiser (1971). My work with science teachers suggests that early elicitation and importance are not systematically related, as far as nomothetic analysis is concerned[69]. The importance of constructs to respondents was determined through the use of resistance-to-change grids (Hinkle, 1965), and the Spearman rank correlations, for 27 grids, between order of elicitation and importance ranged from +0.60 to - 0.38, with a median value of + 0.05. Further, a Spearman rank correlation of only +0.26 was found between the above correlation coefficients for thirteen pairs of pre and post-test grids separated in time by some seven months, implying that whatever order/importance relationship exists

can be at best only weakly stable over time.

As for repetition, Shubsachs (1975) found that frequently produced constructs tended to be rated as more important than other constructs, but his procedure may have cued his respondents into making connections between frequency and importance that might not have been made in other circumstances. Fransella and Bannister (1977) suggest that construct repetition might be a good index of relative superordinacy, and that, if the elicited constructs are to be used in a grid, the 'grid analysis itself will give an indication of superordinacy'[70]. No evidence is presented for either assertion: the first is at best a tenuous connection, and the second was shown in Chapter 4 to be unjustified. However, the issue of frequency of elicitation does not arise in most repertory grid work since the vast majority of researchers ask their respondents for fresh constructs at each stage of the elicitation and do not accept repetitions.

Other work on extremity of rating suggests that extremity and bias[71] both tend to decrease with increasing age - at least, for the range 6 to 17 years (Applebee, 1976); that extremity may relate to the value connotations of scale terms (van der Pligt and van Dijk, 1979); and that extremity may relate to the intensity of emotional state (Caplan et al, 1975), though it must be said that this last finding is the result of observation of a single case.

Reflecting on the underlying determinants of rating extremity O'Donovan (1965) suggested that both personal relevance and psychopathology might be involved. However, Warr and Coffman (1970) found no relationship between extremity of rating and either sex or personality measures. They suggested instead that a compound of the

perceived importance of the stimulus and the personal relevance of the construct dimension might be a critical determinant of rating behaviour, with pathology perhaps playing a significant role under conditions of high personal involvement. Chetwynd (1977) reviews the literature on extreme response style from a different perspective and draws the conclusion that extremity of rating tends to rise with increasing pathological functioning and to fall with maturity and complexity of cognitive functioning.

It is difficult to sum up this literature adequately, since in encompasses a range of terms whose meanings partially overlap (e.g. meaningfulness, salience, importance) and which appear to have been used interchangeably and without definition. Allowing for this looseness in conceptualisation and terminology, it would appear that a tentative general statement could be made on extremity of rating which would draw together the points made by Warr and Coffman and by Chetwynd. One might expect individuals to depart quite markedly from the general trend: my own experience suggests that rating extremity may result from <u>low</u> (rather than high) involvement, the person using the extremes as a way of getting through the procedure with the minimum of effort - effectively converting the rating procedure to a dichotomous allocation of elements.

The point must also be made that whilst rating extremity <u>may</u> imply the importance, meaningfulness or salience of the construct to the respondent, the reverse implication is invalid. Slater (1977) falls into this type of error when he states that a low sum of squares for an element (an indication of a set of ratings close to the mid-points: of the constructs) implies that the respondent is indifferent to the element[72]. This will, in practice, depend upon the psychological

characteristics of the constructs involved: if the construct is evaluatively positive at the centre and negative at each end, a respondent could plump for the mid-point for an important element showing desirable features. Many personality constructs have this E-/E+/E- loading along their dimension (e.g. 'extrovert-introvert', 'bold-shy', 'loquacious-taciturn' and so on), therefore calling Slater's assertion into question. A similar argument can be developed in terms of the variance of ratings on constructs since some constructs may be used dichotomously (e.g. trustworthy- untrustworthy) with extreme ratings producing high variance, and others, equally important to the respondent, may be used in a more graded way and give rise to a much lower variance.

It is clear that the issue of rating extremity is far more complex than many researchers have recognised. Interpretations of extreme ratings must be undertaken with considerable caution, and it would seem highly desirable that researchers try to garner additional information in respect of possible reasons for extremity of rating.

6.6.3 Are elicited constructs superior?

It is perhaps with Warr and Coffman's (1970) work that some doubts began to be raised regarding the superiority of elicited over supplied constructs[73]. They asked 31 students at the University of Sheffield to rate the same twelve stimulus persons on twelve bipolar constructs elicited according to Kelly's Sequential Form of the Repertory Test, and on twelve supplied constructs presented in semantic differential form. They found that the mean extremity scores were almost identical and suggested that their provided constructs (which had been chosen carefully to sample the three semantic differential factors of evaluation, potency and activity) were as meaningful as the elicited

constructs for the assigned task. Bender's (1974) attempt to account for this 'anomalous' finding hinged on the suggestion that the elicitation procedure might have produced less important constructs[74], but his argument does not appear to be sufficiently strong to invalidate completely Warr and Coffman's conclusions.

Adams-Webber (1970) takes the view that, when common adjective labels are used to form dimensions of social judgment, it is likely not to make much difference whether respondents use their own verbal labels or lists of adjectives provided by the researcher. Fransella and Bannister (1977), in a brief discussion of the issues involved, take a similar position in their recognition that supplied constructs have a place in grid methodology, particularly where clinical and educational use is involved[75]. Their contention is supported by work such as that of Ravenette (1975) who, in his work with children, has demonstrated the usefulness of supplying constructs.

As a researcher or clinician one does not have to opt for either elicited or supplied constructs since it is open to combine the two types. Olson (1980a), for example, used ten constructs in his grids, five of which were elicited from his respondents and five provided by himself on the basis of pilot interviews (since there were aspects of science teaching that he particularly wished to investigate)[76]. Olson took the precaution of not showing the constructs he wished to supply before his respondents had produced their own constructs, being aware of the possibility of contamination. The mixing of supplied and elicited constructs offers a compromise between the strengths and weaknesses of both, but safeguards need to be built into the procedure adopted in order to ensure that meaningful communication takes place.

There is another way of combining elicited and supplied constructs, and that is to elicit a number of constructs from a group of respondents believed to be similar to the focal sample of the research and to select constructs believed to be appropriate to the research task using criteria such as frequency of occurrence. Clift et al (1978) used this procedure when they collected the twenty most frequently produced constructs from 130 teachers and used them in a 'supplied construct' grid, assuming consensus regarding meaning. Alternatively, the researcher might conduct an analysis of the list of elicited constructs and 'boil down' the list to manageable proportions, taking advantage of such semantic similarities and overlaps as present themselves. Reid and Holley (1972), for example, followed this general approach in their study of sixth formers' construing of universities, reducing 198 constructs collected from a pilot sample of similar pupils to nine constructs which were supplied in the grid administered to their main sample.

These nomothetic procedures seem inferior to the practice of mixing supplied constructs with individually elicited constructs in grids which are then at least partly idiographic in content. To assimilate constructs which apparently overlap (as Nash, 1973, did in the work reported in 'Classrooms Observed') is to introduce a series of semantic mean or median constructs to which no respondent may claim allegiance and which reflect no more than the researcher's interpretations of the meanings of a cognate group. Neither rater nor analyst can have a clear grasp of the meanings that have been exchanged during the completion of such a grid, which would seem to maximise the risk of misunderstanding.

Although there is in the preceding paragraphs an implicit answer to the question heading this section 'Are elicited constructs superior?'. a formal summing-up is delayed while the issue of communication of meaning is explored a little further. This is crucial to research within the framework of Personal Construct Theory and yet receives little attention in the literature.

I shall return to Bannister's (1973) point that the supplying of constructs implies a need for the researcher to understand the problem area from the perspective of the respondent, to supply a representative list of verbal labels, and to share sufficient understanding with the respondent of their meaning to allow a reasonable interpretation to be made - a tall order. An experienced researcher may well be able to draw on his or her background to offer a reasonable prospect of seeing the problem area from the respondent's point of view - at least, well enough to engage in a meaningful conversation about it. Representative labels may be fairly easily supplied from experience, and this does have the advantage of being a quicker procedure than elicitation: it may be unavoidable where the research is limited by the time available. The assumption of intersubjectivity of meaning, however, is not always justified even for commonplace constructs and it is worthwhile for the researcher to check whether the respondent is sharing the 'supplied meaning' - and, if not, what meaning the respondent is actually intending to convey.

Research using the semantic differential is a fruitful source of information about the problems of shared meaning. In their original presentation of the semantic differential Osgood et al (1957) drew attention to concept X scale interaction[77], a finding subsequently replicated in the cross-cultural studies of Osgood et al (1975)[78].

Here - and translating Osgoodian into Kellian terminology - it is apparent that the meanings of supplied constructs shift according to the element being rated: this is a finding of considerable importance for research conducted within a Kellian framework, and its implications are explored more fully in Chapters 8 and 13. To add to the confusion, Mann et al (1979) reported a concept X scale X person interaction which accounted for a substantial proportion of the variance on scales linked to the three main semantic differential factors, though the 24 concepts used were sufficiently heterogeneous to lead to substantial two-way interactions with scales and persons taken separately[79].

That commonplace verbal labels may not always be consensually valid is indicated by the results of an exploratory study which I conducted. This was grounded in implicit personality theory (Bruner et al, 1958). Twenty-nine lecturers in further education colleges, who were attending a day-release certificate course in education, were asked to assess the probability of implicative relationships between a number of attributes that had been most frequently mentioned by them on an earlier occasion. Among the attributes listed was 'conscientious' which, on dictionary criteria, would not be expected to be more than marginally correlated with attributes such as 'intelligent' and 'creative'. Despite defects in the instrument, it became clear that conscientiousness was, for five of the respondents, systematically negatively related to the various positive attributes being presented, whereas no clear relationships were found for the other 24 respondents (as was expected). For the exceptional five respondents, 'conscientiousness' seemed to connote 'unimaginative' and 'dull intellectually' - yet this finding would not have come to light had the responses not been scrutinised on an individual basis[80].

It is obviously dangerous to generalise too far from the evidence presented above, but it serves as a warning that the use of supplied constructs may sometimes take too much for granted regarding the intersubjectivity of meaning: grid analyses offer some prospect of retrieving a misunderstanding through the appearance of surprising correlation coefficients or distance measures, in which cases it may be possible to return to the respondent and ask for clarification of how a particular construct was actually used.

The problem of intersubjectivity is not avoided by eliciting constructs, for the researcher needs to be sure that he or she has grasped what it is the respondent is wishing to communicate. A critical advantage of elicitation is that the flow of sentences in conversation allows for the cumulative development of understanding of the verbal labels offered when a respondent completes a grid: this interaction is typically absent when constructs are supplied. The process of elicitation may be very slow and may act as a constraint on the number of respondents a researcher can work with, particularly if elicitation is to take place on an individual basis. Olson (1979b) remarks that, in his study of science teachers' responses to curricular innovation, it was rare for constructs to emerge quickly in their final form of words: it appears that negotiation of meaning took place before the verbal label (mutually understood by both parties to the elicitation) was committed to paper[81].

The richness of understanding possible in individual elicitation is severely at risk when grid data is elicited simultaneously from a number of individuals. Simultaneous elicitation requires very clear instructions and also the presence of the researcher who can act as a resource to whom any queries may be directed. In my work with

groups I have found that dyadic elicitation presents fewer problems than the standard triadic procedure - and I can testify as a respondent to some of the difficulties of working through a triadic elicitation sequence. My experience leads me to conclude that simultaneous elicitation is not conducive to the communication of rich and subtly-shaded meaning, but where more abstract aspects of grid work (such as the stability of ratings) are the focus of attention, then the simultaneous approach may prove acceptable in terms of time and adequate in terms of the quality of the data gathered.

Are elicited constructs, then, superior to supplied constructs? My answer is a qualified 'yes'. The first qualification is that elicited constructs treated without taking into account a conversational context may give the researcher too much latitude in the construction of interpretations of what the respondent intended to communicate. The second qualification relates to the purpose of the research: if the researcher is seeking particular dimensions of response then there is a case for introducing supplied constructs, but the meanings of these will probably need to be elaborated in order to convey to the respondent what the researcher had in mind. The third qualification also relates to the purpose of the research and recognises the pragmatic value of supplied constructs in studying matters such as the formal properties of grids.

In making this judgment I am giving the greatest weight to the communication of meaning which I regard as crucial to research within a construct theory framework. To supply a construct means that the respondent must construe the construct and use it accordingly, and that the researcher must subsequently construe this usage: the researcher cannot assume that his or her initial construct is reflected

back without distortion. The elicitation of constructs eliminates one phase of potential distortion, but still leaves open the possibility that the researcher will misconstrue what the respondent is saying. Negotiation of understanding, which can take place in the elicitation conversation, can do much to minimise misunderstanding.

The critical questions for both researchers and respondents are 'Do we know their code?'; 'Do we know they know our code?'; and all their more complex interpenetrating derivatives. On the whole, the elicitation of constructs appears to offer the greater prospect of affirmative responses.

6.7 TYPES OF CONSTRUCT

Kelly distinguishes between different types of construct on the pragmatic basis of the way in which they are used. The distinction between permeable and impermeable constructs has already been made in the discussion of the assumptions underlying elicitation, and it is sufficient to recall here that a construct (which may be based on the construing of a small number of elements, typically three) is permeable if it has the capacity to subsume other elements within its range of convenience, whereas an impermeable construct lacks this capacity. It was noted that impermeable constructs present problems in repertory grid work on the grounds of restricted ranges of convenience.

Kelly also identifies pre-emptive, propositional and constellatory, constructs, of which the last appears to be most frequently mentioned in reports of research involving repertory grids. A pre-emptive construct restricts its subsumed elements exclusively to its own realm. Kelly gives the following example: 'Anything which is a ball can be nothing but a ball'[82], which seems to fit more closely the canons of

Aristotelian logic than those of Personal Construct Theory since it seems to be pre-eminently a statement about class inclusion.

In contrast, a propositional construct is a much more speculative affair in that it extends the range of possibilities regarding an element, rather than restricts them[83]. Here Kelly clearly draws upon the 'As if' philosophy of Vaihinger (1924*), allowing an element to be considered in terms of dimensions that might be regarded as unlikely. It would seem that propositional construing can incorporate the figurative meanings of metonymy, metaphor, symbol and myth, whose importance in human discourse is highlighted, for instance, by Barthes's (1973*)analyses. Yet, as Rowe (1978) points out, these are very difficult to capture in the mesh of a repertory grid[84]. This failing has led Rowe to give the repertory grid only a limited role as a signposting exercise which can indicate potentially fruitful avenues of exploration, and her work with depressives has become increasingly dominated by more free-ranging conversations.

A constellatory construct 'fixes the realm membership of its elements'[85]. In other words, if an element is located at one pole of a constellatory construct, its locations on other constructs are (in the mind of the construer) necessarily fixed. Stereotyping[86] is a good example of constellatoriness: for instance, if a person is construed as female this may - in some minds - entail that the person is sensitive, emotional, timid, unpunctual, and so on[87]. Constellatory construing focuses on the implicative relationships between constructs and it is not surprising that it receives little attention in repertory grid work.

Where constellatoriness has been connected with the repertory grid it

has been operationally defined in terms of high loadings on a common factor (which is assumed to be the constellatory construct itself), but the operational definition seems a long way removed from what Kelly intends by the term. Levy (1956), in a study widely cited as being of constellatory and propositional construing, elicited from each of 55 respondents twenty-two constructs regarding acquaintances. He used Kelly's nonparametric factor analysis to determine both the major factors of the data matrix and the loadings of the constructs upon them. Levy defined constellatory constructs as the five loading most heavily on the first factor, on the grounds of their interlinkedness. Propositional constructs were defined as those five which had virtually no loading on any major factor and which appeared on their own as residual factors. The findings from Levy's study are of no interest here because of the attenuated connection between his operational definitions and Kelly's theoretical position. For present purposes it is sufficient to point out that constructs which are highly correlated in statistical terms are not necessarily related in any constellatory sense[88], and that the first factor extracted in analysis (and any factor, for that matter) is an artefact of analytical procedure and not a superordinate construct. The claims for propositional construing can similarly be shown to be unjustified.

This criticism of Levy's study has been made for two reasons. First, it is widely cited in the literature and is therefore implicitly set up as a model for research into constellatory and propositional construing. And second, it implies that the repertory grid can be used to investigate implicative relationships between constructs - an assumption that has already been challenged in Chapter 4.

It is likely to be difficult to recognise instances of pre-emptive, propositional and constellatory construing during the process of grid administration, and even after reflection and/or analysis these may well not come to light. However, some verbal formulations may alert the researcher to the need to probe a little further. Easterby-Smith (1981), following Kelly, offers a convenient summary of types of verbal labels which are better explored further in order to maximise the meaningfulness of communication[89].

(i) Situational constructs, relating to the physical location of people or things, are unlikely to be particularly informative by themselves.

(ii) Impermeable constructs may well occur with particular triads or dyads, and are incapable of subsuming new elements (e.g. 'works on turret lathes - works on capstan lathes').

(iii) Excessively permeable constructs, which are so general as to be unlikely to offer any new insights except in special circumstances (e.g. 'male-female').

(iv) Vague constructs, such as 'alright - not very satisfactory', offer little in themselves and it may prove useful to investigate the discriminative criteria being used.

(v) Constructs generated by role titles (e.g. from successful/unsuccessful person) simply reflect what is in the grid to begin with.

Stringer and Terry (1978) divided the realm of constructs in a different way when they explored the effects of 'objective' (i.e. concrete, physical) and 'abstract' constructs upon the analysis of repertory grid data.. They reasoned that these different categories of construct would relate to different levels of cognitive functioning in the model

proposed by Schroder et al (1967). Thirty-three of their respondents produced more than one objective construct, and in the majority of cases the dispersion of the elements (as determined by principal components analysis) was virtually unchanged when the objective constructs were eliminated[90]. Stringer and Terry conclude from this that the objective constructs were either trivially significant or so integrated as to be redundant. The matter may not be so simple. It appears that the authors have not taken sufficient account of the fact that, in the Schroder et al model, high integrative complexity (which involves abstraction) subsumes the concrete functioning associated with low integrative complexity. The issue of objective and abstract constructs is not 'either/or' but 'yes/and' in this respect.

In the minority of Stringer and Terry's cases, the removal of objective constructs had marked effects, such as the virtual disappearance of one orthogonal component or a change in the relationship between elements and constructs. These effects are somewhat disturbing since they suggest that the simplification of the data matrix may produce a new data structure that is unstable in that it depends on which constructs (and probably elements, for that matter) happen to appear in the original grid. The consequent danger is that the researcher may over-interpret and reify the factors or clusters output by the analytical procedure adopted.

Stringer and Terry acknowledge, following Kelly, that the issue of objective constructs is more complex than might appear at first sight, and that the ostensibly objective content of the constructs might be treated in an abstract manner in use. This is not an easy issue to resolve: Stringer and Terry conclude that it is preferable <u>not</u> to eliminate objective constructs from consideration since they may prove

to be significant. To which I am led to rejoin that one will not determine the significance from the grid matrix but from the way in which the construct meshes ontologically with others, and that for the researcher to explore this is to commit him- or herself to an investigation of a very different kind.

The return to the theoretical link between Kelly and Schroder et al leads me to mention two other typal distinctions made by Kelly regarding constructs, those between core and peripheral constructs and between superordinate and subordinate constructs[91]. The case has already been argued that the repertory grid has little value in investigating the latter, and it can be argued equally strongly that it is likely to prove inappropriate in respect of the former since the grid is unable to offer any criterion by which one might evaluate the degree of centrality of a construct to a person.

The various types of constructs discussed so far in this section are quasi-logical in character, but the nature of elicitation procedures makes typal recognition very difficult. Put another way, the limitations of the repertory grid as far as implication is concerned seem to make it an inappropriate instrument for the investigation of issues implied in the subsumption of constructs under types.

Attempts have been made to categorise constructs in terms of their social meaning (for example, see Landfield, 1971; Glossop et al, 1975), but these seem to have had little success beyond the boundaries of the problems for which they were developed. Fransella (1981) remarks that when Landfield's system was used to classify constructs from British (rather than American) respondents there was only a partial success, for too many constructs were found to be unclassifiable for

the system to be adopted without modification[92]. Given the diverse ways in which language is used and the immense variety of personal experience, it is perhaps unremarkable that categorisation has been accounted unsuccessful. Classification involves a reduction to the common factors perceived by the researcher and, while it may have its uses in well-defined contexts, it is inimical to the full richness of meaning being offered by the respondents unless precautions are taken to reconnect the nomothetic classification with the idiographic data from which it is derived.

6.8 'AS IT WAS IN THE BEGINNING, IS NOW, AND EVER SHALL BE...' - OR SHOULD IT BE 'OUGHT TO BE'?

There is one typological distinction regarding the nature of constructs that seems to be unrecognised in the literature. A construct can be used to indicate the state of an element

(i) as it was at some time in the past;

(ii) as it currently is (or is believed to be);

(iii) as it is predicted to be; and

(iv) as the respondent thinks it ought to be[93].

Something of this distinction appears when a grid includes elements such as 'self as I was' and 'self as I expect to be', but for 'non-self' elements the ontological status of the element is almost always left unspecified. If the centre of interest is the constructs of the respondent and not their relationship with particular elements, the distinction may not matter.

Kelly's writing emphasises the first three of these categories, in that he describes constructs in terms of interpretations - which would seem to relate to (i) and (ii) - and hypotheses, which capture the

prediction of category (iii). But the distinction between the categories is not always easy to make: Thayer (1969*) makes the point neatly when he remarks that statements like 'this is soft' may be the result of operations producing evidence, or may be predictive in the sense that if certain operations are performed they will reveal the object to be soft[94].

The construing of people in a grid context would seem to be predominantly of the 'was' and 'is' types: people speak as they find (or have found). It is obviously possible to build predictive construing into grids dealing with people, though this would appear to require the researcher to provide a future-oriented context at the outset. Elements such as 'self as I would like to be' or 'self five years hence' would cue the respondent strongly in this respect.

Where 'situations' are the elements - as in some of my own work - the basis of the construing is more uncertain, and there is an increased possibility that the 'ought' type of construct will enter the grid[95]. If the focus of the grid is on the aims likely to be achieved in particular teaching situations, for example, the reality may be very different from the rhetoric: without some empirical cross-validation (either through exploration of what the construct means to the construer or through observations of classroom activity) the status of the construct is likely to remain indeterminate.

Looking through my records of the grids completed by the probationary science teachers, it is possible to detect a number of constructs which are ambiguous in this respect. The ambiguity is not lessened by referring to the tape-recordings made of the grid administration procedure, for at that time the possibility of 'ought' construing had

not crossed my mind and so the ambiguity was not recognised. There are, for instance, a number of constructs in these grids which stress opportunity (such as 'strong.... - little opportunity for feedback from the pupils'), and it is not clear whether the respondent was offering a statement of how the class was actually handled, giving out a received wisdom, or making a statement regarding potential yet unfulfilled. Such ambiguities are not the fault of the respondents; the blame for them lies with a researcher who did not recognise them until he reflected on procedure long after the grids were completed.

In Olson's (1980a) work with science teachers involved in the innovative SCISP[96] curriculum he inferred from the grids he collected that there was an overall dimension of 'influence' (high versus low) on pupils. A reasonable - if perhaps stereotypical - approximation to the distinction between high and low influence teaching would be didactic transmission versus facilitation of pupil learning. Olson followed up the grids by probing this dimension in subsequent interviews, and this gave him a more elaborated idea of the ways in which teachers construed influence and its relationship to their role in the classroom. The elaborations that Olson records suggest that the emphasis on classroom role may have kept them within the realm of the 'is', rather than the realms of prediction or 'ought'.

In both Olson's and my work with science teachers the elements were general teaching situations, rather than specific teaching situations or 'critical incidents'. Construing such elements is perhaps rather like construing a list of role titles, such as those listed in Kelly's Repertory Test, without attempting to personify them. Stringer (1979), whilst recognising the difficulty of construing role titles, nevertheless compared the first factor produced by principal components analysis

of 'role-construed' and 'people-construed' grids, the respondents in each group being 40 students. He found a larger first factor arising from the role-construed grids and interpreted this as showing that the respondents were able to construe people in a more differentiated way than they were able to construe roles. However, roles are probably inherently fuzzier to construe than people, enabling the respondent to move around in the fog and identify similarities that, in conditions of clearer visibility, might be seen to be unfounded. It would also seem reasonable to suggest that role titles might elicit 'ought' constructs, since 'is' constructs would not be available unless the respondent chose to personify the role titles in his or her mind. One might further speculate that 'ought' constructs would tend to be more normative, general and diffuse than 'was', 'is' or 'will be ' constructs - and therefore lead to less differentiation in construing. Tightly specified elements such as acquaintances, presented physical objects and items of memorable experience or of actual behaviour could, in contrast, be expected to produce predominantly 'was' and 'is' constructs.

The argument presented in this section is frankly speculative, since it seems that no attention has been given to this particular problem in repertory grid research. The hypothesis that the production of 'was', 'is', 'will be' and 'ought to be' constructs varies with the elements and the context of their construing is fairly amenable to testing if respondents' construing of their constructs is acceptable as validating evidence. General and specific elements could be construed by respondents (due care being taken in balancing possible contaminating variables such as order of presentation of the two types of grid to individual respondents). Subsequent discussion could take place to establish whether the elicited constructs fall into one of

the four categories mentioned or whether they should be consigned to a fifth category labelled 'indeterminate'.

In summing up this section it is perhaps sufficient to suggest that there are grounds for suspecting that the greater the generality of the elements being construed, the greater the likelihood that 'ought' constructs will be produced, the rhetoric of which may not reflect the reality being investigated.

7 Bipolarity ... or not ?

The factual language dissects and disintegrates experience into categories and oppositions that cannot be resolved. It is the language of either/or...
By contrast, the language of myth and poetry is integrative, for the language of the image is organic language.

A. Watts
1978*: 15-16.

7.1 INTRODUCTION

Kelly's opinion that all construing is, at root, dichotomous in character[1] was strongly challenged in Chapter 3, in which arguments were advanced for a 'mixed economy' of construing. This approach to construing involves the use of traditional concepts for nouns and verbs, whilst allowing that dichotomy might play a considerable (but not exclusive) part in adjectival and adverbial constructs. In most repertory grid work the practice is for noun-type elements to be construed in adjectival terms, although dependency grids[2] involve a matrix of problem events and people who might help, and Ryle and Lunghi's (1970) dyad grid presents interpersonal relationships as elements and requires the respondent to provide verb constructs. In this chapter the focus is on the problems of bipolarity in construing, and discussion concentrates on the most common form of grid in which the constructs are of an adjectival type.

In his original formulation of the repertory grid Kelly assumes that all the elements fall within the range of convenience of the constructs (i.e. there are no 'not-applicable' responses) and can be assigned to one or other of the construct poles. This is a strong form of bipolarity and, as Humphreys (1973) observes, this requires that the poles be mutually exclusive and that together the two poles cater for all possibilities of construing along that dimension[3]. Although Kelly asks his respondents to construe dichotomously, he is aware of the possibility that the 'space' between the poles of dichotomous constructs can be treated in terms of categorical (digital) or continuous (analog) scales[4].

The actual practice of dichotomous construing tends to conceal any difficulties that the respondent might have regarding the allocation

of elements to construct poles. Once the decisions are made and acted upon, the route by which those decisions were reached may well be of little interest to the researcher - the situation is analogous to the objective test item in mathematics which is scored either right or wrong, with no credit being given for good working even though an incorrect answer is due to an arithmetical slip in the final stage. Looking retrospectively from the outcome, the hypothesis of strict bipolarity is seductive: the choice had to be one thing or the other. The respondent's wrestling with the problem of dichotomously construing sweet and sour pork on the dimension 'sweet-sour' becomes practically insignificant once the element allocation has been made.

It is easy to produce examples which challenge the practice of strictly dichotomous construing, and the difficulty of plumping for one side or another has contributed to the development of methods of arranging elements on constructs, such as ranking or rating, which allow for intermediate shadings of opinion. The problems easily visible in respect of dichotomous construing do not go away, however: they are merely transformed into more subtle versions by the changes in procedure. Many of the points discussed in this chapter apply to ranked grids, but the emphasis is on rated grids since these appear to have become the most widely used form.

Rated grids make use of bipolarity in a weaker form than Kelly's proposal of dichotomy, in that each pole subsumes a range of grades of membership. This is clearly shown in the semantic differential where the categories are graded 'extremely', 'quite' or 'slightly' in respect of the quality denoted by each pole. The principle of mutual exclusivity is nevertheless presumed still to hold, since on that instrument an element can belong to one pole or the other, or be

located at the mid-point of the construct. It is worth exploring the semantic differential a little further since it evidences problems which are transferable to the context of repertory grid work.

In presenting their account of the semantic differential Osgood et al (1957) postulate a semantic space whose dimensionality is unknown, but whose underlying metric is Euclidean. This semantic space is composed of a large number of semantic scales, each being defined by a pair of polar (i.e. opposite in meaning) adjectives assumed to be related by a straight line function passing through the origin[5] which is conceptualised as the point of meaninglessness[6].

Any semantic differential bipolar scale P-Q can be represented as in Figure 7.1 in which the mid-point rating of 4 implies 'neither P nor Q' and ratings either side imply increasing 'P-ness' or 'Q-ness' as the extremes of the scale are approached. There are other methods of representing the semantic differential bipolarity in such a way as to preserve the postulated collinearity, but these are less helpful when it comes to discussing the vagaries of rating scales in repertory grids in Chapter 8.

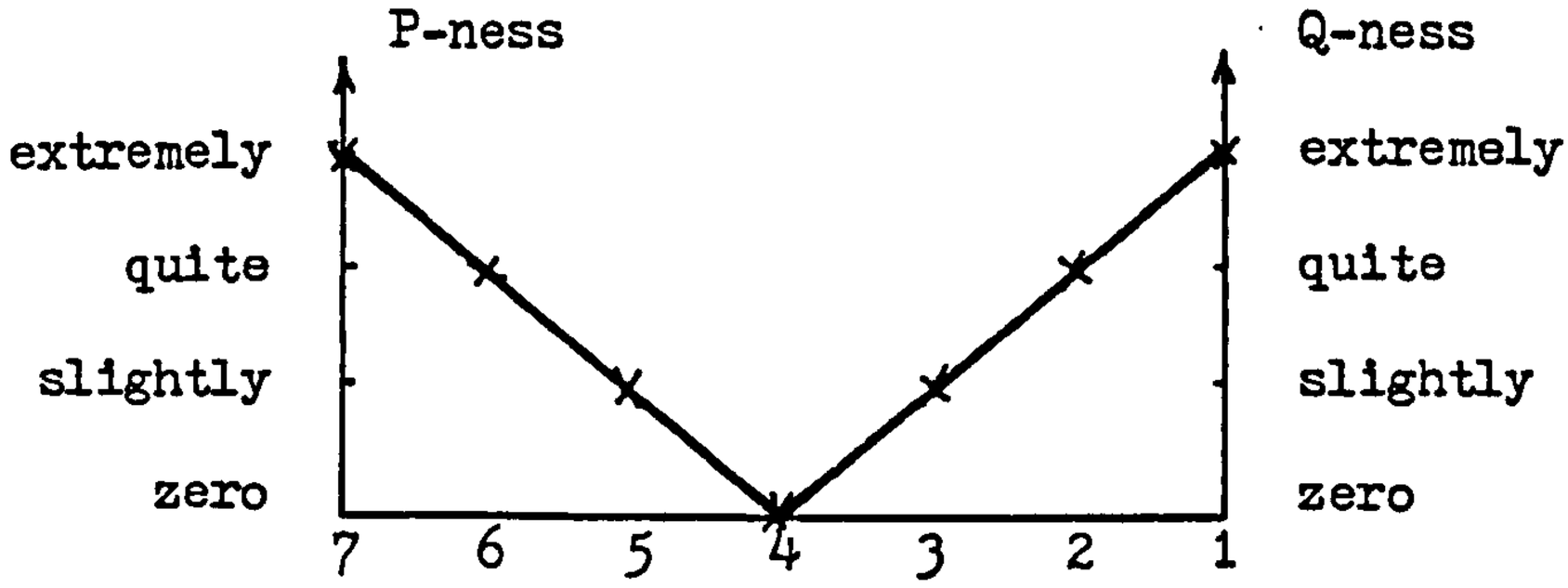


Figure 7.1 A representation of the semantic differential scale P-Q.

Unfortunately Figure 7.1 is not strictly accurate as far as the semantic differential is concerned, since Osgood et al undermine the purity of their theoretical position when they allow the mid-point of the scale to reflect the equal possession of <u>both</u> P and Q. In such circumstances the assumption of mutual exclusivity ceases to hold, and their conceptualisation of bipolarity is irredeemably compromised[7]. The implications of this for the use and analysis of scales are difficult to assess, in that the pattern of non-linearity thus injected is likely to depend substantially upon the particular scale being used and on the concept being rated.

The majority - perhaps even all - of the dimensions suggested by Osgood et al would probably be construed as 'logical' (in Resnick and Landfield's, 1961, sense) rather than 'peculiar' oppositions: indeed, in later work, Osgood et al (1975) chose to leave out qualifiers lacking cross-culturally agreed opposites[8]. In repertory grid work, where the respondent typically supplies the dimensions, 'peculiar' constructs are often elicited - even in response to the 'opposition' approach suggested by Epting et al (1971). The functional antonymy of 'peculiar' constructs is likely to be more questionable than that of 'logical' constructs, thus providing the researchers with a number of problems regarding analysis and interpretation. In this Chapter I concentrate upon the extent to which constructs appearing in grids may be functionally antonymic. The emphasis is on the semantic, problems associated with the location of elements on constructs being deferred to Chapter 8.

7.2 <u>BIPOLARITY: SOME EMPIRICAL EVIDENCE</u>

Kelly indicates at various points in his writing that constructs are

assumed to be dichotomous in form[9], but produces no empirical evidence other than a study by Lyle (1953) to indicate whether his assumption is justified[10]. Lyle's method involved asking students to match a list of words to eight 'target' words representative of four bipolar dimensions, the list of words having been drawn from a thesaurus and found by a prior panel of student judges to be acceptably synonymous to the target words. Lyle recorded the accuracy with which students were able to recognise the target words, given the synonyms, and then used a modified Thurstone factor analysis to produce five factors. The largest of these factors (accounting for 22 per cent of the total variance) appears to be related to general verbal ability: however, Kelly concentrates his attention upon the other four (each of which accounted for approximately ten per cent of the total variance) which showed that, for each bipolar dimension, the two ends both loaded heavily and exclusively on a single factor. From these results Kelly draws the conclusion that, if a person tends to make errors in respect of both separated poles of a construct, this is evidence to suggest that constructs are dichotomous[11].

Reanalysis of Lyle's data using the factor analysis program FTAN from the PMMD suite (Youngman, 1976) produced an oblique solution of excellent factorial simplicity (index = 0.96) with four highly inter-correlated factors. Each of Lyle's original bipolar dimensions loaded heavily and exclusively on one of these four factors showing (more clearly than Lyle was able to do with the analytical routines available at that time) that the word-recognition accuracy scores tended to 'pair up' on the bipolar dimensions used[12].

However, to infer from these results that construing is dichotomous is to over-interpret the analysis. First, accuracy of recognition relates

very much to general verbal ability (shown by Lyle's Factor V and the high intercorrelations in the FTAN oblique factor analysis), and the tendency for accuracy (or error) scores to pair up is a comparatively minor feature. Second, to make errors in recognising 'cheerful' from near-synonyms, and to make errors to a similar extent with respect to 'sad', does not necessarily show that these are bipolar. It can be argued that the four dimensions used by Lyle are representative of different facets of life-experience and, if the argument be granted, each respondent may have been differentially aware of the semantic ramifications of these facets. To summarise: what Lyle may have found are relatively minor perturbations (in particular semantic directions) in the verbal ability of his respondents. Such would seem a more parsimonious interpretation than to claim substantial support for the dichotomous nature of constructs.

Since the development of both repertory grid and semantic differential there have been a number of studies dealing more specifically with the issue of bipolarity. Some of this evidence is reviewed in the following paragraphs.

In reanalysing some of Asch's work on impression formation Wishner (1960) found that dictionary antonyms did not necessarily correspond to psychological opposites[13], and he proposed that antonym pairs be split to form separate (unipolar) scales. Mordkoff (1963, 1965) investigated the distinction between nominal and functional antonymy by splitting semantic differential adjectival pairs and using these as concepts on semantic differential scales, the hypothesis being that the ratings of the two members of each concept-pair would be symmetrical on the scales provided[14]. Taking the two experiments together, Mordkoff found that

fourteen of 28 concept pairs lacked functional antonymy on the dimensions used: for example, both 'masculine' and 'feminine' were rated by his student respondents as 'good', 'beautiful',,'pleasant' and 'clean', whereas the functional antonymy of 'masculine-feminine' would have precluded this. Mordkoff's findings anticipated Constantinople's (1973) point that there are serious doubts regarding the existence of a psychological analogue of sexual bimorphism and hence regarding tests that make the assumption that masculinity and femininity are mutually exclusive. Later, Spence et al (1975) found that, whilst males and females tended to rate themselves according to conventional sex-role stereotypes, the relationship between masculinity and femininity was far removed from simple bipolarity.

Green and Goldfried (1965) drew attention to the fact that respondents to the semantic differential were forced to accept bipolarity, whether or not it suited the way they wanted to respond. Hence the semantic differential, with its insistence on bipolarity, could not be used to investigate the existence of bipolarity in semantic space. They asked 251 undergraduates to rate concepts on a number of unipolar dimensions chosen to reflect oppositeness. Where the oppositeness of unipolar labels was 'obvious' it appeared that bipolarity applied in respect of certain concepts but not others. The ratings for the concept 'sin', for example, on the two unipolar dimensions 'pleasant' and 'unpleasant' correlated a remarkable + 0.39: where this type of effect occurs, the authors suggest that a different subset of attributes of the stimulus is being used for the opposing unipolar dimensions. Picking up the findings relating to 'sin', I would hazard the guess that someone engaged, say, in an extramarital affair might construe adultery as stimulating, exciting, satisfying (and therefore rate it as pleasant), and as dishonest, deceitful, immoral (and therefore rate it as

unpleasant). It may not be so much that 'pleasant-unpleasant' is not bipolar; rather, it may be that the two polar terms each represent such a loose aggregation of meanings and connotations that it is possible both to construe 'sin' (itself an abstract concept) in terms of a range of subsumed types of sin and to construe 'pleasant-unpleasant' in terms of a range of subordinate implications. This would give a notional and indeterminate matrix of 'types of sin' X subordinate constructs within which the respondent is able to move during the task of rating the superordinate element 'sin' on the superordinate construct 'pleasant-unpleasant'. Constructs of a general nature such as 'good-bad' and 'pleasant-unpleasant' would seem much more open to interpretation in terms of varying subordinate implications than more concrete constructs such as 'sharp-blunt'.

There is a problem of logic with Green and Goldfried's study in that there is the need to import an assumptive step to move from the unipolarity that they use to the bipolarity that they are challenging. Heise (1969), reviewing the literature on the semantic differential (including Green and Goldfried's study), concluded that on the whole bipolarity was justified for most semantic differential scales. However, Mann et al (1979) take a less optimistic view of the situation, suggesting that the assumption of bipolarity is only partly supported.

Semantic differential research typically averages the findings from a number of respondents to produce nomothetic findings in which individual variations may be seen as insignificant, or even as error variance. Much work with repertory grids highlights the construing of the individual, and this demands an analysis of bipolarity as seen from an idiographic rather than a nomothetic perspective.

Resnick and Landfield (1961) noticed how many constructs seemed to embody 'peculiar', idiosyncratic ideas regarding contrast, being neither related through some commonality nor obviously antithetical. Examples which they found included 'clever - anti-social' and 'artistic - happy-go-lucky'. On the other hand, some constructs appeared genuinely antonymic: these were termed 'logical' constructs. The authors were interested in investigating whether there was any functional difference between the two types of construct.

They selected one 'logical' and one 'peculiar' construct from grids elicited from 41 undergraduates and asked them to indicate which ten of a list of 40 adjective pair antonyms (i.e. bipolar constructs) would best describe each of the four elicited poles presented separately. The criterion of functional antonymy was operationalised in terms of the extent to which the two construct poles (presented separately) were similar in terms of the dimensions chosen as most appropriate to describe them. 'Clever', for instance, would be deemed to be functionally antonymic to 'anti-social' if there was a substantial similarity in the lists of constructs seen as most appropriate to them. Whilst significant functional antonymy was found for <u>both</u> 'logical' and 'peculiar' constructs, the effect was significantly higher for the former. It would seem reasonable to conclude, on this basis, the more 'logical' the opposition with a construct, the greater the likelihood of functional antonymy[15].

Mair (1967) studied the incidence of bipolarity in the split-half repertory grid[16]. He supplied four adjectives to ten medical students, all of whom had to supply their own contrasts. Each of these oppositional contrasts was presented, together with six supplied constructs of the

'un/like self' type (not of interest here), in unipolar form: the students were asked to assign ten acquaintances from an original list of 20 to each of the fourteen poles. Bipolarity was investigated in respect of the specific oppositions used by each student, perfect bipolarity being inferred when oppositional poles attracted completely different sets of ten acquaintances, giving no 'overlap' and hence a matching score of zero out of a possible 20. However, Mair found that only 22 per cent of matching scores were less than or equal to 2: in the majority of cases there was a substantial overlap. He concluded that, while people may give apparently clear-cut verbal opposites, they may not use the constructs consistently. Mair seems not to have given much weight to the students' remarks that the elements were complex and open to many interpretations (cf the discussion of 'sin', p. 223-224); it would seem that both elements and constructs may have been inconsistently construed during the procedure, and may thus have contributed to the finding of a lesser degree of bipolarity than might have been expected.

Epting et al (1971), whose comparison of elicitation procedures was cited in respect of the 'opposition' and 'difference' poles, used a similar strategy to Mair. They elicited twelve constructs from each of 43 female undergraduates using Kelly's repertory grid procedure, 20 students using the 'opposite' method and 23 the 'difference'. After five days each student was asked to complete a 22 element X 24 unipolar construct grid using the split-half allocation procedure, the constructs being the 24 unipolar labels that the student had provided herself. As with Mair's work, bipolarity would lead to the expectation of a low number of matches. Taking the criterion of six or fewer matches as indicating bipolarity, the authors found that 73 per cent of constructs elicited by the 'opposite' method met the criterion, compared with 57 per cent of those arising from

the 'difference' method. Four per cent of the 'opposite' and six per cent of the 'difference' constructs actually showed a positive overlap, the remainder being in the band of eight to fourteen matches. The results are summarised in Figure 7.2.

METHOD OF ELICITATION	FREQUENCY OF CONSTRUCTS WITH MATCHING SCORES m $m \leq 6$	$8 \leq m \leq 14$	$16 \leq m$	TOTAL
'Difference'	156	105	15	276
'Opposite'	175	55	10	250

Figure 7.2 Summary of findings from Epting et al (1971)

Epting et al concluded that the 'opposite' method produced a greater degree of bipolarity in the constructs elicited, but that both methods left room for improvement in this respect. They speculated that the 'difference' method was more likely to cause the respondent to shift to the emergent pole of another construct, rather than provide a true oppositional contrast to the likeness pole first elicited. One can easily see how the 'difference' method could produce constructs that might, following Resnick and Landfield's (1961) nomenclature, be termed 'peculiar'.

Honess (1978), working with children, found a much weaker bipolarity in ranked grids. Only 29 per cent of 184 constructs from 46 children reached his criterion of bipolarity, a correlation coefficient of ≤ -0.83 ($p < 0.01$) whilst one per cent reached the agreement criterion of $\geq +0.83$. In this case the constructs had been produced by asking the children to supply the opposite poles to four supplied poles that had been selected from essays previously written on liked and disliked same-sex peers. The differences in sample and methodology make it

difficult to compare Honess's results with those of Mair and of Epting et al.

An early study by Ross and Levy (1960) tackled the problem of functional antonymy in a rather different way. They provided their respondents with antonymic terms, such as 'most beautiful' and 'most ugly', and asked them to arrange nine coloured cards (five of one colour, four of another) in patterns appropriate to the provided term. Judging by the patterns produced, Ross and Levy did not find complete antonymy, but it is not clear from their work whether - among other possible explanations - the relatively restricted range of possibilities available within a 3 x 3 card layout gave the respondents insufficient room for cognitive manoeuvre, or whether beauty and ugliness might have been regarded as separate concepts with their own specific meaning - complexes which need not necessarily be fully antonymic[17].

The various researches reported above have produced findings which tend to parallel those of research using the semantic differential. Bipolarity appears to exist in respect of many constructs, but is not guaranteed whichever method of elicitation is adopted. Some possible reasons have been advanced in passing why constructs are less bipolar than one might naively expect, and the final sections of this chapter deal with these in more detail. Though the points below are made with elicited constructs primarily in mind, many are equally applicable to supplied constructs.

7.3 'BENT' CONSTRUCTS

Contrary to Kelly's assumption[18], the construct may be made up from the poles of two different constructs whose contrasts are not articulated. The evidence reviewed above suggests that, especially where the

'difference' method of elicitation is used, constructs may be of the form 'X-Y' where X and Y represent different attributes but have no superordinate rationale to link them in opposition. Such very 'peculiar' constructs might easily be 'disconnected' at their mid-points: Shaw (1980a), for instance, is aware of the problem, observing that two elements rated at the mid-point of a scale could actually be more different than two rated at opposite extremes[19]. Such a situation could arise if the two mid-point ratings reflected different unarticulated contrasts to the stated poles X and Y. Constructs of this sort could be said to be so bent that they have snapped in the middle, even though they might exhibit a considerable measure of functional antonymy.

But the problem of 'bentness' can exist even where there is a super-ordinate relationship. Consider the superordinate construct 'pleasant-unpleasant': this may be verbalised as such, or in terms of any oppositional pairing from its subordinate implications[20]. Further, it is possible that the respondent may select any opposite to the first-identified pole from among his or her repertoire of plausible contrasts, including the superordinate term itself. In the hypothetical example given in Figure 7.3 the respondent has chosen 'repellent' as an opposite to 'friendly', despite the availability of the more 'logical' opposite 'hostile'.

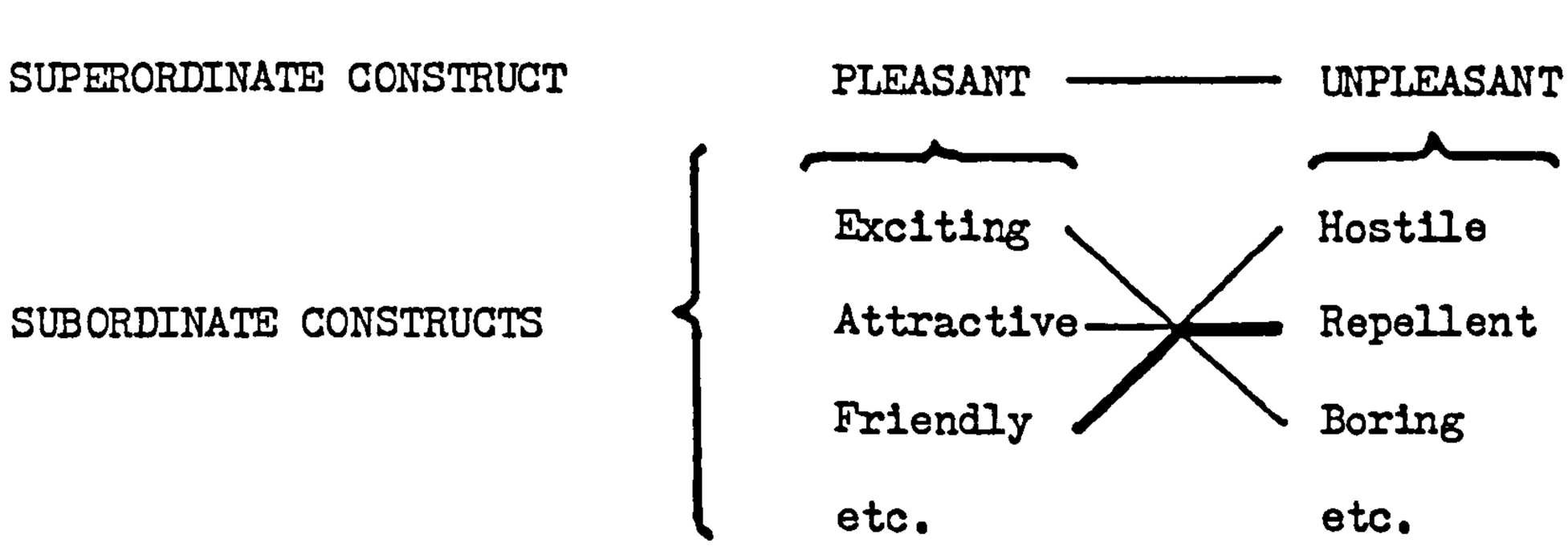


Figure 7.3 An illustration of the possibility of obtaining 'bent' constructs - in this case, 'friendly - repellent'.

Elicitation using the 'opposite' approach is likely to increase the chances of obtaining 'straight' constructs, but the evidence of Epting et al (1971) and Honess (1978) indicates that this should not be taken for granted.

7.4 PSEUDO-BIPOLARITY

Many constructs in grids are bipolar only in the sense that the polarity reflects some gradation of magnitude of an attribute. In its more extreme form this could be represented in terms of a dimension 'X-not X'; less strongly, the situation could be characterised as 'high X - low X'. In neither of these cases is the full oppositional contrast of the Kellian construct present, all reference being made to only one attribute - hence my labelling of these types of constructs as pseudo-bipolar.

A number of writers have pointed to the difficulty that sometimes exists in trying to find a true opposite to a pole (for example, 'jealous' and 'conscientious') without resorting to simple negation[21]. In my experience there is a greater tendency than linguistic usage would suggest for respondents to offer the 'not X' pole as a contrast. It may be that, with an eye to scientific detachment regarding the possible intrusion of my own thinking into the elicitation procedure, I did not press my respondents hard enough when there might have been 'better' oppositions available to them. There is no absolute solution to problems of this sort: the researcher has to make a decision on the approach to be adopted and to be prepared to justify it in reporting the findings.

The point about the 'X-not X' construct is that it runs against Kelly's philosophical position, seeming to have more in common with the

principles of set inclusion than with constructive alternativism. It refers to a <u>concept</u> in the traditional sense, with all that implies about the definition of 'X-ness' and the vagueness of 'not X-ness'. The 'not X' pole is open to a large and indeterminate number of meanings, none of which may be articulated by the respondent. Asking that respondents provide their own opposites to a given polar label provides evidence enough for the potential polysemy of unarticulated contrasts. As a case in point, Carter et al (1968) asked 135 teachers on a summer school to supply opposites to a number of unipolar labels taken from semantic differential scales, among them being 'kind' and 'rugged'. Figure 7.4 indicates the range of contrasts offered: whilst it is doubtful that any individual would be able easily to produce all of those collected by Carter et al, it is possible that, had the respondents used the provided labels as unipolar anchors to scales or had they treated them in 'X-not X' terms, some of the possible options regarding the unanchored ends might have been flickering in their minds. (In passing, and with the use of supplied constructs in mind, it is interesting to note that the antonyms given by Osgood et al (1957) were not the most frequently chosen contrasts).

<u>KIND</u>		<u>RUGGED</u>	
Unkind	40	Smooth	44
Mean	39	Soft	17
Cruel*	28	Easy	11
(21 others)	28	Delicate*	6
		(35 others)	57

Figure 7.4 The range of contrasts given to 'kind' and 'rugged' (Carter et al, 1968).

Note: the asterisks indicate the contrasts given for the semantic differential.

The problem of vagueness remains even when the construct is presented in the 'high X'-'low X' form, and would seem to be particularly acute at

the 'low X' end since this does not provide a focusing contradistinction to the 'high X' end. Whilst 'highly jealous' or 'highly conscientious' might be intersubjectively understood in a general (but unexamined) sense, terms like 'barely jealous' or 'barely conscientious' are much vaguer in their meaning. For the researcher to understand what a respondent means by such terms it is necessary that he or she elicit the implicative constructions of the terms being used. The repertory grid itself does not address itself to tasks of this sort.

Whilst one might be able to <u>rank</u> elements meaningfully on dimensions of this sort, and thereby avoid the issue of the strength of the polar meanings involved, to use ratings on such dimensions seems to be claiming rather more in psychometric terms than the situation warrants.

In summary, then, the extent to which 'X - not X' and 'high X - low X' constructs are bipolar cannot be convincingly assessed from a consideration of the constructs alone. It may be the case that the respondent has in mind a much more specific opposition - or series of oppositions - than is represented in the verbal label made manifest in the grid. It seems best on balance to treat such constructs as unipolar scales, with all this implies for the allocation of elements upon them.

7.5 <u>CONSTRUCT COMPLEXITY</u>

Some constructs, which are at first glance bipolar, may be inherently more complexly structured: if bipolarity is insisted upon, it may distort what the respondent wishes to communicate. The problem does not arise in Kelly's original treatment of constructs as mutually exclusive yet related opposites subsumed under a superordinate concept or construct pole, but the conversion of dichotomy to a form of scale

is not necessarily associated with a smooth gradation of meaning between one pole and its opposite.

Eliciting constructs from the probationary science teachers, I found a number who discriminated within a triad of teaching activities according to a dimension best summarised as 'visual...- verbal activity'. A problem emerged when an element like 'pupil practical exercises' was to be located on this dimension: it was construed by some as both visual and verbal, whilst others wanted to categorise it in terms of manual/physical skills. The former tended to resolve the problem in terms of the visual aspects of practical work, apparently on the grounds that it contrasted with other teaching situations which were predominantly verbal (by which 'oral' was typically meant). This seems to be a distortion in that, during conversation, the emphasis was on the possession of both visual and verbal components - in which case one would have expected ratings near the mid-point. Where the emphasis was on manual skills the rating used was close to the mid-point since this reflected a neutral position on the 'visual-verbal' dimension.

The difficulties found in using the presumed dimension suggest that the presumption of a simple bipolarity may not be justified, and that the construct might be either more complex or more simple according to the level of analysis employed. It will be recalled that, in discussing the Dichotomy Corollary, the argument was put forward for multiple oppositions based on the notion of contrastive sets. It is possible to speculate that, as far as sensory perception is concerned, teaching acts may be construed in terms of one dominant sensory experience which, in the case of practical work, could be expected to be drawn from vision, hearing and touch. Simplistically, the construct could be seen as a contrastive set which might look like

Predominantly visual	Predominantly tactile	(other senses)	Predominantly auditory

In practice, the contrastive set boundaries would probably be fuzzy and the categories could be expected to overlap.

Other similar examples appear in respect of the construing of teaching situations though the problem tends to be masked by the requirements of a bipolar rating scale, and the implicit categorisation underpinning the construct may not be recognised. For instance, a construct like 'pupils working on their own - teacher leading pupils' may subsume at least one category in which co-operation between teacher and pupil is important:

pupils working on their own	pupils and teacher working co-operatively	teacher leading pupils

And 'pupils working as individuals - pupils working as a class' is similar in that it has an implicit intermediate category referring to the pupils working in small groups:

pupils working as individuals	pupils working in small groups	pupils working as a class

In the second of the two examples there is a stepwise gradation in the size of the group involved, but there are difficulties in rating teaching situations on a seven-point scale anchored by the left and right hand categories: despite the gradation, the activities being construed are likely to be qualitatively different and close to being mutually exclusive.

It can be argued that to separate the categories in such a way is to

treat the matter simplistically, and the argument can be supported by pointing to 'pupil practical exercises' which might involve periods of class instruction and debriefing, both before and after group and/or individual work. Such an argument can but be conceded: however, in making it, its proponent is subdividing the event to which reference is being made, for practical work is being construed in terms of its component parts. There is nothing at all objectionable about doing this, but it presses the construing from a macro, gestalt-like level to one that might be more appropriately labelled 'micro'. The complexity of multiple oppositions becomes simplified beyond bipolarity to something more in keeping with the traditional notion of concept.

In such circumstances the categories identified can be treated as a checklist of possibilities, those present in respect of a particular activity being coded 1 in binary terms, absence being denoted by a zero or blank[22]. Each activity will have its own binary code covering the categories being used: 'discussion with class' and 'pupil practical exercises' would be coded according to the size of the group in which the pupils work, as shown in Figure 7.5

Activity	Pupils working as		
	individuals	groups	class
Discussion with class	1	0	1
Pupil practical exercises*	0	1	1
	1	0	1

Figure 7.5 Exemplification of binary coding of teaching activities in respect of a 'contrastive set' type of construct.

*Note: Pupil practical exercises are shown as coded in two senses depending on whether they involve individual or group activity. It is possible, of course, that they demand both.

It might be inferred from Figure 7.5 that the principle of mutual exclusivity associated with contrastive sets has been abandoned. This

is not the case. The fact that 'discussion with class' can involve both the class as a whole and interaction with pupils as individuals merely recognises the subdivision of the element into component 'time-slices'. The serious question that is raised here is the size of the time-slice that should be construed in order to maximise the respondent's communication of meaning to the researcher. It is not appropriate to attempt an answer at this point: hints towards an answer appear in Chapter 8, and a fuller response to this difficult methodological question is attempted in Chapter 13.

'Contrastive set' types of construct are not the only kind of construct that is more complex than Kellian bipolarity would suggest. It is possible for a person to exhibit both polar characteristics simultaneously (e.g. 'logical-intuitive') or an activity to require both intellectual and manual skills together ('thinking-doing').[23] Gaines and Shaw (1981), for instance, recognise that an element may have memberships of the X and Y poles of the construct X-Y ranging from full membership of both to zero membership of both with any intermediate combination[24]. This implies that the construct is effectively 'split in half' and the polar memberships of each element are assessed separately. Roberts (1981) makes this split explicit in her elicitation of a fuzzy grid[25]. The procedure gets over the severe difficulties associated with the characterisation of the mid-point of a bipolar scale (this was mentioned on page 220 in respect of the semantic differential, and is explored in greater detail in Chapter 8, page 293f), but introduces the semantic uncertainty of the unlabelled end of a unipolar scale. If there were some way of 'splitting' a bipolar construct without losing the directionality embedded in X-Y contrasts, this might be of considerable value to users of repertory grids - and of other scales, for that matter. Gaines and Shaw (1981)

outline the nature of the computer program ENTAIL which is capable of processing data of this type, and they indicate that it overcomes the problems found with missing data and multiple meanings of the mid-points of dimensions[26]. This program is, however, founded on formal logic and it is a moot point whether the strictness of its principles is applicable to the relative looseness of much natural-language construing.

Some constructs seem to rest on the possession of both digital and analog characteristics. 'Shut-open' and 'trustworth-untrustworthy' can be construed as having a semantic metric like the'off-on and volume' knob on a radio: the left hand pole describes a particular state, but once the element is 'moved' from this position it is on a graded oppositional scale. A door is shut or not and, if not, there is a range of degrees of openness; a person is trustworthy or not and, if not, there is a range of untrustworthiness available on which the person can be located[27].

Some constructs seem to be underpinned by an implicit behaviourally-anchored scale. A construct like 'innumerate - can solve complex algebraic problems' could, as far as rating is concerned, be sub-divided into a series of categories representative of increasing mathematical ability[28]:

Innumerate	Can perform basic arithmetical operations +-x÷ on integers	Can perform basic arithmetical operations +-x÷ on decimals and fractions	.....	Can solve complex algebraic problems

As presented, this example conceals a difficulty - the uncertainty that behaviours or skills nest sequentially within each other to

produce a graded scale[29]. For instance, different pupils may learn mathematical skills in different sequences: one may become proficient in tackling fractions and decimals before basic algebraic problems, whereas for another the reverse may be the case.

As regards problem-solving, a self-sufficient person is clearly superior to one who has to be told both what the problem is and the appropriate method for its solution. But which intermediate position is superior - being able to identify the problem but not the method of solving it, or being able to solve a problem once it has been identified? A reductionist analysis would press the researcher towards the binary characterisation of the presence or absence of each skill or behaviour in question.

The arguments presented in this section are suggestive rather than conclusive. Empirical evidence is presented in Chapter 8 that bipolarity may be a more complex matter than a superficial analysis might suggest. There are implications both for repertory grid methodology and other methodological approaches falling within the ambit of Personal Construct Theory: these, together with the implications for analysis are discussed in Chapter 13.

7.6 INHERENT VAGUENESS

Many writers have pointed out that concepts tend to be imprecise, apart from those used in highly specialised and formalised languages such as those of the sciences and technology. T.S.Eliot sums up the situation with considerable economy in 'Burnt Norton'.

> 'Words strain
> Crack and sometimes break, under the burden,
> Under the tension, slip, slide, perish,
> Decay with imprecision, will not stay in place,
> Will not stay still.'[30].

In repertory grid work using natural language, polysemy and multiple connotation abound, indeterminately. The construct labels are overdetermined. The terms used to identify the discriminations required by repertory grid methodology tend to be short and unqualified, and provide fertile soil for ambiguity to flourish. The respondent may set out with a particular discrimination in mind but, during the location of the elements on the construct, may shift the meaning of the terms used in order to accommodate new elements which were not part of the original discrimination procedure. In other words, like the often-found scale X concept interaction in the semantic differential, there may be an element X construct interaction in the repertory grid - and, as Mair (1967) suggested, people may sequentially use a variety of contrasts. Duck (1973) provided some evidence to support this argument when he found, in his work on friendship formation, that his respondents experienced difficulty in locating elements on constructs since the similarity pole required different contrasts when the construct was used in connection with different people[31]. In Kellian terms, it seems that a different construct might have been used for each individual person in the grid, yet the verbal label may have remained unaltered.

A further case in point is Jane's grid (Shaw, 1980a). In this grid Jane articulated in some detail a number of construct poles to which elements were allocated in the original Kellian dichotomous manner. Jane's grid is of particular interest in that it highlights the problem of compromising between specificity of the construct meaning on one hand, and general applicability(or'permeability') on the other; a problem which seems to me often to remain unnoticed or unacknowledged in repertory grid research.

Any of the eight constructs elicited from Jane would have served to illustrate how construct poles are very frequently components of larger meaning-complexes: construct 6 (see Figure 7.8) is taken as an example, the argument in respect of which can be generalised to the remaining seven.

Enjoy intellectual discussion Difficult to understand initially City livers Seek challenges* Insecure backgrounds	versus	Affectionate Humble Sensitive Musical Involved with those immediately around Compassionate* Philosophical

Figure 7.6 Construct 6 from Jane's grid (Shaw, 1980 a:79).
Note: Jane summarised the construct in terms of the asterisked labels.

From the grid presented by Shaw it appears that Jane was able to assign eight people to the left hand pole (summarised as 'seek challenges') and four to the right hand pole ('compassionate'). It seems extremely improbable that all the people construed by Jane fell unambiguously into one or other of the two categories, for the meaning-complexes represented by the polar labels are very broad and are probably not mutually exclusive. It is difficult to see what analysis of this grid would reveal about Jane's construing of each of the people-elements as an individual, since it can be persuasively argued that they have all been force-fitted into a mould appropriate for perhaps only the three involved in the triadic elicitation[32]. To have used a rating scale would have done nothing to resolve the difficulties posed by the manifest portmanteauism of the construct poles and the allocation of elements to such wide-ranging dimensions.

If, as I am led to conclude, Jane's verbal labels are attempts to

specify the characteristics of individuals or particular pairs of people, the nomothetic use of such comparatively precise labels is likely to be invalid for the remainder of the element sample. The vagueness of the typical repertory grid'sverbal labels has been exchanged for inaccuracy in construing the majority of the element sample.

The vagueness of verbal labels seems to be an approximate function of the distance of the construct from the realm of sense-data. Physical attributes such as those involving the units of mass, length and time can be expressed with considerable precision, and are very likely to obtain intersubjective agreement. Some functional or descriptive constructs such as 'compulsory-voluntary' which, whilst not referring to physical attributes, have the force of rule to underpin them and are also precise in that respect.

Ogden (1967*)made a broad distinction between two classes of adjectives - those related to sensation and those relating to the functional or descriptive features of events[33]. He pointed out that the latter class could be treated as analogous to 'sensation' adjectives, the result being a variety of vague contrasts. Moscovici (1981) seems to be standing in the same area of territory when he makes the point that the image associated with an idea tends to become reified and endowed with a quasi-physical, quasi-autonomous existence[34]. In other words, <u>concepts</u> slip towards becoming <u>percepts</u>: psychological states, for example, can be 'observed' and 'recognised' as if they were physical attributes. In classrooms, children may be identified and described to others as 'extrovert' or 'introvert': the currency of such exchanges very often passes unquestioned, no-one calling at the bank to cash the notes. It is only when the idiosyncratic

meanings of these terms are explored that the quasi-physical construct is revealed in all its imprecision.

There is an implicit claim in respect of the repertory grid that, as a systematic idiographic methodology, it is appropriate for the rigorous investigation of personal construct systems. Yet if the grid is used in respect of subject areas in which vague constructs are likely to emerge - the area of personality is but one example - it would appear that it cannot, on its own, sustain the implicit claim of methodological rigour.

7.7 SUMMARY

Kelly's assumption that construing is fundamentally dichotomous, involving mutual exclusivity of construct poles under some superordinate construct, is a very strong statement of bipolarity which is open to challenge on a number of grounds even when it is weakened slightly (as in the semantic differential) by allowing polar membership of elements to be graded. Although the empirical evidence drawn from research into both semantic differential and repertory grid is not entirely satisfactory, it is probably strong enough, when taken in conjunction with a conceptual analysis of bipolar construing, to suggest that many constructs do not fit the semantic differential's model of functional antonymy. Membership of construct poles is shown to be neither necessarily mutually exclusive nor even complementary: a variety of possible construct types appears to exist, ranging from pseudo-bipolar gradations of a single attribute to constructs whose underlying metric seems to be at minimum not a simple linear function and may be indeterminately complex. Some constructs seem to come close to existing as contrastive sets, and if they do, it is ironic

that the best way of characterising elements upon them may be in terms of binary categories - thus returning, but at a much finer level of analysis, surprisingly close to Kelly's original thoughts on construing.

8 Completing the grid matrix

Practical and conceptual issues

Between the idea
And the reality
Between the motion
And the act
Falls the shadow

T.S. Eliot
The hollow men

8.1 INTRODUCTION

In the previous chapter the challenge to the Kellian view of dichotomous construing was extended to embrace the broader notion of bipolarity, the case being made that assumptions of linearity, continuum and mutual exclusivity of two antonymic poles may frequently not be justified. In this chapter I take up the theme from the point of view of the practicalities of completing a grid matrix, blending theoretical discussion with empirical evidence. Relatively little attention has been given in the literature to the task of completing a grid matrix: the argument presented in the previous chapter suggests that a detailed consideration of the issues involved is overdue. The greatest attention is given to rated grids in this chapter (on the grounds of their wide use in research), but many of the points raised have general application to both dichotomous and ranked grids.

8.1.1 Dichotomous grids

The completion of the grid form of Kelly's original Role Construct Repertory Test requires the respondent to place each element at one or other pole of each construct, making the assumption that all of the elements fall within the range of convenience of the constructs being used (an assumption about which he himself entertains some doubts[1]). Where an element is located at the emergent pole of a construct, Kelly puts a check mark in the appropriate cell in the grid matrix: the cell is left blank if the element is not assigned to the emergent pole. (There is a further assumption at this point - that 'not allocated to the emergent pole' implies allocation to the implicit pole). Kelly places no restrictions on the number of elements to be assigned to either pole, save that the emergent pole must subsume a minimum of two elements because of the nature of the triadic elicitation procedure adopted.

This procedure results in the production of many 'lopsided' constructs in which the allocation of elements to construct poles often falls considerably short of a rough equivalence[2]. In itself, this is not a problem - indeed, the very presence of lopsidedness may be of distinct psychological interest at the individual level.

Lopsidedness has proved to be of considerable nomothetic interest to Adams-Webber[3]. He and his associates have been forcibly struck by the tendency for elements to be allocated to construct poles in the proportion 62:38 rather than 50:50. He notes that this nomothetic lopsidedness is very close both to the Pythagorean concept of the 'golden section' and to the more recent notion of 'strikingness' derived from information theory, and suggests that by allocating elements to the negative pole approximately 38 per cent of the time negativity is made maximally striking as a 'figure' against a 'ground' of positiveness. The argument is aesthetically appealing, but it conceals issues which offer some threat to the establishment of a kind of golden section theory of judgment. First, the proportion quoted as 0.62:0.38 is a mean value: standard deviations for the data are of the order of 0.10, implying that individuals may differ quite markedly from the allocation of elements according to a 'golden section' criterion[4]. Second, the research so far has used personality constructs and it is unclear whether a similar tendency towards lopsidedness may obtain in respect of other realms of experience. Third, the personality constructs used have been 'logical' oppositions; yet (as was argued in the preceding chapter) many constructs may have structures that are less straightforward, and again it is unclear whether the golden section hypothesis - if valid for the area of personality - could be applicable more generally[5]. The empirical

evidence regarding the golden section hypothesis is considered in more detail later in this chapter when some psycholinguistic aspects of scales are discussed: suffice it to say at this point that the hypothesis rings a little hollowly - perhaps faintly echoing the music of the spheres.

Lopsidedness has analytical implications. Kelly's original non-parametric factor analysis is based upon the number of matches (tick with tick, and blank with blank) between rows or columns. However, the use of matching scores can lead to some spurious associations, as Bannister and Mair (1968) make clear[6]. Recognition of this problem led to the suggestion that the elements be allocated to construct poles on a 'split-half' basis[7]. Lopsidedness is thus defined out of such a grid, and false associations arising from the statistical properties of matching scores are eliminated.

The trouble with the 'split-half' form of the grid is that it forces the respondent to divide the elements equally between the two poles, whether or not this reflects the particular discrimination that he or she wishes to make. Bannister and Mair are aware of the difficulty that this causes, giving the example of the construct 'male-female' where the numbers of male and female elements are unequal. This forces them to admit that constructs which are unavoidably lopsided in the respondent's eyes would have to be eliminated from consideration[8].

This defect in the 'split-half' grid is sufficiently serious to warrant its rejection except where only the construct inter-relationships are of interest and elements can be provided which do not demand a 'lopsided' response. Examples where the use of the 'split-half'

grid can be justified include the Grid Test of Schizophrenic Thought Disorder (Bannister and Fransella, 1966) and Salmon's (1969) investigation of differential conforming and the developmental processes of children. In both of these examples the respondents were asked to evaluate photographs of unknown people against a set of supplied adjectival constructs, and 'split-half' allocation presented no problems[9].

8.1.2 Ranked grids

The rank order form of the grid avoids a number of the difficulties associated with the dichotomous allocation of elements, but brings in a new problem as far as analysis is concerned. Lopsidedness ceases to be an issue when elements are ranked, and ranking allows the respondent to provide more information in the location of elements along constructs. It is possible to rank elements according to the extent to which they possess a (unipolar) attribute X, which amounts to much the same thing as ranking with respect to the pseudo-bipolar construct 'X - not X'. Ranking appears to be best achieved by presenting all of the elements, asking which of them has the most 'X-ness', removing that element and repeating this sequence until the elements have been ordered according to their 'X-ness'[10]. Fransella (1972) exemplified this approach in her research on stuttering, the elements being twelve situations involving talking and the constructs being nine unipolar statements of the type 'situation in which I am most likely to stammer'. Sequential elimination of elements gave Fransella a picture of the respondent's construing of the relationship between situational context and stuttering[11].

'Not applicable' elements present a problem in ranking on an 'X - not X' construct. If one ranks from the 'X' pole, where are such elements

to be located? My suspicion (lacking any evidence on the matter) is that they would probably be placed towards the 'not X' pole on the grounds of lacking X-ness, thus coalescing the meanings of 'not applicable' and 'not X'. It seems unlikely that they would appear in the middle of the order (as might be the case with rating, where respondents are often instructed to use the central category for such elements), unless the ranking can be done 'inwards' from the two poles of the construct. Ravenette (1975), for instance, asks children to rank on the basis of 'the situation when it is most most/least likely that you would be pleased with yourself'[12].

Ranked grids allow the use of rank correlations between constructs, enabling their interrelationships to be presented in a variety of ways, but Humphreys (1973) points out that ranking along constructs (the rows in the usual grid) implies that the element-similarity data (pertaining to the columns) loses nearly all its meaning[13]. This represents a loss in comparison with the dichotomous forms of grid in which the matches between elements and between constructs are as meaningful as each other.

The ranked grid, therefore, would seem to have some advantages where the focus of interest is upon construct interrelationships and where unipolar or pseudo-bipolar constructs are to be employed. There are practical advantages in using ranking, particularly where successive elimination of elements allows the respondent to make discriminations sequentially, rather than <u>en masse</u>[14]. It may be the case that asking respondents to discriminate in this way elicits subtleties of gradation that would be obscured in dichotomous or rated grids. Clift et al (1978), who used rated grids, later felt that rating had been a mistake since some discriminations appeared not to have been faced by

their respondents. On the construct 'likeable - less likeable', for instance, discrimination was poor, and the authors suggest that the forced preference of ranking would have avoided the social desirability effect and would have spread out the people-elements with respect to their likeableness[15]. This procedure, however, does not get over the problems of the meaning of the 'not X' or 'less X' poles.

Ranking tends to become more difficult as the number of elements rises and 'bunching' on the construct occurs (often - but not always - at the middle of the dimension being used). If fine discrimination is not deemed to be of paramount importance it may be more appropriate, and less stressful to respondents, to use a rated grid instead.

8.1.3 Rated grids

Widely used in research involving repertory grids, rating overcomes two of the problems associated with rank ordering. These two advantages of rating are first that it allows the respondent to indicate the extent to which each element fits one or other of the construct poles and, second, that it once again becomes possible to compute fairly meaningful element similarities.

The use of rated grids raises a number of issues for discussion, both methodological and conceptual. Many of these apply to other forms of grid and they are treated here under the heading of 'rated grids' merely as a matter of convenience. The practical, methodological issues are discussed first, the conceptual issues being treated in the latter part of the chapter.

8.2 PRACTICAL AND METHODOLOGICAL ISSUES

8.2.1 Problems of rating

Hopwood and Keen (1978) present their respondents with a TARGET grid blank in which the elements are labelled and the triadic sorts are specified[16]. The respondent then provides a bipolar construct in respect of the first triadic sort and rates all the elements, beginning with those of the triad. The process is repeated for each row of the grid until either the grid matrix has been completed or the respondent is unable to produce further constructs.

In responding to such a grid (not as a naive respondent, but as one who had begun to study Personal Construct Theory and had administered some rated grids in pilot work) I found myself influenced in the following ways - all of which could be expected to be detrimental to the validity of the grid which I was producing.

(i) Unfamiliar - and unremarkable - elements too vaguely recollected for construing to be clear.

(ii) Some awareness of 'pattern' as the grid matrix was building up, leading to the possibility of intrusion of a 'halo effect'.

(iii) Accidental reversal of the rating scale, mentally switching from 5 = high to 1 = high (perhaps because 'five points' on the five-point scale and 'first' are both ways of indicating high quality). The problem seemed greater when negative (or negatively-valenced) terms were used in respect of the emergent pole, to which ratings of 5 and 4 apply in the TARGET protocol.

It is worth recording that accidental reversal of the rating scale can happen in respect of a subset of the elements as well as with the full

set, as I have found from test-retest reliability checks which I have carried out. The effect can be suspected when the differences between test and retest ratings tend to be multiples of two, and is particularly well exemplified by the two sets of ratings drawn from a pilot grid which I administered, and which relate to the construct 'irrelevant...-relevant to students': this was treated as a five-point scale with the negative 'irrelevant ...' scoring 5.

(a)	Original ratings	4	2	2	1	4	2	4	2	3	2	2
(b)	Retest ratings	2	4	4	5	2	4	2	4	3	4	4
	Difference: |(a)-(b)|	2	2	2	4	2	2	2	2	0	2	2

Suspecting reversal, I asked the respondent to re-rate the elements on the construct, and he produced a set of ratings very close to the original set:

(c)	Second retest	4	2	2	1	4	2	4	1	3	1	1

In my experience of grids I have found two further problems.

(iv) Variation, during the administration of the grid, of the context within which the elements are construed.

(v) Failure to follow the rules of the rating procedure. (An example presented by Hopwood and Keen, n.d., of a TARGET response shows this clearly. In this grid, where the similar pair had to be rated towards the high end of the five-point scale and the singleton at the low end, triads are rated as 5,4,4; 1,1,2 and 1,2,4[17]. Errors of protocol such as these must inevitably call in question at least the validity of the construct rows involved).[18]

Most of the problems (i) to (v) above have shown up in grids

administered to groups of individuals, and it would seem probable that elicitation of grids in these circumstances, or in the absence of the researcher[19], would increase the risk of unsatisfactory results for three reasons:

(i) the researcher may not become aware of errors or difficulties in responding, as a result of 'remoteness' from the respondent;

(ii) the respondent may be unaware that he or she is failing to follow the instructions correctly; and

(iii) even if the respondent is aware of difficulties, he or she may not wish to draw attention to the fact by asking for help.

It would seem prudent, therefore, to avoid the administration of grids to groups of respondents unless elements and constructs are supplied, and a simple procedure is adopted for the location of the elements on the construct dimensions (cf. Salmon, 1969); or, if elicitation is to be involved, the procedures should be somewhat simpler than is often the case with triadic protocols.

Where the grid is completed by an individual respondent in the presence of the researcher many, if not all, of the problems noted above can be overcome. In my work with the science teachers I used the triadic 'opposition' method of eliciting constructs (sampling from a set of fifteen supplied elements), probing for dimensions which were seen as meaningful within the fairly tightly defined context of teaching science to first year secondary school pupils. I tried not to 'lead' the respondents during the elicitation procedure, and used a portable cassette recorder to check whether the administrative procedure had introduced an unacceptable level of bias.

I placed in front of each respondent a large sheet of paper divided so as to make a 'ladder' representing the seven grades of rating, an eighth 'step' being available for use when an element fell outside the range of convenience of the construct being considered. The high and low ends of the rating scale were then labelled according to the elicited construct and the respondent was asked to locate each element (which was written on a small card) in an appropriate position, as is shown in Figure 8.1. Having previously recorded the elicited construct poles on another form (Figure 8.2), I wrote down the initials of the element against the appropriate rating category. The whole process was repeated for each construct elicited and, when convenient, the record of ratings and constructs was transferred to a standard grid blank in order to have it in a format suitable for analysis.

It was important to treat the whole procedure outlined above as a conversation, for this enabled potential errors and misunderstandings by the respondent to be detected and corrected immediately, as proved necessary on a number of occasions. By undertaking all the recording of data myself, the respondent was able to spend more time on the construing and rating than would otherwise have been the case (an important point where the time available is limited), and interference between the cognitive aspects of elicitation and the mechanics of writing down responses was eliminated.

8.2.2 Distribution of ratings

There is little in the repertory grid literature regarding the distribution of ratings given by respondents. Such as there is suggests that there is a tendency to rate at the ends of the scale or

Requiring children to use verbal (oral) skills

7	12. Discussion with individuals 11. Discussion with small groups 2. Discussion with class
6	
5	5. Pupil project work
4	6. Outside visits 10. Showing film or video-recording
3	4. Pupil practical exercises 15. Display 9. Science club 13. Using models and analogies
2	7. Pupil individual work in school
1	8. Homework 14. Achievement testing

Requiring children to use writing and drawing skills

X	3. Teacher demonstration 1. Teacher exposition

Figure 8.1 Rating sheet, showing the location of element cards for a construct from H.T.'s second grid. All cards were of equal size, and much larger, for the experimental procedure. The rating sheet was also larger.

This rating sheet exemplifies a number of problems with grids.

CONSTRUCT										
Pair	Singleton	7	6	5	4	3	2	1	X
Used a lot in my teaching	Not used — . — ..	D.L. P.Prac		HW	T.D. D.I. M.A.	T.Exp. P.Proj F.V. D.S.G.	P.Ind. A.T. Disp	O.V. S.C.	
Individual exercise	Collective experience	HW D.I.	P.Prac P.Proj P.Ind.		S.C. D.S.G. Disp.	M.A.		T.E. D.C. T.D. O.V. F.V. A.T.	
Concrete	Abstract	O.V. M.A. P.Prac	P.Proj. P.Ind	F.V. Disp	T.E. T.D. D.S.G. D.I.	H.W. D.C.	A.T.		S.C.
School-based	Home-based	T.E. D.C. T.D. P.Ind F.V. D.S.G. D.I. M.A. A.T. Disp		P. Prac S.C.	O.V.	P. Proj		H.W.	
Requiring children to use verbal (oral) skills	Requiring children to use writing and drawing skills	D.C. D.S.G. D.I.		P.Proj	O.V. F.V.	P. Prac S.C. M.A. Disp	P.Ind	HW A.T.	T.E. T.D.

Figure 8.2 An illustration of the way in which grid ratings were recorded, using data from the first part of H.T.'s second grid. The fifth row corresponds with the data in Figure 8.1.

near the centre, but the sampling is very thin in respect of the vast number of grids that have been reported.

Orley (1976), using a five-point scale, found ratings of 2 and 4 to be less used than an even distribution would require, the other ratings being more frequent than would be expected[20]. The percentage frequencies, with expected frequencies in brackets, were

1 and 5 : 45% (40) ; 2 and 4 23% (40) ; 3 : 32% (20).

Slater (1977), without giving any data, states that raters tend to ignore ratings of 2 and 6 on a seven-point scale[21].

An analysis of the distribution of the ratings in 37 of my own grids suggests that no general conclusions can be drawn. Figure 8.3 summarises the findings from 27 grids in which the 15 elements were aspects of science teaching and from 10 grids in which the 10 elements were pupils.

Grid Elements	No. of grids analysed	No. of constructs	Rating level 7	6	5	4	3	2	1	Not applicable or omitted
(a) Aspects of science teaching. At beginning of probationary year.	14	205	22.3	14.1	11.7	13.5	7.7	10.9	18.4	1.4
(b) Aspects of science teaching. At end of probationary year.	13	184	20.0	14.5	12.2	14.0	8.3	8.6	18.9	3.5
(c) Pupils	10	126	15.8	16.4	13.9	14.9	11.7	12.0	12.9	2.4

Figure 8.3. Percentages of rating categories used in 37 grids.

Rows (a) and (b) refer to the same group of teachers on two separate occasions, whilst row (c)refers to a completely different group.

There is, for each group of grids, a tendency for the highest three ratings to be used more frequently than the lowest three, with ratings 2 and 3 being the least preferred categories. This may be an artefact of the 'rating ladders' used (see Figure 8.1) in that the top of the scale may catch the eye more easily and therefore act more strongly as a referent[22]. It would be unwarranted to assume that this bias would be present when other rating strategies are used[23].

However, the distributions in respect of the science teaching elements differ from those when pupils are the elements, in that the extreme ratings are more favoured. This could reflect a tendency to dichotomise ratings in a situation where some constructs did not lend themselves as easily to the shadings of meaning possible with personality and behavioural constructs. Where the construing of pupils was concerned, there was a tendency for evaluatively positive poles to be given first (and hence be presented at the top of the 'ladder'), and the asymmetric rating distribution here might be a reflection of a weak 'Pollyanna hypothesis' bias[24].

These nomothetic trends, however, mask considerable differences between the ways in which individuals apportioned their ratings. The patterns of some were dominated by extreme ratings whilst others tended to avoid extremes, and there is evidence that some respondents assign ratings roughly evenly across the full range of categories. The evidence from the science teachers (who completed grids at the beginning and end of their probationary year) points quite strongly to the stability of personal rating style[25].

Whilst the broad trend of my findings supports most of what little has been claimed, the variance that is hidden by averaging and the

diversity of rating formats available make a general statement on rating behaviour unwise.

8.2.3 Range of convenience

Elements falling outside the range of convenience of constructs present a considerable problem for the analytical routines widely available, as is discussed in the next chapter. It is often suggested to respondents that they use the mid-point of the scale if they cannot locate an element on the construct under consideration. This is an expedient to get the researcher off the hook of analytical limitations and, although its effects may on occasion be small (as Olson, 1980a,found[26]), it is a dubious practice that subordinates meaning to methodology.

Keen (1979) invited his respondents to use a zero when an element fell outside the range of convenience of a construct, but he gives no indication of how zero ratings were handled during the analysis of the data which he collected[27]. Lipschitz (1972), reflecting upon her grid methodology, remarks that it was probably a mistake not to have a 'not-applicable' category since this omission made it difficult to determine the limits to the range of convenience of a construct[28].

The problem seems more serious in grids in which the elements are situations[29] rather than people. In the latter case constructs relating to personal characteristics are typically elicited, and it would appear both from the literature and my own empirical experience that respondents find little difficulty in subsuming a range of people under such constructs. Where situations are the elements it is quite possible for there to be a 'range of convenience problem' with

'X - Y' constructs, though not with 'X - not X' constructs. Olson (1980a),using classroom activities as elements, found on average ten per cent of missing values, similar to Forthman's (1973) finding that, if people were given the opportunity to strike out particular semantic differential scales which they felt were inapplicable to the concept being rated, the proportion was of the order of ten per cent.

In my work with science teachers the overall average percentage of 'gaps' in the grids was 2.7, the vast majority arising because elements were deemed to be out of the range of convenience of particular constructs. A small proportion of these gaps were due to errors or omissions in the rating procedure: not surprisingly, the bulk of these appeared in retest grids which the respondents completed in their own time. For the record, the percentages of gaps for each occasion of administration are given in Figure 8.4

OCCASION	Number of Grids	Total Number of Constructs	Total Number of Grid Cells	Total Number of Gaps	Percentage of Gaps
First grid; researcher present	14	205	3075	42	1.4
Retest of first grid; researcher absent	14	202*	3030	94	3.1
Second grid; researcher present	13	184	2760	98	3.6
Retest of second grid; researcher absent	11	157	2355	65	2.8
All	-	748	11220	299	2.7

Figure 8.4 Incidence of 'gaps' in grids in which the elements were classroom activities.

Note *Indicates three occasions in which respondents failed to re-rate on a particular construct.

The percentages of gaps shown in Figure 8.4 must, if anything, be viewed as an underestimate of non-applicability. As was noted in the previous chapter, the insistence on bipolarity clearly constrained

the construing of respondents such that there were occasions on which they wished to tell me more than the bipolar format would allow - for instance wanting to construe 'pupil practical exercises' in terms of manual skills when the construct insisted on 'visual-verbal' as the dimension. A mid-point rating in such circumstances is equivalent to stating that the element lies outside the range of convenience of the construct. There were a number of such instances whose implications escaped me at the time, but which press themselves forcibly upon me in the light of hindsight.

The tenor of this evidence and discussion is towards the recognition and acceptance of the fact that, in certain types of grid, some elements can be expected to lie outside the range of convenience of the constructs being used. As far as grid elicitation is concerned, this presents no particular problems. It is in respect of analysis that major difficulties arise: discussion of these is held over till Chapter 9.

8.2.4 Interim appraisal and reorientation

As far as obtaining a completed grid matrix is concerned, the problems mentioned in the preceding pages are not insuperable. Careful procedures of elicitation, and a willingness to accept 'gaps' in the grid, can go a long way towards ensuring that a data matrix is technically satisfactory. However, it must be remembered that a grid matrix is predicated upon a number of assumptions which the previous chapters have indicated should not be taken for granted, and whose validity may need to be assessed anew in respect of each grid that is designed.

In comparison with the conceptual issues that remain to be discussed, the methodological matters considered in the first part of this chapter are trivial. In what follows, necessarily the lengthiest section of this chapter, attention is directed to further aspects of the psychology of rating.

8.3 CONCEPTUAL ISSUES IN THE PSYCHOLOGY OF RATING

8.3.1 Social judgment in the light of construct theory

Eiser and Ströbe (1972) point out that most explanatory models of social judgment are derived from psychophysical principles assumed to operate in judging physical objects in terms of sensations - for instance, the estimation of weights. Such a model is shown in Figure 8.5, and is of a mediated stimulus-response type.

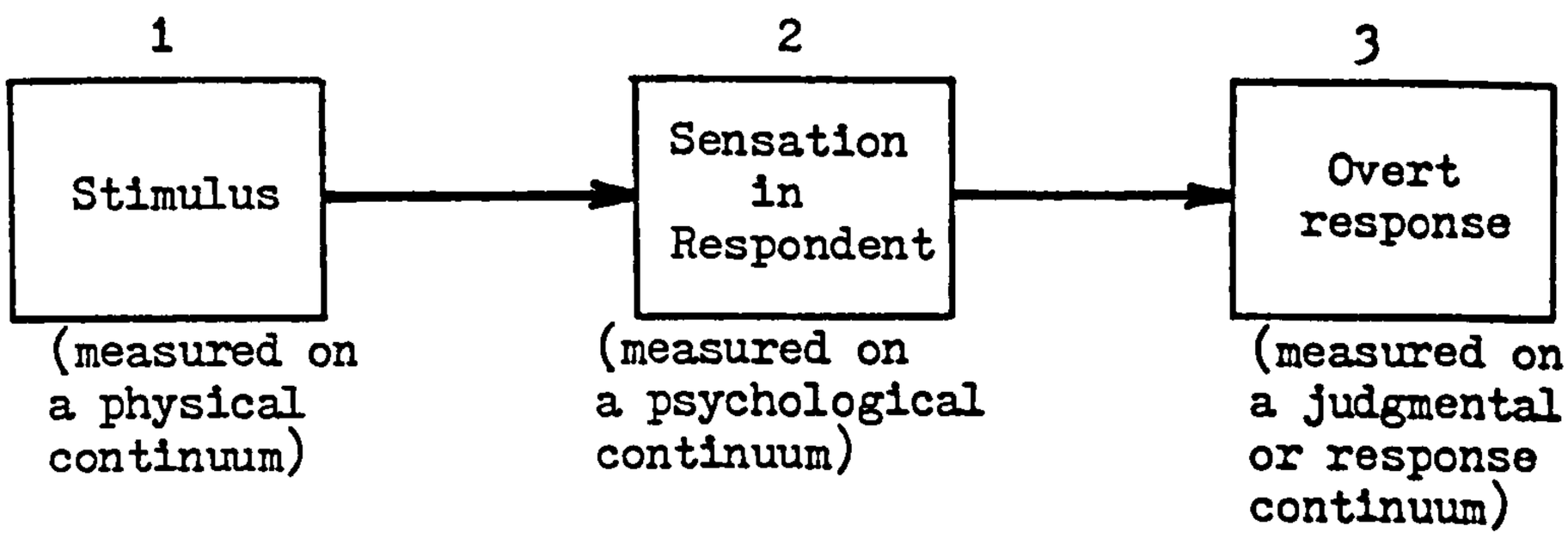


Figure 8.5 A mediated stimulus-response model of judgment (from Eiser and Ströbe, 1972:7)

In psychophysics 'hard' measurement is available for stages 1 and 3, and Eiser and Ströbe indicate that most theories of absolute judgment in that field attempt to account for the relationship between the data gathered in those particular stages. In social judgment (and I use the term to subsume the construing of a wide range of social phenomena), often the only 'hard' measure available is at stage 3 - the judgmental response[30]. In the present context this is the rating which the

respondent gives to an element in respect of a particular construct.

From a personal construct perspective the model represented in Figure 8.5 is inadequate. A personal construct theorist would wish to expand stage 2 to indicate more clearly the involvement of the respondent's construct system in the process of judgment. It was argued earlier that new constructs were unlikely to be formed during the process of elicitation in a repertory grid[31], and it seems reasonable to extend that argument to cover the rating phase. The element, then, is construed within an already-established framework in which the construct being used is also located. It is possible to envisage an iterative process in which the 'match' of element to the construct dimension is progressively refined from an initial 'yes/no' decision regarding the construct's relevance to the decision as to whether the rating should be one value or another. The argument can be advanced that a construct does not exist in isolation as regards rating and that the refining process that eventually produces the actual rating inevitably calls into the proceedings such 'neighbouring' constructs in the system as are relevant[32].

In effect, the judgment process is closely related to the assimilative cycle shown in Figure 3.1, though the behaviour is internal during perhaps a number of refining cycles before it is made overt in the production of a rating. This personal construct perspective on the process of judgment is presented in broad outline in Figure 8.6

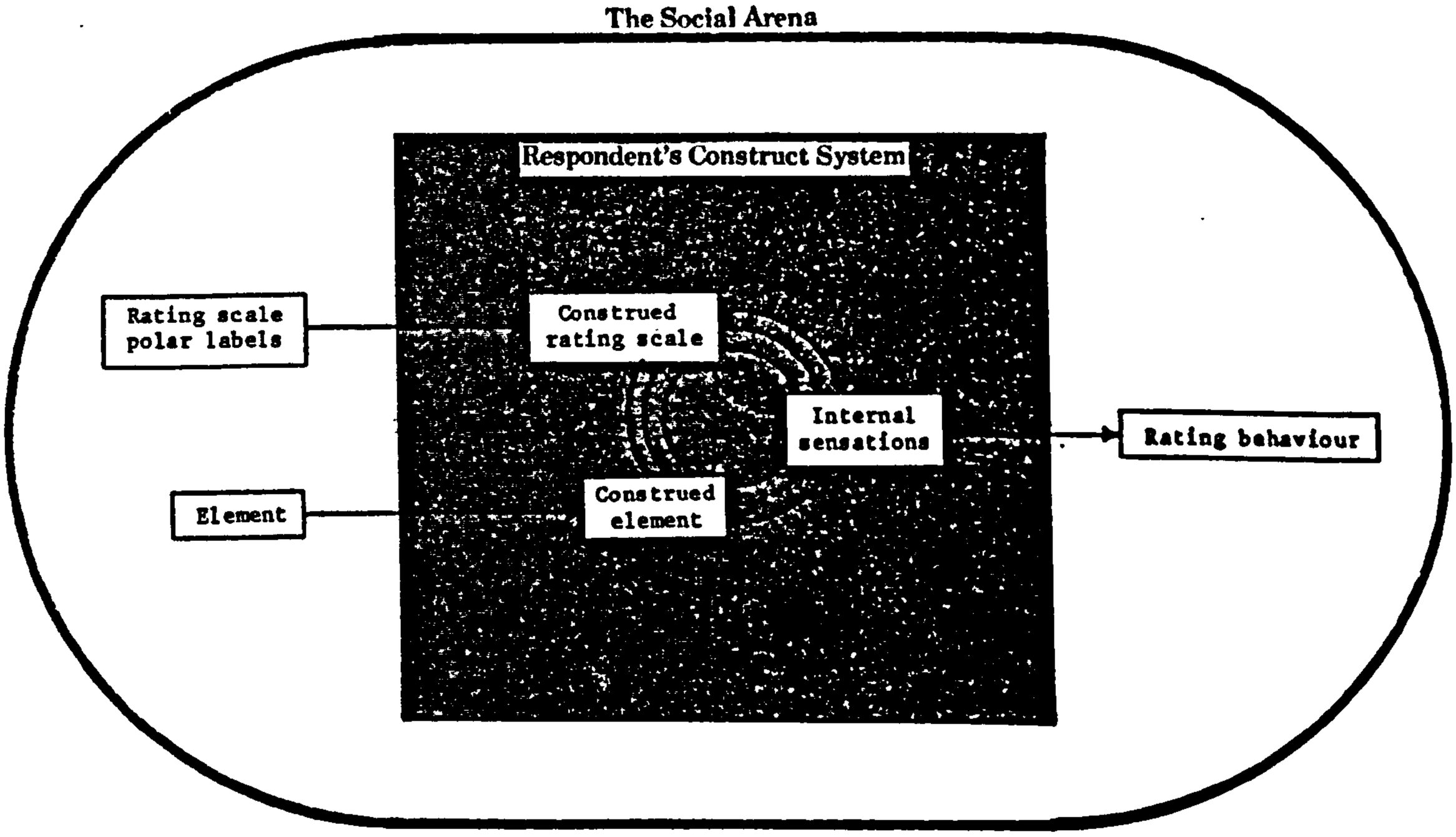


Figure 8. 6 A model of judgment deriving from a Personal Construct Theory perspective.

In Figure 8.6 the concentric arrows <u>symbolise</u> the cycles of refining of judgment: the actual mechanism is likely to be indeterminable. The possibility exists of sub-loops within the construct system as implicative relationships of the construed element and/or rating scale are explored, not necessarily consciously. In Kellian terms internal sensations would be regarded as discriminations of one sort or another and hence as sub-sections of the whole construct system of the respondent (of which the act of rating the element E on the construct X - Y involves only a part).

What is the point of offering a model of social judgment from a personal construct point of view? The point I wish to make, and strongly, is that rating probably involves very complex processes within the respondent's construct system. An element is not 'just an element' like the standard kilogram, stable for all time. It is construed according to an initial cognitive set - to limit the domain of the

argument - but may be significantly reconstrued as its implications are revealed or different dimensions of construing come to mind. Similarly, a rating scale is not likely to be a fixed entity like the standard metre: its meaning is likely to shift with reconstrual of the element, as a result of calling other elements to mind, or as a result of related constructs coming to mind.

The danger is that the researcher will read too much into the rating being given. As one moves further from 'concrete' elements and physicalist rating scales the probability of intersubjective understanding between respondent and researcher will decline, though neither of them may realise how approximate their mutual understanding may be in such circumstances. It is only by cross-checking understandings using other modes of inquiry that the breadth of meaning of, say, a rating of 5 on a seven-point scale labelled 'extrovert-introvert' is likely to be narrowed down sufficiently for meaningful interpretations to be made.

8.3.2 Rating Scales: the problem stated

It is surprising that rating scales are so widely taken for granted. Whilst standard texts on psychometrics - such as Guilford (1954) and Remmers (1963) - deal with the practicalities of their use and adduce a mass of empirical evidence regarding systematic errors in rating behaviour, there is little attempt to question the assumptions of rating scales or to set them in any kind of theoretical perspective other than the quasi-physical.

From a personal construct perspective the question of rating scales has to be approached from a different direction - that of the individual's method(s) of using them. Little seems to be known

about how scales are actually used in responding to grid protocols - or in cognate situations, for that matter. Evidence was presented in Chapter 6 suggesting that extremity of rating tends to be higher with increasing personal meaningfulness of the construct being used, irrespective of whether the construct is elicited or supplied - though elicited constructs are likely to be more personally meaningful. Pathology may exert a significant effect under conditions of high involvement. In presenting his interaction model Bonarius (1971) goes somewhat further, suggesting that rating extremity is a function of the element being rated, the construct, and the rater him- or herself.

The above may well act as influences upon the practice of rating, but they are at a level of generality above the existential experience of actually assigning ratings to elements. In inquiring how a respondent may use a rating scale it is necessary to look at a number of theoretical matters such as the metrics of scales, anchoring, psycholinguistics, construct complexity and interaction effects.

There is far more here than can be encompassed by the present work. All that I can attempt in the space-time at my disposal is to discuss briefly some of the issues that have emerged whilst I have been working with grids: it will be apparent that much of what follows can be applied to other contexts in which rating scales are used.

8.3.3 Stevens's work on scaling

A rated grid treats each construct as if it is linear and bipolar; that is, the poles are oppositionally related and mutually exclusive. Further, it is assumed that the rating scale itself is consistent during use, and - as Humphreys (1973) points out - that it possesses

equal intervals[33]. This last assumption is deeply embedded in the statistical routines used in the widely-available methods of grid analysis.

Stevens (1957) draws attention to the distinction between qualitative ('metathetic') and quantitative ('prothetic') scales. The former is essentially <u>substitutive</u> as one moves from one scale position to the next: he gives the example of musical pitch. The latter is <u>additive</u>: the further up the scale one moves, the more of the attribute is present - for instance, as in a scale of loudness. There is an important difference from the point of view of the subjective use of scales. The 'just noticeable difference' for metathetic scales tends to relate to a constant interval on the scale irrespective of the section of the scale involved, whereas for a prothetic scale the magnitude of the 'just noticeable difference' depends upon the scale position. For instance, Stevens points out that it is easier to discriminate 0.5 from 1.0 seconds than it is to discriminate 3.5 from 4.0 seconds. In the prothetic case the subjective scale of judgment may be markedly different from the physical, objective scaling of time.

It is difficult to transport Stevens's distinction into the realm of the repertory grid since bipolar labelling makes the scaling more complicated (see Chapter 7, and below). It would seem that pseudo-bipolar constructs can be discussed in Stevens's terms, though the actual distinction between prothetic and metathetic cannot be settled at the level of abstraction - it must rest upon the way in which the respondent actually uses the dimension. 'High... - low frequency of assessing pupils' might plausibly be argued to be prothetic, whereas 'Lead to high ... - low pupil self-reliance' would seem to be more

metathetic in character in that the psychological substrate relates to the qualitative aspects of the classroom activities being construed. The point being tentatively put forward here is that, if the researcher is using scales which are susceptible to analysis of this sort, it is possible that some of the subjective continua used by the respondent may systematically deviate from the metric being implicitly assumed by the researcher (see Figure 8.7). The relationship between such subjective continua and assumed metric will be monotonic, but the deviations may have implications for the method of analysis to be used and for the interpretation of the data.

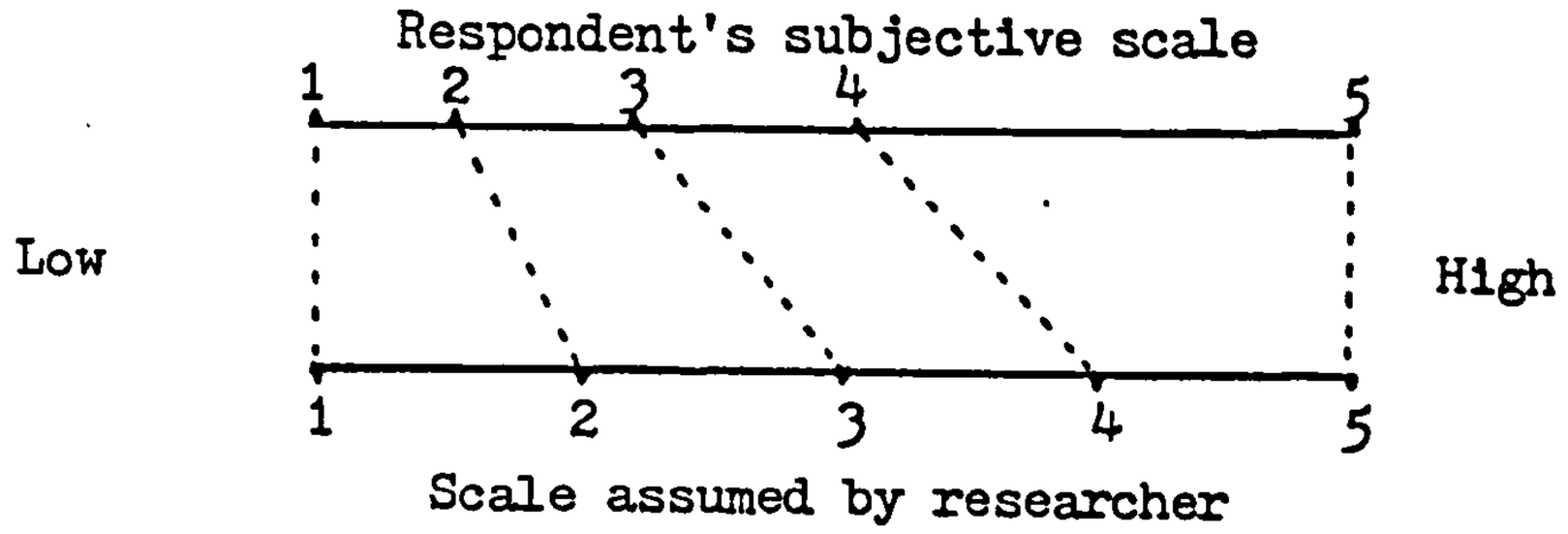


Figure 8.7 An illustration of the possible systematic deviation between the respondent's subjective scale and the researcher's assumed scale when the dimension is 'prothetic'.

Where scales are labelled antonymically (such as in the semantic differential) the mid-point <u>may</u> be the point of meaninglessness, in which case the scale could be 'cut' in the centre with each half being treated as a separate entity: the above conceptual analysis could then be applied. But, as will be shortly illustrated, many bipolar scales are likely to be more complex (I would argue so in respect of many semantic differential dimensions), involving cross-currents of meaning between the headlands of the verbal labels. In this choppy and confused sea it may be difficult to resolve what is observed on the surface into its pattern of undercurrents.

8.3.4 Anchoring

The typical rating scale in a repertory grid is labelled at each end but not at any of the intermediate positions, and is thus similar to what Sherif and Hovland (1961) describe as a scale without a graded stimulus series. The point of the comparison is that Sherif and Hovland assert that such a scale 'is markedly influenced by internal anchors devised by the individual, if he judges alone...'[34]. It would seem more likely that if anchoring does occur in grid dimensions, the anchor will lie within the dimension (or, at worst, be slightly outside it) - in contrast with some of the extremely distant anchors used in psychophysical research (for instance, in the judgment of weights). If this is the case, it follows that any distortion in rating is probably assimilative (i.e. towards the anchor) rather than contrastive (away from the anchor).

Assuming the possibility of anchoring effects, what might serve as an anchor? Volkmann (1936) suggested that any psychophysical referent could be selected by the observer and used as an anchor, and Eriksen and Hake (1957) took a broader view in their suggestion that a person might select a few stimuli in the series being presented (in their case, colours) and use them as standards for judging the rest.

Physical entities are relatively accessible as reference standards, but there is only the indirect evidence of inferences from observed outcomes to support the 'internalised standards' argument. With intangible elements, such as recollections of specific instances or people, the chain of inference is longer, and it is not surprising that little attention has been given to anchoring in such circumstances: the quasi-physicalism dominant in psychology may have discouraged the use of other approaches to the problem. I do not have the space

here to explore all the possible avenues and concentrate on a brief consideration of the possible contribution of 'self' in grids dealing with people, and of the verbal labels themselves. Some other observations on the practice of rating will be found later in this chapter during the discussion of possible element x construct interaction.

Landfield (1951) based a study of threat upon an early formulation of Personal Construct Theory and made use of rating scales (the repertory grid presumably not having been invented then) on which scale points were explicitly defined by people. Some scales were functionally dichotomous and some dimensional. The construct poles were 'fixed' by named people and, where the scale was dimensional, the mid-point was treated in a similar fashion. The anchoring seems to have been a matter of practical convenience rather than a reflection of a theoretical standpoint regarding rating methodology, and Landfield does not elaborate on the possible effects.

Taking the more general issue of triadic elicitation in the context of rated grids related to people, the elements in a triad are aids to the identification of a scale, and hence can be hypothesised to exert an anchoring effect upon it. If such an effect exists, it is likely to be at its strongest when the grid is built up row by row and all of the elements are rated on the construct in question before moving on to the next elicitation.

The heterogeneity of the elements in TARGET grids makes the possibility of anchoring more difficult to assess. The 'self', 'most effective' and 'least effective' teachers are likely to be stronger anchors than the 'video-elements' - and all, some or none of the three personally

meaningful elements may appear in the various triads. Anchoring here may be differentially powerful depending on the nature of the triad.

More generally, if 'self' is an element but not in the triad being construed, there may be an identification of the 'self' with at least one of the triad - and this could lead to an implicit, almost ghostly, self-anchoring effect.

A clue to the possible role of 'self' as an anchor is given by Leventhal's (1957) suggestion that similarity to the self might be a feature of judgmental behaviour[35]. Leventhal was only suggesting this of 'simple' judges, but if self-reference does occur in judgment there seems no good reason why it should not occur with the most 'complex' of judges. In many grids 'self' appears as an element, implicitly creating an ambience in which self-reference could flourish[36]. The ambience is strengthened when 'self' appears in each triad in order to make the elicited construct specifically meaningful to the respondent[37]: 'self' is therefore bound to be oppositionally related to at least one element in a triad, and it may be that all of the triad members (which have, after all, 'defined' the construct) act as anchors in their own right.

Adams-Webber (1979) reviews a number of studies in which people have been assigned dichotomously to the poles of the 'like self - unlike self' dimension, and finds a consistent tendency for approximately 62 per cent of the assignments to be made to the 'like self' pole[38]. He interprets this as showing some support for assimilation: Benjafield and Adams-Webber (1975) suggest that assimilative projection seems to underpin a number of grid indices where the dichotomous allocation of

elements is permitted to be lopsided. The situation may be more subtle, however, since respondents given a free choice may tend to select as elements those with whom they share a positive relationship and hence who are more likely to be 'like self' than 'unlike self': assimilative projection may well have taken place much earlier (at the stage of friendship formation, for instance: Duck, 1973), and hence not necessarily have resulted from some form of anchoring in the grid.

The situation is complicated still further where rating scales are being used in grids since the 'self' may be located at different positions on different scales, making it difficult to assess any influences on the underlying metric. Where the 'self' appears towards one end of a construct the respondent may add another, contrasting, anchor at the other end[39] - perhaps from the triad of elements used in the elicitation of the contrast: following Sherif and Hovland (1961) and Adams-Webber (1979), one might expect a measure of displacement of ratings towards the ends - perhaps particularly towards the presumably more salient 'self'.

Isaacson and Landfield (1965) conducted a study which suggested that there might be a connection between self-reference and personal construing, though the data are not presented in a form suitable for a definite conclusion to be drawn. They gave 99 psychology undergraduates Kelly's Role Construct Repertory Test, eliciting 15 constructs. 40 of the students returned for the subsequent task which was to rate, on a nine-point scale 'least... - most like me', 30 items constructed by splitting the individual's constructs and rewording them in a self-referential form (e.g. 'I am friendly') and 74 statements from

the Butler-Haigh Q-sort procedure. A re-test of the rating procedure was administered three weeks later. The results from both occasions of rating showed a strong tendency for items derived from personal constructs to be rated 'most like me', whereas the ratings tended to be only moderately 'like me' for the Butler-Haigh statements[40].

Although the example Kelly gives of his test only uses 'self' as an element three times in the triadic elicitation procedure[41], it would seem distinctly possible that there would be a measure of self-reference in the elicitation of each construct, a construing of others with reference to the respondent's own experience of the characteristics or behaviour associated with the elements. In Isaacson and Landfield's study such self-reference would have probably been more distanced in respect of the Butler-Haigh statements, where there seems to have been little, if any, bias associated with casting the items in the form 'I am X'. It is odd, however, to find that more than half of the respondents were able to place at least half of the 'split' personal constructs in the category 'most like me', since one would expect - if construing is broadly speaking oppositional - only one of the poles would be very 'like me', and hence a <u>maximum</u> of roughly fifty per cent of the split constructs would be expected to fit this category[42].

It is possible that 'self' may be identified with certain of the construct poles in the grid, investing them with a psychological charge. Evidence from resistance to change grids shows that construct poles tend to be valued differentially, and that on certain constructs respondents are very reluctant to switch to the opposite: in other words, it would appear that certain poles are highly ego-involving. If this is the case it is difficult to see how some sort of anchoring

could not take place for the most important poles. Where the respondent indicates only a marginal preference it is unlikely that there would be sufficient ego-involvement for any significant anchoring to take place.

Rosch and Lloyd (1978) take a broader view of anchoring when they suggest that category 'prototypes' may act as salient reference items in respect of the dimension of judgment[43]. Their suggestion is somewhat reminiscent of Landfield's (1951) explicit anchoring, but in this case the anchors are likely to be internal constructions. If category prototypes are constructed by the respondent in response to the presented scale labels it is likely that each construct in a grid will have different anchors and it will be difficult to tease out the effects involved.

The whole issue of anchoring is clearly complex, and it is difficult to assess the extent to which it is likely to influence the ratings given in repertory grids. The literature on anchoring suggests that any assimilative or contrastive distortion will not cause the ratings (assuming sufficient scale points) to deviate from monotonicity with ratings given in the absence of an anchor (though this may be an unapproachable yardstick in the light of the possibility of the respondent 'inventing' his or her own anchor).

8.3.5 Lexical marking

Boucher and Osgood (1969) put forward the so-called 'Pollyanna hypothesis' which encapsulates a supposed human tendency to use evaluatively positive (E+) words with greater frequency, diversity and facility than evaluatively negative (E-) words. Eiser and Mower White (1973) set out to test the Pollyanna hypothesis by asking

60 children to allocate twenty nonsense words to the poles of twenty bipolar constructs, treating the nonsense words as if they were the names of people. The twenty constructs were formed by splitting antonymic pairs (similar to those used in semantic differential work) and negating each adjective with the prefix 'not'. Each child was thus presented with a set of 'X - not X' constructs in which the 'X' adjectives were E+ and E- to an equal extent. However, Eiser and Mower White found virtually no difference in the proportion of elements allocated to the E+ and E- poles. Adams-Webber (1978) replicated this study with 60 Canadian undergraduates and obtained similar results, supporting the suggestion that the Pollyanna hypothesis was inadequate.

Adams-Webber (1978) also reanalysed the data gathered by Eiser and Mower White and found that allocations to 'X' poles were markedly more frequent than to 'not X' poles, irrespective of whether 'X' was E+ or E-: the same pattern was found in his own results, and the relative proportions of 'X' to not X' allocations in the two studies gave strong support to the golden section hypothesis[44]. But what Adams-Webber seems to have overlooked is the lexical significance of using 'X - not X' constructs in each study. The 'X' pole defines the dimension, and 'X-ness' decreases as a function of distance from the 'X' pole (a loose analogy can be made with the decrease in gravitational potential the greater the distance from a star). If 'X' is more meaningful than 'not X', one would expect some measure of bias in that direction.

The 'X' pole in an 'X - not X' construct is linguistically unmarked[45]; that is, in a bipolar construct it serves to define the superordinate concept under consideration, whilst the marked pole is a negation of some sort which is not used in the naming of the superordinate.

Marking is not restricted to 'X - not X' constructs, but can apply to 'X - Y' constructs as well. If the dimension 'happy-sad' is taken as an example, the superordinate concept would normally be taken as 'happiness' (with a range of values subsuming 'unhappinnss' as well), 'happy' being the unmarked pole and 'sad' being marked. Marking is often accomplished by a minor transformation of the unmarked term such that it remains evident in the marked form - for example, in 'happy - unhappy'. The situation regarding 'X - Y' constructs is not as straightforward as that described above, for two reasons. First, if the 'X - Y' opposition is 'peculiar' (in Resnick and Landfield's, 1961, sense) it will be unclear what the position regarding marking will be, since a 'peculiar' construct may be formed from the union of two separate emergent poles both of which could be said to be unmarked. Second, it may be the case that the pole assumed on linguistic criteria to be marked may in practice be used as an unmarked pole: if the dimension is 'happy - sad' there is nothing to prevent a melancholic construing this in terms of sadness. Markedness, in practice, could be a function of the psychological constitution of the rater. That psychological constitution may influence marking is suggested by the work of Adams-Webber and Benjafield (1973), discussed below.

Using twelve comparatively 'pure' bipolar scales from semantic differential research, Adams-Webber and Benjafield (1973) found a significantly greater tendency for respondents to rate people more extremely on the marked, rather than unmarked, poles when the construct poles were presented separately. In only two of these constructs was the marking 'obvious' (involving a negative prefix): 'unfair' and 'unpleasant' being the marked poles. The authors also

found that the more preferred the construct , the greater the tendency to locate the elements dichotomously at the unmarked pole, though there were some occasions when the majority of the acquaintances being construed were located at the marked pole. More extreme ratings[46] tended to be given in respect of the marked pole irrespective of the direction of asymmetry of the dichotomous assigning of elements. As far as rating extremity is concerned, Adams-Webber and Benjafield see asymmetry of element allocation as an individual parameter which seems to override the general linguistic norms of lexical marking.

Adams-Webber and Benjafield found women to be inclined more often than men to assign acquaintances to the marked pole, though this occurred in a small minority of cases of the twelve constructs used. They were unable to suggest why this may have happened, but two possible explanations can be advanced - if only speculatively, since the authors do not give the details on which a convincing argument might be based. The first possible explanation derives from the nature of the twelve scales used, which were designed to sample the three major factors of the semantic differential - evaluation, potency and activity - and it is largely with respect to the quartet selected as representative of potency that the argument is advanced. These four dimensions were 'strong - weak', 'bold - timid', 'hard - soft' and 'rugged - delicate' and, given traditional cultural stereotyping, it would not be surprising to find women using at least some of the four supposedly marked poles in an unmarked way: I would hazard the guess that 'soft' and 'delicate' would be the most probable candidates in this respect. The guess has some support from Osgood et al (1957) themselves, since they noted that both 'rugged' and 'delicate', when used separately, appear to be evaluatively positive[47]: this suggests

that there might be <u>two</u> superordinate constructs present ('ruggedness' and delicateness') each with its own unmarked pole.

The second possible explanation derives from the sampling of the elements. It is not made clear in the account given by Adams-Webber and Benjafield whether the twelve role titles selected from Kelly's Repertory Test were balanced for sex. If men tended to select men (and women tended to select women) to fit the role titles, then some differences in the allocation of elements might be expected, particularly if markedness is partially related to sex. Although no overall main effect was found for sex, the greater frequency with which women assigned elements to the marked pole (averaging 2.2, versus 1.3 for men in respect of the twelve constructs used) was presented as statistically significant.

The inference can be drawn from the work reported here that the relationship between lexical marking, evaluative valence and rating behaviour is more complex than is implied in the experiments cited. The suggestion is made that the sex of the respondent may be an influential factor in rating behaviour in respect of an indeterminate minority of bipolar constructs where sub-cultural norms could implicitly reverse the more generally-accepted marking. Further, the type of bipolar construct cannot easily be disregarded. In practical terms, 'X - not X' constructs can be differentiated from 'X - Y' constructs on the ground of cognitive implications for rating behaviour, and 'X-Y' constructs themselves may be further differentiated into three categories: antonymic pairs in which the marking is 'obvious', antonymic pairs in which the marking is not immediately 'obvious', and 'peculiar' constructs in which lexical marking is problematic. The distinctions are summarised in Figure 8.8.

RELATIONSHIP		TYPE OF CONSTRUCT	EXAMPLE
NEGATION	STRICT NEGATION	X - not X	happy - <u>not happy</u>
	NEGATION BY MODIFICATION OF UNMARKED POLE	X - Y	happy - <u>unhappy</u>
OPPOSITION	'LOGICAL' OPPOSITION	X - Y	happy - <u>sad</u>
	'PECULIAR' OPPOSITION	X - Y	happy - businesslike

Figure 8.8 Types of construct and lexical marking. The underlined poles would normally be considered to be lexically marked. The marking of 'happy-businesslike' is problematic.

So far research in this area has concentrated upon 'simple' bipolar constructs of the semantic differential type. Multiple opposition has not been given any attention, and the fact that elicited constructs (often, but not always, adjectival in character) can be expressed in comparatively complex phrases[48] makes the possible influence of lexical marking very difficult - perhaps impossible - to assess: consider, for example, the construct 'pupils being active with the teacher looking on - teacher being active with the pupils looking on', which is drawn from my work with the science teachers.

Research has been sparse to date regarding lexical marking in repertory grids, although linguistic theorists have given the issue rather more attention[49]. There would seem to be a need for a further exploration of the practical aspects of marking, setting it clearly within a perspective of linguistic theory. As far as current repertory grid practice is concerned, probably all that can be suggested at this stage is that the researcher look very carefully, in the light of the points made in the preceding discussion, at the verbal labels being used (whether supplied or elicited) with a view to gaining some insights into the grid responses that are made.

8.3.6 The complexity of meaning

One of Kelly's assumptions regarding the elicitation of constructs is that the respondent does not shift ground between the labelling of the oppositional poles - in other words, the verbalisations do not reflect the emergent poles of two different constructs[50]. The rating form of the repertory grid presses this assumption further in that it assumes the invariance and approximate linearity of the dimensions that comprise it. The range of meanings possible in respect of the mid-point (see page 296 below) casts some doubt on the tenability of these extended assumptions, a doubt which increases the further the construct is from the realm of the physical or quasi-physical (such as the physicalist criteria for judging the similarity of cars: Green and Carmone, 1969).

The question of the complexity of construct poles was raised in Chapter 7 during an initial discussion of Jane's grid (Shaw, 1980a), and it was argued then that it was highly unlikely that a meaningful dichotomous allocation of acquaintances could have been made to the construct poles, given the concatenation of meanings ascribed to each. The conversion of Jane's dichotomous grid to a rated grid would not have eliminated the problem because she would have been faced with the problem of striking a balance between the various attributes listed for both construct poles in order to provide a series of scale ratings. It is likely that a separate balance would have been necessary for each acquaintance being construed, implying the possibility of element and construct interaction (the theme of the following subsection).

But even if it were possible for Jane to identify, say, fixed meanings for five different scale positions (relating to specific complexes of verbal labels), it is doubtful whether such a scale would approximate a smooth continuum of meaning. And even if the 'espoused' scale were

linear, the 'scale in use' would seem in practice to be influenced by the differential salience of the attributes listed for each individual; the gradation would be rendered unstable by the establishment of a series of new meaning-complexes, a number of which could, for example, plausibly be rated at 2 on a five-point scale, but for very different reasons.

Rated grids typically employ bipolar labels assumed to define a linear scale between them, but the assumption of scale in them has hitherto been left untested. The argument presented in respect of Jane's grid (which is a particularly obvious example of construct complexity) can probably be advanced in respect of many scales in that they are likely to subsume unarticulated, idiosyncratic meaning-complexes even though they may be stated in apparently simple terms.

As part of a study into the metrics possible with construct dimensions (discussed below, page 284f) some evidence was collected regarding the meaning that a number of respondents gave to supplied construct poles and to the mid-points of the dimensions being considered. The findings suggest that the meanings given to dimensions do not necessarily fall into the steady gradation assumed in much psychometric scaling. Examples of non-linearity included those presented in Figure 8.9, in which the supplied labels are presented in capitals and the respondents' elaborations in lower case.

RESPONDENT	LEFT-HAND POLE	MID-POINT	RIGHT-HAND POLE
No. 40 Physics PGCE	OVERCONFIDENT Unthoughtful and lacking in judgment but Aggressive and Possessing a reserve of relevant skills	Normal Able to communicate but Not boorish	LACKING IN CONFIDENCE Thoughtful and considerate Knowledjable but Introvert Good integrity but shy
No. 14 English/ History PGCE	LOGICAL .... accepts nothing that cannot be supported by palpable evidence and has the ability to structure his own thought and draw own conclusions	Accepts some things on face value but is not afraid to stand by instinctive judgment on occassions	INTUITIVE .... has an inherent understanding of the subject and is sensitive enough to appreciate ideas.
No. 18 English PGCE	CONSCIENTIOUS ...who attended school regularly, handed homework in on time, took care over books and work.	...who took an average interest in school work. If continually chased, he would hand in homework etc.	NOT CONSCIENTIOUS Careless. didn't hand in homework. played truant.

Figure 8.9 Examples of the elaboration of supplied constructs, showing their multidimensionality.

The multidimensionality of meaning evidenced in Figure 8.9 in respect of supplied verbal labels was similarly apparent with elicited constructs when, in pilot work, a small group of M.Ed. students were asked to construe (using the triadic method) six acquaintances and to indicate what their own verbal labels meant to them. Two examples are shown in Figure 8.10; in this pilot work no attempt was made to

collect any background information about the respondents.

LEFT HAND POLE	MID-POINT	RIGHT HAND POLE
STRONG MINDED Always takes a positive stand even if he is wrong: Loves argument Have strong Right or Left wing views	Will listen to other points of view. Always willing to help people in need Weighs the evidence and then decides Always willing to get involved	TENDER MINDED Always willing to take other person's point of view Does not like active positive approach Loves the underdog.
INTRAVERT ... prefer own company, never take lead, remain in background, wear acceptable clothes-never stand out. Colourless conversation if at all.	... enjoys the company of others but has no desire to be a leader, ... will join in group activity but also indulges in activities alone, ... sometimes motivates others to do things, rarely draws attention to themselves by dress, voice pitch, actions.	EXTRAVERT Lively, interesting, outstanding in terms of dress, conversation, voice pitch, often initiating group activity and leading discussions, enjoys the company of others. "People who wear odd socks deliberately".

Figure 8.10 Examples of the elaboration of elicited constructs, showing their multidimensionality.

Many other examples of meaning-complexes were produced by the respondents in elaboration of both elicited and supplied constructs, though some respondents were content to provide a synonym or synonymous phrase. Some of the meaning-complexes were more easy to see in terms of a rough grading scale than others. In Figure 8.9, for example, there appears to be a coalescence of three dimensions reflecting care, attendance at school and the completion of homework under the broad concept of conscientiousness as far as Respondent No.18

is concerned. On the other hand, Respondent No.40 elaborated 'overconfident-lacking in confidence' in such a way that any subordinate dimensions were fairly well obscured.

The evidence presented here suggests that constructs may well be far more complex than might be supposed from a simple inspection of the verbal labels.

Construing the respondents' construing of the supplied or elicited verbal labels (shades of Romanoff and Juliet here[51]) leads me to doubt whether scales reflecting meaning-complexes such as those given in the examples can be meaningfully used in a graded manner[52]: in fact, it would not seem improbable that, as the verbal labels are used in practice, the respondent would flick from one to another of the meanings (or groups of meanings) subsumed under the verbalised label. Care must be taken not to claim too much for this argument, however, since it is by no means certain that an 'unprompted' respondent would elaborate the verbal labels in the same way or to the same extent. All that is shown here is that the <u>potential</u> for so doing must exist, since the respondents were able to produce their elaborations freely.

8.3.7 <u>Rating scales: a multiplex of metrics?</u>

Elicited scales - and, to judge from the literature, a number of provided scales - in repertory grids seem to be hurriedly constructed. Many grids contain mixtures of some or all of the following: 'logical' and 'peculiar' scales, dichotomies and continua, unipolar and bipolar scales, and some scales whose underlying metrics would appear to defy simple description. The implications of this state of affairs for analysis and interpretation are far-reaching, and it is necessary to

examine the problem of typal variation in constructs a little further in order to see whether it might offer a serious threat to the validity of grid methodology.

An empirical investigation (fully reported in Appendix 4) of the potential variation in construct type suggests that the problem may be substantial. Fifty-one students in teacher education were each presented with seven bipolar constructs previously elicited from a class of in-service teachers. They were asked to write down what they would mean in using each pole and mid-point, assuming they were rating pupils on these dimensions. Having given their meanings, they were asked to select, for each construct, the most appropriate from eight diagrams designed to represent different ways in which a construct might be used in practice (see Figure 8.11) - or to supply their own diagram. For those unhappy with diagrammatic presentation, a verbal explanation was also provided. In the event, only five out of 351 responses were the students' own diagrams.

In this investigation the problem of the size of the 'time-slice' was ignored. Models 1 to 7 (Figure 8.11) all indicate ways in which attributes may be construed in terms of gradations, but this conceals a potentially serious confounding of the frequency and the strength with which a particular attribute is exhibited. Thus, referring to the dimension 'kind-cruel', would the mid-point of the scale imply 'neither particularly kind nor particularly cruel' being maintained consistently; flickering between kind and cruel actions (but these roughly balancing each other over a period); or the exhibition of kindness to some and cruelty to others - say, in a classroom - in a very short period? Rating scales in themselves cannot resolve such

issues: the implications of the question are discussed later in Section 8.4.

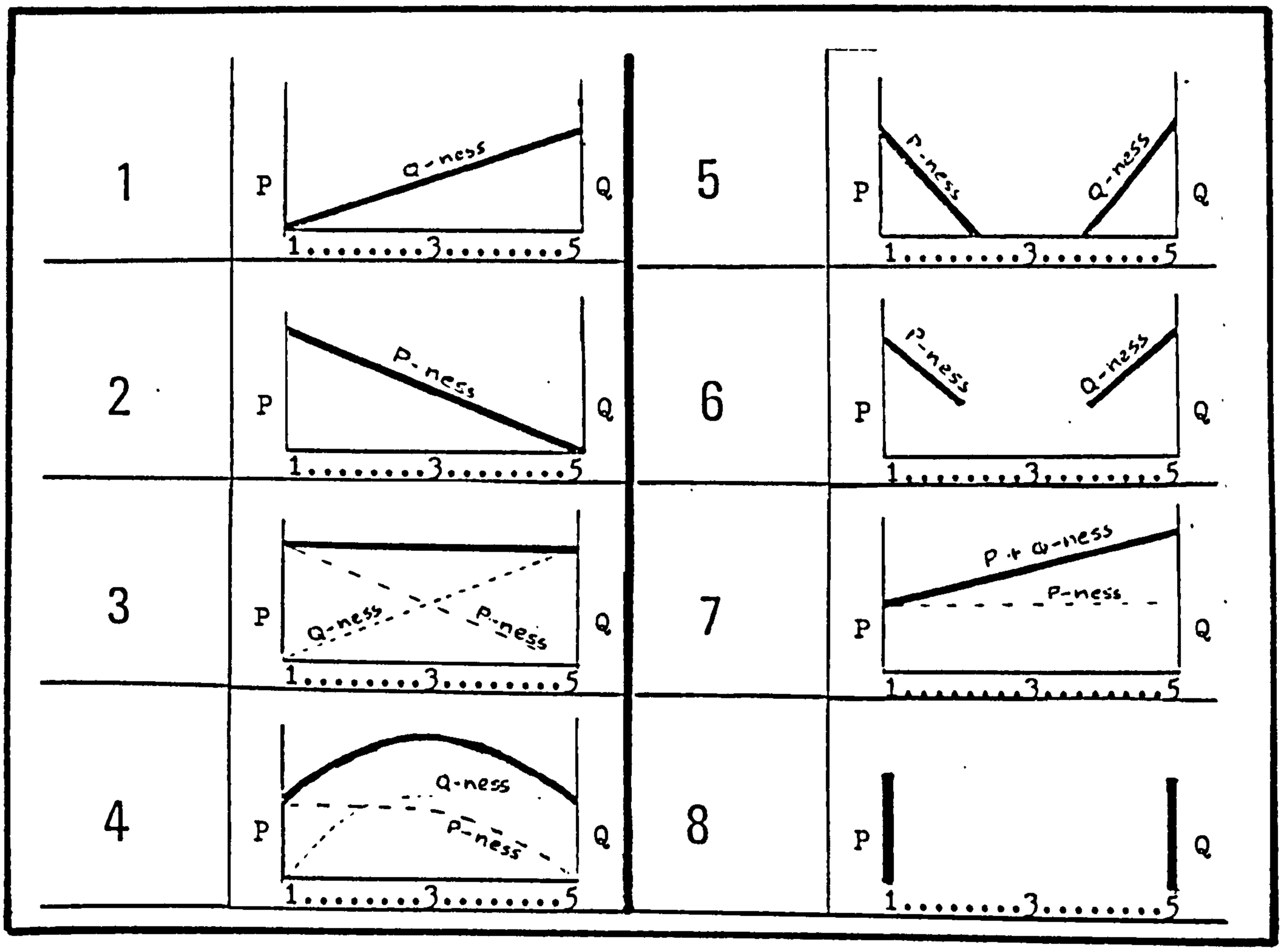


Figure 8.11 The eight diagrammatic representations of construct use presented to the 51 students. In the diagrams P and Q refer to the particular dimension being considered: thus, if the dimension being considered were 'Logical---Intuitive', P would refer to 'Logical' and Q to 'Intuitive'. The full version of this Figure is given in Appendix 4.

CONSTRUCT	FREQUENCY OF CHOICE OF DIAGRAM									TOTAL
	1	2	3	4	5	6	7	8	Own	
Logical---Intuitive	2	2	19	13	0	5	4	2	3	50
Extrovert---Introvert	4	9	19	5	5	6	0	2	0	50
Trustworthy---Untrustworthy	6	15	11	3	3	7	0	8	0	53
Overconfident---Lacking in confidence	5	10	20	5	6	0	0	2	2	50
Inarticulate---Articulate	19	3	14	3	2	6	0	2	0	49
Thinks in concrete terms---Thinks in abstract terms	3	0	13	14	0	6	10	2	0	48
Conscientious---Not conscientious	7	13	15	4	1	6	1	4	0	51

Figure 8.12 Frequency of choice of diagram from Figure 8.11 for each of seven constructs. Underlining indicates a frequency greater than would be expected on average.

Note: The totals for the rows vary because of incomplete response. Where the total exceeds 51 two students each chose two diagrams as appropriate to the construct.

The main findings are summarised in Figure 8.12, and reveal no massive consensus regarding the most appropriate model for any of the seven constructs. While one might expect a measure of 'error variance' in the performance of what is an unusual and complex task, the results suggest that there was a tendency for more than one model to attract more choices than would be expected on an assumption of randomness - in fact, the distribution of choices on each construct was significantly different from random (all p values less than 0.001; one sample chi-square test).

The most immediately striking feature of Figure 8.12 is that each of the seven constructs attracted the choice of Model 3 more frequently than would be expected on average, though the proportion of choices

for this model in no case exceeded 40 per cent. It is perhaps worth recording that the overall mean proportion of choices of Model 3 was 32 per cent, but this cannot be used as a basis for comparisons with the other models since to do so would overlook the effects of the heterogeneity of the particular constructs being considered.

The choice of Model 3 suggests that both poles of each construct are being treated as representative of attributes meaningful in their own right, even though three of the constructs ('trustworthy-untrustworthy' 'inarticulate-articulate', and 'conscientious-not conscientious') are explicit negations. In terms of linguistics, though 'untrustworthy', 'inarticulate' and 'not conscientious' are lexically marked terms (which should logically derive their meanings through negation of the corresponding unmarked terms), it would appear that - for some individuals at least - each marked term may exist within its own meaning-complex (possibly of unmarked terms).

Model 3 implies complementariness: as the degree of membership of one pole declines, so the degree of membership of the other increases. It may be that the comparative popularity of this model is due to the potential it offers for 'averaging out' attributes inferred from behaviour over an indeterminate 'time-slice'. For example, if I am construed as inarticulate in committees and as articulate in the lecture theatre, Model 3 would allow the rater the options of rating me at 2, 3 or 4 (weighting my in/articulateness as appropriate) on the dimension concerned.

Model 3 also seems to be comparatively attractive where the mid-point of the dimension can be construed to be the most positively-valenced point. The mid-points of both 'extrovert-introvert' and 'overconfident-

lacking in confidence' were noted by some respondents as being connected with a notion of 'normality', and Model 3 catches the sense of the ideal person as 'a bit of both' better than any of the other models presented. It can further be argued that an ideally balanced person would be roughly equally logical and intuitive: the task then is to decide whether this is a complementary, substitutive matter (implying Model 3 - which received 19 choices) or an additive matter (implying Model 4 - 13 choices). It is a weakness of bipolar rating scales that they are insensitive to such differences.

There would probably be a wide measure of agreement that 'extrovert-introvert' is a good example of the 'logical' opposition typical of semantic differential scales. If, as Osgood et al (1957) suggest, the mid-point of such scales reflects meaninglessness (disregarding their instruction to rate at the mid-point when both poles apply to a concept), it would be expected that Model 5 - or perhaps Model 6 - would be chosen fairly frequently. However, for the dimension 'extrovert-introvert', both of these models were chosen less frequently than average. Whilst it seems reasonable that a person construed to be at the mid-point of this scale would be 'a bit of both' in respect of the two poles (rather than 'neither', which does seem meaningless here), the evidence leads me to wonder whether the assumption of polar mutual exclusivity which underpins the semantic differential is dubious - at least, as far as actual rating behaviour is concerned.

Model 8 (the simple dichotomy) represents the principle of mutual exclusivity in its strongest form, while Models 5 and 6 retain mutual exclusivity but do allow gradation of membership of each pole. It is not surprising to find a number of respondents opting for Model 8 in respect of 'trustworthy-untrustworthy': a person simply is or is not

trustworthy in some eyes. The choice of Model 6 is interesting, perhaps indicating a recognition that there are shades of grey regarding trustworthiness and at the same time a realisation of the difficulty of placing a sensible interpretation on the mid-point of what is fundamentally a dichotomy. The more general choice of Model 6 may reflect the vagueness with which mid-points often seem to be associated (see page 293f).

It is interesting to find Model 2 being chosen with a frequency above average for both 'extrovert-introvert' and 'overconfident-lacking in confidence'. This suggests that respondents may construe some personality dimensions in terms of the more 'outgoing' pole, rather than in terms of the complementariness of the two poles: if so, the more 'outgoing' pole is acting as an unmarked pole (though both poles may be negatively valenced) and the lexical marking introduces a subtle semantic asymmetry to the dimensions concerned. As would be expected from the criterion of lexical marking, a substantial number of choices were given to Model 2 in respect of 'trustworthy-untrustworthy' and 'conscientious-not conscientious', and to Model 1 in respect of 'inarticulate-articulate'. One would not, at first sight, expect Model 1 to be chosen in respect of 'trustworthy-untrustworthy' and 'conscientious-not conscientious' since this appears to imply an inversion of the lexical marking of the scale. But if the rater uses the positively-valenced term as a starting-point (perhaps as a gestalt 'ground') for rating, it is possible for a negatively-valenced attribute to stand out (as figure) increasingly with its distance from the contrasting pole. As has been noted, there is some evidence to support this view[3] which has a grounding in information theory.

It must be emphasised that I have not addressed the problem of actually

giving ratings to elements on the seven constructs provided. The study discussed above was not designed to investigate actual rating behaviour, but it is possible to make some tentative points in this respect. If the preceding discussion is valid, it is clear that there are a number of possible influences on the rater such as lexical marking, valence and semantics of the polar labels. I would hazard the opinion that, for the majority of the models concerned, the ordering of the elements on the constructs would be at most slightly affected by such influences, which might be anticipated to distort the subjective judgment metrics systematically in relation to the implicit equal-interval scale defined by the categories of rating. In other words, it would seem preferable to use ordinal rather than interval statistics in the analysis of the numerical data, though to do so might still be to assume too much in respect of more 'peculiar' dimensions. This would cause some difficulty for the researcher interested in the element similarity matrix (see page 248f). Apart from Bannister's consistency score, the most widely-known methods of analysis make use of the interval assumption, on which the bulk of the discussion in Chapter 9 is necessarily founded. Goodge (1979) also advocates ordinal statistics, but his method of cluster analysis is not entirely clear from his article.

It is with constructs exemplified by Model 7 that the greatest difficulties may be found in respect of both rating and analysis. These difficulties can be highlighted by considering the construct 'thinks in concrete terms - thinks in abstract terms', for which three models received more choices than would be expected on average.

The most frequently chosen model (No. 4; 14 choices) allows both attributes to be maximally present together at the mid-point, in

contrast with the third most frequently chosen model (No. 7; 10 choices) in which this happens at the right hand pole. Given the sample of respondents, it is perhaps surprising that this model did not recieve more choices, since it is built upon assumptions consonant with Piagetian developmental psychology. Model 3 (chosen by 13 respondents) implies that concrete or abstract thinking are two distinct abilities; the more one is concrete, the less one is abstract, and vice versa.

Consider for a moment these three models being applied to a group of children, and suppose that some of these children are capable of formal operational (abstract) thought. Model 4 would probably place such children at the mid-point unless there was no observable evidence of concrete thinking, in which case the right hand pole might be preferred. Model 3 might be used in a roughly similar way. But Model 7 assumes a platform of concrete thinking and only takes account of the abstract thinking evidenced in behaviour: the most highly abstract thinkers would be rated 5 at the right hand pole. It is therefore possible for raters to produce different orderings of pupils along the dimension according to the way in which the scale is used. The relationship between ordinality and semantics is not necessarily simple.

The argument presented above bears out Gaines and Shaw (1981) in their claim that membership of construct poles is not necessarily determinable on the basis of mutual exclusion[53]. It also suggests that the ways in which people actually use scales may not be those dictated by the rigour of logical and mathematical thinking. Further, the empirical evidence gives reason to suspect that the method of use may vary with the scale and/or be a function of personal cognitive style.

It is important to stress that this investigation does not prove that people actually use scales in ways such as those presented here - it only suggests what might be the case. A different type of study would be needed to explore the ways in which scales are used in practice. For example, a number of logically antonymic contrasts might be elicited from respondents in the context of 'person-perception', and the respondents could then be asked to rate the elements whilst 'talking through' the reasons for the ratings. A content analysis of the transcripts of the procedure would provide some evidence as to whether the construct is or is not being used in scalar fashion - though it has to be acknowledged that, at root, any grounding hypothesis would be metaphysical and not open to rigorous empirical test.

8.3.8 The 'mid-point problem'

The empirical work described in the preceding section draws attention towards the meaning of the mid-point of a bipolar rating scale. As part of that study, the 51 respondents were asked to indicate (having provided their own meanings) which of the two poles and mid-point of each of the seven constructs listed in Figure 8.12 were clearest, and which the vaguest, in their minds. More than one position on each construct could be assigned to either category of clarity. The results are shown in Figure 8.13.

RATING	NUMBER OF OCCASIONS WHEN SEEN AS	
	CLEAREST	VAGUEST
1	233	62
3	71	217
5	194	80

Figure 8.13 Clarity and vagueness of ratings 1,3, and 5 of seven constructs treated as five-point rating scales.
(The seven constructs are those listed in Figure 8.12)

On the whole, the poles tended to be placed in the 'clearest' category and the mid-point in the 'vaguest' category, and this tendency applied for all of the seven constructs used. Where the tendency was weakest was where the mid-point could be (and was explicitly noted by some as being) associated with some notion of 'normality'. This obtained for the two constructs 'extrovert-introvert' and 'overconfident-lacking in confidence'.

The evidence presented here dismisses the assumption often made (as in the semantic differential) that the mid-point is necessarily devoid of meaning. Keen (1979) did not go quite so far in his work with repertory grids, but his weaker assumption that the mid-point (the 'origin' in his terms) represents the 'point of total uncertainty'[54] is similarly called in question.

It is readily conceded that for functionally antonymic constructs the mid-point may be a haven for total uncertainty. Given the limits enforced by bipolarity, it could also represent a balance between oppositional attributes: half the time a person may be manifestly 'happy', and half the time 'sad'. It is doubtful whether either interpretation can be made so strong as to claim meaninglessness for the mid-point. But many constructs may have very meaningful mid-points - take, for instance, the commonsense construing of 'extrovert-introvert'. Extremes of both extroversion and introversion are likely to have a negative valence, and 'normality' - however difficult to specify - would be represented at or near the mid-point: the respondent can presumably recognise a 'normal' person when he or she sees one. It would seem likely that any construct of the general form 'hypo X - hyper X' could be discussed in similar terms, even though the implied valence may be well concealed, such as in 'loquacious-taciturn'.

The point can be pressed a little further by drawing upon a physicalist parallel. Consider the dimension 'oblate... - prolate spheroid'. There are degrees of oblateness and prolateness, as is shown in Figure 8.14, and the mid-point is defined with the utmost precision as a sphere.

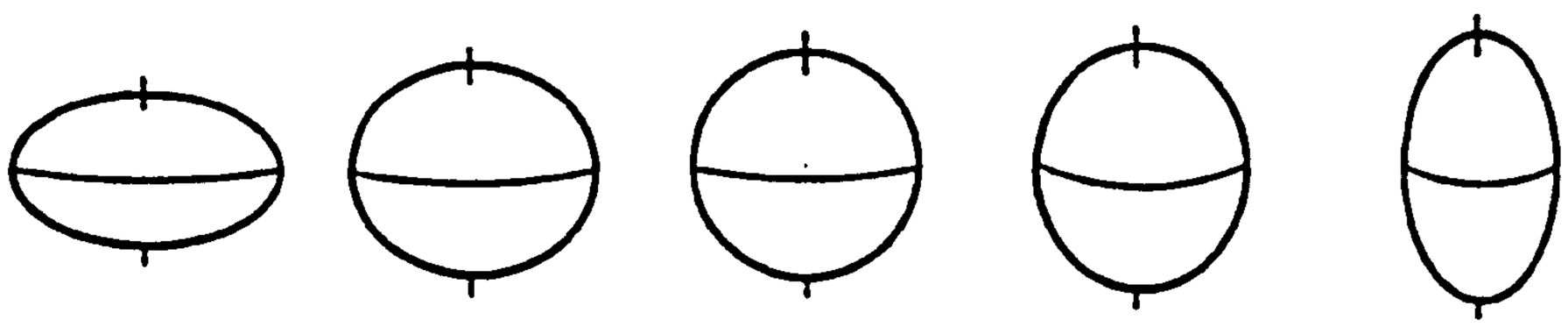


Figure 8.14 The 'oblate... - prolate spheroid' dimension.

Finally, when a scale represents a gradation of a property from low to high ('strength', for example), the mid-point represents some notion of a medium, or moderate amount of the property.

It is clear from the foregoing analysis that a repertory grid can include a variety of scales and that no blanket conclusion can be drawn about the meaning of the mid-point. Each construct has to be treated on its own merits, and the researcher may have to enquire during the elicitation of the grid matrix the way(s) in which the respondent is using each scale.

As was mentioned a few paragraphs ago, the constraint of bipolarity forces a variety of meanings into the mid-point category on a rating scale. Shaw (1980a) is one of the few writers who have addressed this issue, but makes the mistake of confounding the problem of the meaning(s) of the mid-point with that of 'bent' or 'peculiar' constructs[55]. In using rated grids researchers have very frequently offered the mid-point of the scale as a 'dustbin' for a range of responses that would have otherwise been difficult to locate: this

may not have been unconnected with the problems which 'missing' data pose for the statistical analysis of grids. Writing of survey research, Galtung (1969) criticised the practice of lumping together a variety of responses at the mid-point of a scale, arguing that to do so failed to differentiate between meanings such as 'don't know', 'not applicable', 'not certain', and 'neutral'[56]. More recently Gaines and Shaw (1981) have looked at the problem from the point of view of logic and point out that to allow the mid-point to subsume a variety of meanings is to assign a single truth value in respect of a range of different circumstances[57].

There seem to be many possible ways in which the mid-point can be used. The list given in Figure 8.15 is not claimed to be exhaustive, nor is it intended to imply that the various possibilities suggested are either equi-probable or of equal importance: the letters symbolise an element E being located at the mid-point of a bipolar construct X-Y.

1. E is neither X nor Y.
2. E is half X and half Y.
3. E is equiproportionately X and Y
4. E is wholly X and wholly Y simultaneously.
5. Half of E is entirely X and half of E is entirely Y.
6. Half of the time E is entirely X and half of the time E is entirely Y (i.e. E 'flickers' between X and Y)
7. E is Z where Z is 'sort of' half-way between X and Y.
8. E is sometimes X, sometimes Y and sometimes neither.
9. E is sometimes X, sometimes Y and sometimes both.
10. E is sometimes X, sometimes Y, sometimes neither and sometimes both.
11. E's position on the construct X - Y is uncertain.
12. E's position on the construct X - Y is unknown.
13. X - Y is irrelevant to the construing of E.
14. It is not wished to construe E in terms of X - Y.
15. The respondent does not care about the location of E on X - Y.

Figure 8.15 Fifteen ways of construing the location of element E at the mid-point of the dimension X - Y.

Provided that more than one of these ways of using the mid-point stands up to scrutiny, there is a problem for the personal construct theorist.

To allow the mid-point to be a portmanteau of possibilities is to limit the capacity of the researcher to receive the information which the respondent is willing to provide - and this must surely be counter to the philosophy underlying Personal Construct Theory. Further, the possibilities made manifest in Figure 8.15 give a clear indication that different types of dimension, though superficially similar, are based on very different psychological assumptive structures.

8.3.9 Element x construct interaction

Throughout this and the preceding chapter there has been a ground-swell setting in the direction of a possible element x construct interaction analogous to the concept x scale interaction detected in a number of studies involving the semantic differential. A priori, there would seem to be good grounds for expecting such an interaction. Most elements are complex (whether they be people or situations), and the possibility of constructs subsuming meaning-complexes (rather than existing as unique, isolable and scalable attributes) was discussed earlier[58].

Verbally-labelled constructs reflect discriminations made between elements, and the range of constructs elicited reflects the range of different features of the elements involved. If cardboard packing-cases are to be construed, it is likely that constructs would cover aspects such as the toughness and thickness of the cardboard, the strength of the staples holding the case together, the size and shape of the case and the printing on the surface. In all but the last of these it could be expected that there would be a fair measure of consistency in the use of constructs, since they reflect physical attributes regarding which people are likely to have at least an intuitive and reasonably replicable subjective measuring scale.

It is when the notion of scale is less certain, such as when judging the aesthetic quality of the printing, that matters become more difficult. As elements become more complex, judgments become more difficult and the attention of respondents may flicker to different parts of the stimuli in turn, as Hollingworth (1913) found to be the experience of his judges of handwriting[59].

Even relatively simple 'complex' stimuli have been shown to be salient to respondents in respect of their component parts rather than their wholeness (cf Underwood et al, 1962)[60]. As far as repertory grid work is concerned, attention has already been drawn to the observations of Mair's (1967) respondents that the people they were called on to judge were complex, and Mair suggested that variation in attention to the elements might have been a contributory factor in his finding of a weaker bipolarity in constructs than might have been expected, though the possibility of variation due to construct complexity may also have been involved[61]. In contrast to Mair's findings, Duck's (1973) respondents felt that they needed a range of contrasts to a particular similarity when construing a number of acquaintances, implying the need for a different construct for each element even if one pole was nominally constant. Duck suggests that some parts of construct systems are relatively specialised and have specific relevance to only one or two of the elements being construed[62].

The evidence in favour of the argument for an element x construct interaction is scanty since what one might loosely call the phenomenology of rating has attracted little attention. What empirical support exists for the argument has to be drawn from the work which shows a consistent concept x scale interaction in the semantic differential. It is tempting to translate this effect directly into

the context of the repertory grid, but to do this would be to fail to take into account the contribution to scale x concept interaction of attempts to construe stimuli on inappropriate dimensions[63]. Some inferences can be drawn from the field of personality research and, within that field, the few repertory grid studies in which the problem seems to have surfaced.

Features of elements that might contribute to an element x construct interaction include complexity, familiarity and affectivity, all of which overlap to some extent. The argument that elements are often complex was presented in Chapter 5 [64] and a couple of paragraphs ago (when reference was made to Mair's (1967) findings), and will not be repeated here.

That familiarity of people interacts with trait-perception is suggested by the work of Koltuv (1962) who found that familiarity and personal relevance of 'social objects' appeared to influence the inter-correlations between traits. Construing her findings from the perspectives of a possible element x construct interaction leads me to speculate that, if construct intercorrelations are 'disturbed' by familiar elements (to keep to the grid lexicon), then it might be the case that the constructs are being used in a different way when such elements are being considered[65]. In the light of the argument presented earlier, it could be that familiar elements are construed with reference to one or more specific incidents belonging to the construer's experience whereas less familiar elements may be related to less well-remembered incidents or to a more generalised framework derived from semantic memory. The construct label used would be the same in each case, the meaning qualitatively different.

The affective 'loading' of a person is likely to be related to familiarity. Mueller (1974) selected 64 students on the basis of high or low cognitive complexity as judged from scores on the Bieri test (Bieri et al, 1966). Two sets of role titles, differing in affective loading, were presented to the students who were asked to 'supply' appropriate people as elements. There were two tasks, which in brief were: to provide similarity judgments relating to the elements and to rate the elements on twenty supplied constructs presented as semantic differential scales. Mueller's results showed that affectivity appeared to influence the judgments made on both tasks, the effect being greater for those students identified as 'cognitively simple'.

Mueller deliberately built the dimension of affect into his lists of role titles: in other words the construct 'like-dislike' suffuses the elements selected. Collett (1979) makes the more general point that, where role titles imply constructs, the salience of those constructs attaches to the elements involved[66]. The question here is whether this leads to an element x construct interaction when elements of differential affectivity are being construed in other ways, or whether a construct x construct interaction is taking place. This last point is taken up in the following subsection.

8.3.10 Construct x construct interaction

At first sight Mueller's (1974) work points towards the notion of element x construct interaction but examination of the nature of the elements reveals the 'like-dislike' dimension clearly, and his findings implicitly relate to the interaction of this dimension with the others which he used. In his article Mueller does not give the construct labels, but if they were person-oriented constructs elicited from pilot

work with University students it seems reasonable to assume that they would have shown a number of substantial intercorrelations with 'like-dislike' had that construct been administered in the same semantic differential format. The question at issue is not the existence of such intercorrelations, but whether the intercorrelations are distorted as a result of the respondent's construing on one dimension being affected by construing in respect of others. Or, put another way, do constructs which are 'naturally' positively correlated interact to produce a spuriously inflated degree of association?

Tajfel's (1957) accentuation theory proposes an interaction effect between dimensions of judgment which Eiser and Ströbe classify into focal and peripheral dimensions[67]. Translated into the Kellian lexicon, this theory suggests that if elements vary concurrently along two or more dimensions (i.e. intercorrelate), discrimination along the focal dimension should increase. The theory rests on the assumption that where correlation exists 'peripheral' cues are available to assist discrimination on the focal dimension. Thus judgments of the lengths of lines could be expected to be aided by drawing the lines in different colours in such a way that colour is correlated with length. Eiser and Ströbe (1972) remark that any accentuation effect will depend upon the relative discriminability along the dimensions involved. Where discriminability along the peripheral dimension is so low as to be insignificant, accentuation of judgment on the focal dimension will not occur; and when discriminability on the focal dimension is low, peripheral cues will be unable to assist: in other words, there has to be a subjectively perceptible intercorrelation for accentuation to occur[68].

Tajfel and Wilkes (1963) tested the accentuation theory in an experiment

which required respondents to judge line lengths. Eight lines, varying successively in length by five per cent, were presented on cards. Three conditions were established: (i) a superimposed classification in which the four shortest lines were labelled A and the four longest B; (ii) a random allocation of labels A and B; and (iii) no superimposed labelling whatever. Tajfel and Wilkes found a large increase in estimate of line length in condition (i) when respondents moved from class A to class B, but not in either of the other two conditions. The importance of peripheral cues was suggested in a further experiment by Tajfel et al (1964) in which two Canadians and two Indians were interviewed for eight minutes regarding books or films they liked. The authors found that, when the audience was asked to rate the interviewees on a series of bipolar constructs, the Canadians and Indians were construed in terms of established cultural stereotypes rather than in terms of intra-class differences. Whilst they admit that the interviewees might actually have been like the stereotypes, Tajfel et al suggest that the 'peripheral' cues relating to inter-class differences had overridden the 'focal' intra-class differences - but here 'focal' and 'peripheral' would seem to be referring to the researchers' view of the problem, and not necessarily to that of the respondents.

It would appear that some aspects of stereotyping might be subsumed under accentuation theory. W Mischel (1973) suggests that in daily life people construe each other as being highly consistent even though the fragments of behaviour which give rise to the construing are inconsistent[69]: differentiation may be overridden by the cognitively more simple act of focusing upon such similarities as present themselves, and perhaps upon the assimilation to 'prototypes'.

What, then, are the implications of accentuation theory for the repertory grid? First, it is possible that a respondent could become

'locked' into a meaning-complex when constructs are being elicited. That is, the respondent might produce a construct such as 'active-passive' during elicitation, continuing in subsequent sorts to produce constructs related to 'active-passive' and perhaps derived from the meaning-complex of which that construct forms a part. This would be a kind of accentuation during the elicitation of constructs (connecting with the problem of sampling adequacy discussed in Chapter 6) which does not fall under Tajfel's original formulation: it might be termed 'semantic accentuation' in contrast to Tajfel's theory which deals with judgmental accentuation. Semantic accentuation could be expected to produce highly associated constructs in the grid, and these could lead the researcher to interpret the analysis as showing the construct system to be dominated by one (or more) factors without realising that the dominance may merely be an artefact of the elicitation procedure itself.

Second, a subjectively-construed correlation between two constructs could lead to accentuation in the rating patterns being used: there is a relationship here with Newcomb's (1931) 'logical error'. I would hazard the guess, based on Koltuv's (1962) work, that where an element (such as a person) is well-known to a respondent, this knowledge would tend to override any accentuation - in contrast to the less well-known elements which might be more affected by the implicit theories held by the respondent. For instance, a relatively unfamiliar element might be viewed as 'highly intellectual' and, because the respondent tends to correlate intellect with being stimulating, give the same element a high rating on the 'stimulating' end of the dimension 'stimulating-boring' even though there was little 'hard' evidence to justify the rating. This (a facet of implicit personality theory) would seem to come close to Tajfel's theory, but

without drawing the distinction between focal and peripheral dimensions which - in the absence of evidence from something like a resistance-to-change grid - would be difficult to ascertain. In any case, there does not seem to be any clear reason why accentuation should not take place in either direction (from peripheral to focal dimensions, and vice versa) even though a strictly logical approach might suggest a unidirectional relationship: in practice, focus and periphery are likely to be dependent on the way in which the particular individual (researcher or respondent) sees the problem.

Third, there may be the pervasive influence of an evaluative construct like 'good-bad'. If people construe others using a confirmatory strategy, then general evaluative ambience may induce deviations from the 'true' rating in the direction of the evaluative preconception. As with semantic accentuation, the net outcome in grid terms is likely to be the strengthening of associations between constructs in respect of which judgmental accentuation takes place. The result would be an overtightening of the factors produced in analysis. Such judgmental accentuation is not new, long ante-dating Tajfel's (195) formulation: E.L. Thorndike (1920) called it the 'halo effect'.

8.4 THE PRACTICE OF RATING: CONCEPTUALISATION REVISITED

The long preceding section of this chapter has pointed to the possibility of a number of interaction effects associated with repertory grids, but so far it has skirted around the question of how a respondent actually uses a rating scale when he or she construes a series of elements upon any particular dimension[70]. That question, the crux of this whole chapter, must now be addressed.

Gregson (1975) remarks that human judges are susceptible to cognitive

overloading in respect of both the amount of information to be processed and the number of logical operations needed to combine information. He postulates two possible strategies which may enable a judge to simplify the complexity of the task of judging overall similarities of entities.

The first of Gregson's strategies is to estimate the differences between the entities in terms of each dimension separately, and subsequently to combine the judgments to arrive at an overall decision regarding similarity. The second is to take each entity in turn and to estimate its position on the several dimensions before combining the information in order to make an overall judgment. Gregson sees the distinction as important since the second strategy avoids a susceptibility to intransivity intrinsic in the first, the self-chosen order in the second being influential in determining to some extent the goodness of fit of behaviour to theory[71].

As far as repertory grid work is concerned, almost all the reported research requires the respondent to rate all of the set of elements on one construct before proceeding to the next. In the typical grid arrangement, the matrix is built up row by row. Little use is made of the second, column by column strategy. It is less convenient in terms of administration (when time is often at a premium) and it is more open to the halo effect than the row by row strategy. It seems doubtful whether Gregson's preference for a column by column strategy can be sustained in the context of the practicalities of typical grid administration, the intransitivity issue notwithstanding.

Given, then, that rating in grids is typically done on a row by row basis, how does one actually go about it? Does one use the particular

scale as some kind of ruler, or is the actual rating derived from cognitive operations of a very different order? Have attempts to approach respondents' construing beoome stuck in the ruts of a track leading in the wrong direction?

The literature having been no more than marginally helpful in considering these questions, I am thrown back upon my own resources. Introspection leads me to suggest that one possible way of assigning ratings is to scan my experience for evidence pertinent to the dimension under consideration: if the elements are people, to take each person in turn, recall what I can from memory, and locate this information on one or more constructs which are subsumable under the actual verbal labels to which I am making explicit reference. I can exemplify this approach in respect of the construct 'generous-mean' and of four people I have known in the past fifteen years.

Thinking of these four people in the context of 'generous-mean', I recall that

A stood back and let me take the credit for something which I had initiated, but on which we had collaborated;

B issued an open invitation to stay whenever I needed to retreat from the pressure of events;

C - only a casual acquaintance - offered me the loan of a newish Volvo car in order that I might find out if I liked the model enough to buy one myself; and

D was inclined to attach 'D' prominently to work for which I was largely responsible.

To me these are, in the language of experimental psychology, 'the salient aspects of the stimuli' to which I am required to respond: they are, in this case, related to events in episodic memory - though A, B, and C are connected with individual events whereas D is connected

with a number of events spread over a substantial interval of time, and my response in respect of D is necessarily more generalised. The construct 'generous-mean' is being operationalised differently for each element being construed, using alternative combinations of poles drawn from the meaning - complexes subsumed by the poles 'generous' and 'mean'. My ratings of A, B, C and D on the construct are shown in Figure 8.16, from which it is apparent that I construe A and B as equally generous (but <u>for quite different reasons</u>), that I discriminate between A and D in respect of a meaning-complex connected with giving and taking credit, and that I discriminate between B and C in terms of the complexes associated with the sharing of belongings. This is a far cry from the physicalism of the ruler and is much closer to the tenets of existential phenomenology.

'Sharing belongings' Complex	B offers use of house in-definitely	C offers loan of Volvo car				'Selfishness' complex
'Giving credit' complex	A allows others to take credit				D takes more than fair share of credit	'Taking credit' complex.

GENEROUS ← 5 — 4 — 3 — 2 — 1 → MEAN

Figure 8.16: Differential construing on the dimension 'generous - mean'. It is a simple matter to recast the data in the format of a typical repertory grid matrix of elements and constructs, though, given the content of the example, the matrix will be incomplete.

Now, if time and stamina were no object, it would be possible to devise a grid which would allow me to display all the nuances of meaning of 'generous-mean' and of all the other constructs that could be elicited from me. Such an enormous grid would almost certainly be remarkable for an overwhelming number of vacant cells as the definition of

constructs becomes progressively finer and each individual construct becomes applicable to progressively fewer elements: the beginnings of such a situation are evident even with the comparatively coarse discriminations in Figure 8.16. In other words I am drawn towards the conclusion that, as distinctions become finer, the element x construct intersection looms larger in importance than the more general notion of scaling to which the greatest attention has hitherto been given.

No doubt many would wish to respond quickly to the preceding paragraphs and to say that in the real world one has to do with approximations and to try to perceive something through the haze of meaning surrounding each construct: the danger is that a construct - often only a 'simple' label - stands for a complex web of ideas whose differential salience is never explored.

In discussing the rating of elements on constructs I have taken the considerable liberty of indulging in introspection[72], and it is necessary to provide some support for the position I have taken. Forty-two teachers from Further Education on a day-release Certificate in Education course were asked to identify three people whom they knew well (and who came from the same life-context) and to rate them on one or other of the constructs 'generous-mean' and 'bold-cautious'. When they had all done so, they were asked to write down what had come into their minds at the time of giving the ratings.

A small minority of respondents referred to aspects of the focal person's behaviour, for instance:

'Profiteer in business to an unacceptable degree but tight in company, esp when buying the beer. £1 lasts all night yet he is never without a beer' (Respondent No.22);

'Despite a total income equivilent to many other families constantly pleads poverty, in domestic decoration or modification or in personal clothing etc. always buys cheapest materials etc, often to detriment of final product.

Is quite loving in nature but this does not translate into much direct help to others'(Respondent No.15); and

'A round dodger in the pub, goes to the toilet when its his turn to buy the beer' (Respondent No.17)

It is interesting that aspects of behaviour relating to acts of meanness were the most sharply recorded - perhaps negativity standing out against a ground of positivity.

The more frequently found response was to write down some more generalised dispositional statement such as might be provided in a character reference. Examples here include

'Careful, slow to form opinions - qualifies opinions. Very reluctant to take risks or to offend. "Safe" lifestyle - would be unlikely to make drastic changes in this' (Respondent No.25);

'Forthcomming with ideas. Will accept any task. Uses own initiative. Will speak out opinions on any subject. A little disruptive' (Respondent No. 38); and

'Constant help and assistance, thoughtfulness, unselfish nature, plays a major role in my life, a good relationship between us, never let me down, always available if needed, easy-going nature and considerate'. (Respondent No.3).

Other responses of this generalised kind were more laconic, the most terse being descriptions of three people as 'outspoken', 'diplomatic' and 'reserved', respectively (Respondent No.23). Parsons et al (1983), in a detailed study of a single teacher, note that much of her

professional knowledge appeared to be stored in generalised memories of events.

Immediately after the task was completed I probed further the way in which the ratings had been given, and nearly half of the group said that they had rated on the basis of specific remembered actions even though these had often been written down in more general terms. A further quarter of the group had used specific actions in respect of one or two of the three people they had rated, whilst nearly one third had not consciously referenced specific actions[73]. These results are summarised in Figure 8.17

CONSTRUCT	Reference to specific actions for all three people being rated	Reference to specific actions for one or two of the people being rated	No reference to specific actions	Total
Bold-cautious	10	2	7	19
Generous-mean	9	8	6	23

Figure 8.17 Methods of use of the constructs 'bold-cautious'and 'generous-mean' by 42 teachers in Further Education.

This very limited empirical enquiry provides some evidence for the argument that rating may be done by reference to specific circumstances- and also for the complexity of construct dimensions. I am not seeking to claim that all constructs are as capable of subsuming a range of meanings as are the two instanced above. Nor would I wish to suggest that people necessarily construe with respect to critical incidents recalled from memory: first, specific actions may well have been forgotten and a generalised characterisation may be all that is available to consciousness; and, second, knowledge gained indirectly may be stored

in semantic memory as might be the case with labelling and stereotyping.

Despite the thinness of the empirical evidence adduced in support of the phenomemological[74] approach to rating, the argument is not easily dismissed - and in some ways it is closer to Kelly's views of both construing and events than the (currently more popular) ranking and rating forms of grid allow.

Kelly's use of the term 'events' is realist: events are 'out there' to be construed[75]. In discussing the Fundamental Postulate I argued that an event only has meaning in so far as it can be construed (however 'defectively' that might be done), hence - for me - an event is the 'intersection' of an instance (element) with, at minimum, a single construct and is therefore represented as an individual cell in a grid matrix. If constructs are associated in meaning-complexes rather than exist as isolable entities, the theoretical precision of the individual cell is lost, and the event is 'smeared' down the appropriate column of a grid which has an indefinitely large number of 'meaning - possibilities', i.e. rows. These rows would be representative of the contents of both semantic and episodic memory, and may well be dichotomous in their reference to presence or absence of - say - an attribute. ·It is a truism to observe that different events are associated with different patterns of 'meaning-possibilities' in their respective grid columns. In other words, each event could in theory be represented by an enormously long binary number indicative of the constructs being used to give it meaning[76].

The 'cognitive algebra'[77] involved in processing meaning is of truly staggering proportions, and far beyond the capacity of rational

analysis - and it should be noted that the binary number for each event would be dominated by zeros. Gregson's (1975) postulated strategies offer the possibility of reducing cognitive overload to manageable proportions but they are based on the implicit assumption of a relatively limited number of dimensions, as would seem to be the case with all grid work whether the elements be allocated dichotomously, ordinally or by some form of grading.

I would, in the spirit of constructive alternativism, put forward a different view based on the phenomenological argument exemplified a little earlier. It is possible to aggregate meaning-complexes under verbal labels, but the constitution of meaning-complexes varies with the instance being construed. To subsume a set of instances under a bipolar construct implies that the bipolar construct must weight heavily (logically, restrict itself to) those aspects of meaning-possibility that are common across instances: in binary terms, matching patterns of check-marks in the (yet to be developed) grid matrix. In other words scales - if they are to approximate to the qualities of physical measurement - emphasise the normative at the expense of the distinctive. As exemplified in the simplistic example of Figure 8.16, the gradations of rating may well allow the possible discriminations of meaning to emerge. It is probably inadequate to assume that the uniqueness of instances will emerge from idiosyncratic patterns of ratings on broad construct dimensions: the picture that is shown by such a grid would seem likely to be about as fuzzy as a close-up photograph taken with the camera lens set for a distant panorama.

If, instead of emphasising the rows of a grid matrix one emphasises the columns, a qualitatively different picture begins to appear.

The instances begin to emerge in sharper detail because of the combination of both the normative and the differentiating in the acts of construing. Similarity and contrast are both present, but in a subtly altered form in comparison with Kelly's use of the terms. The focus is on the 'thing - as-it-appears', but there would seem to be a sharp contrast with the 'bracketing' of Husserlian phenomenology, since explicit use is made by the construer of the full richness of his or her construct system. With events being described in greater detail than is possible within the framework of a grid's scales (and one need look no further than Rowe's, 1978, book for evidence) it becomes unimportant to try to compare all the elements in terms of all the constructs. (I have the strong suspicion that many grids collect the answers to questions which are irrelevant to the researcher's true intentions, merely because a completed grid matrix is necessary for the analytical method to be used: a partially-complete matrix might suffice in terms of the information necessary for the problem at hand.)

As was argued in my discussion of the Dichotomy Corollary, construing of the instance becomes 'prototypic' under the particular phenomenological rubric outlined here[78]. It allows the exploration of the constructs used rather than restricts their use to the broad generalities typical of grid methodology. Like the Cheshire cat, the relevance of the grid as an entity begins to disappear though columnar concatenations of cells remain in view, intersected by fragments of rows reflecting the comparative aspects of construing - the whole taking on the appearance, perhaps, of a sparsely lighted crossword puzzle.

From such a standpoint rating and ranking decline to take on merely subservient and clarificatory roles. The quasi-physicalist assumptions of rating (here construed as involving the highly complex balancing of

semantics and action, and hence more problematic than is often taken to be the case) are thereby rendered comparatively unimportant. Pressed to its limits, the argument from phenomenology suggests that rating scales are at best a convenient fiction. At worst they are grossly misleading in that they purport to indicate a gradation of attribute which can be used without modification across a range of elements in a manner akin to a ruler measuring length and scales measuring weight. If, judged from the criterion of validity of communicated meaning, rating scales other than physical do not exist, those researching from a construct theory perspective will find that the implications are profound.

8.5 SUMMARY

This chapter has explored in some detail the nature of the processes of completing a grid matrix, focusing primarily on the practice of rating. A number of practical problems were identified, including the failure to follow the rating procedure accurately, accidental reversal of scales and uncertainties associated with the nature of elements and their context. The case was argued that most of these could be resolved by an approach requiring the researcher and respondent to engage in a systematic, structured dialogue in which the researcher takes the responsibility for recording the grid data. It was pointed out that, in many grids, the researcher can expect to find 'gaps' in the data: whilst this causes no problem in the completion of the grid matrix, the demands of analysis can lead to the dubious filling of the gaps by using the mid-point value of the rating scale.

The solution of the practical problems left a number of important conceptual issues unexamined, and these were taken up in the third section. Eiser and Ströbe's mediated stimulus-response model of

judgment was adapted to fit a Personal Construct Theory perspective on judgment, the point being made that most rating scales in repertory grid work are unlikely to be grounded in the assumptions of physicalist measurement, and may not be adequate vehicles for the communication of meaning between respondent and researcher.

The physicalist assumptions of the naive user of subjective rating scales were challenged in a number of ways which suggested that, if constructs were viewed as continua, the level of measurement could be regarded as at best ordinal in view of the range of possible manners in which metrical distortion could occur: possible influences here included anchoring, lexical marking, complexity of meaning and interaction effects. Empirical evidence was produced which suggested that individuals might construe a given bipolar construct in terms of various models, not all of which could claim to entail mutual exclusivity of poles: some of these models implied rating behaviour inconsistent with the limitations of bipolarity. The potential variation in scale use was supported by showing that the mid-point is open to a variety of interpretations, many of which are inconsistent with the principal of polar exclusivity.

Finally, the validity of rating scales in repertory grids was challenged from a phenomenological standpoint. It was suggested that an 'event' is an intersection between an instance and a complex of constructs, the complex being different for each instance being considered. If construing of elements in the typical grid rests upon what can be called up from episodic or semantic memory of experience rather than from some sort of internalised scale, then rating scales are called seriously into question, with potentially profound implications for methodology.

9 Analysis

...a dark
Inscrutable workmanship that reconciles
Discordant elements...

W. Wordsworth
The prelude (1805) I: 352-354

9.1 INTRODUCTION

Notwithstanding the challenge to rating scales made in the previous chapter, completed repertory grids have to be subjected to analysis and interpretation. This is not as simple a process as much of the literature implies, there being a tension between statistics and semantics.

A repertory grid produces a vast amount of information. A 10 x 10 grid, for instance, provides responses to one hundred different questions. The problem for the researcher is how to reduce this bulk of data to manageable proportions whilst retaining as much of the information as possible[1] (the 'minimax problem', as Kelly put it): the two aims are contradictory, and compromise is inevitable. In this chapter some of the issues pertinent to this compromise are discussed, particular attention being given to two analytical approaches: Slater's Grid Analysis Package and the suite of programs based upon Thomas and Shaw's FOCUS algorithm.

The analytical routines currently available gloss over a number of important issues, and these are addressed in the latter part of the chapter. The chapter ends with an echo of the radical critique of rating presented at the end of Chapter 8, in which the analytical correlate of an 'events-based' methodology is outlined as a 'trailer' to a fuller treatment in Chapter 13.

9.2 GRID ANALYSIS BY HAND

Kelly developed a nonparametric factor analytical method for the grid form of his Repertory Test[2], making full use of the assumption that the elements (dichotomously assigned to the poles of each construct) all fall within the construct's range of convenience. The factor analysis

routine, applicable to both rows and columns of the grid, is based on the construction of a hypothetical row[3] of ticks and blanks corresponding to the average distribution of ticks and blanks for all the rows taken together. This hypothetical row is then compared with each row of the grid taken in turn: where the number of 'matches' falls below half of the maximum total possible, the construct is 'reflected' (i.e. the verbal label and the pattern of ticks and blanks are reversed). This process is iterated until a maximal match between the hypothetical row and each individual row is obtained. The finalised hypothetical row then becomes the first factor to be extracted.

The rows which exhibit a significant match[4] with the first factor are assumed to load heavily upon it. These rows are then removed from consideration and further factors are extracted in the same way from what remains of the grid.

Although Kelly's procedure is simple and straightforward, it is also tedious and time-consuming in operation. The power of modern computers has transformed analysis from a laborious chore into an almost effortless exercise. This is not without its dangers, as Cronbach (1956) points out. Whilst he acknowledges the indirectness and exceptional flexibility of the repertory grid, he observes that the complexity of the data has lured Kelly's students into ' analysis so involved as to obscure serious errors in reasoning'[5]. Little (1977), reviewing the collection of papers in Slater's (1976) book 'Explorations of intrapersonal space', expresses concern that the sophistication of the statistical methods used is leading researchers away from the main focus of Kelly's philosophy, the problems of the real world. A broadly similar view is held by Fransella and

Bannister (1977), Watts and Pope (1982), and Adams-Webber (1979), the last of whom points to the 'confusion surrounding the meaning of many repertory-grid-based measures currently used in both research and clinical assessment'[6]. It would be unfair to lay all the blame for the problems of analysis upon computerisation, but the point can fairly be made that it is easy to slip into inferring that computer output is hard, objective and 'right'[7] - and just as easy to overlook the quality of the information loaded into the computer and the assumptions upon which the analytical routines are founded[8].

But not all analysis requires the sophistication of the computer. With small dichotomous grids 'hand focusing' may be adequate, all that is needed being a pair of scissors and a roll of sellotape in order to rearrange first the rows and then the columns (or vice versa) to maximise the match between the neighbouring strips of grid in both horizontal and vertical directions. Hall (1978) gives an example of hand focusing without explaining the steps needed to get to the required layout.

Fransella and Bannister (1977), whose main interest appears to lie in the relationships between constructs[9], describe how a ranked grid may be analysed by computing Spearman's rho for each pair of constructs taken in turn, the rank correlations then being used as the basis for a kind of factor analysis which can be represented graphically. They indicate that, in the example presented, there were high correlations between the first two axes of their analysis and the first two principal components of Slater's (1964) INGRID program, though the latter has the advantage that it allows the relationships between the elements to be explored. It would seem (having worked through their procedure) that Fransella and Bannister's method - if used in hand-

calculation form - suffers from a degree of laboriousness equivalent to that of Kelly's original method of analysis.

Another analytical approach which does not require the use of the computer, but which attempts far less than Fransella and Bannister's method, is the classification of constructs according to criteria which may be constructed a priori or developed from the information provided by the respondent. Landfield (1971) describes a category system designed to cater for the construing of people[10]. Fransella (1981) reports that it has proved less successful in Britain where a considerable number of constructs have been found not to fit Landfield's system[11].

It may be that classification is more successful when it is done on an a posteriori basis to suggest further lines of inquiry. The potential fruitfulness of the approach is suggested by Thompson's (1975) classification of constructs elicited from teachers at two different primary schools[12]. Whilst the teachers in one school produced a 'spread' of types of construct, those in the other emphasised personal qualities but not social behaviour. It would be dangerous to draw conclusions about the aims of education or the curricular structuring in the two schools from such limited evidence, but it does suggest potentially valuable lines of further inquiry.

9.3 COMPUTER ANALYSIS

There are several computer-based routines available for the analysis of repertory grids, and choosing from among them is not an easy matter - unless the range of choice is limited by what is available on the computer systems to which the researcher has access. Where choice exists, there is the danger that the researcher might give inadequate

attention to the assumptive bases of analytical methods[13], possibly even choosing a particular method of collecting data because it fits a particular mode of analysis rather than because it fits the questions which the researcher wishes to ask. That it is important to look inside the packaging of an analytical method is indicated by the disagreement between Rump (1974) and Slater (1974) over the relative merits of cluster analysis and principal components analysis, and Sneath and Sokal's (1973) comment that different methodological approaches require different analytical routines[14]. These last authors point out that principal components analysis is suitable for major clusters (but is 'notorious for falsifying distances between close neighbours'[15]), whereas agglomerative cluster analysis is reliable for within-cluster similarity though the reliability falls as the cluster size increases[16]. In other words, there is no single, obvious 'best buy' on the market: each situation has to be appraised anew regarding data analysis. Whatever routine is selected, it would be naive to assume that it will deliver a correct and fully meaningful output reflecting the way the respondent views the world - or, at least, that part of it providing the focus of the investigation. Even a part-world view is more complex than can be caught in the broad sieve of grid method.

Computer algorithms are sets of rules by which data is reduced to manageable proportions, and the patterns output by them depend on the ways in which these rules are combined. Shaw (1980a), for instance, shows how a change in an algorithm can result in a different clustering of elements[17], and (on a larger scale outside the area of repertory grid research) similar variation of clustering with algorithm is shown in the contrast between Bennett's (1976) original analysis of teaching

style and the subsequent reanalysis of the same data[18].

Evidence such as this tends to cast a cold shadow over Hope's (1969) somewhat breezy assertion that

> '... if he has prepared the data matrix in conformity with the requirements of the field of study, the research worker knows that his interpretation is not subject to the distorting influence of artefacts introduced by the computing techniques employed'.[19]

Hope's opinion also conceals the question of the quality of the data, a point clearly exposed in Shepard's (1972) recognition that some data in the social sciences are too fragmentary or unstable to support the determination of a well-defined spatial representation. And evidence has already been presented which suggests that the data in a grid matrix are likely to be surrounded by an atmosphere of uncertainty, and that different constructs may have been scaled in very different ways. Cattell (1957) recognises the problem in the context of research into personality, where scales reflecting both antonymic opposition and degree of possession of an ability may coexist[20]. However, Cattell seems to be unaware of the potential variations in oppositional scaling when he states that a single mathematical convention will enable the outcome of analysis to be mathematically correct. Correctness of mathematics there may be, but it may be making erroneous assumptions about the scaling actually being used - and the semantic aspects of the data are ignored, aspects to which a Kellian approach would probably give primacy.

Analytical techniques, then, all have their drawbacks. Blashfield (1976), writing of the emergent cluster analysis routines, advocated a sceptical attitude to the results from them: whether his scepticism implied outright dismissal is an open question, but scepticism's connotation of critical appraisal might well be taken up by the user of any technique for analysing repertory grid data. It is possible

to run a data matrix through a series of different routines and to identify the most consistently-obtained patterns (as Everitt, 1974, suggests, in respect of cluster analysis[21]), but that presupposes the availability of programs and time - both of which are often in short supply. A number of writers - among them Anderberg (1973), writing of cluster analysis, and Fiske (1971), writing of factor analysis - suggest that the program output should be used as the basis of further exploration rather than be treated as an end in itself: the output can be discussed with the respondent both to check whether it represents what he or she wanted to communicate and to allow elaboration of the original information which was given.

9.4 SPECIFIC METHODS FOR ANALYSING REPERTORY GRID DATA

Two methods for analysing repertory grids have become well established in Britain; Slater's Grid Analysis Package and the suite of programs based on Thomas and Shaw's FOCUS algorithm. Other methods do exist, though they have yet to become as widely used as the two mentioned above: Olson (1980a) used the principal components analysis PA1 from the SPSS suite; Boxer (1981) uses LITTLE NIPPER, his own creation; Pope and Keen (1981) describe Keen and Bell's GRIDDLE package; Smith and Leach (1972) used Johnson's (1967) hierarchical cluster analysis; Ravenette (1975) uses McQuitty's (1966) single linkage analysis; and Leach (1981) describes a method based on single linkage analysis which appears close to Thomas and Shaw's FOCUS but has the advantage of being able to cope with missing data. Such is an indication of the range of methods being made available to the grid analyst. The range is too large to allow most of the methods more than this brief mention, and discussion is limited to the central algorithms of the packages devised by Slater and by Thomas and Shaw: fuller details are given in Slater (n.d. and 1977) and in Shaw (1980a) respectively.

9.4.1 Slater's INGRID program

Slater's INGRID program, first published in 1964, has been developed over the years and in its currently-available form INGRID 72 forms the heart of the Grid Analysis Package. INGRID 72 applies principal components analysis to the grid matrix, and is capable of handling dichotomous, ranked or rated grids. For simplicity, the discussion presented below will concentrate upon rated grids.

Slater (n.d.) sees the most important advantage of principal components analysis as lying in the orthogonal nature of the components produced[22], the first component accounting for the largest proportion of the variance, the second accounting independently for the largest proportion of the variance remaining after the extraction of the first component, and so on. However, Fiske (1971), writing generally about factor analysis, points out that in personality research many observed dimensions exhibit considerable covariation and he leaves as a moot point whether the greater conceptual simplicity of orthogonal structure will prove to have greater theoretical advantages than oblique structures[23].

The heart of the INGRID program is the product-moment correlation coefficient. This has been widely criticised as a similarity measure[24] on the grounds that it fails to take into account both elevation and scatter, Gregson (1975) being particularly pungent in his dismissal of the use of this coefficient for indicating similarity as 'a statistically illiterate practice'[25]. The weakness of the Pearson r is simply demonstrated by examining the correlations of a set of ratings 1,2,3 with other sets 4,5,6 and 2,4,6: 1,2,3 correlates + 1.0 with each of the others although it is obvious that the inter-set distances are

different and quite marked[26]. As Brennan (1972) remarks, the Pearson r is only a partial measure of similarity[27], and Keen (1979) makes the further point that there is no guarantee that the construct ratings are wholly independent[28].

Slater's approach is also open to criticism on the grounds that, to produce a table of correlation coefficients and angular distances between the constructs, his method normalises the sets of ratings given on each of the constructs. Whilst there are statistical advantages in minimising skewness in the distribution of scale ratings, such a procedure inevitably removes some of the original meaning from the data. The researcher has the option to extend the use of normalisation to the full principal components analysis, and Slater indicates that the arguments for and against exercising this option are finely balanced[29]. On one hand, the rating scales are arbitary and likely to be incommensurate despite any restriction on length imposed by the researcher. Further, extremity of rating may not reflect a carefully measured distinction, for the respondent may in effect be dichotomising the ratings - in which case the variance contributed by such a construct would be disproportionate, a point of which Slater himself is aware[30]. On the other hand, 'one should not tamper with the evidence'[31]; the respondent has exercised choice in the manner of responding and this should be respected - and imbalance in response may be of signal interest. In his doctoral research Keen (1979) used the normalisation option, but records in a handwritten marginal note[32] that he would subsequently have reversed this decision, a position which reflection upon the issues involved would lead me to endorse.

A problem with the INGRID program is the decision regarding the number of components to be considered significant. The program includes a

routine devised by M.S. Bartlett which tests, via chi-square, whether the residual variation following the extraction of components is distributed randomly in the component space. The variation in the smallest component is tested against the penultimate component, and the routine sequentially tests the total variation in the n + (n-1) + ...(n-k) components against that in the k^{th}. Slater recognises that this procedure can give enigmatic results[33] (and I have found in one of my grids that nine components were significant according to this test, the ninth contributing a mere 0.52 per cent of the variance in the matrix[34]), and it seems reasonable to treat this test with a great deal of caution.

A further problem with Slater's method is the difficulty it has in coping with missing data. Slater points out that a single missing value prevents that construct from being located in element-space (or the element to which the same 'void' intersection applies from being located in construct space). He indicates that either the relevant row or column must be deleted[35], whichever would do least damage to the researcher's needs - and he does not 'take the easy way out' by suggesting the interpolation of a value such as that of the mid-point of the scale. In grids in which people are the elements there is often no problem with missing data since people all tend to fall within the range of 'people' constructs. In my experience (using the method described on page 254 for completing the grid) there have been very few cases indeed of missing data, and those few seem to have arisen by accident.

However, once the researcher moves away from grids dealing with people and moves on to use situations as elements, the range of convenience problem appears. It was noted earlier that Olson's (1980a) work with science teachers using classroom activities as elements gave rise

to some ten per cent of 'blanks' in his grids, whereas my broadly similar work was accompanied by 2.7 per cent of blanks. Appropriately dispersed, ten per cent of blanks could eliminate an entire 10 x 10 grid from INGRID analysis, using Slater's procedure! The problem increases with grid size; only seven per cent of blanks could eliminate a 15 x 15 grid. The inescapable conclusion is that for grids in which there is more than a very small proportion of blanks, INGRID analysis will be unsuitable if rows and/or columns are eliminated, since the program will only be able to output an analysis based on that part of the grid that has survived evisceration.

The great virtue of INGRID is its ability to present the various element-construct relationships in a single composite diagram. Whilst computing centres are now able to print plots of the location of elements and constructs with respect to <u>pairs</u> of components (e.g. Component I versus Component II), the presence of an orthogonal third component which accounts for a substantial part of the variance makes these two-dimensional plots difficult to interpret. Slater remarks that the first three components often account for over 90 per cent of the variance in grids of typical size, though in my 27 'science teacher' grids the percentage fell within the range 71-91 (mean 80; s.d. 5.6). Built into INGRID is the capacity to compute polar co-ordinates, which allows the location of elements and constructs with respect to the first three components to be displayed on the surface of a globe or, more usually, on some form of geographical projection. Allowing for the reservations expressed in the preceding paragraph (a few of my grids were noticeably pruned by Slater's procedure for missing data), I have found geographical projection to be helpful in obtaining a synoptic view of the various element-construct

relationships whilst being aware that to compress hyperspatial information into three dimensions is bound to introduce some distortion.

Element and construct orientations are specified by horizontal and vertical angles akin to longitude and latitude. Also computed is a radial measurement which indicates the extent to which each function is represented in the three-dimensional subspace of the hyperspace defined by all the components. Where this radial measure is comparatively small it suggests that the relevant function is inadequately represented within the first three components and that its relationships within the grid might be better revealed by inspection of the minor components and the original grid data.

Whilst the researcher can, by this device, obtain a space-satellite view of the respondent's world, the latter does not have the advantage of such detachment. Shaw (1980a)points out that the presentation of data which have been processed in a manner which mystifies the respondent can be very alienating[36]. It also runs the risk of presenting the respondent to him- or herself through the constructs implicit both in the computer program and in the interpretation, which is not in the Kellian spirit. In my work with the science teachers a few found difficulty in relating to geographical projections based on INGRID when I went back to them to discuss the outcomes of their analyses. The existential view at the surface of the world is not always commensurate with that of the astronaut.

9.4.2 Thomas and Shaw's FOCUS program

The problem of presenting computer output in a form intelligible to the respondent was taken into account by Thomas and Shaw in the

development of the FOCUS program from earlier work at Brunel University. The aim of FOCUS is to rearrange the rows and columns of a grid in such a way as to maximise the similarity between any row or column and its neighbour(s). The similarity measures used are based upon the Minkowski 'city block' distance metric, which is transformed by mappings which recognise the differences between elements and constructs. If d is the aggregated difference, disregarding sign, between two columns of elements, the similarity measure used in FOCUS is given by

$$\text{Element similarity} = \frac{-100d}{(n-1)c} + 100$$

where n is the maximum value on the rating scale (assumed to run from 1 to n), and c is the number of constructs. Element similarity values can range from zero (no match) to 100 (complete match).

Construct similarity has to take into account the possibility, because of bipolarity, of negative matching. The above formula is adjusted as below to give a range of similarity values from - 100 (complete negative match) to + 100 (complete positive match):

$$\text{Construct similarity} = \frac{-200d}{(n-1)e} + 100$$

where e is the number of elements and the other symbols retain the meanings from the preceding formula[37].

The FOCUS program 'reflects' constructs (and the relevant ratings) where an improvement in the positive matching between constructs can be achieved by so doing. This enables a dendrogram of positive inter-construct matches to be attached to the grid, similar to that possible in respect of elements. The practice of reflecting constructs is assumed to be valid, yet it might be questioned whether a respondent would produce exact reflections of ratings when the order of

presentation of the construct poles is reversed, i.e. the construct is presented as Y - X rather than as X - Y. The point might be more sharply pressed where the construct is 'peculiar' or 'X - not X' in character. However, the investigation reported in Appendix 6 indicates that this assumption is tenable in practice.

Shaw (1980a)gives the choice of the 'city block' metric a somewhat sketchy justification[38]. In preferring the city block metric to the Euclidean distance, she sees it as an advantage that the elements or constructs can be considered to be the same distance apart whatever the <u>pattern</u> of differences between the two sets of data. I would argue to the contrary: Figure 9.1, although a fraction of a hypothetical grid matrix in which the data represent an extreme instance, nevertheless illustrates the point I wish to make.[39]

CONSTRUCT	ELEMENT a	b	c	....	CITY BLOCK DISTANCES a-b	b-c	EUCLIDEAN DISTANCES a-b	b-c
A-A'	2	1	7		1	6	1	36
B-B'	3	2	2		1	0	1	0
C-C'	4	3	3		1	0	1	0
D-D'	5	4	4		1	0	1	0
E-E'	6	5	5		1	0	1	0
F-F'	7	6	6		1	0	1	0
					1	1	0.41	1
					Mean character differences		Mean Euclidean differences	

Figure 9.1 An illustration of the difference in effect of using the city block and Euclidean metrics in the computation of difference values (and hence similarity scores in the FOCUS program).

It can be argued strongly on psychological grounds that the difference between elements b and c is likely to be of greater interest than that between a and b because of the marked distinction in respect of

construct A-A': one might be tempted to place relatively little psychological significance on the one-point differences between elements a and b (which would, incidentally, show a product-moment correlation coefficient of + 1.0). The Euclidean distance, by summing squares of differences, would seem to capture more successfully than the city block metric the psychological distinction likely to inhere in the data[40].

In practice, however, it seems to make little difference which metric is chosen for cluster analysis. This appears to arise from the fact that ratings on constructs very rarely show such marked differences in the pattern of distances as is exemplified in Figure 9.1. Work reported in Appendix 14 shows that there appears to be little difference in the patterns of clustering produced by FOCUS (using the city block metric) and the cluster analysis program CARM[41], when the Euclidean distance is used, the rank ordering of element distances computed by the two metrics correlating very highly (Spearman rho ranging from - 0.95 to - 0.99). Unfortunately, it was impossible to undertake a direct comparison since FOCUS does not have an option to use the Euclidean distance and CARM does not offer the city block metric. Indirect comparison was possible by using the similarity and distance matrices output by the respective programs as inputs into the nonmetric multidimensional scaling program MINISSA - 1B (Roskam and Lingoes, 1970; Roskam, 1975). Two-dimensional plots were produced from six original repertory grid matrices for the FOCUS element similarities and the CARM Euclidean element distances, and only minor differences in configuration were observed.

Despite the conspicuous similarities of configuration resulting from the use of the city block and Euclidean metrics, it nevertheless seems

preferable on psychological grounds to use the latter. Where complex stimuli are involved, there is some evidence from studies of perception that the Euclidean distance is more appropriate where attributes not easily measured physically are involved. In contrast, for attributes such as size and form, the city block may be the more appropriate metric[42]. In most repertory grids it would seem that judgmental criteria are used which would suggest the use of the Euclidean metric, and it is worth noting that Burgoyne (1981) found a decision model based on this metric to be superior to both the city block metric and a metric based on the summation of the square roots of the individual differences[43].

Whilst the choice of metric is under discussion it is worth pointing out that the Euclidean distance could be useful where differences are being computed by both element and row in respect of two nominally similar grids[44]. A shift of one scale-point might not occasion much comment, but marked shifts might be of considerable importance. For instance, two shifts of three scale-points are likely to be of more psychological interest than six shifts of a single point - but such matters are easily available to direct inspection if each grid cell is divided diagonally with the first rating placed in one corner and the second in the other (cf Kevill et al, 1982). Coloured inks could be used to highlight changes, using a colour-code for both direction and intensity of change.

Like INGRID, the FOCUS program is unable to cope with 'gaps' in the data. Thomas and Shaw state explicitly that omissions or 'not applicables' must be entered at the mid-point of the rating scale[45]. The authors do not discuss the implications this might have for

clustering but, as was noted earlier, some types of grid are prone to 'gappiness' - in which case the arbitrary use of the mid-point might distort the relationships and meanings which the respondent wishes to convey.

McQuitty's (1957, 1966) linkage analysis[46] has been the basis from which the FOCUS clustering algorithm has been developed via an intervening program written by Thomas and Garnons-Williams (1973). The virtue of FOCUS is that it reproduces for the respondent, in an intelligibly rearranged form, the data that he or she originally supplied. The hierarchical clustering dendrograms show clearly what fusions have taken place, but the authors offer no guidance regarding levels of significance associated with the stages of clustering, and this makes the interpretation of clusters somewhat arbitrary. The separation of element and construct dendrograms makes it difficult to see quickly what the main relationships between elements and constructs actually are: this has latterly been improved in the SPACED variant which lays out the data in such a way that clusters of elements and constructs are separated by matrix spacings that are wider than average, thus giving an improved visual display[47]. In the two-dimensional display printed by FOCUS or SPACED, the interrelationships between <u>all</u> the elements (and, similarly, between all the constructs) cannot be depicted, and it is easy to infer, possibly erroneously, that the separation of one element from another by a number of intervening elements in the printout implies a large distance between the two elements being considered. It is important not to overlook the information available in the tables of element and construct matching scores, since it is in these that a truer picture of the interrelationships may be found.

The mathematical validation of FOCUS[7] is dependent upon a number of

assumptions which are not made available for scrutiny. If scaling metrics vary as much in practice as the previous chapter suggested might be the case, mathematical validation is an extremely complex process indeed, and involves a consideration of the semantics involved in scaling. Shaw and Thomas (1978) get the matter in better perspective when they write:

> "As an articulator of conversation, the focused grid is a crude but useful tool. It is the <u>beginnings of a psychological reflector</u> which can reflect back to a person a view of himself as seen with his own eyes." [48]

9.4.3 <u>Cluster analysis and profile</u>

If the primary interest of the researcher lies in the elements it may be preferable to use cluster analysis to identify groups of like elements which can then be presented in profile form against the constructs[49]. In the examples given by Youngman (1979) a number of groups are shown superimposed on the same profile blank: however, my own interests have lain in the characterisation of individual groups of elements (rather than with inter-group comparisons), and I have deliberately separated the profiles of the identified groups. I have used the cluster analysis program CARM with the error sum metric option, judging group membership by inspection of the printout (in particular taking into account sharp increases in the error plot, which signify the clustering of rather dissimilar sub-groups). I plot each group's average ratings on a separate, previously prepared blank profile form, indicating where intra-group variations occur (see Figure 9.2). In the context of working with the science teachers, this provided a convenient summary of element-construct relationships. It will be noted that there is little emphasis on

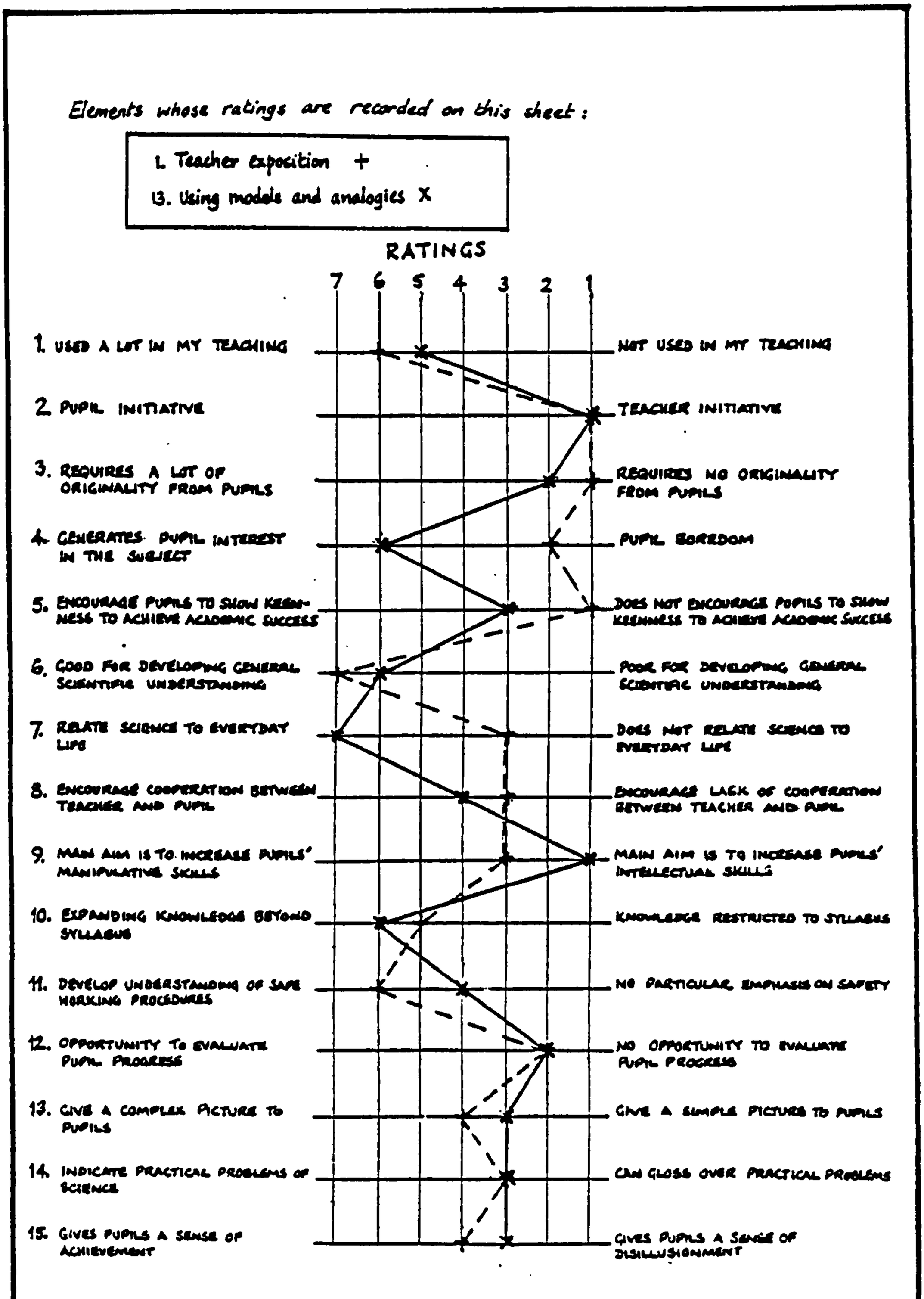


Figure 9.2 An example (size reduced) of the profiles produced from the analysis of N.X.E.'s first grid. The CARM cluster analysis had suggested the closeness of Elements 1 and 13, and their separation from the remaining elements.

construct interrelationships: this arises from my interest in the ways in which respondents construed the elements, aspects of science teaching. Where construct interrelationships are also important, Hope's (1969) suggestion of a combination of factor analysis and cluster analysis might be worthy of exploration. Hope advocates factor analysis for the variables (i.e. constructs), allowing their interrelationships to be shown by contiguity in the profile (somewhat akin to FOCUS and SPACED), and the use of taxonomic techniques (such as cluster analysis) for the entities (elements).

Cluster analysis followed by profiling overcomes some of the problems associated with INGRID and FOCUS, in particular that of missing data. With one exception which was too 'gappy' for this approach, clustering of the elements was performed with respect to all constructs for which there was a complete set of ratings. Constructs with missing data could be added to some group profiles where all the members of the group had been rated, but the 'gappiness' of such constructs excluded them from contributing to the determination of group membership by cluster analysis.

If the centre of interest lies in the similarities between elements, Youngman's set of randomly-generated distances provides a yardstick against which the significance of any observed distance may be tested. He has available a range of distances for different sizes of grid, with cut-off points at the 1 and 5 per cent levels for the smallest distances, and 95 and 99 per cent levels for the largest distances: a sample is given in Youngman (1980).

Cluster analysis procedures continue to be developed and, despite

Blashfield's (1976) scepticism regarding the available cluster analyses, he suggests that the centroid relocation method (of which CARM is a representative) is superior to the linkage method used by workers such as Rump (1974), Ravenette (1975) and Thomas and Shaw (n.d.). However, this superiority is based on moderately large samples and it is unclear whether it extends downwards to the clustering of small numbers of instances typical of repertory grid data.

9.5 MINIMISING THE PROBLEM OF MISSING DATA: ONE APPROACH

The problem of missing data is not solved satisfactorily in any of the three methods of analysis discussed above. Though the approach using cluster analysis and profile offers a partial solution, it is still vulnerable to a wide dispersion of blank cells in the grid matrix. Frane's (1976) method of interpolation, based on regression techniques, does not seem to offer much practical help to the grid user since the three assumptions which it requires are difficult for the grid user to sustain:

(i) data must be missing at random;

(ii) each missing variable must be highly correlated with one or more of the available variables; and

(iii) the amount of missing data should not be large.

Given the likely quality of much grid data, it seems improbable that satisfactory interpolations can be made where data is missing. It would seem better to accept the fact that some gaps exist and to adopt an analytical strategy which minimises their disruptive influence. One strategy to do this is to compute, as before, the distances - the same applies to similarities - between elements but to build into the

program the ability to delete from consideration any pair of entries containing one or two blank cells. If this deletion is applied only to pairs of entries instead of the whole construct row, some pair-data is retained instead of being lost and the distance measure is minimally impaired. The hypothetical example of Figure 9.3 illustrates the point.

CONSTRUCT	ELEMENT a	b	c	....
A-A'	1	1	5	
B-B'	5	-	2	
C-C'	-	6	1	
D-D'	4	6	7	
E-E'	2	4	1	
F-F'	7	3	2	
....				

Figure 9.3 Illustration of a 'gappy' grid.

The distance between elements a and b is calculated with respect to constructs A-A', D-D', E-E' and F-F'; between a and c with respect to constructs A-A', B-B', D-D', E-E' and F-F'; and between b and c with respect to constructs A-A', C-C', D-D', E-E' and F-F'. The deletion of whole rows of the matrix would have meant that B-B' and C-C' would not make any contribution whatever to the distance measures. Having been faced by me with this problem Youngman developed a program to compute distance measures on this basis and to output a triangular matrix of interelement distances[50] together with the number of paired observations taken into account in computing each interelement distance. This I have found very useful in the processing of 'gappy' grids.

If the distance measure is calculated as the 'mean Euclidean difference'

(or in city-block terms, Sneath and Sokal's, 1973 'mean character difference'), the individual inter-element distance measures are rendered comparable, and the distance matrix itself may be input into a multidimensional scaling program to enable a visual display of clustering to be presented. Alternatively, the distance measures can be checked for significance against the tables produced by Youngman (1980), care being taken to allow for the number of paired observations relevant to each distance measure.

9.6 PROBLEMS OF INFERENCE

The analytical techniques briefly considered above have in common the ability to present their particular forms of data-reduction in a structured format. The point was made earlier that different methods of presenting the output of analysis are likely to prove differentially meaningful to the respondents. But what of the researcher? What is legitimate for him or her to infer from the output?

In most circumstances it is difficult to decide what constitutes a cluster of elements or constructs[51]. Shaw (1980a) appears to leave the decision to the researcher's subjective opinion, following further discussion with the respondent. The 'error plot' shown in the CARM cluster analysis output offers some guidance to the composition of clusters but is not always a clear enough aid to decisions. In Slater's INGRID analysis tables of correlations and angular distances between elements and constructs are provided, the former allowing tests of significance to be made. However, it should be noted that the geographical projections can distort the hyperspatial relationships quite considerably and that visual inspection of the projections can therefore be misleading[52].

Whichever analytical approach is adopted, there seems to be a strong argument for treating the analytical output as the basis for further exploration rather than as an end in itself. To accept this would seem to accord with Kellian philosophy and such writers on analytical approaches as Fiske (1971), Anderberg (1973) and Shaw (1980a).

9.6.1 Individual grids

In an ambitiously-titled paper Smith and Stewart (1977) suggest that the repertory grid is an appropriate technique to render people's mental maps objective and explicit. Disregarding the embedded assumptions, it is instructive to examine the empirical evidence of Joe's grid from the point of view of the inferences made. Joe construed predominantly a variety of workmates in terms of constructs which were largely work-oriented. The INGRID analysis shows Joe as closer to his ex-mates than to his wife, and the authors interpret this in terms of his being 'a very traditional working class male' without considering the likelihood of contextual bias in the grid and the narrowness of the basis of the judgment[53]. To interpret thus is to jump to conclusions, and in the following paragraphs evidence is presented showing that this is not an isolated case. The point being made is that interpretation on the basis of grid evidence alone is likely to be inadequate unless it is linked to further correlative information and, perhaps, exploration.

An opportunity for the further exploration of grid analysis output is available in the TARGET teacher self-appraisal grids of Hopwood and Keen (1978), and the authors report that about 60 per cent of their respondents availed themselves of this option[54]. Judging by the example given, this is not surprising. However, the irony

here is that the respondents, whose construing has been processed, are asking the consultants what they themselves meant. The exploration is an inversion of what a researcher would need to do.

In its feedback to respondents TARGET presents three bar charts which are 'weighted profiles' of effective teaching, 'self' teaching and ineffective teaching. Each bar indicates the percentage contribution of each of the components found to be significant (in the example given there are four which together account for 65 per cent of the variance), and the four constructs listed against each bar are listed in descending order of capacity to define the component in question (although the visual display suggests equivalence in this respect). Some bars appear as negative weightings, yet where this occurs it is by no means clear whether this represents some kind of <u>lack</u> in terms of the construct poles listed, or whether it represents the qualities denoted by these poles. How the bar chart helps the respondent to define and clarify his or her superordinate construct system is not made clear[55].

The analysis, based on Slater's principal components analysis, can produce strange components. For instance, in the example provided by the authors, Component 3 indicates both poor subject knowledge and nervousness as part of the profile for effective teaching. Whilst it is possible to envisage a positive correlation between nervousness and effective teaching[56], it is hard to see how poor subject knowledge can be positively associated with effective teaching. The suspicion exists in my mind that this outcome is an artefact of grid methodology and/or analytical approach rather than a reflection of existential reality. I have, for the purposes of making this point in a seminar,

concocted a grid whose analysis by INGRID 'showed' Margaret Thatcher to be a man - all that was necessary was to construct high inter-correlations between (stereotypically) masculine-feminine constructs relating to 'action' and the construct 'male-female': Margaret Thatcher, as a 'woman of action', was distorted by INGRID and the three-dimensional geographical projection to come close to the 'male' pole of the 'male-female' construct.

This critique of the TARGET output accords with Shaw's (1980a)desire to avoid mystifying the respondent when presenting grid feedback, and to use feedback as an opportunity to elaborate on what has already been contributed to the grid.

There is a danger in repertory grid analysis that the researcher will seek to identify and label factors or components from analyses, in a sense reifying them. Hope (1966) suggests that Slater's use of polar co-ordinates avoids the reification of factors on the grounds that the geographer's reference lines are arbitrary[57]: the thought surfaces that this hope might be arbitrary, too.

Kelly was very much aware that to interpret the outcomes of analysis was to construe, with respect to the researcher's own construct system, the construct system of another[58]. He was also very much aware that, whatever technique of data-reduction was used, the factors arising from it only related to the elements and constructs in the grid itself, and that generalisation beyond the framework of the grid was unwarranted[59]. Yet once an interpretation is offered for any set of data, that is tantamount to naming factors - however provisionally - and to establish-ing a cognitive 'set' in the mind of the researcher (and perhaps in the

mind of the reader of the research report). Even if the interpretations are themselves construed as speculative hypotheses, they nevertheless tend to define and delimit the boundaries of future exploration. And people are not in general unbiased scientists (if extrapolation is made from Snyder and Swann's, 1978, findings) in that they show a distinct tendency to search for supportive rather than refutational evidence regarding hypotheses in the field of interpersonal relationships.

A number of writers appear to assume that clusters of constructs can be subsumed under a superordinate label which may never have been articulated by the respondent[60]: that is, factors derived from analysis can be identified as representing superordinate constructs. Others label as superordinate that construct which has the highest number of significant correlations[61]. Makhlouf-Norris et al (1970) develop the latter theme by claiming that correlational analysis of grid data allows the hierarchical level of constructs in the system to be assessed[62]. Correlational analysis, as was pointed out in Chapter 4, does nothing of the sort. Whilst it may indicate the degree of 'overlap' between constructs in terms of shared variance, of itself it provides no direct evidence whatever about the implicative relationships involved, nor does it necessarily carry any implications about the relative importance of constructs to the respondent. Simply put, these authors are confusing structure and cluster.

Sneath and Sokal (1973) make the point that clusters are generally based on phenetic resemblances and have no necessary phyletic connotations[63]. Patterson(1982) makes it clear that the set-inclusions in cladistics (comparable to the situations where clustering dendrograms are 'cut off' at particular levels) can only give clues as

to possible phyletic relationships but make no specific claim for phylogeny: there can be many different 'trees' underneath any particular pattern of set inclusion. As far as biology is concerned, the structural relationships are almost invariably more complicated than the phenetic approach of Linnaean classification would suggest. The complexity of construct systems would suggest an extension of Patterson'spoint to repertory grid data.

Pattersonmakes the further point that cladograms (cladistic 'trees') vary according to the level of the data that are used. If hominoids are classified at the level of observable physical features one obtains a different cladogram from that based on molecular data. In other words, there is a problem of deciding on the relevant features to be considered - a problem that users of repertory grids gloss over in their treatment of constructs as equivalent irrespective of their degree of superordinacy.

A further confusion in interpretation derives from the failure of some authors to recognise the ambiguity inherent in the word 'hierarchy'. Applied to construct systems it implies the notion of superordinacy/ subordinacy, whereas when it is applied to cluster analysis it refers to the sequence of operations (divisive or agglomerative) undertaken in the search for the most informative way of clustering the cases being considered. With respect to cluster analysis, 'hierarchy' contains no necessary implications for superordinacy/subordinacy in the psychological sense. This confusion lies at the heart of Smith and Leach's (1972) investigations into cognitive complexity. They assume, following Kelly, a hierarchical organisation of construct systems but then use this as a justification for using Johnson's (1967)

cluster analysis on their correlation matrices because it 'provides a hierarchical picture of the way in which the constructs group together'[64]. In fact, they have fallen into the same trap as Makhlouf-Morris et al (1970), for Johnson's method is no more than a method of reducing the complexities of a correlation matrix to manageable proportions. Smith and Leach assumed that pairs of constructs were functionally equivalent when the correlations between them were significant at the 5 per cent level, and that such pairs could be collapsed into new single constructs. The new reduced correlation matrix was then subjected to cluster analysis, the assumption being made that if the fine detail of construct systems were more important for complexity the impoverishing of the grid matrix would have a more dramatic effect on the clustering of the elements for cognitively complex respondents - and thus lead to a new measure of cognitive complexity. The shaky foundations of this exercise, together with the slender relationship between their results and those based on the Bieri et al (1966) measure of cognitive complexity (itself a dubious index), give few grounds for confidence that they found what they sought[65].

9.6.2 Aggregated grids

Interpretation presents considerable problems when grid data collected individually is subjected to a nomothetic analysis. Both Slater (1977) and Shaw (1980a)have developed programs which can compare individual grids and which can analyse aggregated grids[66]. Slater's PREFAN analysis, for example, aligns all the grids by element to give, in Harrison and Sarre's (1975) term, a 'supergrid' composed of a set of common elements and all the constructs produced by the respondents. The use of the routine implies that the elements are identical for

the respondents, but the output can still prove difficult to interpret when a long 'tail' of components each accounting for a small proportion of the variance is obtained. Riley and Palmer's (1976) study of the construing of 25 holiday resorts by 60 respondents resulted in a 25 x 672 grid which produced 24 principal components, the first six of which accounted for 16, 11, 10, 6, 5 and 4 per cent of the variance respectively (52 per cent of the total), and it is evident from their account that they were struggling to interpret this diffuse mass of data.

Where elements are not identical for all respondents, such as when idiosyncratic nominalising of role titles is used (e.g. Langrish and Smith, 1979) the validity of PREFAN analysis is very much open to question.

The extension of the 'supergrid' horizontally, by keeping the constructs constant whilst allowing the elements to vary from respondent to respondent, can be accommodated by Slater's ADELA program. Here the assumptions seem to be on shakier ground than in PREFAN analysis, since the constructs (which are already abstractions) are assumed to be commonly construed by all respondents. The polysemic character of language, and the individual differences of experience acknowledged in Kelly's Individuality Corollary, together make ADELA analysis a dubious undertaking.

Grids in which both elements and constructs are supplied also make the massive assumptions of common experience and meaning, and these assumptions are extended if the analysis seeks to extract regularities from the superimposed data matrices.

The problems associated with combining data are exemplified in a study, reported by Perrott et al (1976), which explored teachers' reactions to a self-instructional microteaching course on 'Effective questioning'. The teachers attended the course at one or other of two centres and, as part of the feedback procedures, were asked to complete a repertory grid in which seven people nominated in respect of role titles constituted the elements (Figure 9.4).

1. Your own teaching as you see it now.
2. Any teacher whose work you generally respect.
3. Any specific teacher whose work you think is in general rather poor.
4. Your own teaching as it was before the microteaching course.
5,6. Two other specific teachers whose work you know.
7. The sort of teacher the microteaching course seems to try to produce.

Figure 9.4 Role specifications for the elements used in Perrott et al's (1976) evaluation of a microteaching course.

The seven elements were then rated on fifteen supplied constructs which were presented in a semantic differential format. It was noted earlier that the constructs appear not to be closely related to the aims of the microteaching course, but rather to teaching at a more general level[67].

Disregarding the problem of treating individual teachers and teaching as elements at the same level of abstraction, considerable difficulties present themselves when the authors' analytical procedure is examined. They first produced two 'average' grid matrices from the two centres and, on finding a high correlation between these 'average' grids,

decided to pool all the data into a single 'consensus' matrix which was then submitted to principal components analysis. It was found that the first two components accounted for 88 per cent of the total variation in average ratings, and these were interpreted as relating to 'professional competence' and 'degree of formality'.

It is not surprising to find such a high percentage of the variance explained by the first two components, for the averaging procedure will have concealed variation at the individual level: the'average' grid represents the means of the ratings, but the importance of the standard deviations has been overlooked. Perrott et al are also assuming the respondents' consensus regarding their understanding of the constructs and the equivalence of the respective elements. The former is questionable, the latter invalid. None of the seven elements is common to all the teachers - in fact, it is highly likely that each of the seven elements was unique to each individual. The diagram purporting to show the interrelationships between elements and constructs is therefore meaningless: there is surely no one point at which 'respected teacher', for example, can exist for all who completed the grids. A diffuse cloud drifting in the wind of social desirability would seem more likely to represent the reality of this data than a single point in component space, though it would be probably no more helpful to the interpreter. And what is meant by the two points on the diagram representing elements 5 and 6 (two other specific teachers)? These completely defy interpretation, for they can in no way be termed the average teacher, or teachers in general[68]. Their function in this research is unnecessary and totally redundant: perhaps they serve no more useful purpose than a coccyx on the evolutionary rump of Kelly's role titles.

Had Slater's ADELA program(which aligns grids by common constructs) been available to Perrott et al, some of the above errors would have been avoided[69] - though the conceptualisation of the problem which they sought to investigate via grid methodology and the associated research design would still have remained open to question.

It is perhaps unfair to single out this research paper for detailed criticism, save that it does highlight a number of traps for the grid user with a penchant for the nomothetic approach. Harrison and Sarre (1975) who investigated the construing of Bath by twenty female residents, also used a consensus matrix based on supplied elements (locations in the city) and constructs. In contrast to Perrott et al, their elements were common to all respondents. Not surprisingly, principal components analysis revealed a massive first component accounting for 79 per cent of the variance. Subsequent PREFAN analyses (using idiosyncratic constructs) gave a more meaningfully-interpreted output, the details of which are not of concern here.

The tendency to work with averaged data has existed for a long time in psychological research and its normative nature is often unquestioned. This is not to attack 'mass' statistics as such, but to point out that the focus of the research is often ill-served by averaging the numbers collected: whilst the mean may be important, there is considerable interpretive value in the standard deviation. Thus Watson et al (1976), investigating stress in long-term prisoners, might well have obtained more understanding of the problem by looking at individual construing instead of aggregating ranked grid data to produce two major components accounting for 92.5 per cent of the variance. Similarly Jones et al (1980), in asking students and teachers to rate (on 33 supplied constructs) 16 vignettes of models of teaching derived

from Joyce and Weil (1972), lose the richness of the data gained by only analysing a matrix of median values.

Though there have been sporadic papers in the literature pointing out the fallacy of averaging (e.g. Sidman, 1952 ; Bakan, 1954; Baloff and Becker, 1967) it is rare for a writer to have looked into the variations contained in the data and to have realised the implications of a nomothetic analysis. Stringer (1972), for instance, observes that the nomothetic analysis of grid data 'hid' a substantial number of deviant idiographic responses - in fact, ten out of 34 - but he does not pursue the issue; and Katz (1982b), in a smaller-scale study of the relationship between physiology and construing, gives evidence showing that one response of five ran counter to the more general trend[70]. Whilst acknowledging that not all of their samples ran true to the trend of the majority, neither of these writers look for a superordinate explanation.

At the level of theory du Mas (1955) is interested in the single case and discusses what he terms 'idiographic nomothesis', which amounts to the generalisation by averaging from the number of individual cases involved[71]. In Chapter 13 I argue that this type of approach to generalisation is inappropriate to studies in which the individual's construing or behaviour is the centre of interest and that research based on a personal construct position requires a radically different concept of generalisation.

9.7 SOME FURTHER UNRESOLVED PROBLEMS OF GRID ANALYSIS

A number of the problems already mentioned in this chapter remain unresolved in repertory grid work even though they have been

recognised by some researchers. There are, however, three further problems - there may well be more - associated with analysis that seem to have received little attention to date: the 'hairpin construct', highly correlated constructs and the inequivalence of constructs.

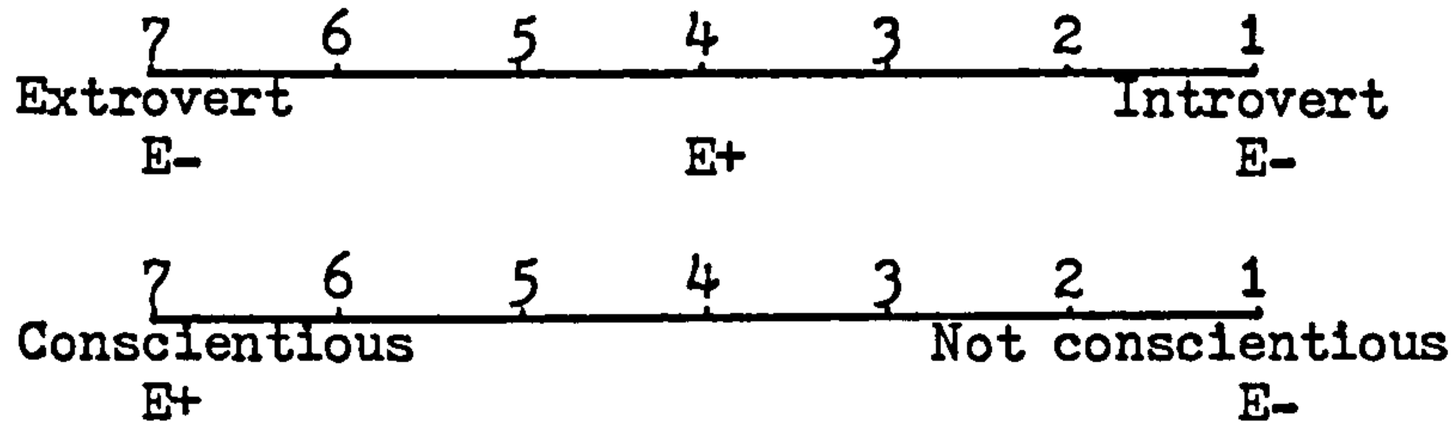
9.7.1 The 'hairpin construct'

The 'hairpin construct' is a construct which would probably find general acceptance as a 'straight' oppositional construct (likely to represent the scaling of an attribute) yet which can appear to be sharply bent when set alongside other constructs. Consider the construct 'extrovert-introvert', for example. Many people, to judge by the study of rating scales described in Chapter 8 (page 284f), appear to give the most positive valence to the middle of the construct on the ground that it (E+) represents some notion of normality, whilst both ends are negatively valenced (E-) in that extremes of personality are considered undesirable and perhaps pathological. This E-/E+/E- 'evaluative overlay' would seem to apply to many constructs (especially personality constructs) in which a 'happy medium' could be said to exist.

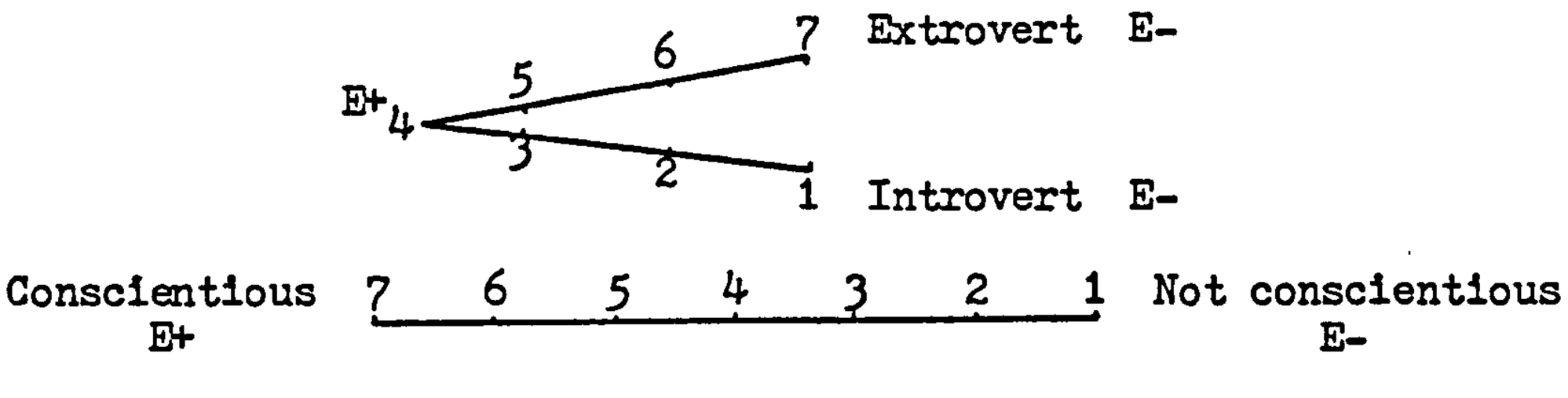
Suppose now that the grid contains a number of constructs whose evaluative overlay is monotonic with the rating continuum (for examples 'conscientious - not conscientious'). Analysis treats the constructs as mathematical entities, aligning them according to the co-distributions of ratings: the situation is represented symbolically in Figure 9.5(a).

But place an E-/E+ construct together with an E-/E+/E- construct in

the grid, and the <u>evaluative</u> relationship is more complex than the mathematical, yet the grid analysis will treat the E-/E+/E- construct as if it were E-/E+. Aligning the two constructs in terms of their valence shows how the E-/E+/E- construct bends to a hairpin under the pressure: this situation is symbolised in Figure 9.5(b).



(a) Symbolic representation of mathematical alignment



(b) Symbolic representation of evaluative alignment.

Figure 9.5 Alignment of constructs according to the criterion of (a) mathematics; (b) valence.

The mathematical analysis will not - it cannot - reveal the evaluative relationships latent in the data. An orthogonal relationship is insufficient on its own to indicate whether there is a complex evaluative relationship underlying the mathematical calculations.

It might be felt at this point that the argument is trivial, but it seems distinctly possible that the problem may pervade grid work involving 'person perception'[72] and could extend further. Repertory grids dealing with children or teachers often focus on qualities, many of which have a simple evaluative structure monotonic with the

numerical scale being used. However, there are likely to be many instances where the mid-point is the most positively evaluated (perhaps particularly with respect to extremes of behaviour) and thus would render evaluative interpretations problematic.

As cases in point, a number of the constructs supplied by Perrott et al (1976) come into this category - for example, 'realistic-idealistic'; 'pupils usually talking-pupils usually listening' and 'teacher controls course of discussion-pupils control course of discussion'[73]; it will be noted, in the light of Chapter 8, that none of these constructs has an easily-discerned oppositional evaluative structure.

The problem of the 'hairpin construct' is, as far as I know, unresearched in repertory grid work since I first mentioned it (Yorke, 1978). A tentative solution advanced in that paper was to split the 'hairpin construct' at the middle and treat it as two separate halves at the cost of obtaining two vaguely defined poles at the point of severance. That solution did not go far enough, however, since it implicitly assumed that the researcher would know when the respondent was supplying a 'hairpin construct'. It would therefore seem desirable to ask the respondent where the most positively valenced and most negatively valenced points are on any construct, perhaps splitting potential 'hairpin constructs' into component parts. Though this practice remains somewhat crude the researcher can obtain some idea of how the respondent feels about the construct and the elements located on it, and it is likely to be rather simpler to use in practice than a technique such as Coombs's (1964) unfolding procedure[74].

9.7.2 Highly correlated constructs

Highly correlated constructs pose a problem for the grid analyst since it is difficult to determine the extent to which they are replicating the same information content, and hence to determine the degree of redundancy in the grid.

When highly correlated constructs are present in the grid principal components analysis will reveal this through their loading heavily on to particular components (usually the first). There are obvious advantages in reducing the data matrix to a form that is more easily interpreted. The price to be paid may be the presence of a first component accounting for a large proportion of the variance, leaving other components to share out the residue. In the process of sharing out this residue such components (which might largely represent single constructs uncorrelated with the majority) could become underemphasised. The problem for the researcher is that he or she does not know whether the respondent really does construe the part of the world under investigation in a predominantly unidimensional way, or whether such unidimensionality has arisen as an artefact of methodology.

Hierarchical cluster analysis of elements presents the problem less obviously. The closer the relationship between a group of constructs, the more this group will influence the dendrogram showing the element clustering. Not only could a cluster of elements appear 'tighter' than it might really be, but also the effect of construct inter-correlation could cause elements to be 'misclassified' into groups during the clustering routine. As with principal components analysis, there is no way of telling from statistical analysis whether the picture revealed by the computer is a fair representation or a distortion of the respondent's world. The two-way cluster analysis in Thomas and

Shaw's FOCUS program does reveal the high associations between constructs and between elements (the SPACED version highlighting the clusters visually), allowing the thoughtful researcher to ponder upon, say, the clustering of the elements in the light of construct intercorrelations. However, the determining effect of intercorrelations upon the output of the cluster analysis algorithm remains beneath the surface.

The same problem exists with the 'cluster and profile' approach, more severely if no attempt is made to look at the interrelationships between the constructs. Hope (1969) argued for the variables (constructs) to be subjected to factor analysis and for taxonomic techniques to be used for the entities (elements). 'Cutting off' the hierarchical taxonomic analysis at a suitable point would enable profiles to be drawn, the variables being rearranged according to the outcomes of the factor analysis (in the interests of maximum clarity), This does not, of course, remove the problem of interpretation when highly correlated constructs are present in the grid.

The essential point to be made here is that the grid, coupled with whichever form of statistical analysis is adopted, does not in itself provide the answer to this dilemma of interpretation. Again the need is emphasised for a more extended dialogue with the respondent than is likely to occur during the completion of a single grid, for the researcher will probably be unable to detect and probe high intercorrelations as the elicitation of the grid develops unless he or she is using an interactive computer program with the respondent.

9.7.3 Construct equivalence/inequivalence

The analytical routines discussed in this chapter all weight the

constructs equally, yet it is apparent both from life-experience generally and from specific techniques such as resistance-to-change grids that some constructs weigh more heavily than others. There is a further problem in that, as Kelly recognises, construct systems are not rigid and fixed in their relationships but labile, their component constructs varying in importance depending on the circumstances[75].

Honikman (1976) found that occasionally a subordinate construct proved more important than others which were superordinate according to the canons of laddering [76]. Bearing in mind my earlier criticism of Hinkle for laddering in the abstract, it may be that Honikman's respondents were switching from an abstract (espoused) conceptual framework to a specific action-oriented framework (theory-in-use)[77] in which certain 'low-level' constructs emerged as salient and potential determinants of action - that is, the context promoted them to superordinacy[78]. Reflection on human action at points of crisis suggests that this is not an unreasonable interpretation - and that the 'elevated' constructs may be given a comparatively superordinate position in the 'espoused' system long after the precipitating event has passed.

It is doubtful whether a useful distinction can be drawn between salient and superordinate constructs if both are seen in terms of having more implications for others than the reverse. One cannot draw a sharp contrast between 'espoused theory' and 'theory-in-use' in terms of action, for the giving of an espoused theory to an interviewer is itself an action governed by that particular context. If espoused theory and action are discrepant - as they may turn out

to be - then it is the case that construct systems are more unstable with respect to circumstances than the personal construct theorist, looking for something other than shifting sands in which to anchor a theoretical position, would prefer to admit.

Although various attempts have been made to rank constructs in order of importance, or to rate their degree of importance, the information seems not to have been incorporated into the analytical routines themselves. Some use has been made of weighting where research has been undertaken in the field of decision-making. McKnight (1977) managed to get seven respondents to work through an extremely lengthy protocol for determining their preferences for long-playing records in a variety of circumstances. Having completed this protocol, his respondents were asked to rank their elements in order according to each of the various purposes: from the four who completed the research task without interference from extraneous matters, McKnight obtained a series of correlations between 'protocol preference' and subjective ranking ranging from + 0.782 to + 0.976. He concluded that these correlations supported the preference-determining protocol he had used. One might be tempted to conclude, on the basis of these figures, that the simple subjective ranking procedure was (on the criterion of the effort needed to acquire the information) equally valid and distinctly preferable.

McKnight's work related to element preference. In contrast, Gardiner and Edwards (1975) describe an approach to social decision-making which takes into account construct preference. They asked their respondents to rank a number of constructs relating to property development (such as the area of the development in square feet, and the aesthetics of the development running on a subjective scale running from poor to

excellent), and then to transform the rankings to ratings on a scale ranging from 0 to 100 whilst trying to preserve preference ratios. Each construct's raw weight was then divided by the sum of weights for all constructs, giving each construct an individual percentage weight. Each proposed development (i.e. each element) was then rated on each construct (scaled from 0 to 100), and the utility of each proposal was calculated by summing its weighted ratings on each construct, each rating being multiplied by the weightings computed for the constructs. As with McKnight's work, the procedure is complex and time-consuming and is likely to find few adherents in repertory grid work.

Preference does not necessarily imply superordinacy since the constructs could be subsumed ordinally (at the same level of abstraction) by the dimension of preference. Some connection may exist in a rather general way which may not be sustained in respect of individual cases. On the whole the issues of superordinacy, salience and preference have been neglected by the grid analysts - perhaps because of the complexity of the problems involved. Yet to fail to grapple with these problems is once again to miss an opportunity of enriching one's understanding of the data contained in the grid matrix.

It may be that the answers to the problems of salience and super-ordinacy are not to be found in ever more sophisticated methods of analysis. It is necessary to treat analytical routines with a measure of circumspection, and to support these by an exploration with the respondents of the importance of elements and constructs. Though such information may be no more than qualitative, it may be sufficient

to give the researcher a level of understanding adequate to the purpose of the investigation.

Sneath and Sokal (1973) make the point that a priori weighting is objectionable since it presupposes a sound basis for the weighting[77]. Unless the researcher can obtain a set of meaningful weightings in the course of administering the grid, it is probably least damaging to assume equivalence of weight. It could, however, lead to the sort of distortion that was earlier noted in respect of highly correlated constructs, and it would seem highly desirable that the 'hard' assumption of weighting equivalence be softened for the purposes of interpretation as a result of using of appropriate questioning techniques during grid administration.

9.8 CONCLUDING HERETICAL POSTSCRIPT

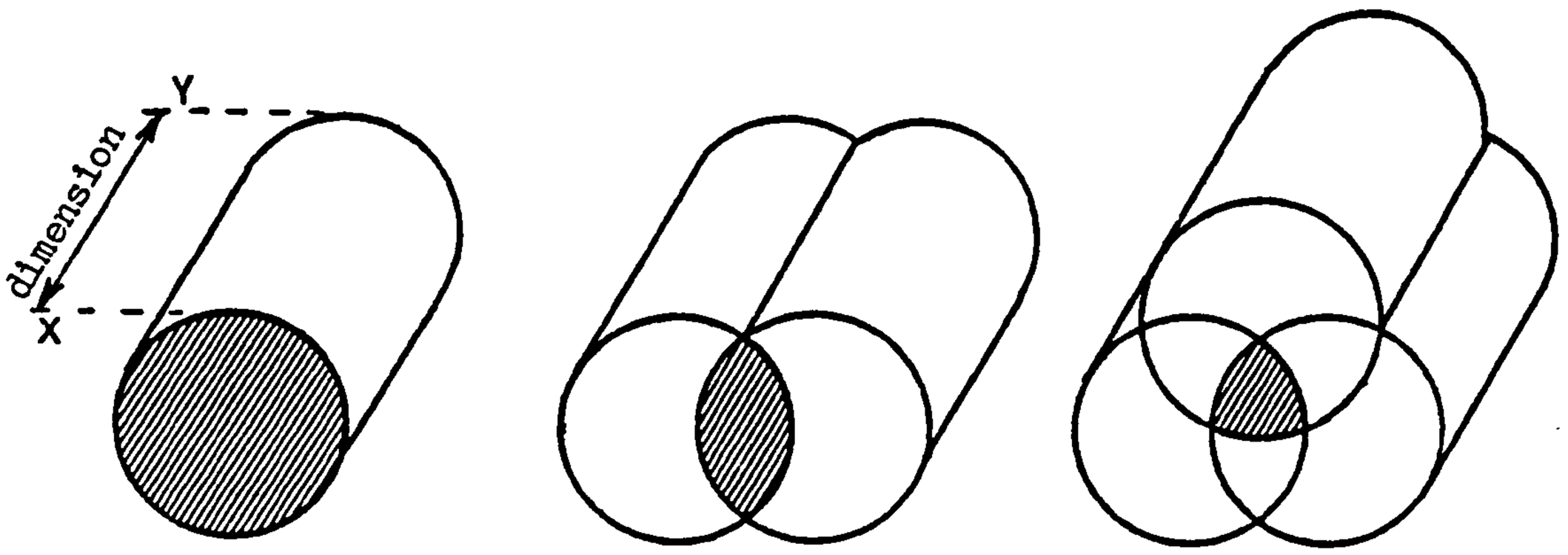
As with the previous chapter it is possible to conclude by proposing an alternative view of analysis that threatens to undercut the typical discussion of the technicalities of analysing rating data, by questioning the assumptions upon which it is predicated.

In all the analytical techniques considered in this chapter there is an idiographic nomothesis[78] in the sense that the individuals' constructs are treated as quasi-physical scales upon which an element can be located with a fair degree of accuracy. The problems this poses for meaning are strongly implied by Giorgi (1966) who writes that '... the transformation of an essentially qualitative phenomenon into a quantitative expression such as is done in many scaling techniques, does not capture the essence of the qualitative as such'[79]. The argument set out at the end of Chapter 8 develops Giorgi's thesis

and suggests that the normative assumptions of scaling may be untenable, and that each construct may take on a meaning specifically related to its context: in other words, the event (i.e. the intersection of element and construct system) is of critical importance in determining meaning.

If the position is adopted that the meaning of a construct may only be understood by considering its contextual relationship with other constructs[80], then it is not legitimate to talk of correlations (or other indices of association) between constructs since in repertory grid terms an event (the singularity defining the context) is represented by a single column in the matrix and hence no meaningful correlation or association between rows can exist. The rejoinder might be made that constructs are generalisable across elements, but this can be countered by arguing that the broader the range of elements the less the commonality of meaning across them.

The situation is symbolised in Figure 9.6.



(i) Construct meaning in respect of a single element.

(ii) The same nominal construct applied to two elements showing the core of common meaning

(iii) As (ii), but involving three elements

Figure 9.6 A symbolic representation of the decline of commonality of meaning as a nominally-constant construct is applied to an increasing number of elements.

Correlations computed to any number of significant figures cannot ultimately conceal the vagueness of such an indeterminate yardstick as that of a non-physical construct treated as if it were a physical scale.

If events are given primacy over the generalising of constructs, it becomes much more important to focus upon the implicative relationships between the relevant constructs in a system. Hinkle's implications grids fail in this respect since they treat constructs nomothetically (and as if they held their meaning static) in an indeterminate context: I would argue that relationships between constructs cannot generally be established in the abstract[81] and that in order to complete an implications grid the respondent has to invent a context - and this may vary markedly during the course of filling the implications matrix.

It is at this point that Eden's work makes a strong claim for attention. Working as a management consultant, Eden has perforce been faced with his clients' construing of particular problematic situations and so his method of approach reflects the uniqueness of the 'event' or situation. Constructs are elicited with reference to the specific problem at hand and to institutional factors such as norms and constraints. This is not to claim that the individual is construing without a personal nomothetic frame of reference, for if - following Rosch - construing makes reference to 'prototypes' abstracted from experience, any new experience will be related to a greater or less extent to the relevant prototypes contained in long-term memory (and will ultimately exert its own influence upon the store of prototypes).

Eden's work clearly treats each situation being construed as a historical event with a unique and complex concatenation of antecedents. It implies that repertory grid methodology *per se* is fundamentally misconceived, and hence that the currently available modes of analysis are grounded on a fiction too insubstantial to sustain the pressure of forensic examination. Insufficient evidence is available to come to a firm decision as to which pole of the opposition 'construct nomothesis - construct singularity' is the more appropriate for the researcher to choose, or whether some pragmatically intermediate 'ideal point' can be found. The challenge posed by this postscript to the main discussion in this chapter highlights the taken-for-grantedness of much that is done in repertory grid methodology. It will have served a useful purpose if it stimulates the grid users' critical senses.

9.9 SUMMARY

Rather than concern itself with a detailed appraisal of the currently available methods of grid analysis, this chapter has sought to explore some of the underlying problems. Slater's INGRID, Thomas and Shaw's FOCUS, and 'cluster and profile' were examined from the perspective of an inquiry into their assumptive bases: each was shown to embed a different subset of analytical assumptions and to exhibit particular strengths and weaknesses. Computer analysis of grids is generally vulnerable to the effects of missing data, and it also runs the risk of being treated as 'hard' and objective despite the researcher's awareness of the limitations of whichever diagrammatic representation of relationships is employed. The conclusion was drawn that analytical output in itself is of limited value and requires inference and further exploration of the maximum is to be extracted from the data available.

As far as the interpretation of analytical output is concerned, this needs to be undertaken with more care than some researchers have hitherto given it (as witness the failure to distinguish between the meaning of 'hierarchy' as applied to construct systems and to cluster analysis). Problems which have gone largely unrecognised are those of the 'hairpin' construct, of high intercorrelations between constructs, and of indeterminacy of the hierarchical levels of any particular construct.

The problems of analysing and interpreting idiographic data are compounded when they are aggregated to produce nomothetic conclusions, and it was argued that such a procedure is at best simplistic and at worst misleading.

The chapter concluded by advancing a view of analysis (based on the 'alternative' perspective on rating put forward at the end of Chapter 8) which challenges grid methodology and analysis at a level much deeper than would arguments about how to tinker with statistical techniques in order to produce a superior output. This challenge, whilst consistent with the broad brush strokes of Personal Construct Theory, suggests a dimension of debate hitherto neglected in repertory grid methodology.

10 Stability

Since 'tis Nature's law to change,
Constancy alone is strange.

J. Wilmot, Earl of Rochester
A dialogue between Strephon and
Daphne, I: 31

10.1 INTRODUCTION

It is invalid to talk of 'the reliability of the grid'[1]. Writers on the subject of reliability point out that the variations possible in grid content preclude the establishment of a reliability coefficient for the grid. The reliability of a grid must necessarily be determined with reference to the context of its use. This statement is no different from statements on reliability made in other contexts: as Guilford and Fruchter (1973) put it, writing with psychological testing in mind, 'One should speak of the reliability of a certain instrument applied to a certain population under certain conditions'[2].

That there is a wide range of conditions under which repertory grids have been used is attested by the research discussed in previous chapters: however, the notion of 'population' merits further consideration. At one level it can be taken to refer to the totality of a group presumed to be homogeneous, such as thought-disordered schizophrenics or head teachers in comprehensive schools. To use 'population' in this way is, in research terms, often to imply the drawing of a sample from that population with a view to generalising the research findings from the sample to the population from which it was drawn. Snow (1974), for example, has emphasised practical difficulties of sampling and generalisation[3], and the arguments he develops call in question the uncritical use of psychological instruments for nomothetic purposes. As far as repertory grids are concerned, it will have become clear from the preceding chapters that there exists a case for a radical scepticism, based upon both conceptual and practical considerations, regarding the nomothetic use of grids in most circumstances.

But in repertory grids 'population' can be construed in a more limited,

idiographic sense. It can be argued that any construct system must contain a finite - if indeterminate - number of constructs, and that a repertory grid samples this idiosyncratic population. As with the nomothetic aspects of sampling and generalisation, the validity of the sample of elements and/or constructs appearing in any grid is indeterminate[4]. However, in the case of individuals there is a greater opportunity to cross-check the sample (for instance, through conversational approaches) to enable some qualitative estimate of face validity to be made.

Bannister and Mair (1968) use the term 'population' in a still more restricted sense when they treat the <u>responses in a grid</u> as a population[5]. There would seem to be nothing wrong in so doing, provided that the data is treated as complete in itself and is not used as the basis for inferring generalisations. However, the practical realities of grid use would suggest that inferential interpretation is a likely consequence of a grid analysis. It seems reasonable to argue that the validity of interpretation of a grid analysis would generally be higher when the respondent's construct system is allowed to contribute to the interpretation, rather than when the statistical relationships are subsumed solely by the construct system of the researcher.

The reliability of grid measurements reflects this last notion of population. Guilford and Fruchter (1973) identify three different approaches to assessing the reliability of a test.

(a) computing the internal consistency of a test taken on a single occasion;

(b) comparing results from parallel (alternate) forms of the test; and

(c) comparing results from the same test taken on more than one occasion[6].

Of these three, internal consistency would seem to be inapplicable to repertory grids on account of their likely heterogeneity of content and hence the impossibility of identifying comparable 'split halves'. A case can be made for 'parallel form' reliability calculations, particularly where equivalent sets of elements are being employed[7], but to judge from the literature the main emphasis on reliability in grid methodology has related to the stability of various measures over time; that is, test - retest reliability.

10.2 RELIABILITY AND PERSONAL CONSTRUCT THEORY

Before examining some of the literature relating to reliability in repertory grids it is worth spending a little while looking at the issue of reliability from the standpoint of Personal Construct Theory. Kelly sees people as 'a form of motion'[8], implying that physical and mental change are a normal state of affairs - in sharp contrast with homeostatic conceptions of the person. It is not implied here that change is random, for such would destroy the purposive, anticipatory character of human behaviour: as Hampshire (1959) observes, there must be a certain minimum of consistency and regularity in behaviour for human action to be accounted intentional [9].

The implications for reliability are clear. It is not to be expected that identical - or, acknowledging the likelihood of 'experimental error', nearly identical - grid responses will be obtained whether either parallel forms of the grid are used or the same grid is used on two or more occasions, for both would reflect a homeostatic perspective. As Bannister and Mair (1968) put it, the notion of reliability should be construed in terms of predictable stability and

predictable change[10]; in other words, very high reliability coefficients should not generally be anticipated in research based upon Personal Construct Theory. The problem then is whether a low reliability coefficient implies unreliability in the use of the grid technique itself, or whether it is a valid indicator of change in the person concerned[11].

Be that as it may, it does not alter the fact that, within a Personal Construct Theory framework, both stability and change have their place. Mair (1970) suggests that stability might be studied by encouraging movement and seeing what stays still[12]; there are arguments both for and against such an idea. But it has to be remembered that stability itself can be deceptive. Slater (1972) makes the point that inter-correlations in the grid can remain stable provided that change on the related variables has taken place in a consistent manner across all of them[13]. Stability may be present according to one index whilst the test-retest reliability may be low. Slater's point calls to mind the constant level of the Liberal Party's support in Britain during the mid-1960's, in which opinion poll findings were actually masking a dynamic equilibrium between a large influx and a large efflux of support[14].

It is possible to consider the reliability of ratings across a grid as a whole - indeed, this is discussed at various points during this chapter - but from the point of view of understanding what the respondent is saying it may be preferable to 'partition' the grid responses into areas of stability and areas of change. The partition-ing might highlight change on particular elements and/or constructs, but is not necessarily limited to the rows and columns of the grid: individual grid cells may exhibit psychologically significant change

against a background of stability.

Reliability can be construed in terms of the respondent's ability to reproduce on a subsequent occasion the lists of elements and/or constructs already elicited. Some discussion of this aspect of reliability was presented in Chapter 5 (elements) and Chapter 6 (constructs) and the matter will not be considered further here.

In this chapter attention is given to 'test-retest' and 'parallel form' stability since these can be used (with a range of validities) in grid-based research. Missing data continues to present problems for the grid user, and an alternative approach to overall grid stability is discussed. A small-scale comparison of grid stability indexes is presented, and the chapter ends with a discussion and summary of factors which are likely to affect stability.

10.3 TEST-RETEST STABILITY

Although the majority of research into the reliability of grid measures has been devoted to test-retest reliability, the evidence is rather difficult to weigh. This is because a wide range of grid-based indexes has been used, making meaningful comparison problematic.
A full analysis of the issues involved is beyond the scope of the present work: what is offered in this section is a summary of some of the main findings from the literature together with a preliminary attempt to draw some general - albeit tentative - conclusions from what is a very diverse collection of information. Given the argument presented in the preceding section, I use the term 'stability' rather than 'reliability' in that it does not have the latter's somewhat pejorative connotations regarding the undesirability of change.

Bannister suggests a consistency score pertaining to the stability with which the rank order of construct interrelationship is maintained over time[15]. In an experiment reported by Bannister and Mair (1968) respondents were asked to rank fifteen common objects on six supplied bipolar constructs relating to physical attributes, and ten photographs of people on six supplied unipolar constructs relating to personal attributes. On retesting after a period of six weeks the mean consistency score for objects was + 0.93 and for photographs + 0.86. However, in a further study involving fifteen photographs the mean consistency score over six weeks was only + 0.56. Bannister and Mair surmise that many of their respondents may have been unable to use personal attribute constructs on a task requiring such fine discrimination[16]. It might also be suggested that the ranking of photographs of unknown people is open to considerable 'error variance' on the grounds of 'element vagueness' even though the constructs being used might be fairly well internalised. Whilst some elements may 'stand out' as extremes in relation to particular constructs (and thus be consistently ranked low or high), those in the middle of the range may well be subject to shifts in construing and hence their rank ordering might tend towards random. The greater the number of 'middling' elements in the set, the more likely it seems that the high stability in construing the 'end' elements would be undercut by near-random shifts in construing in the middle of the order: an explanation of this type would appear superior in accounting for the results.

Although most work on consistency has referred to the pattern of construct interrelationships (perhaps because of the predominant interest in construct systems as such[17]), there is no reason why such an approach to stability should not be used in respect of element interrelationships. Fransella and Joyston-Bechal (1971) used both measures in a year-long

study of a psychotherapy group which involved five administrations of a grid whose eight elements were (for each person) 'self' and the seven other patients, and for which twenty constructs were supplied. The authors found that, although there was no significant change in construct pattern consistency, element consistency showed marked changes; the suggestion is made that the construing of individuals had changed markedly with time whilst the individuals' construct systems remained relatively unchanged.

Reid (1976) suggested that the issue of element and construct stability might be clarified by classifying elements and constructs in terms of their stability over a number of administrations. He proposed three categories of element or construct : stable, transitional and unstable. Stability was associated with the sequential replication of ratings to a statistically significant level, whereas instability was associated with sequential inconsistency as judged by the same criterion. 'Transitional' elements and constructs were defined as those which, whilst exhibiting consistency in early and late grid administrations, nevertheless showed a discontinuity in the middle of the sequence. Using this classification (which appears to be something of an over-simplification), Reid showed that constructs relating to people tended to be more stable than the construing of the individual people themselves[18], thus paralleling the findings of Fransella and Joyston-Bechal (1971).

Gathercole et al (1970) also used consistency scores as the index of test-retest stability. Using 8 x 8 grids with either acquaintances or photographs of people as elements, and either supplied or elicited constructs, they found a range of consistency scores from + 0.08 to + 0.94 (median + 0.72) from their 53 respondents. They also found

that test-retest consistency was higher for supplied rather than elicited constructs (though this is confounded with the types of respondent used), but unfortunately they did not investigate whether the type of element used might have had an influence on the consistency scores. Gathercole et al's procedure, which also incorporated a study of parallel-form stability, involved each of their subjects in completing three grids in a single session: in other words, the interval between test and retest was very short. Lansdown (1975), who investigated consistency over time using 8 x 8 grids with 59 children aged nine to eleven, found that two thirds of his sample had consistency scores greater than 0.70. Approximately one third of Lansdown's sample were retested on the same day (it was more often than not the case that the retesting took place within 30 minutes), the retesting of the remainder being largely completed during the subsequent week. As one might expect, Lansdown found a negative rank correlation (rho = - 0.35) between consistency and time to retesting.

Slater (1972) devised a coefficient of convergence designed to be more generally applicable than Bannister's consistency score: where both can be used Slater indicates that the two indexes are likely to have similar values. Slater's DELTA program, on the other hand, computes (as the general degree of correlation between two grids) a kind of element consistency score rather than one of construct consistency. Watson et al (1976) used this latter index in their study of long-term prisoners, finding (for the 32 who completed two grids) a range of element consistencies from + 0.30 to + 1.00 (mean + 0.74). Orley (1976) reported that seven out of ten DELTA consistencies were higher than + 0.60, the remaining values being + 0.55, + 0.21 and + 0.03.

Stability indexes based on the Bieri et al (1966) grid measure of

cognitive complexity have frequently appeared in the literature despite the considerable uncertainty regarding what cognitive complexity might be. Bieri's (1955) original study of cognitive complexity involved retesting at the end of the same experimental session in which the original grids were completed. In this study ten bipolar constructs were elicited, the elements being ten acquaintances selected on the basis of role titles[19]. Bieri found an overall reliability of + 0.78, which seems to have been broadly replicated in other studies in which the retest took place after an interval of one week (e.g. Tripodi and Bieri, 1963: reliability + 0.86 for supplied constructs, + 0.76 for elicited constructs; Tripodi and Bieri, 1964: average reliability + 0.71). Vacc and Vacc (1973), using a modification of the standard Bieri grid with children aged seven to eight, found a reliability of + 0.82 - though when the instrument was tested on two samples of college students the cognitive complexity scores correlated only + 0.51 and + 0.55 with those from the standard Bieri grid. Schneier (1979) showed that for a substantial sample of students (N = 176) and managers (N = 37), results from the Bieri grid (modified in the light of Vannoy's, 1965, critique) gave reliability coefficients of + 0.54 and + 0.82 respectively. The sting in the tail of these findings was that the Bieri/Vannoy measures only correlated + 0.08 with those from the standard Bieri grid!

Bavelas et al (1976), in a study of a wide range of grid-based scores, found a reliability for the standard Bieri grid of + 0.67. The reliability of most of the other grid-based measures employed by these researchers did not decline over a three-week period (judging by cross-sectional evidence) and ranged between + 0.26 and + 0.66 (median + 0.53). However, it is not clear to what many of these measures were actually related. The difficulty of interpreting this study is compounded by

the authors' idiosyncratic reading of Kellian theory and by errors in their statistical reasoning.

The findings of Bavelas et al however, seem in broad agreement with those obtained earlier by Mair and Boyd (1967) when they collected test-retest data in respect of both split-half and ranked grids, the interval between the two grid administrations being two weeks. Mean test-retest correlations for the relationships between each of the three whole-figure constructs 'like self', 'like father' and 'like mother' and the remaining constructs in the grid ranged from + 0.43 to + 0.72. However, at the level of the individual construer there was a range of stabilities from - 0.56 to + 0.93.

Other work which has a bearing on test-retest stability tends to support rather moderate values. Hayden et al (1977) used a repertory grid with thirty emotionally disturbed boys in residential treatment, in which the elements were photographs of a boy working on a block-design puzzle. Constructs were elicited regarding the boy's feelings or thoughts, and the photographs were then rated on each construct. Repeating the ratings with the same elements and constructs produced a reliability of + 0.48, which the authors described as 'satisfactory' in the circumstances. Caine and Smail (1969) found an average element rank stability of + 0.68 in terms of hysteroid-obsessoid traits paralleling those contained in the Hysteroid-Obsessoid Questionnaire. Whilst there is a problem in comparing the two instruments since the actual tasks differ, the grid stability compares unfavourably with the + 0.93 of the Questionnaire. Finally, Frost and Braine (1967) report 'satisfactory' levels of test-retest reliability in the field of market research. No figures are given in their account which is of interest only in that the retest procedure required the sorting of the constructs

in terms of the elements - although the test and retest grid matrices would have been in the same form, the route by which the retest grids were obtained was very different from that leading to the production of the original grids.

On the evidence presented here mean test-retest stabilities would seem to be moderate, spreading about a median value of the order of + 0.60. At the level of the individual the variation in test-retest stability is very wide, and this again calls in question the practice of averaging data nomothetically.

10.4 PARALLEL-FORM STABILITY

There are two practical ways of assessing parallel form stability in repertory grid work: either one may 'fix' the constructs and vary the elements according to some criterion of parallelism, or one may 'fix' the elements whilst varying the constructs[20]. Taking the latter first, it seems unlikely that an argument for 'construct parallelism' can sustain a scrutiny which would point to the variations due to both idiosyncracy and cultural polysemy: validity in these circumstances would appear to be irredeemably compromised. At best, this deficiency can be partially rectified by allowing the respondent to select the parallel forms to be used, but it remains improbable that such nominal parallelism would be justified in practice. Referring to 'equivalent form' reliability, Bannister and Mair (1968) are surely overoptimistic when they remark that it is 'reasonable to regard different sets of constructs as equivalent, in the sense that they are a "fair sample" of the system'[21]. Whilst they are aware that stability coefficients calculated on this basis would have structural rather than content implications, it seems unlikely that such an index would have a great deal of meaning.

Calculations of stability would seem to be on less uncertain ground where the constructs are 'fixed' and parallel sets of elements are used. In repertory grids parallel sets of elements often seem to consist of photographs of unknown people or people nominated by the respondent according to lists of role titles. Evidence regarding such stability coefficients is relatively sparse.

Bannister and Mair (1968) report a reanalysis of work conducted by Bannister (1965) in which eighteen 'normal' respondents were asked to rank, on six supplied unipolar constructs, ten photographs of people unknown to them; the whole process was repeated twenty times over ten days with different photographs being used on each occasion. Bannister was interested in the effects of validation and invalidation on the 'construct pattern' (i.e. the interrelationships between constructs), and showed that the validational fortune of respondents was reflected in the stabilities of their construct patterns. However, it is Bannister's third group of six respondents (who were given no information regarding the accuracy of their judgments) whose results are of particular interest here. Bannister calculated for each individual a set of 19 mean reliability coefficients by relating each test to its immediate successor, and he found a range of stabilities between + 0.04 and + 0.78 (grand mean + 0.585).

Work which Bannister and Mair (1968) conducted using as elements 15 common objects (ranked on six supplied 'physical' bipolar constructs) and 10 photographs of people (ranked on six supplied unipolar constructs relating to personal attributes) gave mean parallel-form stabilities of + 0.92 and + 0.72 respectively, the interval between testing being six weeks. In another study involving 15 photographs of people the stability fell to + 0.54: reasons for this drop in stability were

suggested in the previous section[22].

Gathercole et al (1970), whose work regarding test-retest stability was also mentioned in the preceding section, investigated parallel-form stability using 8 x 8 ranked grids in which the elements were either photographs of people or persons named in response to a list of role titles. Where those of their 53 respondents were 'normals', construct were elicited; where they were psychiatric in-patients, the constructs were supplied. Parallel-form stabilities based on consistency scores ranged from - 0.16 to + 0.88 (median + 0.66), these being marginally lower than those found in respect of test-retest stability. It was also found that, for both parallel-form and test-retest stabilities, the consistency scores were significantly higher when the constructs were supplied rather than elicited. The authors make the suggestion that, as the supplied constructs were highly evaluative, they may have been easier to construe. No evidence is presented regarding the comparison of the supplied with the elicited constructs, but if the former happened to stimulate greater 'evaluative spread' in the respondents' minds (untypical of research on extremity of rating, but not impossible), then ranking of the elements might have been more definite and less susceptible to 'error variance'[23].

Element consistency was measured in a different way by Sperlinger (1976), whose reliability figure of + 0.95 after an average interval of 7.7 months has been widely quoted in the literature. Unfortunately it is not always made clear that this coefficient relates to the distances (computed by an early version of Slater's INGRID program) between 'self' and eleven other elements based on the role titles and that a fresh set of constructs was elicited on the second occasion. One would expect intuitively that distances would be high and stable

with respect to elements such as 'former close friend' and 'person who is hard to get along with', and low for elements of positive affect. And when it is realised that Sperlinger is presenting not the individuals' distance measures but a mean value from 25 respondents in the first test and from 18 in the second, it is probable that a lot of individuals' variance has been left by the wayside - in fact, the standard deviations of each set of distances (of the order of 0.25 where the means themselves average slightly less than 1.0) suggest that this is the case. It seems probable that Sperlinger has identified a reliable nomothetic tendency for people to be psychologically close to those they like and relatively distant from those with whom they feel uncomfortable.

The more important of Sperlinger's findings seem to have been consistently overlooked, these being based on the correlations (Pearson r) between the two sets of 'self' - 'other' distances for each of the 18 individuals who completed both grids. Sperlinger reports here an average correlation of + 0.57. It is to be regretted that he does not give a fuller set of data: he indicates that eleven correlations were significant beyond the 0.05 level, that three others were 'quite high' and that the remaining four ranged between - 0.06 and + 0.18. It would seem that Sperlinger's findings are of a similar order to those obtained by Gathercole et al (1970) although the two studies used different grid measures and stability indices.

As with test-retest stability, parallel form stabilities seem, on the whole, to be moderate and to centre upon a median value of the order of + 0.60.

10.5 TEST-RETEST AND PARALLEL-FORM STABILITY: A CONCLUDING COMMENT

The heterogeneity of construct dimensions would appear to rule out of consideration any form of internal consistency measure, and hence attention has been focused upon test-retest and parallel-form stabilities. The advantages and disadvantages of each are well documented in the psychometric literature[24], and it is not proposed to repeat these here.

The results discussed in the two preceding sections are not markedly inconsistent with Fransella and Bannister's (1977) review of the literature regarding construct pattern consistency which suggests a typical range of + 0.60 to + 0.80 for both test-retest and parallel-form measures. The implication of these findings is that generalisation from one set of elements or constructs to another is suspect.

The stability levels reported here would appear to be lower than those recorded for many psychometric tests, and this may reflect a tendency for grid elements and constructs to be less stable than test items, even when they are nominally identical on successive occasions. The potential for variation would seem to be greater for constructs than for elements, since constructs are more open to 'movement' within their meaning complexes. Where elements are not physical entities (such as photographs or recollected people), their potential for variation would appear to approach that of constructs. A grid stability index, then, is open to the indeterminate effects of element, construct, and element X construct interaction variance: whether <u>reliabilities</u> of the order quoted are satisfactory is a matter for the researcher's judgment.

As a form of test which elicits aspects of the respondent's world view,

this level of reliability would appear insufficient to sustain any conclusions based on a single grid administration. It would seem preferable to treat each grid as a limited snapshot needing supporting information to provide a qualitative check on the reliability of its contents. With such an approach the researcher might choose to use the grid as an indicator of potentially fruitful avenues of inquiry rather than as an end in itself.

Where a respondent completes more than one grid, the overall stability index may be of less interest than the patterns of stability and change. This may require analysis at the level of row or column, or even at that of individual cells. The difficulty is that the researcher will find it difficult to partition the observed change between that due to the administration of the earlier grid in any pair, and that due to those events extraneous to the grid procedure whose meanings the researcher is trying to tap. Respondents sometimes claim that the administration of a repertory grid causes them to think in new ways about the subject under investigation, as both Olson (1980a) and I have found. Reid (1976) goes a little further when he remarks that grid procedures are as likely to precipitate changes in construing as to measure them[25].

Further, it is difficult to determine whether (in test-retest terms) the respondent has reappraised the elements in the light of unchanged constructs or has reconstrued the constructs (despite using the same labels) and relocated 'fixed' elements upon them. (It may be, of course, that both types of shift contribute to grid instability). Where the respondent provides a different set of constructs on a readministration of a grid, there is the additional difficulty of trying to determine whether the second set of constructs represents

a new and distinctive sample from a stable system or whether instability in inter-element distances is an indicator of important changes in the construer's system.

Where feedback is provided after a grid has been administered it is likely to diminish stability: Hopwood and Keen (1978) claim to have observed this in respect of TARGET grids, though the incompleteness of their statistical data does not allow an estimate to be made regarding the magnitude of the effect.

If there is one main point to be made regarding stability in repertory grids, it is that stability cannot be treated in a simplistic manner.

10.6 THE PROBLEM OF MISSING DATA

This section runs the risk of being construed simplistically in that it focuses on technique rather than on meaning. However, it is a risk that has to be run if the problem of missing data - which is largely overlooked in the literature - is to be addressed.

Elements falling outside the range of convenience of a construct or capable of being assigned to both poles pose a threat to stability coefficients. More generally, the wide range of possible mid-points on a bipolar scale (see page 293f) is likely to present difficulties - in terms of meaning, if not in terms of mathematics. Fjeld and Landfield (1961) did recognise problems of this sort in their study of personal construct consistency and used, in contrast to Kelly's dichotomous procedure, a four-category scheme for the allocation of 15 elements (acquaintances) on 15 elicited constructs. As well as allowing respondents to allocate the elements to whichever pole best described them, Fjeld and Landfield made available the additional

categories 'not able to decide' and 'not applicable'. Retesting after two weeks produced 83 per cent agreement in allocations and a high contingency coefficient of + 0.80.

Fjeld and Landfield's approach would seem to be valuable where there are 'gaps' in the grid matrix, whether these gaps arise deliberately or by accident. It suggests that, in these circumstances at least, a stability index based upon comparisons of equivalent cells in 'test' and 'retest' grids might be generally useful. This could provide a general indication of stability in the grid as a whole, and could be applied on a 'per-element' or 'per construct' basis to indicate the pattern of stability and change in the grid.

It might be suggested (following Fjeld and Landfield's study) that the contingency coefficient might be suitable, using an (n + 1) x (n + 1) contingency table to relate the two distributions of ratings, where n is the number of rating intervals. It would often be necessary to include an extra row and column to cater for blank cells. This presents the first problem: often the number of blank cells is small and this would produce a table with too many low expected frequencies. The second problem is similar in that, unless the number of elements and/or constructs is large compared to the number of rating intervals, there will be too few instances per cell. This second problem can be countered by 'collapsing' the rating intervals - but at the expense of some sensitivity in the data. On the whole, these points militate against the use of the contingency coefficient in most circumstances.

In her study of student teachers' construing before, during and after teaching practice Pope (1977) used an index based on comparisons

between individual grid cells, to which she gave the acronym CHAT (changing after teaching). She calculated the change for each cell in each grid, ignoring the sign of the difference, and summed these changes in order to obtain a percentage change score in which actual change was related to maximum possible change[26]. Reanalysis of some of Pope's data shows that she overlooked the influence of the ratings in the original grid on the maximum change possible.

The maximum possible change on a scale is a function of the extremity of the original rating. Thus, for a five-point scale such as Pope used, the position can be shown as below:

ORIGINAL RATING	5	4	3	2	1
MAXIMUM POSSIBLE CHANGE	4	3	2	3	4

Pope's assumption that all possible changes were at the maximum value led her to underestimate slightly the proportion of change in her respondents' grids (and, for the purposes of the discussion here, to overestimate the stability)[27]. The point is illustrated in Figure 10.1 which draws on data provided by Pope's Student No.19[28]. The coefficients are converted to decimal proportions here in order to remain consistent with standard practice in reporting stability coefficients.

Measure	Grid 2-Grid 1	Grid 3-Grid 2	Grid 3-Grid 1
CHAT change score	0.10	0.09	0.14
Corrected CHAT change score	0.13	0.13	0.17
Corrected stability	0.87	0.87	0.83

Figure 10.1 CHAT scores and stabilities for Student No.19. (Original data from Pope, 1977: 225).

Pope appears to have had few, if any, problems with 'gaps' in her grids. This may have been because she required the elimination of

constructs from consideration if they could not subsume the disparate ranges of elements elicited. In my own work using situations as elements there were a number of instances in which an element fell outside the range of convenience of a construct. Rather than eliminate rows or columns in the analysis, and wishing to gain an indication of the stability of the data collected, it was necessary for me to devise an index of stability that could be used where 'gaps' existed in the data (and the gaps were not necessarily in the same cells in repeat grids). Accordingly, and independently of Pope, the stability index based on 'per cell' comparisons with original grid ratings was used as an indication of the overall stability of the grid ratings.

The design of my research required that an initial grid be completed as soon as possible in the first term of the science teachers' probationary year, a second grid (using the same elements but a new set of elicited constructs) being completed as late as possible in the third term. The test-retest stabilities were determined for both of these administrations after roughly four days by asking the respondents to complete a postal version of their original grid. This clearly could not be administered in an identical manner to the original grid, but the style of presentation adopted was as close as possible to that previously used despite the absence of a live administrator.

At the beginning of the probationers' first academic year, stabilities calculated on the basis of change per cell (eliminating any pair of cells in which there was at least one blank) produced fourteen coefficients ranging from 0.66 to 0.81 (median 0.75). At the end of the year the same group of teachers (apart from one who had left his school and two who did not complete the second retest grid) produced

eleven coefficients ranging from 0.69 to 0.87 (median 0.81). Treating rating consistency as the dependent variable, and ignoring the changes in elicited constructs between the beginning and end of the year, the change in stability for the eleven respondents who completed all four grids was statistically significant (Wilcoxon test: T = 8; N = 10; $p < 0.05$)[29]. No a priori hypothesis had been advanced in respect of any change in stability being a function of time. One possible a posteriori explanation is that familiarity with the grid procedure might have reduced the 'error variance' of what to the uninitiated might initially have appeared as a complex and perhaps confusing task. Another possible explanation is that a year's experience might have crystallised the teachers' construings of science teaching and so have enabled them to be more definite and consistent in their ratings. Repeating the procedure with experienced teachers might allow a judgment to be made between the two suggested explanations (and others), but the problems of meaning-shift in nominally identical constructs or of using a different set of constructs on each occasion would not make that judgment easy.

As a side-issue to this investigation into grid stability it was recognised that test-retest comparisons could be undertaken on a 'per-element' or 'per construct' basis. In Chapter 9 it was argued that the city-block metric (effectively the metric used in the calculation of whole-grid stabilities) was less suitable than the Euclidean distance when looking for the more important changes in grids. It might be felt that this position is inconsistent with that adopted in respect of whole-grid stability, but the distinction is defended here by pointing to the two different types of question being asked of the data. The whole-grid stability index is a response to the question 'Given a set of grid ratings, to what extent are these replicable?', whilst the

identification of element or construct change is a response to the question 'In respect of which elements (or constructs) is the respondent tending to fluctuate most widely?'

The question of whether different types of element may be associated with different levels of grid stability was discussed in Chapter 5 (page 193f) in respect of an investigation which I conducted. It suffices here to reiterate the finding that when the elements were pupils chosen by their teachers the stability index was significantly higher than when the elements were either specific teaching situations from personal experience or general teaching situations (Smirnov test: both p values less than 0.05, one tailed.)[31] The results are summarised in Figure 10.2.

TYPE OF ELEMENT	NUMBER OF RESPONDENTS	RANGE OF STABILITY COEFFICIENTS	MEDIAN STABILITY COEFFICIENT
Pupils	9	0.83 to 0.89	0.87
Specific teaching situations	7	0.74 to 0.90	0.83
General teaching situations	4	0.55 to 0.83	0.81*

Figure 10.2 Stability coefficients for different types of element, constructs having been elicited from respondents. (* signifies the arithmetic mean of the two central coefficients)

The coefficients in respect of pupils shown in Figure 10.2 are broadly similar to those obtained in an earlier investigation whose primary focus was not upon stability although the data collected allowed stability coefficients to be calculated[32]. The respondents in this case were ten serving teachers undertaking an in-service B.Ed. course who were asked to name ten pupils as elements, from which sets of constructs were elicited. The respondentsthen rated the pupils on each of their constructs, the rating procedure being repeated seven

days later. The stability coefficients ranged from 0.75 to 0.92 (median 0.855), one of the lower coefficients being suspected to be an underestimate in that it appeared that reversal of the poles of one construct had occurred unnoticed during the second rating.

10.7 STABILITY INDEXES: A SMALL SCALE COMPARISON

In the previous section I argued for the advantage of a stability index based upon 'per cell' comparisons on the grounds that it could maximise the use of the data when there were blank cells in one or other (or both) of the grid matrices. The collection of the ten pairs of grids mentioned in the preceding paragraph made it possible to undertake a comparison of a 'per cell' index with three indexes currently available; namely, Bannister's consistency score[15], Slater's coefficient of convergence (as computed in his COIN and NEW COIN programs), and Slater's general degree of correlation (as computed in his DELTA program)[33].

The above four stability indexes were calculated for each pair of grids, fourteen of the 126 constructs being eliminated from consideration because of blank cells[34]. The indexes are given in Appendix 12 : it is sufficient for my purposes here to present in Figure 10.3 the correlation matrix for the four sets of indexes.

	Bannister consistency	'Per cell' index	Coeff. of convergence
'Per cell' index	0.51		
Coefficient of convergence	0.96	0.53	
General degree of correlation	0.75	0.92	0.80

Figure 10.3 Correlations between four stability indexes.

Whilst there is a high positive correlation between the Bannister consistency score and the coefficient of convergence (congruent with the value of 0.943 quoted by Slater in respect of 46 grids[35]), their separate low correlations with the 'per cell' index necessitate a further analysis.

One of the ten grids exhibited what was almost certainly a reversal of the rating scale for one construct, giving it a markedly low Bannister consistency and coefficient of convergence, 0.50 and 0.49 respectively. In contrast, the 'per cell' index for this grid was 0.84 and the general degree of correlation was 0.71, the former being in the middle of the range of 'per cell' stabilities[36]. On removing the 'offending' construct from that grid, the four indexes became

Bannister consistency	0.96
Coefficient of convergence	0.96
'Per cell' index	0.89
General degree of correlation	0.89

The evidence suggests that the Bannister consistency and the coefficient of convergence are particularly vulnerable to inadvertent construct reversal because of their methods of computation. Briefly, a half-matrix of inter-construct rank correlations (Bannister) or distances (convergence) is produced for each of the two grids, the relevant index being computed as the correlation between the two sets of rankings or distances. If, as in this case, one construct out of eleven is reversed on one of the two occasions, its ten relational measures with the other constructs will be out of kilter with the remaining 45 (with eleven constructs in all, there are 55 inter-construct measures). With such a high proportion of 'aberrant' values, both the Bannister consistency and the coefficient of convergence are likely to be sharply lowered[37]. The cumulation of variance and covariance of the aligned

constructs in computing Slater's general degree of correlation seems to make this index less vulnerable to the effects of construct reversal. However, the 'per cell' index, which takes each pair of ratings as separate, is the least vulnerable of the four indexes considered.

Since reversal can occur with part-constructs as well as with whole constructs it would seem that neither the Bannister consistency score nor the coefficient of convergence are sufficiently robust to give the grid user a meaningful indication of the general degree of stability in the grid. The evidence presented here indicates that the argument for the 'per cell' index has the greatest strength, particularly when its capacity for coping with blank cells is taken into account.

10.8 SOME FACTORS WHICH MAY AFFECT STABILITY

The less than definite title to this section is indicative of a number of aspects of repertory grid methodology which may have a bearing on the stability to be expected in grid measures, yet to which little attention has been given. In consequence, the discussion at this point is speculative rather than definitive.

10.8.1 Rapport

Where grids are completed on an individual basis the degree of rapport between researcher and respondent may influence the contents of the grid. To draw on personal experience, I found it easy to establish a rapport with some of the science teachers and yet with others I would have to acknowledge that the relationship was more distant whilst still being friendly. One woman (Ms. A) found the grid administration an opportunity to 'talk through' aspects of science teaching which appeared to be of considerable concern to her. She was very willing

to co-operate in completing the grid and embellished the constructs and elements with comments (given in confidence) on the school and individual colleagues. A lot of thought went into this particular 15 x 15 grid (which took 70 minutes to complete), and I have to admit that I am uncertain as to what proportion of it represented new structuring of constructs as opposed to the reiteration of previously stored structuring.

In contrast, Mr. B completed a grid of 15 elements x 17 elicited constructs in the short time of 50 minutes. The rating pattern showed high extremity, but my subjective feeling here was that this reflected a lack of involvement, a 'hurrying to get it over', rather than the high involvement that the literature on extremity of rating would suggest. On readministering the grid at the end of the year, a similar reaction was obtained.

These two instances represent, subjectively, the extreme cases among the fourteen science teachers with whom I worked. The stabilities of the three grids involved (Ms. A did not complete a 'retest' of her second grid) were:

Ms. A	Grid I	+ 0.81		
Mr. B	Grid I	+ 0.76	Grid II	+ 0.69

These figures are an inadequate basis on which to claim a positive relationship between degree of rapport and stability, though one might speculate that a link might exist from rapport through commitment to the task to stability. It is worth noting that the stability of Mr. B's second grid was markedly lower than any other second grid stability, and that he was one of only two teachers whose stability coefficients were lower on the second occasion.

Unlike many inventories and schedules, the repertory grid is comparatively difficult to 'fake', and it seems improbable that much deliberate faking occurs during the completion of grids. This is a different matter from the carelessness born of indifference, to which some reference has already been made. It would seem reasonable to conclude at this stage that the quality of the interaction between the researcher and respondent is likely to have an influence on the latter's commitment and carefulness, and hence indirectly upon the stability of the measures contained in the grid.

10.8.2 Loose construing

Loose construing is likely to contribute to grid instability. Whilst most researchers (following Kelly[38]) have looked at loose construing from the point of view of the movement of elements on constructs (perhaps because the latter were insufficiently specific), the case was argued in Chapter 5 that 'element fuzziness' or complexity might contribute to the problem as well[39]. Pope (1977) provides good examples of both from her research with student teachers: a sample is given below[40].

> Student No. 4. '... sometimes look at it from one view and put it in one place, and then from another viewpoint you could put it somewhere else.'
>
> Student No.15. '... in the gaps between one grid and the next you tend to forget if some of the constructs were ambiguous.'
>
> Student No.24. '... sentence too short - couldn't remember what was meant by each construct.'

Looseness of construing is also reported by Tippett and Silber (1965) who found their respondents to flick between alternative facets of the definition of a personality trait[41]. Interpreting 'traits' as 'constructs' indicates the polysemic character of the latter and hence

implies the possibility of instability in their use.

Where elements are relatively unfamiliar it is likely that they will be unstable components of the grid. Though the issue is (as far as I know) unresearched, one might anticipate that the ratings of relatively unfamiliar people chosen to fit given role titles would be less stable than those of people who are very familiar to the respondent. If such an effect exists it might be expected to appear at its greatest in dichotomous grids where no mid-point provides a haven for all those elements not easily subsumed under one pole or the other. Photographs of unknown people would seem prone to a similar indeterminacy, particularly where their features are not construed as representative of extreme positions on constructs. The sharp decline in stability of the rankings of photographs when the number of elements increased from ten to fifteen (Bannister and Mair, 1968) was earlier interpreted as being possibly due to uncertainty of rating elements which did not appear to be extreme in respect of the construct being used[42]. It is not at all evident (in constrast to the authors' suggestion) that it is the respondents' powers of discrimination that are at fault, since it can be counter-argued that photographs provide inadequate evidence for anyone to make meaningful and stable judgments regarding personality. A task such as this is in no way commensurate with the task of ranking fifteen named common objects on six constructs relating to physical attributes[43].

10.8.3 Extremity of rating

The literature suggests that extreme ratings tend to be more stable than those nearer to the mid-point of a scale, and the discussion in the preceding paragraph is consistent with such a thesis. Evidence for the thesis comes not only from experiments with 'physical' stimuli

(such as Eriksen and Hake's, 1957, work on judgments of the size of a sequence of squares), but also from the field of social judgment (exemplified here by Miller and Bieri, 1963, who provided trainee social workers with brief case vignettes and found that judgments were more reliable in respect of two 'extreme' cases[44]),

Support from grid work for a connection between extremity of rating and stability is given by A. Ryle (1975), who tested the hypothesis that constructs with the highest variance in a grid would be more resistant to change (i.e. be more stable) than those of low variance. Nineteen respondents completed pairs of grids, the intervals between the grids ranging from twelve to twenty months. Instead of testing the stability of all the constructs, Ryle chose a questionable method of testing this hypothesis in that he took the three most stable and the three least stable constructs (judged on the basis of the consistency of the elements' ratings) and looked back to the first grid to see the proportion of the variance for which each accounted. He found that the more stable constructs accounted for a higher proportion of the grid variance in fourteen of the nineteen cases considered[45].

Bannister and Mair (1968) discuss, under the heading of 'Construct variance', what at first sight appears to be cognate evidence. However, a closer inspection reveals that their discussion (of a reanalysis of work by Bannister and of a subsequent paper by Mair and Boyd, 1967) refers to the stability of inter-construct relationships and not to the stability of constructs themselves[46]. Further, it is probably inappropriate to compare these two studies since Bannister and Mair appear to have used a "whole matrix" approach to relationship stability, whilst Mair and Boyd employed as a criterion the stability

of match of each construct with the single referent construct 'like I am'[47]. It may be the case that some relationship does exist between the variance associated with individual constructs and that of construct matching scores, but on the evidence presented by the respective authors its strength is indeterminate.

The limited evidence reviewed above suggests that there might be some validity in the extremity-stability hypothesis in respect of repertory grid work. It can be plausibly argued that high variance (or extremity of rating[48]) on a construct implies that the respondent wishes, and is able, to make a substantial number of discriminations using the full range of the scale. Those elements allocated to extreme positions are likely to remain there or thereabouts over time, whilst those in middling positions might (but would not necessarily) fluctuate in a manner approaching randomness. An element at the end of a scale, although theoretically 'free' to move the full length of the scale, is only able to move in one direction (and typically this movement is small) since there is no scope to give a rating more extreme than the end of the scale. The ratings of elements near the end of a scale might fluctuate randomly where all the elements are clustered together - that is, when rating extremity is high and construct variance is low.

On the argument presented here, the construct with high variance is likely to be 'locked' into high variance by stability at the ends, irrespective of fluctuations around the mid-point. Where the construct's variance is low, such 'end-locking' is likely to contribute in smaller proportion towards the overall stability of that construct, making the stability coefficient more vulnerable to ratings near the mean.

However, the argument is an oversimplification because it neglects the influence of meaning upon the construct dimension. Central ratings do not necessarily fluctuate in a near-random manner: for instance, if the mid-point of a construct is meaningful for the respondent[49], ratings in that region could remain stable and the construct as a whole exhibit high stability even if it did not account for a particularly large amount of the variance in the grid. In Section 8.3.7 it was argued that there were many possible ways in which bipolar scales might be used, each with a different "semantic metric", and it may be that a connection between extremity of rating and stability exists in respect of some types of construct but not others (depending on the way the construct is used in practice). Moreover, extremity does not necessarily imply stability: Mr B's two grids were notable for the extremity of their rating (the most extreme of the sample on both occasions), yet the stability of his ratings on the second occasion was the lowest by a substantial margin[50].

So what is the status of the extremity-stability hypothesis?

Two sets of grids which I elicited shed a little light on the problem[51]. High positive correlations were found between the average construct variance per grid on two occasions of testing for

(i) 11 science teachers, grids approximately seven months apart, same elements, different constructs ($r = +0.92$, $p < .01$); and

(ii) 10 in-service B.Ed. students, grids one week apart, same elements, same constructs ($r = +0.95$, $p < .01$)

These findings strongly suggest a consistency of rating style, taking each whole grid as the unit of measurement. Within individuals, the B.Ed. students exhibited more variation in construct variance when the construct was taken as the basis of measurement, the correlations

between the sets of construct variances on the two occasions ranging from - 0.51 to + 0.74 (median + 0.51).

When, for each individual, the correlation between construct variance in the first grid and an index of change per construct was calculated, in each of the three sets of grids a wide range of coefficients was obtained centring near to zero (see Figure 10.4)

SET OF GRIDS	N	RANGE	MEDIAN
Science teachers, occasion 1	14	- 0.61 to + 0.30	- 0.04
Science teachers, occasion 2	11	- 0.74 to + 0.62	+ 0.04
In-service B.Ed students	10	- 0.71 to + 0.68	- 0.10

Figure 10.4 Correlations between construct variance in the first grid and change per construct.
Note: negative correlations are consistent with the extremity/stability hypothesis since 'change' is an inverse function of stability.

The data presented in Figure 10.4 conceal the levels of the two variables concerned - a well known fault of the product-moment correlation coefficient. Thus grids with high overall construct variances could be highly stable without this being detected in idiographic correlations, since individual construct variances could be consistently high yet not be systematically related to individual construct stabilities (or changes). At the nomothetic level, however, a non-significant association was found (for the in-service B.Ed. students) between average construct variance per first grid and whole-grid stability ($r = + 0.33$). A weaker association was found in respect of the first pairs of grids completed by the science teachers ($r = + 0.23$, n.s.), but even this low figure was not sustained on the second occasion ($r = - 0.05$, n.s.). In other words, in the circumstances of these empirical studies there is little evidence for a

general relationship between extremity of rating and stability.

The evidence from this limited study suggests that, whilst respondents are likely to be consistent in their style of rating, on the whole they may be much less consistent in the way in which ratings are allocated. Only one of the eleven science teachers who completed two sets of pre/post grids evidenced stable rating behaviour that is consistent with the extremity-stability hypothesis. It would appear that the connection between extremity and stability - if such exists - is much more complex than may be commonly supposed, perhaps involving parameters such as rating style, number of elements to be rated, length of the rating scale and the respondent's construing regarding the semantic and psychometric characteristics of the scale.

It is clear that, as far as the stability of grid data is concerned, the researcher needs a finely-tuned ear to detect the signals among the noise.

10.9 STABILITY: SOME TENTATIVE CONCLUSIONS.

It is difficult to draw firm conclusions from what is a diverse collection of stability coefficients obtained under very varied circumstances: the variation between the studies reported here and the range of different grids used militate against trying to establish some kind of norm for repertory grid work. The objectives associated with repertory grid use also influence the way in which a stability coefficient is used: for example, if the research involves the study of change in an individual or group over a period of time, then overall stability may be less important than the identification of where change takes place. To make such a point is not to dismiss the importance of stability per se for, as was observed at the beginning of this

chapter, some stability is essential to provide a ground against which the figure of movement may stand out. In other words, stability has to be construed within the context of the research being undertaken.

That said, what are the main trends to be detected in studies which bear on the issue of stability? Since most of the evidence relates to short-term stability, the question can only be answered within that further limitation. The following points are suggested by the evidence.

(i) Stability tends to be raised when the context of the rating procedure is clearly defined.

(ii) Stability tends to be raised when the elements are familiar or relatively simple to construe. (Familiar elements include well-known acquaintances and relatives, and common objects. Common objects may also be simple to construe, particularly with respect to 'physical' dimensions.)

(iii) Stability tends to be raised when constructs are personally meaningful.

(iv) Stability tends to be raised when constructs are deeply culture-embedded; that is, when there appears to be a broad consensus regarding meaning (even though this consensus might break down under the pressure of probing).

(v) Stability tends to be raised when respondents are allowed to indicate that elements fall outside the range of convenience of constructs (for example, by leaving blanks).

(vi) Grid administration on an individual basis is likely to reduce errors in completing the grid matrix and hence enhance stability.[52]

(vii) Test-retest and parallel-form stabilities tend to be of an equivalent order of magnitude under equivalent conditions - cf points (i) to (vi) above. Median values appear to be moderate.

(viii) Individuals appear to exhibit stability in rating style.

(ix) Extreme ratings do not appear to be directly related with stability (in contrast to the 'received view' in the literature).

When the conditions (i) to (vi) do not obtain, stability is likely to be lowered. The demands of research may require a compromise to be reached regarding some of the points listed above, in which case the identification of areas of potential weakness may enable some estimate of error - albeit subjective - to be made. A question which seems to be rarely asked of stability data is that of the <u>psychological significance</u> of the change and/or stability detected in the grid matrix, yet this might be more important than attaining stability values normally deemed appropriate to a good psychometric test.

Consistency for consistency's sake, as Emerson might have put it, is the hobgoblin of little minds.

11 The intolerable wrestle

Words, meanings, and the repertory grid

The imperfect is our paradise.
Note that, in this bitterness, delight,
Since the imperfect is so hot in us,
Lies in flawed words and stubborn sounds.

W. Stevens
The poems of our climate

11.1 INTRODUCTION

In the discussion of the Construction Corollary[1] Kelly was taken to task for his failure to take account of the crucial role of language in construing (though he does acknowledge, in passing, a cultural relativism[2]). It may be that language was, for Kelly, a truly transparent templet through which he viewed his world[3]. Perhaps it was such a commonplace that he failed to recognise it for what it was, like the 'blind' observers of the postman-murderer in G.K. Chesterton's story.

The theme of this chapter, then, is the exploration of the relationship between language and construing, particular emphasis being given to the standing of that relationship in the context of grid methodology. It is only now that the hints and guesses contained in the earlier chapters can be drawn together; faint beams of light brought to a focus by the lens of linguistic theorising. Properly, this is a large undertaking, and it is not possible here to do more than give a few indications of the linguistic factors which I believe research within the framework of Personal Construct Theory needs to acknowledge. For me, this chapter is the necessary link between the critique of theory and method already set out and the possible developments essayed in Chapter 13.

11.2 LINGUISTIC RELATIVISM

At a number of points in his writing Kelly claims that people create their own ways of seeing the world which they experience[4]; these are their frameworks for construing, or construct systems. Construct systems are developed through experience. Kelly's emphasis here is on the anticipatory, purposive aspects of behaviour, and the determinism of a person's history is noticeably underplayed[5]. Though Kelly does

not use the terms in 'The Psychology of Personal Constructs', his view that a person can use construing to gain freedom for the self and can make free choices suggests that he is close to the notion of the autonomous ego or the self-actualizing person[6]; his is a Romantic position that is open to challenge in the light of sociolinguistic theorising.

Kelly never fully explores the implications of the cultural relativism whose existence he recognised, perhaps because his opposition to behaviourism led him to reject looking at a person as 'a lump of matter shaped by the happenings of the past'[7]. Yet at the same time as he was writing, linguistic relativism was well established in America through the work of linguists such as Sapir and Whorf[8]. Chomsky's transformational grammar had not yet threatened the position of the linguistic relativists (a threat which in retrospect seems to have evaporated with the realisation that Chomskian syntagmatics and the semantics of linguistic relativism deal with largely different aspects of linguistics)[9].

Some writers have sought to weaken the Sapir-Whorf theory of linguistic relativism. Brown and Lenneberg (1954), for instance, argue that Americans can - like the Eskimos - discriminate between three types of snow if they are forced to do it[10]. The point they miss is that normatively the discrimination is unimportant to most Americans: that they <u>can</u> discriminate if they need to (but typically do not) is evidence <u>for</u> relativism, rather than against it. Cole and Scribner (1974) are prepared to accept only a weak relativism on the grounds that language is a less powerful factor in constraints on perception and thought than Whorf believed to be the case[11]. More recently J.R. Anderson (1980) has taken a stronger position: though he

recognises that language can influence people he sees this in terms of the communication of ideas and not as a determinant of the ideas that can be thought about[12].

Concurrently with attacks on linguistic relativism, others - for instance, Gadamer (1975*) and Lacan (1977*) - have pressed its cause much more strongly than Whorf. Gadamer, following von Humboldt, writes:

> 'Language is not just one of man's possessions in the world, but on it depends the fact that man has a world at all.' [13]

and Lacan (surprisingly laconically for him) puts it:

> 'The world of words creates the world of things'.[14]

Corroboration of the power of language to determine the ways in which people construe events comes from work across a broad range of settings[15]. Freire (1972) shows how language is involved in the oppression of subjects within state systems; feminist writers such as Daly (1979*), Miller and Swift (1979*) and Spender (1980) point to the power of male-dominated language to place women in an inferior position in society; Berger (1972) shows how a radically different critique of art can be mounted once the traditional language categories of appreciation are replaced; du Preez (1979) indicates the National Party's subtle shift in political terminology from that of racial discrimination to that of national self-determination of ethnic groups in South Africa; Kuhn (1970) chronicles the power of 'normal science' to influence the content and methodology of scientific research: the list could be extended indefinitely[16]. The balance of evidence presented here is consonant with Peters's (1974) opinion that 'our minds are mainly social products'[17], a perspective which underpins Berger and Luckmann's (1967) thesis of the social construction of reality[18].

At the level of individual construing an early experiment by Dearborn (1910) is of interest, in which he conducted an introspectionist investigation of similarity and difference in the perception of inkblots. He recorded that all but one of his twenty subjects 'made no progress in characterising the norm-blot to themselves until words had arisen in their minds to make its character or characters definite and sharp'[19]. He further remarked that the process of cognition 'does not ordinarily find issue until the actual word symbolic of a concept is fairly clear in consciousness'[20]. Dearborn's findings bear out Giddens's (1976) conclusion that intentionality (in the phenomenological sense) cannot be seen as an expression of an ineffable world of private mental experience, but should be construed as necessarily grounded in the communicative categories of language[21].

Language is a 'third universe' midway between the phenomenal reality of empirical experience and the internalised structuring of consciousness[22]. Its formal symbolism is thus a limiting factor in respect both of what people can construe and of what they can communicate to others[23]. This is not to claim some kind of Platonic purity for language, for it is a socially constructed universe open to the influence of those who use it - as witness the propagandist who talks of 'clean bombs', thus allowing the juxtaposition of terms to imply the attribution of moral and physical integrity to destruction[24].

This 'third universe' subsumes the series of commonsense constructs that Schütz (1962*)sees social scientists as using to pre-select and pre-interpret the world; it provides the 'givenness' of the world[25]. Schütz contrasts the social scientist with the natural scientist who (in Schütz's view) defines - according to the procedural rules of the

science - and determines the facts, data and events relevant to the purpose in hand. The contradiction expressed here exposes the autonomy of the scientist as a fiction: pressed hard, it is difficult for the notions of autonomous ego or transcendental self to be sustained in any human context. Wilden (1980) makes the point tartly that 'the traditional epistemology of the life and human sciences is founded on an essentially religious belief in the real existence of such popular fictions as the "antonomous ego" .... [which] necessarily generates a further fiction, essential to its own survival'[26].

Seen in this light the importance of language to Personal Construct Theory is obvious. Personal construing is influenced strongly by language[27]: in its strongest form this is stated in the structuralist claim that 'the language speaks us'. Research within a construct theory framework neglects at its peril the contribution of language. The fragmented nature of the linguistic components of repertory grids poses a palpable threat to the validity of conclusions drawn from grid-based research.

11.3 THE REPERTORY GRID: A LINGUISTIC PERSPECTIVE

In Chapter 4 it was suggested that the labels typically used in repertory grids were impoverished fragments of language, and possessing an indefinite set of connotative meanings[28]. Since researcher and respondent will hold different meaning-complexes for each fragment, the researcher's inferences may not correspond to the respondent's communicative intention. In such circumstances the amplification of verbal labels by conversation (or other means) was seen as essential.

My critique of the repertory grid would be seriously deficient if the linguistic aspects of construing were neglected. In this section I

consider the semiotics of individual words and then broaden the discussion to semantics. Taken together, these two strands of argument lead to a further consideration of grid methodology.

Peirce suggested a threefold division of signs in terms of

(i) icon (directly representative of that being signified);

(ii) index (indicative of the presence of the signified); and

(iii) symbol (learned, rule-like associations between the signifier and the signified). [29]

As far as grid work is concerned, the last of these - the symbol- is of greatest importance, for most constructs are related by convention to the qualities, attributes and such that they signify and do not generally possess iconic or indexical features.

Saussure (1974*) emphasised the distinction between the 'signifier' (the acoustic image) and the 'signified' (the concept). Both of these are abstract psychological representations, the former of the phonetic aspect of the word and the latter of - in Kellian terms - the element or construct: together they make up a unified piece of knowledge - the unity of the sign. Saussure realises that, on its own, this is not enough, for the <u>meaning</u> of any signification is determined by its relationship with other terms. Part of the meaning of, say, 'dinghy' resides in the fact that it is <u>not</u> a liner, a power boat, a houseboat, and so on: it is a member of a 'contrastive set' (Miller and Johnson-Laird, 1976). But a dinghy also has features appropriate to a dinghy, such as sails, centreboard and tiller - here a link with Rosch's work on prototypes suggests itself. A 'concrete' object such as a particular dinghy is not at first sight the best basis for an argument which will need to consider adjectival 'signifieds'; however, it serves well enough as a point of departure.

The difficulty in dealing with abstract terms is that the relationship between signifier and signified becomes more complex than a one-to-one denotative correspondence. The adjectival signifier is very often not anchored to a single sense-datum but is defined by its location in a connotative mesh of similarly functioning terms. These terms may stand in different implicative positions vis-à-vis the articulated signifier: in a weak sense they may be regarded as 'signifieds', though once they are brought to consciousness they become in effect further signifiers and riffling through the pages of the personal construct dictionary becomes a limitless exercise. For the purpose of the present argument the articulated signifier will be treated as connoting an indefinite set of signifieds whose further relationships are acknowledged but will not be discussed.

Even this oversimplification is sufficient to indicate the problem facing the grid user. If, for example, I construe someone as introverted I may have used this term to signify a meaning-complex of attributes including being broody, self-analytic, introspective, solitary, withdrawn, egotistic .... It may be that only a sample of this indefinite set is sufficient for me to make the judgment of introversion. Alternatively, different samples of the indefinite set (or differential weightings of all set members) may be sufficient to characterise a range of acquaintances as introverted: 'introverted' is therefore overdetermined by its connotations.

Constructs may well overlap. 'Selfish' might connote self-indulgent, egotistic, mean, stingy ... and so on. If it appears in the same grid as 'introverted', there could be an overlap of meaning via 'egotistic'. This would result in a 'natural' correlation whose value would be governed by the subjective influence of 'egotistic'. Not all

constructs will overlap (even though they may derive their meanings contrastively from each other), but work in the area of personality such as implicit personality theory) indicates that a complex overlapping system of constructs is very likely to exist. As was noted earlier, the network relationships are not necessarily on a single hierarchical level. Having articulated 'introverted' and used it in a grid, the respondent may become aware of the connoted 'broody', name this as a new signifier, and use it in the grid. Implicative relationships of this sort become very difficult to identify by statistical analysis, and it is necessary to probe further through the use of language (it being an irony that the complexity of language produces the problem in the first place).

Accepting into the grid labels such as single adjectives or short phrases, the researcher is faced with a problem of meaning, which can be subdivided into the following components.

(i) The verbal label is a cover for a range of connotative meanings.

(ii) The researcher cannot tell which of the connotative meanings underpin the use of the construct in the circumstances being studied, nor assess the weightings which might apply to each.

(iii) The connoted meanings and their weightings might well vary according to the element being construed (i.e. element x construct interaction).

(iv) 'Connotation', in common usage, is an ambiguous concept capable of subsuming both associative meanings and defining attributes[30].

(v) The researcher has his or her own set of connotative meanings, which places a limitation on inference from

the label provided by the respondent.

(vi) Both researcher and respondent operate within sets of socially constructed meanings (which may or may not be coextensive), over each of which has been laid a symbolic idiolect born of individual experience within the social world.

The richness of language derives from connotative rather than denotative meaning[31] and, as Richards (1942) once put it, a word can be seen as a set of possibilities of understanding[32]. The analysis laid out above bears out Campbell (1969) in his assertion that the smaller the linguistic fragment, the more equivocal is the meaning[33].

Saussure made the important distinction between '<u>langue</u>' and '<u>parole</u>', the former being the culturally-established language and the latter being (somewhat loosely) language-in-use. In the grid the link between respondent and researcher is made through the normative aspects of language, filtered by the researcher's own construct system (and limited by social conventions and expectancies)[34]. This can be expected to introduce a measure of distortion whose extent is likely to be a function of the cultural matching between the two individuals involved.

I shall give the grid the benefit of the doubt at this point and ignore the problem of cultural disjunction between researcher and respondent whilst recognising the potential difficulties involved. The point I want to make here is that the grid is likely to lead the researcher to interpret the information provided by the respondent in normative terms. 'Extrovert', 'happy', 'loving', 'stingy', 'cynical', for example, are likely to be construed in terms of cultural norms of

meaning, the current language structure - or, in Saussure's terms 'langue'. As Bergson (1910) pointed out, the language-norms only fix the 'objective and impersonal' aspects of meanings[35]. If this is the case (and synchronically it seems not unreasonable), the grid user is faced with the knotty problems of working out how a quasi-physicalist rating scale can indicate grades of objective meaning between functionally antonymic poles, and how the subjective aspects of meaning can be conveyed.

Whatever is subsequently done with the data in the interests of generalisation, the construct theory researcher places great value on the individual respondent's construing. This implies that the focus of interest is the individual's language-in-use ('parole') rather than its normative framework, since it is that which enables the individual to stand out as figure in contrast to the ground of generality. The respondent may well offer constructs that are idiosyncratic, but the researcher will never detect the idiosyncrasies unless further exploration takes place: 'parole', filtered through the mesh of the grid, becomes 'langue' - or (worse still) the 'parole' of the researcher.

The linguistic reductionism in the grid is, in my view, a strongly contributing factor to the method's failure to do other than scratch the surface of meaning. By restricting verbal labels to brief tags the grid comes close to locking its content into semiotics rather than semantics, when it is clear from the discussion of connotative meaning that semiotics (based on the relationship between signifier and signified) is inadequate on its own to deal with the complexity of human discourse. Natural language goes beyond the semiotic into the semantic. It is the combination of signs into lengthier units of

language (such as sentences) that allows intentions and meanings to be made manifest[36]. As Ricoeur (1978*) puts it, the semantics of discourse is not reducible to the semiotics of lexical entities[37]; the whole is both more than the sum of the parts (in that meaning emerges from the interaction of signs) and less than the sum of the parts (in that a sentence, say, closes off certain possibilities of meaning contained within the signs being used).

There is a difference between the linguistic impoverishment of the grid and the highly elliptical (and on the face of it, impoverished) communication of much day-to-day interaction that has been highlighted by the work of Garfinkel (1967)[38]. Constructs in grids are often less firmly grounded than the fragmented communication typical of much conversation, the latter being very heavily context-based and dependent upon a mutual awareness of context on the part of the protagonists. If the grid construct is applied across a range of elements it falls somewhere between being well contextually-grounded in respect of a few (and therefore strictly speaking irrelevant to the rest) and operating at a diffuse level of generality across them all. In neither case (nor in any of the possible intermediate positions) can a construct maintain a close contextual grounding across all the elements.

The pressure of the argument is in the direction of acknowledging context to a greater extent than most grid-based work would allow. It urges in the direction of speech act theory in which Searle (1969) sees contextually-grounded sentences as the units of communication[39]. This is not inconsistent with the Kellian notion of a relevance-conditioned contrast theory of meaning: contextual grounding provides the relevance framework, and the meaning of terms in the sentences can

be seen in terms of both 'positive' matching to prototypes and the contrasting sets of alternative possibilities.

But contextual grounding implies an increased emphasis on the event; on action past, present and future. The researcher becomes more interested in exploring the event through dialogue with the respondent, listening carefully to what is said, the way in which it is said - and noticing what is left unsaid. For the purposes of this chapter I shall concentrate on the manifest rather than the latent (which is taken up in Chapter 13).

What is said may or may not be true as far as the respondent is concerned: untruth may range from downright lies through the presentation of self in a socially desirable way to an idiosyncratically distorted picture of the 'reality' seen by the world at large. The researcher may see the respondent as self-deluding. Extended conversation will offer clues to the value positions held by the respondent and may suggest contradictions that would pass without notice in grid administrations (where the researcher is often under pressure to get the matrix completed).

The failure to cater for <u>how</u> things are said is a crucial weakness in grid methodology. Tone of voice, the pattern of stress in sentences, and 'body language' can convey a great deal of meaning when taken in conjunction with the actual words uttered. For instance, "I'm not accusing X of bending data" can be delivered 'straight' (as it were) or with an inflection leaving the listener in no doubt of the irony intended.

The particular choice of words is important. Synonyms are never

identical[40]; each has its own connotative net giving it a distinct colouring - 'dread', 'fear', and 'be afraid' all convey different meanings, and it may be possible for the researcher to estimate the significance of the words chosen. Qualifiers change in meaning according to what is qualified: 'attractive' has a different connotative 'ring' when used in connection with buildings as opposed to people - and people may be 'attractive' for very different reasons[41].

Where the qualifier is an adverb, its precise location in the sentence can give a subtle shading to the meaning being conveyed: contrast, for instance, 'she deliberately walked down the pavement', 'she walked down the pavement deliberately' and 'she walked deliberately down the pavement'. Though no absolute readings can be given, the first sentence foregrounds the deliberateness of the choice to walk down the pavement, whereas the third emphasises the manner of the walking. The second sentence seems to leave 'deliberately' hovering between the possibilities of meaning present in the other two sentences: without further evidence it is difficult to resolve - even probabilistically - the ambiguity, and one's attention switches back and forth in a linguistic analogue of the phenomenology of the Necker cube[42].

In other words, different forms of linguistic expression are not cognitively equivalent[43]. Yet a repertory grid itself minimises the chances of emergence of forms of expression which give clues to the meaning intended by the respondent[44]. The researcher using the grid may be so set upon the task of completing the grid matrix that its contextual embedding - the conversation by means of which the administration of the grid takes place - is overlooked. Method may be given priority over meaning.

The argument so far presented in this chapter has much in common with the position of Austin and Searle regarding ordinary language philosophy. The sentence in its context is taken as a more appropriate unit for the conveying of meaning than its constituent words, notwithstanding the contribution that semiological analysis is able to offer. Although - as Labov (1973) points out - a limited number of words can produce an almost unlimited number of meanings, the problem of semantics is lessened when the influence of both rules of language and context of use are taken into account.

And whilst on the subject of words it is worth mentioning that people may well have a far greater number of words available than the typical repertory grid investigation tends to suggest. Elicitation normally seems to produce between ten and twenty constructs before the respondent 'dries up'. The elicitation procedure, however, usually requires the respondent <u>not</u> to produce constructs broadly similar to those already elicited, on the grounds that the researcher is seeking to explore the various boundaries of discrimination within which the respondent operates. Thus grosser distinctions may be emphasised at the expense of the fine discriminations which the respondent could make if asked, and this seems to be an unacknowledged feature of much grid research. Whilst individuals probably do not have an infinite personal dictionary[45] of constructs available to them, Oswalt's (1974) work suggests that the number of constructs capable of being articulated by an individual may be large. Using a range of open-ended elicitation procedures (i.e. not triadic or dyadic elicitation) Oswalt obtained from eight university students roughly 3,800 words to describe 380 people, though the total number of <u>different</u> words seems to have been of the order of 400. Oswalt does not indicate the number of different words used by each of the individuals in characterising people, but an

estimate of around 100 would appear not unreasonable. This would represent something between five and ten times the number of constructs appearing in typical repertory grids even allowing that many of the words elicited by Oswalt could probably be 'collapsed' into antonymic pairs[46].

11.4 DOES FUZZY SET THEORY OFFER A POSSIBLE SOLUTION TO THE PROBLEMS OF GRID METHODOLOGY?

A section on fuzzy set theory might, at first sight, seem out of place in a chapter dealing with linguistics. However, the work of Zadeh (1976) and his followers makes a specific connection between the two which is beginning to make its influence felt among the avant-garde of construct theorists. It is therefore necessary to consider some of the issues involved.

The empirical work presented in Chapter 8 regarding the psychological structuring of scales suggested that it is possible that a scale appearing to be a simple opposition (in the semantic differential sense) could in practice have a complex psychological structure. In other words, membership of the construct poles may not be mutually exclusive (one can, for example, think in both concrete and abstract terms even though the bipolar construct elicited might imply opposition of concreteness and abstractness). In such circumstances one could, say, have a full membership (1.0) of the 'concrete' pole and a partial membership (0.3) of the 'abstract' pole. A 'standard' repertory grid is incapable of coping with the problem.

Roberts (1981) has been experimenting with the assessment of the degree of membership of each pole of a bipolar construct taken separately; for instance, showing that an in-service F.E. teacher construed another

teacher as being wholly student-centred (membership 1.0) and partially subject-centred (0.4). She therefore collects two matrices, one for all the left-hand poles and one for all the right-hand poles. These are superimposed to give a 'D matrix' which indicates the degree of overlap between the poles. Whilst the analysis and interpretation of the D matrix present the researcher with problems, there are difficulties with the initial assumptions. What Roberts is in effect doing is to convert a bipolar construct into two eleven-point unipolar constructs and to ask for ratings of elements on each. Whilst the labelled ends of such unipolar constructs may be clearly construed, the unlabelled ends lack definition, leaving the researcher with the problems of unipolar scaling noted in Chapter 8. Responding to such a grid I found difficulty in assigning meaningful ratings once the element was noticeably distant from the 'full' membership of the pole being considered. Ranking could be a way out of that problem - but at the cost of the statistical meaning of the D matrix. Even if one is prepared to accept the numerical basis of Roberts's fuzzy grid, there still remains the vital question of meaning.

Zadeh (1976) emphasises the importance of linguistics in humanistic systems where, as the complexity of the system increases, the ability to make precise and yet significant statements about its behaviour decreases. Precision has to be sacrificed in the interests of significance, leading to the substitution of the fuzzy set for the unit of measurement typical of 'hard' science. This requires the use of linguistic modifiers such as 'extremely', 'very', 'quite', 'more or less' and so on within a framework of fuzzy logic[47].

Gaines (1976), following Zadeh, takes the view that there is no difference in principle between physical and abstract constructs as far as reasoning

based upon fuzzy set theory is concerned: 'tallness' and 'beauty' are as primitive as each other in everyday reasoning despite the greater ease of explication of the former and the multi-dimensionality of the latter[48]. However, at the level of rating (or degree of membership) practice there would seem to be a difference. Elements can be fairly easily ordered on physical constructs such as those connected with mass, length and time - and measurements can often be made with considerable precision. Where the construct is less precise (for example, 'big') the degree of membership of the pole implies measurement on a scale whose nature lies somewhere between ordinal and interval, whether membership is presented in terms of numerical proportions or linguistic qualifiers. One would expect a monotonic relation to hold between the two modes of describing membership of such a pole. It is not surprising to find that fuzzy set theory is successful in applications such as the control of an industrial process whose complexity defies the operator to fine-tune each parameter continuously and where a measure of slack can be permitted in the cybernetics of the technology[49].

'Beauty' and 'generosity', to take two examples, present greater problems. It is possible, as Zadeh does, to classify women in terms of their membership of the construct pole 'beautiful' (Helen's membership being 0.8 and Jillian's 0.85) - but this provides no information as to the grounds on which these judgments have been made, unlike the pole-membership of physical construct. Helen's rating may have been given largely because of her skilful use of taste and style in dress whereas Jillian's rating might reflect physiological parameters such as skin, bone-structure and body-proportions. Further, this approach does nothing to disentangle the confounding of frequency of occurrence with strength of occurrence - a critical problem for grid methodology in particular, and rating procedures in general. Put another way, the

fuzzier the basic concept, the greater is the possibility for element x concept (construct) interaction and the less is the likelihood of monotonicity between numerical and linguistic representations of the membership of construct poles. It seems that manipulations of either form of data — presentation will be making very considerable assumptions about the basic nature of the data, and the legitimacy of operating on data such as '0.5 (or fairly) generous' is very much open to question. A Kellian perspective could be expected to put more emphasis on the semantic implications of the terms used than upon statistical or logical manipulations and, with the element x construct issue surfacing yet again, the thrust of construct theory is given a further nudge in the direction of construing 'the events as they appear'.

Concepts in ordinary language exist in a haze of fuzziness, as has already been demonstrated by Rosch's work on category membership. G. Lakoff (1972) draws upon Rosch's early work in his presentation of the idea of 'linguistic hedges' which are imprecise but indicative qualifiers of category membership. Lakoff emphasises the 'hedging' possible in statements of noun-category membership (X is [to some extent] a member of category Y), but the potential for fuzziness seems rather greater in the situation pertaining to fuzzy grids where the statements are typically couched in the form of an adjectival relationship (X is [to some extent] Z). The range of connotations embedded in the adjective Z would seem to give elements more latitude to move within the meaning of a particular degree of pole membership.

Linguistic hedges present further problems for the analyst seeking some form of monotonic transformation from hedges to numerical relationships capable of statistical analysis. Mosier (1941) and Sapir (1944), for instance, indicate that in ordinary language qualifiers exhibit complex

and subtle effects which cannot be accommodated in monotonic transformation[50]. There may be semantic confusions. Lakoff (1972) shows how these can arise from the metaphorical nature of language (contrasting 'Esther Williams is a regular fish' with 'John is a regular bachelor'), and makes it plain that the truth value of the sentences as a whole depends not only on the literal meanings of the predicates involved, but also upon their contexts and connotations[51].

Lakoff sees hedges as offering a strong challenge to logicians and indicates that their characteristics render them unamenable to operations within the framework of fuzzy logic. Given the complexity of signification and reference in language and the state of the art of fuzzy logic as applied to linguistics, it is hard to avoid the conclusion that the ways in which people construe their worlds are unlikely to be greatly elucidated through the use of the fuzzy set theorist's high technology.

11.5 IMPLICATIONS FOR METHODOLOGY

The analysis presented in this chapter offers a serious challenge to grid methodology at the level of the semiotic, in that construct labels may be an inadequate sample of those available and that their meanings may remain ambiguous if they are unrelated to a broader linguistic context. It will be recalled that some of Pope's (1977) students were unable to remember the meanings of their own constructs during a subsequent grid administration, the verbal tags on their own having proved insufficient to 'lock' the meaning[52].

Further, the analysis challenges grid methodology at the level of the semantic on the grounds that its inherent limitations make it difficult - if not impossible - for subtle shades of meaning to be conveyed. Where

sentences are used, they tend to be treated <u>adjunctively</u> to the task of completing the grid matrix, rather than to constitute evidence in their own right: the researcher's attention may be focused upon the grid at the expense of the total communication of the respondent.

What, then, are the implications for methodology?

It is clear from the preceding argument that methodology must go beyond the semiotic, important as the semiotic is for understanding[53]. The linguistic signifiers - which I am here loosely equating with construct labels - are likely to be indeterminate unless they are embedded in a framework of utterance which indicates the bounds of meaning intended by the respondent.

Structuralist analysis of grid data might be considered to be a way out of the problem, but - as Wilden (1980) points out - structuralist thinking confuses the building blocks from which meaning is constructed with meaning itself:

> 'In spite of the important contributions it has made... structuralism fails in the life and social sciences in exactly the same way and for exactly the same reasons that both structural linguistics and information science fail in those areas. They are all anti-semantic in that they substitute the supposed characteristics of a theoretically neutral INSTRUMENT OF ANALYSIS (the 'bit') for the USE to which it is put, as an INSTRUMENT OF COMMUNICATION, at given levels in a given goalseeking system, where no information is ever neutral. Meaning - the goal - becomes bounded not by the structure of the context in which it occurs, but by the structure of 'science'. As a result the methodology implicitly becomes an ontology'. 54

Ricoeur (1978*) seems to be making a broadly similar point when he identifies signified meaning with the semiotic order and intention with the semantic[55]. Critical weaknesses of structuralism are its implicit assumption that language and communication are isomorphous,

and its consequent failure to develop a theory of levels of communication[56].

If personal construing relates to a nexus of meaning, intention and action - as I so construe it - the reductionism of the grid would seem to eliminate a considerable proportion of the information which the researcher grounded in construct theory might be seeking. At the end of Chapter 8 I suggested, parenthetically, that the grid might actually be collecting answers to questions in which the researcher had no interest. Putting this another way, it is unlikely that the researcher will be interested in the location of each element on each construct (suspending for the moment my reservations regarding the physicalist nature of scales), and might prefer to concentrate the inquiry on those cells of the grid matrix most pertinent to the problem at hand. It would appear that a number of researchers working from a construct theory base have allowed themselves to become locked in a prison cell, the outlook from which is barred by the framing of grid methodology.

To be fair to those researchers, the developments in methodology have tended to be concerned with refinements of grid technique and analysis rather than with a fundamental reappraisal of what the grid seeks to achieve. In practical terms, these relate to the semiotics of verbal labels rather than the semantics of meaning, intention and action. It is perhaps now appropriate to begin to develop an alternative conceptual framework for the methodology of research founded upon Personal Construct Theory.

If this research is to draw upon the 'parole' of the respondent, it follows that he or she must be given the opportunity to articulate as full a response as possible in order that the possibility of misconstruction

on the part of the researcher might be minimised. To expect this level of exploration in respect of each cell in a typical repertory grid would clearly be unrealistic: the choice has to be made between a superficial skimming over the whole grid (the present position) and a study in depth of those aspects of the grid believed to be most salient to the investigation. The argument I put forward leans heavily to the latter, though in practice the selection of events from a formalised grid framework might not take place - instead, the researcher might seek to build up a partial grid whose boundaries would be progressively defined through interaction with the respondent.

This latter approach emphasises the contextual embedding of sentences in the attempt to gain an understanding of the respondent's construing and, as such, is closer to speech act theory than to semiotics. The meaning of an event is likely only to be communicable through a set of sentences which are oriented in its general direction and which progressively refine the boundaries of possible meaning: there is a loose analogy here with the notion of successive approximations to a limiting value which can never be precisely defined or reached.

The methodological implications of this position are twofold. First, a deep study of selected events is likely to require extended conversation in order that their features and meanings be explored as fully as possible. The evidence from conversations will inevitably be extensive, making some form of electro-mechanical recording almost essential if some of the paralinguistic aspects of communication are to be retained. There is a real danger that conversations will ramble round the foci of interest without fully exploring at least some: the researcher needs a systematic framework for the investigation which is

not a set of imposed prescriptions but which derives from what the respondent is wishing to communicate. It is at this point - rather than at the point of analysis - that the computer may be able to assist the research given its capacity to act as a very efficient administrator[57]. Eden's work at the University of Bath is suggestive of ways in which an interview might be structured along the lines of the respondent's thinking without allowing too many conversational 'loose ends' to remain. This line of thought is explored a little further in Chapter 13.

Eden's work is also of interest in respect of the second methodological implication I wish to pursue, that of analysis. It is transparently obvious that the standard statistical approach to analysis is inappropriate for the type of evidence collected as a result of conversation. The outcome of a conversational approach - the 'sparsely lighted cross-word puzzle' relating elements and constructs (to return to the metaphor used at the end of Chapter 8) may, however, allow a relational analysis to be undertaken. This would relate elements and constructs by way of implicative networks such as those produced by the COPE program developed by Eden and his colleagues. Examples of this type of analysis are given in Eden et al (1979, 1982)[58], but the computer is not essential to the approach since statements made by the respondent can be converted 'by hand' to an implicative flow diagram. However, this rapidly becomes a very complex process as the number of statements increases - and there are many problems associated with the need to make inferences, and with the validity and reliability of any inferences that are made.

The mesh of sentences minimises the potential for polysemy in each one taken in isolation. Each speech act is accompanied by a contextual linguistic framework in which it is located, enabling the researcher to

maximise the chances of arriving at a meaning consistent with its setting - a setting which, in addition to the words used by the respondent, is

> '.... made up of all he knows about the person who pronounces it, about his past experiences, his plans, about what the author of the phrase knows and thinks about those for whom the phrase is intended, and so forth ... ' 59

This is not to imply that each component of the mesh of sentences is completely consistent with the remainder - an assumption that Kelly is at pains to dismiss when he writes that the Modulation Corollary 'tolerates inconsistency between subsystems'[60]. The statements made by the respondent may well reveal contradictions, double-binds and suchlike which are of importance to the researcher. Of equal significance may be the dogs that didn't bark - those matters which might have been mentioned, but which were not.

The point I have been working towards, and which can now be articulated, is this. Even though the fundamental unit of the speech act may be the contextually grounded sentence, this is likely to be too molecular a unit for the purposes of research aimed at elucidating meaning. Although the sentence may serve as the fundamental unit in some contexts, I would wish to press - from the point of view of the researcher - that the paragraph, or perhaps even the text <u>in toto</u>, should be considered the fundamental unit of meaning[61].

Meaning unfolds over time in an interview or a conversation. In addition to setting the respondent's statements against an external context, the researcher has the opportunity to test the 'text' for internal consistency as it is sequentially revealed, and a comprehensive reading requires its component parts to be tested against its wholeness (and <u>vice versa</u>). Such an approach will draw on semiotics and semantics, but its conceptual framework is superordinate to these. It is towards

hermeneutics that the attention of the construct theorist is inexorably drawn.

11.6 SUMMARY

In this chapter the 'transcendent self' implicit in Kelly's theoretical formulation was strongly challenged from the standpoint of the power of language to influence construing, evidence being adduced from a wide range of contexts.

The case was argued that the fragments of language typically used as verbal labels in repertory grids impoverish the communication of meaning between researcher and respondent, and that they leave it unclear as to whether what the researcher subsequently reports reflects the respondent's language-in-use ('parole'), the normative aspects of language ('langue') or the researcher's own 'parole'. Signifiers are over-determined by meaning, and the communication of specific meaning requires more than the simple adjectives or adjectival phrases often found in grids. Semiotics, though necessary (and illuminating), is insufficient to deal with the semantics of discourse.

The claim of ordinary language philosophy that the 'speech act', or contextually grounded sentence, is the fundamental unit of communication advanced the argument into the realm of semantics. The important distinction was made between language and communication for the purpose of showing that structuralism is unable to do other than contribute to an understanding of communicated meaning.

The development of the methodological implications foreshadowed in previous chapters led to the view that the researcher operating from a construct theory standpoint might well gain a greater understanding of

the topic being investigated by trying to probe certain element X construct relationships in depth whilst ignoring those that were peripheral or irrelevant. This position was sharply contrasted with the blanket coverage of all element X construct relationships in the repertory grid. The consequence of this argument was that it would be necessary to examine relationships in terms of implications rather than statistics, and that this would enable contradictions, double-binds and 'silences' to be explored. Such an approach, which would explore part-whole relationships, is fundamentally hermeneutic in character.

This chapter is the point of transition between the retrospective analysis of theory and method contained in Chapters 3 to 10 and the speculations in Chapter 13 regarding possible developments in both theory and research practice. It signals a rereading of Kelly the scientist as Kelly the historian, but before this is done, it is necessary to draw together in Chapter 12 my main conclusions regarding repertory grid methodology.

12 Sieve or riddle ?

The question of validity

And what you thought you came for
Is only a shell, a husk of meaning....

T.S. Eliot
Little Gidding

12.1 VALIDITY AND REPERTORY GRID METHODOLOGY

Beneath all the preceding discussion of repertory grid methodology has lain - ever present, but largely unacknowledged until now - the question of validity. There is no single determinant of what constitutes the validity of repertory grids, for the validity of any grid will depend upon the purposes for which the grid is being used.

Cronbach and Meehl (1955) draw the now familiar distinction between four ways of construing validity:

> predictive validity - in which future performance is the criterion against which the test data are set;
>
> concurrent validity - which depends upon the correlation between the test and another, established test;
>
> content validity - in which the validity depends on the demonstrations that the test is composed of a representative sample from the relevant universe; and
>
> construct validity - referring to the extent to which the test can be taken as a measuring instrument for a particular attribute or quality. 1

Of these four aspects of validity, the discussion regarding the sampling of both elements and constructs (Chapters 5 and 6, respectively) has indicated that content validity of the grid is indeterminate and open to some doubt. Construct validity is difficult to assess since the grid user is often not seeking to measure a particular attribute or quality - and, when he or she is attempting such a measurement (such as cognitive complexity or intensity) the meaning of what is to be measured proves to be as elusive as plankton in a shrimping net.

Kelly is often quoted as saying that validity 'refers to the capacity of a test to tell us what we already know'[2]. This is tantamount to defining validity in terms of concurrency - the extent to which the

information derived from the repertory grid correlates with some accepted yardstick of the attributes, characteristics or suchlike being measured. But he was well aware of the circularity of construing validity in terms of concurrency, and preferred to view validity in terms of usefulness[3] - thus keeping faith with pragmatism (or, at least, that version espoused by William James). Clearly, embedded in the notion of usefulness is the idea of prediction which is, after all, a central tenet of Kelly's theorising.

Fransella and Bannister (1977) point out that usefulness is not seen by Kelly in terms of the prediction of 'some arbitrarily chosen and relatively trivial aspect of human behaviour'[4], but they evade the discussion of the actual nature of usefulness by presenting a number of examples culled from the literature and covering a range of repertory grid investigations. This is not entirely satisfying, since any application of a grid can be claimed to have produced something useful - if not (say) to the clients in psychotherapy, then to the clinicians or researchers involved. Fransella and Bannister offer the ironclad definition of the validity of a technique as 'its capacity to enable us to <u>elaborate</u> our construing'[5], and they continue by pointing out that this elaboration may be in terms of extension or definition, in accordance with the Choice Corollary.

Validity depends upon context. If the purpose of administering a grid is to acquire information in order that the researcher may predict the respondent's future behaviour, then the criterion of validity is the extent to which behaviour matches prediction. The situation can be inverted, in that the researcher may already have data (such as information, or behaviour) recorded about a respondent and use the grid to try to seek underlying cognitions that might help to provide an

adequate explanation for what is known. In this circumstance the grid is valid to the extent that it can contribute to the explanation. In each of these two examples there is, embedded in the notion of validity, an unacknowledged 'concurrent' component which is the normative conceptual framework within which the researcher develops predictions or deems explanations adequate or otherwise[6]: the grid helps the researcher to focus on those parts of the vast cognitive reservoir likely to prove relevant to the purposes at hand. In that sense, the grid - as any other method of inquiry - will, if valid, 'tell us what we already know'.

What is missing from Fransella and Bannister's definition of validity is a relativistic perspective. Methods other than the repertory grid may be superior for particular types of investigation without necessarily implying the need to abandon a personal construct perspective. It will not always be possible to give a numerical coefficient to the validity of research based on Kellian theory, and validity may have to be assessed on the qualitative grounds of procedural evidence and argument. If validity is treated in this way, this 'procedural validity' will have to be established afresh for each piece of research and its justification will need to be included as an essential section of any research report in order that the reader may properly evaluate it.

Fransella and Bannister assert that 'it is not sensible to dispute the validity of <u>the grid</u> as such'[7]. In so far as there is a wide variety of grids constructed to fulfil a variety of purposes, this is fair enough: they, too, make the point that the validity of <u>a</u> grid can reasonably be evaluated only with reference to the grid user's purposes. They go on to suggest that, if a particular grid were to prove to have no predictive value, they would look for flaws in the format rather

than make any general statement concerning grid methodology.

However, the evidence and argument presented in the preceding chapters suggest that there are generic problems regarding grids as investigative instruments. A grid is of little value if its conceptual base has not been thought through - and their remark that grid use 'is only limited by the user's lack of imagination'[8] tends to divert the reader from the task of rigorously conceptualising both problem and method.

My appraisal of the repertory grid leads me to suggest that the problem of validity is much larger than Fransella and Bannister imply, and that many general points can be made which bear, explicitly and implicitly, on the validity of grid methodology[9]. These are presented in the following section and represent the main inferences I have drawn from the evidence and argument contained in the preceding chapters.

12.2 FACTORS INFLUENCING THE VALIDITY OF A GRID

The statements listed in this section are a distillation of Chapters 4 to 11, and are based on an appraisal of the evidence in the literature, empirical enquiry, and reflection upon principles and practice. Whilst all the statements are either explicitly or implicitly concerned with the validity of grid methodology, they also act implicitly as recommendations for practice. Each statement is supported by a brief outline of the evidence and/or argument adduced in support, and cross-references are given to the main discussions given in the body of the text.

12.2.1 The grid itself

A. Validity is compromised when the context of the grid is not defined.

In some of the studies surveyed it has been possible for respondents

to slip from one context to another during the completion of the grid.

(Section 5.2)

B. <u>The larger the segment of time explicit or implicit in the context of the grid, the greater the scope for the 'averaging-out' of responses</u>.

In grids dealing with personality, for example, there is typically no explicit limit on the duration which respondents are expected to bear in mind when they complete the matrix. Thus a person who is arrogant in some situations and humble in others may be construed at either pole or in an 'averaged' way, depending on how the respondent construes the task. This potential vagueness is reduced as the 'time-slice' to which the grid applies is made smaller: in the limit, the grid is reduced to single events which may no longer be commensurate.

(Sections 8.3.7; 8.4)

12.2.2 <u>Elements</u>

A. <u>The validity of a grid is a function of the match between the elements and the context in which the grid is set</u>.

If extraneous elements are allowed to enter the grid they may influence the elicitation of constructs and the subsequent rating procedure. It then becomes difficult to interpret the findings in respect of those elements which are appropriate to the context of the grid. Where no control is exercised over the elements included in the grid, the meaning of the procedure approaches the vanishing-point.

(Section 5.2)

B. <u>Elements which are irrelevant to the respondent may have an adverse effect on validity</u>

There is a danger that elements which seem relevant from the researcher's point of view may be irrelevant to the respondent -

particularly when elements are provided. Such elements merely provide 'noise' in the grid.

(Section 5.3)

C. 'Homogeneous' elements tend to enhance the validity of a grid matrix

If the elements used are a diverse collection, it becomes difficult for the respondent to provide constructs which apply to all of them. This leads to outcomes such as grids with a limited range of - probably trivial - constructs, grids with many vacant cells, or (if a rated grid is used) the allocation of many elements to the mid-points of the scales being used.

(Sections 5.2; 5.3)

D. The validity of grid procedures depends upon the representativeness of the sample of elements

The representativeness of elements is a difficult criterion to adopt. The sampling of elements has to be justified in qualitative terms where circumstances are such that the population of elements is indeterminate. Where the population is isomorphic with the sample (such as when a class of pupils is used), the question of representativeness does not arise.

(Section 5.3)

E. The more specific the elements, the greater the validity of the grid procedure is likely to be

The more precisely specified are the elements, the less scope there is for the respondent to move around in a cloud of vagueness whilst construing. Such evidence as is available suggests that there is a descending order of specificity from concrete objects through well-known individuals and specific incidents to generalised aspects of life-experience.

(Section 5.5)

F. The decision whether to supply or elicit elements (or to use a combined approach) is a factor to be taken into account in evaluating the validity of a grid.

There is no simple relationship between the origin of the elements and the validity of the grid. Any statement regarding validity must take into account the researcher's purposes and the justification given for the particular procedure adopted.
(Section 5.4)

12.2.3 Constructs

A. The decision whether to supply or elicit constructs (or to use a combined approach) is a factor to be taken into account in evaluating the validity of a grid.

The argument is similar to that given in point 12.2.2F above, save that although the researcher may choose to specify the elements (and justify doing so in the interests of defining the problem to be addressed) the grounds for specifying constructs are typically much less firm.
(Section 6.6)

B. The validity of grid procedures depends upon the representativeness of the sample of constructs.

This point parallels that of point 12.2.2D above, but is much more difficult to justify in practice. Within a grid there is no guarantee that the researcher has gained a representative view of the way the respondent construes that part of the world which forms a focus for the research: it is naive to assume that, because a respondent has apparently exhausted his or her stock of dimensions, an asymptote of communication has been approached. Though it is impossible to assess the extent to which the ideal of representativeness has been reached, this ideal is likely to be more closely approached where the researcher seeks to cross-validate the constructs elicited. This may be attempted by referring to other available information such as recordings of the elicitation procedure (which can contain constructs that do not reach the

formalisation of the grid), and the actual behaviour of the respondent.

(Sections 6.3; 6.4)

C. <u>The nature of the sample of constructs obtained may be determined by the particular elicitation procedure adopted</u>.

In other words, the method of elicitation may facilitate or block the respondent's production of constructs. Whilst the 'triadic' form of elicitation appears to be the most widely used in repertory grid work, it does not always facilitate the elicitation of constructs, for some respondents appear to find the cognitive demands of the procedure alien to the way they think, or would prefer to respond. 'Dyadic' elicitation has been used on a number of occasions, apparently successfully, without compromising the Kellian assumption of similarity and contrast. Freer forms of elicitation (such as interviewing) have been used in clinical situations, but in general these have not been used to provide constructs for use in grids. Their potential richness suggest that further exploration in this direction may prove profitable to researchers working within a construct theory perspective.

(Section 6.5)

D. <u>Without careful specification of the grid procedure the existential status of a construct may be unclear</u>.

There is little evidence in the literature that researchers have given attention to whether the constructs they elicit are oriented to the past, the present or the future - the verbal labels actually supplied by respondents may imply any or all of 'what was', 'what is', 'what is expected to be ', 'what ought to be' and their subvariants. Further, the more the elicitation relates to the construing of the unfamiliar (i.e. is a problem-solving activity), the greater is the possibility that the respondent is elaborating

his or her construct system (and the extent of recourse to pre-existing construct relationships would be lessened)
(Section 6.8)

E. <u>No valid claim may be made, on the basis of a repertory grid, for the hierarchical level of any construct.</u>

Assuming, with Kelly, a hierarchically organised construct system, it is necessary to use instruments oriented towards structure in order to gain an indication of the hierarchical level of any construct and its implicative relationships.
(Sections 4.6; 4.7; 6.5.4; 9.6.1)

F. <u>No valid claim may be made, on the basis of a repertory grid, for the salience of any construct to the individual.</u>

Whilst there is likely to be some positive relationship between the salience of a construct and its hierarchical level, it is doubtful that this is monotonic (in any case, both salience and level may vary with context). The probable lack of monotonicity may reflect the tension between - referring to Hinkle's methods - the logical entailment of an implications grid and the ego-involvement of a resistance-to-change grid. Evidence relating to this would have to be sought beyond a repertory grid. Anticipating the conclusions regarding analysis (Section 12.2.6), it is a mistake to relate salience to the size of factors derived from principal components analysis or from the grouping of constructs by cluster analysis.
(Sections 3.3.3; 9.7.2; 9.7.3)

12.2.4 <u>Bipolarity</u>

A. <u>It is doubtful whether Kelly's claim for the universality of dichotomous construing can be sustained.</u>

Linguistic qualifiers frequently have meaningful antonyms, though there are examples where the only possible opposition is a

negation (if 'peculiar' constructs are to be avoided). Noun and verb constructs frequently seem to have no specific opposites, but stand instead in opposition to 'contrastive sets' of terms. Further, evidence has been presented which suggests that construing may, in many circumstances, be of a positive pattern-matching type.

(Sections 3.3.2; 8.3.7; 8.4; Chapter 6, passim)

B. Some constructs appear as negations (i.e. X - not X) rather than Kellian relevance-conditioned contrasts (of the X-Y type).

Although this point is made in 12.2.4A it is worth emphasising because of the difficulty the researcher may have in interpreting the 'not X' end of the construct. Negations also call in question Kelly's assumption that all construing is limited to the relevant range of convenience, and also his insistence on rejecting the tenets of classical logic.

(Sections 3.3.2; 8.3.5)

C. There is a considerable risk that elicited constructs (especially 'peculiar' constructs) will be 'bent'.

'Bentness' may arise when a contrast pole is selected from a group which includes both a superordinate contrast and its subordinate implications. In effect, this amounts to the use of the emergent poles of two different constructs - and it may cause difficulty when the respondent attempts to locate elements on a dimension whose extremes are defined in such a way.

(Section 7.3)

D. There appear to be loose associations between

(i) the 'oppositional' method of elicitation and 'logical' constructs, and

(ii) the 'difference' method of elicitation and 'peculiar' constructs.

Whilst the 'oppositional' method appears to elicit a greater proportion of constructs that are functionally antonymic, the 'difference' method may produce the more personally meaningful discriminations on the part of the individual - even though the

oppositions elicited are 'peculiar'. There is a particular threat to validity of the grid (not necessarily to - say - triadic discrimination) when 'peculiar' constructs are used, since they are more likely to exhibit restricted permeability than 'logical' constructs: there is an increased likelihood that elements will fall outside their range of convenience.

(Section 7.2)

E. <u>Oppositional construct poles are not necessarily mutually exclusive</u>.

The point is openly acknowledged in respect of semantic differential work and seems equally valid in respect of repertory grid investigations. Although an opposition may be elicited as a result of considering - say - a triad of elements, it is possible for other elements to claim membership of both poles within the context of the grid. Whilst the problem is reduced by decreasing the explicit or implicit time-slice to which the grid procedure refers, the cost is likely to be the 'collapsing' of the grid.

(Section 8.3.7; Chapter 6, <u>passim</u>)

F. <u>Construct poles tend to be inherently vague, and vagueness compromises bipolarity of construing</u>.

Many construct poles - especially those at levels of abstraction beyond sense-experience - are portmanteau terms. It is unlikely that respondents maintain a fixed content in each portmanteau as they construe a range of elements in terms of the relevant dimension. Thus a person construed as highly extrovert may nevertheless exhibit a few introvert characteristics: the generality of many grid dimensions may fail to resolve the fineness of such distinctions.

(Section 8.3.6)

12.2.5 <u>Completing the grid matrix</u>

A. <u>Dichotomous allocation of elements to construct poles does not reflect the shadings of reality and produces problems for the grid analyst</u>.

Shades of meaning are often important, and the crudity of dichotomous allocation can do violence to what the respondent wishes to convey. 'Lopsidedness' creates problems for analytical routines which are based on matching scores: the 'split-half' grid suggested as a possible remedy solves some analytical problems - but often at the expense of meaning.

(Section 8.1.1)

B. <u>The ranking form of the grid is useful when unipolar scales are being used, but tends to limit analysis to interrelationships between constructs</u>.

There are advantages in using ranking techniques with respondents (for instance, children) who might find rating too complex, but ranking becomes more difficult as the number of elements increases. Rank correlations between constructs can be computed, but element similarity scores derived from the grid matrix lose most of their meaning.

(Section 8.1.2)

C. <u>The labels given to the ends of construct dimensions (presented as rating scales) are likely to be more clearly construed than the mid-point of the scales.</u>

This is not to claim that the mid-point is necessarily unmeaningful: it <u>may</u> be unmeaningful, but it is not necessarily so.

(Section 8.3.8)

D. <u>There are many ways of construing the mid-point of a rating scale</u>.

The underlying metric of each rating scale is rendered problematic to an extent determined by the number of possible constructions of the mid-point.

(Section 8.3.8)

E. <u>There is empirical evidence to suggest that a bipolar rating scale is open to use in manners which vary between individuals and which have differential metrical implications</u>.

In conjunction with earlier points which question the validity of the assumptions of mutual exclusivity (12.2.4E) and functional

antonymy (12.2.4C and .4F), the evidence strongly implies the need for the researcher to find out <u>how</u> the respondent actually uses the scales contained in a rating grid. Further, it may be possible to make some assessment - even if only qualitative - regarding any possible anchoring effects.

(Section 8.3.7)

F. <u>Whilst there is evidence to suggest that extremity of rating is a function of personal meaningfulness and perhaps psychopathology, it is possible that the respondent's attitude to the research may, on occasion, be a dominant factor</u>.

A fair measure of research has probed extremity of rating from the point of view of personal meaningfulness and psychopathology, but little has been done with regard to the way the respondent approaches the task of completing a grid matrix. Personal experience suggests that extremity of rating could be a strategy through which the respondent gets through the task in the minimum of time and with the minimum of effort, though the evidence for this proposition is at present extremely thin.

(Sections 6.6.2; 10.8.1; 10.8.3)

G. <u>There is some evidence to suggest that rating procedures are vulnerable to human error, especially when grids are administered to groups of respondents</u>.

The most probably error would seem to be that of reversing the rating scale, particularly when the pair (in triadic elicitation) is labelled in negatively-valenced, or implicitly negatively-valenced, terms. It is easy to fail to notice the occurrence of scale reversal and other errors, and some procedures have been suggested which appear to be successful in overcoming the main difficulties.

(Section 8.2.1)

H. <u>Rating behaviour appears to be influenced by social aspects of language</u>.

Given the present state of knowledge, all that can probably be said here is that the effects of lexical marking, evaluative affect and cultural stereotypy may all interact in a complex way to affect rating behaviour. Unexplored to date are the complications introduced by the range of possible oppositional and negational constructs.

(Section 8.3.5)

I. <u>Element X construct interactions are likely to be a prominent feature of the rating patterns in grids.</u>

There is substantial evidence for this claim in the annals of research using the semantic differential. The argument has been presented that the effect is highly likely to feature in repertory grids since the associative meanings of the construct poles are likely to vary according to the particular element being construed. The assumption of the stability of the meaning of a construct, irrespective of element, seems to be based on a positivistic theoretical position which is at variance with other aspects of Kelly's theorising.

(Sections 4.8; 8.3.9)

J. <u>There is a possibility that construct X construct interactions may feature in the rating patterns in grids</u>.

This point is advanced rather tentatively since it seems to be unresearched in repertory grid studies. Tajfel's accentuation theory gives some grounds for considering the possibility of such an interaction.

(Section 8.3.10)

K. <u>A phenomenological view of the use of rating scales suggests that the tacit physicalist assumptions often underlying grid work may be untenable</u>.

The argument has been presented that in construing elements in terms of a rating scale a respondent may be making reference to

personally-experienced events and to mediated knowledge stored in memory, rather than using the scale in a quasi-physical manner. If the argument is justified the implications for the validity of grid methodology are profound, and require a radical reappraisal of methodology subsumed under a Personal Construct Theory framework.

(Section 8.4)

12.2.6 Analysis and interpretation

A. Each of the analytical routines generally available has its particular strengths and weaknesses in respect of its capacity to deal with grid data.

Each approach to grid analysis contains its own embedded assumptions regarding the statistical quality of the data with which it deals. The validity of analysis is raised when the researcher maximises the congruence between the purposes of the research and the characteristics of the analytical method.

(Sections 9.4.1; 9.4.2; 9.4.3)

B. Missing data causes difficulties for the grid analyst

None of the methods of analysis that have been discussed completely overcomes the problems of missing data. Whilst the use of the mid-point rating or the elimination of rows or columns containing blank cells allows statistical computations to be undertaken, there are potentially serious implications for the meaning of the output of such analyses.

(Sections 9.5; 9.4 passim)

C. The analysis of combined, or averaged, grid data exaggerates commonality at the expense of idiosyncrasy.

In other words, the variance between individuals is ignored. Aggregation of data becomes meaningless where dubious assumptions of equivalence are made, such as when verbal labels are treated

as identical across a range of respondents or when a role title is implicitly taken as a constant despite the fact that respondents may choose to construe different individuals who fit that specification.

(Section 9.6.2)

D. Statistical analyses are not necessarily concordant with meaning relationships.

Statistical analysis relates solely to the mathematical properties of a grid matrix and does not take into account its semantic characteristics. This is illustrated when the 'evaluative overlays' to constructs are not monotonic with the mathematical scaling: in such circumstances what appears, in logical terms, as a 'straight' construct may become sharply bent when its E-/E+/E- overlay is aligned with that of a construct with a simple E-/E+ gradation of valence.

(Section 9.7.1)

E. Analysis of a repertory grid gives no indication of the implicative relationships between constructs.

A similar point was made in 12.2.3E above. Relationships between constructs cannot be inferred directly from grid analysis, but require the importing of other evidence if any such claim is to be made. A number of researchers have made the mistake of confusing the phenetic with the phyletic, and some have confused the concept of 'hierarchy' in the separate fields of construct theory and cluster analysis. Further, large clusters and factors do not necessarily imply dominance in a construct system despite the claims often made in the literature.

(Section 4.6; 9.6.1)

F. If the challenge to the quasi-physicalism of rating scales in grids is sustained, the assumptive base of standard analytical methods is undercut.

Following 12.2.5K above, if the respondent's use of a rating scale

is grounded in the memory of personally-experienced events and mediated knowledge, then the emphasis in the grid lies towards the matrix columns and the unique way in which each construct is applied to each element. In these circumstances standard analytical methods are revealed as inadequate, as is 'standard' grid methodology.

(Section 9.8)

12.2.7 Stability

A. The evidence reviewed suggests that the stability of grid data is highest when the elements are specific, familiar and 'simple', and lowest when elements are general, unfamiliar and complex.

Concrete objects construed in terms of constructs of high inter-subjectivity tend to produce high stabilities. Stabilities tend to be somewhat lower when people are construed in terms of personality and other attributes, and to decrease further with with increasing generality of the elements.

(Sections 5.5; 10.3; 10.4)

B. It is open to doubt whether there is a consistent relationship between extremity of rating and test-retest stability.

Whilst respondents appear to exhibit consistency in the extent to which they rate towards the extremes of scales, such consistency appears not to obtain in respect of correlations between construct variance and stability within individual grids, and (nomothetically) between indexes of construct variance and stability which are based on 'whole-grid' data.

(Section 10.8.3)

C. A stability index for the grid as a whole can be based on the cumulated change per cell and this minimises the problems associated with missing data and instability of rating in a small minority of constructs.

Stability calculations based on standard correlational coefficients

are at the mercy of gaps in the data, and some are markedly affected by instability in rating in a single construct. A 'per cell' coefficient allows stabilities to be computed for a grid as a whole or for its constituent rows and columns without them being inordinately influenced by either of these problems.

(Sections 10.6; 10.7)

12.2.8 Language

A. Short verbal labels in grids are likely to be ambiguous.

Everyday language contains a considerable degree of redundancy, but often that redundancy allows its near-synonymity to triangulate the meaning intended. A single short verbal label has no capacity for triangulation and hence is likely to be ambiguous - and there is evidence that respondents, faced with their own labels, may find it difficult to recall what was originally meant. The ambiguity present in verbal labels allows the respondents the opportunity to move within their range of associative meaning as each element is successively construed.

(Sections 6.2.3; 10.8.2; 11.3)

B. The typical grid, with its emphasis on producing markedly different constructs from their predecessors, may constrain the the respondent from using the full range of meanings available to him or her.

There is some evidence that respondents have a wider range of verbal labels available than is typically evidenced in repertory grid work. Whilst this evidence is derived from the field of personality characteristics, it would seem not unreasonable to expect the same to be the case in grids dealing with other subjects.

(Section 11.3)

C. <u>The use of short verbal labels may make it difficult for the researcher to understand what the respondent wishes to convey.</u>

In a sense this is a corollary to 12.2.8A in that it focuses specifically on the problems that construct labels pose for the researcher. Whilst culturally-embedded terms may, to a first approximation, be intersubjectively communicable, there is some evidence to suggest that labels which appear straightforward to interpret are occasionally used in distinctly idiosyncratic ways.

(Sections 6.6.3; 11.3)

D. <u>Short verbal labels for constructs are very likely to be an impoverished version of what respondents could, or do, convey.</u>

The grid has been seriously called in question regarding its linguistic impoverishment, and it offers minimal scope for expression of, for example, the metaphoric, metonymic and mythic aspects of communication - often those aspects which convey with the greatest vividness what the respondent is seeking to indicate. Further, the precise location of words in sentences offers additional clues as to the meaning intended by the respondent.

(Sections 4.4; 6.2.3; 11.3)

E. <u>It seems unlikely that fuzzy set theory will prove to be helpful to repertory grid methodology.</u>

The complexity of language (in respect of multiple connotations, metaphoric relationships and suchlike) render problematic the degree of membership of a construct pole, whether this is expressed numerically or in terms of linguistic hedges. Logical operations upon such problematic data can only produce outcomes which are themselves problematic.

(Section 11.4)

Many of the points summarised in the preceding pages of this chapter pose a serious challenge to the validity of the repertory grid as a research

instrument. Lest users of other research methods take undue comfort from this, it has to be said that many of the problems confronting repertory grid methodology also confront other approaches to psychological investigation and may, in those milieux, prove more intractable.

Kellian theory and its methodological progeny the repertory grid have at their centre the construing individual and seek to expose the ways in which individuals 'see' the world, rather than to try to fit them into a framework representing the researcher's construction of reality. Whilst this theoretical position is defensible in ontological terms, it is clear from the evidence and arguments deployed at length in this study that the portcullis in the outer wall of methodology is vulnerable to attack. Far from being an accurate and precise method of gaining access to people's contruct systems, the repertory grid is revealed as partial and vague, and as requiring considerable support from cross-validational inquiry and from the researcher's interpretive skills. On its own, it is manifestly inadequate to bear the burden of the purposes for which some researchers have employed it - but the same may be said of other methods of investigation such as the semantic differential, personality inventories and attitude scales whose connections with theoretical formulations may be rather more tenuous.

12.3 THE REPERTORY GRID AND BEHAVIOUR

One aspect of grid methodology that lies outside the strict terms of reference of the present study nevertheless needs to be considered briefly here - the issue of the relationship between the repertory grid and behaviour. This has implications for validity in terms of external criteria rather than in terms of the internal criteria with which this study has been primarily concerned.

At the beginning of this chapter I suggested that Kelly's construing of validity in terms of usefulness implied prediction. The grid can be evaluated in terms of predictive success in at least two ways. First, the grid may reveal sufficient of a respondent's construct system to enable the researcher or clinician to predict (on the basis of his or her own construct system) the future behaviour of that respondent. Second, the grid data might suggest ways in which the researcher or clinician could deal with the respondent in order to facilitate a particular behavioural outcome. In either case, there is a danger that too much may be claimed for the grid if the predictions are successful (the fact that I see a black raven does not strengthen the thesis that all ravens are black). And if the prediction does not turn out successfully it is indeterminable whether the grid itself is to blame or whether the fault lies in the chains of inference and behaviour that links the grid data with the ultimate observation of the 'target' behaviour. The use of the grid in predictive studies has resulted in both successes and failures (on this criterion), ranging from Fransella and Bannister's (1967) study of voting behaviour (which can broadly be accounted a success) to Bannister and Bott's (1973) failure to elicit crucial constructs related to a couple's sexual relationship.

I also suggested that the grid can be evaluated in terms of its explanatory capacity. Here the grid may be employed in the hope that it will elicit sufficient data to enable the researcher or clinician to make some sense of behaviour that has already occurred and regarding which no satisfactory explanation is yet available. Fransella and Adams (1966) report an interesting study of this type in which repertory grid technique was used in order to discover how an arsonist himself

construed the act of fire-raising, for which a number of competing constructions existed (such as illness, crime and distorted sexual activity). In the event Fransella and Adams found evidence from grid technique to suggest that the patient construed the act of arson in terms of retribution for sinners, and that this unexpected outcome of the grid-based investigation might provide the framework for an explanation of his poor response to treatment.

The majority of studies in which repertory grid methods and behaviour have been linked appear to lie in the clinical field, although examples of this connection exist in other settings, for instance, voting behaviour (Fransella and Bannister, 1967), friendship formation (Duck, 1973) and social adaptation (Hayden et al, 1977). Adams-Webber (1979) remarks that the last of these examples is 'one of the few studies which have succeeded in demonstrating a direct relationship between formal characteristics of an individual's construction processes and his "real-life" behaviour'[10].

It is possible to infer from Adams-Webber's words a criticism of a Kellian psychology perhaps too concerned to look inward at construction systems and insufficiently outward-looking towards their relationship with behaviour. Or, in Argyris and Schön's (1974) terms, has personal construct psychology given too much emphasis to 'espoused theory' at the expense of 'theory in use'? Ought not research to be more explicitly focused on the link between construction systems and behaviour? If the answer is in the affirmative, there are serious implications for research methodology. Problems in the world outside the psychology laboratory are messy, multifaceted and difficult to solve, yet, if psychology is to define its range of convenience to include the life-world, it is with issues of such complexity that it must engage.

12.4 THEORY AND METHOD: A REPRISE

Before I attempt to sketch an approach towards linking construct theory with research in the life-world it is necessary to revisit the relationship between theory and method in order to indicate why I am led to the conclusion that the repertory grid stands in fundamental contradiction to its parent theory (and hence why its validity is compromised at the level of theory rather than of practice). And therefore why a reappraisal of method - and to some extent theory - is a necessity.

It will have been evident at a number of points in my writing that there is a tension between my 'readings' of Kellian theory and of repertory grid methodology. I take the view that the theory, in essence, stresses the primacy and uniqueness of events which - in grid terms - are captured in the interactions between elements and constructs. A strong case has been put forward that constructs, in the vast majority of circumstances, cannot be treated as if they were physicalist scales but alter their meaning according to the particular element being construed. Thus if the elements - i.e. the grid columns - are clearly defined (by no means always the case) the same cannot be said of the constructs, or grid rows. The theory, in my view, emphasises the columns of the grid and tends to treat the rows qua rows as of little value, whilst the method makes no such distinctions despite the problems of validity associated with both 'within row' and 'between row' relationships.

Put another way, the notion of similarity and contrast (which is explicitly captured in the rows of the grid) is far more complex than is acknowledged in the naive nomothesis and latent positivism of a repertory grid's construct dimensions. Whilst it seems wholly reasonable that contrast plays a role in the determining of meaning it is perhaps being over-reductionistic to press this line of argument too far when there is

evidence to suggest that elements may well be construed in terms of prototypic instances or attributes, in which case the contrast - not necessarily dichotomous - is implicit rather than explicit (or, in Kelly's term, is 'submerged').

The argument briefly put forward here (which was presented at greater length in Section 8.4)leads me to suggest that the notion of similarity and contrast is elevated to too important a position in repertory grid methodology, with the result that it tends to blur, rather than sharpen, the researcher's construction of the respondent's construct system. At first sight such a claim appears to be a heresy in Kellian terms, yet I hope to demonstrate that - at a deeper level - there is no need for my committal to the heretic's pyre.

In order to substantiate my claim it is necessary that I develop the argument at some length (drawing upon points made, passim , in Chapter 8) since it is crucial both to the question of grid validity as such and to the direction in which research in the Kellian tradition might wish to take. Although I present my case in terms of a specific example (in the interests of making my intentions plain) it is a case which can be construed in terms of superordinate generality.

Consider the rating of a teacher T on the dimension 'kind-unkind': what is actually involved on the part of the rater (whom I take to be myself, for convenience in presentation)?

If the dimension arises from triadic elicitation I will have characterised, say, two teachers as kind and one as unkind. In other words, the dimension relates to my construing of the behaviour of real people and not to the abstractions at the ends of supplied constructs or semantic

differential dimensions (though I acknowledge that I might have to relate such poles to my experience of kind and unkind actions in order to use supplied scales). Given that I possess some existential understanding of what 'kind' and 'unkind' mean[11], I have three initial problems to solve in determining a rating for T:

(a) I have to come to some decision about the frequencies of kind and unkind actions on the part of T;

(b) I have to take into account that the various actions of T have differing <u>intensities</u> of kindness and unkindness; and

(c) I have to find some way of combining the information from (a) and (b) in order to produce a single rating value.

But step (c) involves me in a comparison of T with those teachers whose behaviour is effectively acting to 'anchor' the dimension: there are no absolutes of kindness and unkindness, and the dimension is - strictly speaking - 'relatively kind - relatively unkind' in practice.

Having anchored the scale in essentially behavioural terms, there is a problem associated with the differing contexts in which the triad of 'anchoring' teachers were construed as being kind or unkind. There is no justification for assuming that these contexts are commensurate - indeed, they must inevitably differ from each other and from the context in which T is being rated. (At this point Kelly might well have wished to argue that the problem demands the abstraction of behaviour to a superordinate level in the construct system: this point is taken up a little later in this chapter and more fully towards the end of Chapter 13.)

In comparing T with the anchors I am making the assumption that I know all four of them - and any other elements I wish to involve - equivalently

well. It may be that the triad of elements that happened to frame the elicitation of the construct 'kind-unkind' were not particularly well known in respect of this distinction, which would then rest upon an extrapolation of the little I know of them: put another way, it is possible that the triad - and T, for that matter - may be more vaguely construed than the ratings appear to suggest.

For the purposes of my present argument, it is important to point again to crucial failure in the rating procedure to give any weight to the problems of the frequency and intensity of actions - see (a) and (b) above. Frequency and intensity depend upon the ability to construe acts according to the dimension being used ('kind-unkind') and to recognise gradations along the scale. Such recognitions are complex, and would seem to demand the construal of each act within its context, rather than the application of stereotypical labels. It seems probable that I need to have already stored in my memory an appropriate schema ('prototype' appears to have insufficient dimensionality here) to enable me to undertake these tasks: if this be the case it seems unlikely that I have any need to make specific reference to the 'anchoring' elements, save in so far as they already appear in the schema. To return to the issue of abstraction raised in parenthesis a couple of paragraphs ago, it would seem that a 'schematic' approach to construing is broadly concordant with Kelly's ideas on the subject, but does not require the anchors implicitly provided by triadic elicitation and its simplistic substrate of similarity and contrast.

Summing up the argument I have advanced, it is doubtful whether overt comparison with others helps me generally to construe the kindness or unkindness of T in the classroom - which is not to deny the possibility (or the value) of direct comparisons in certain circumstances (T_2, for

instance, might be known to behave in a very different way from T under broadly similar conditions). The difficulty of comparative construing increases according to the number of elements involved, and all that a rating scale seems to offer is a relativistic evaluation of T against $T_2 \ldots T_n$ in terms of a meaning-criterion which varies according to the reference person. The meaning of each rating is reduced to a glimmer of light shining through the mist of uncertainty. And such would be the stern critic's conclusion regarding grid validity.

12.5 REORIENTATION: FROM FORMISM TOWARDS CONTEXTUALISM

What, then, can be salvaged? Does the phenomenological orientation towards 'the things as they appear' offer a way round the repertory grid impasse without giving away too much of Kelly's theory? The remainder of this study attempts to sketch a justification for some movement in the general direction of existential phenomenology (captured by Roche's, 1973, phrase about people being 'in a state of society'[12]) in order to offer a new orientation to research conceived within a personal construct psychology framework.

I do not wish to pre-empt what is presented in the following chapter, save to reiterate the claim for the primacy of events and their uniqueness. The approach which I seek to map out relies upon the researcher's observation and construing of events, and/or upon the respondent's (re)constructions of the events that he or she has experienced. In each case, the heart of the methodology is the schematic construing of events (as presented a few paragraphs ago), stripped of the obfuscation introduced by the misplaced precision of rating (or other allocative procedure). The Kellian insistence on similarity and contrast still remains, overtly in the case of direct comparisons and more obliquely when the respondent is construing in terms of schemata sedimented

through experience. The grid, however, fades into insignificance.

An approach based upon events has the potential to allow the detail of events to be more closely observed and more richly interpreted than seems possible with the repertory grid. There is a price to be paid, but it is a price that many researchers might be prepared to pay - that of a much greater mass of information and the consequent need to find a way of reducing it to manageable proportions without losing its richness.

Pepper's (1942) world hypotheses of 'contextualism' and 'formism' capture something of the distinction I have been trying to make. These should not be regarded as mutually exclusive propositions[13], but as representative of contrasting emphases regarding events. Contextualism takes as its point of origin the historical event, with all that implies about the complexity of elements and their interrelationships: it is also dynamic, in that it presupposes change and novelty. Formism, on the other hand, subsumes within its root metaphor the notion of similarity, and Pepper's explication foreshadows Kelly's interest in similarity and contrast and - more indirectly - the notion of dichotomous construing.

The contrast between contextualism and formism can be expressed in terms of such polarities as holism versus reductionism, and synthesis versus analysis. Pepper sees them as similar in so far as the facts dealt with underneath their assumptive structure are 'taken one by one from whatever source they come and are interpreted as they come and so are left'[14]. From this standpoint the universe 'is not in the end highly systematic'[15]. The implication is of imprecision, in contrast to the root metaphors of mechanism and organicism which are more limited in

their scope[16].

Sarbin (1977) refers to Kellian theory as contextualist. At the general level this seems a reasonable characterisation, particularly in the light of Pepper's linking of this root metaphor to the pragmatist tradition of American philosophy, but it should be pointed out that traces of the other root metaphors are detectable in Kelly's writing. On the other hand the repertory grid, with its strong emphases both on analysis and on similarity and contrast, seems best located under the heading of formism. Granted this distinction, there exists a philosophical tension between theory and methodology which is only partly dissolved by grafting contextualist methods (such as 'freer' forms of interview, or observation) on to the repertory grid rootstock.

The changes I wish to make to Personal Construct Theory do not diminish what I accept as a contextualist stance - if anything, they strengthen it. In this redefined context it can be expected that the repertory grid will have little, if any, part to play as a research instrument - which is not to deny its utility as a tool for clinical exploration, that being an issue with which I have not concerned myself here.

But even within its current theoretical framework the validity of the repertory grid has come under severe questioning, and it is open to doubt whether it can stand up to the pressure to which it has been subjected. The arguments developed and sustained through the preceding chapters lead towards the general conclusion that, in terms of the collection and interpretation of constructions of the world, the repertory grid is more a coarse riddle than a fine-meshed sieve. Fransella and Bannister (1977) come to much the same evaluation when they write

> 'To the extent that a grid gives us a map of an individual's construct system, it is probably about as accurate and informative as the maps which Columbus provided of the American coastline'. 17

They continue, however,

> 'At that, it may be a good deal more sensitive to the nature of the person than the kinds of psychological instrument we have tended to use to date'. 18

The evidence of this study indicates clearly that grid users have no grounds for complacency. Cartography has advanced to a high degree of sophistication and accuracy since the fifteenth century. Can research within the framework of Personal Construct Theory make similar progress? In the following exploratory chapter I offer a programmatic sketch, a propaedeutic, of the direction which I believe theory and method should now take.