

Expert/Novices differences in Case Analysis

G. Easton
Department of Marketing, Lancaster University

T.C. Ormerod
Department of Psychology, Lancaster University

Final report to the European Case Clearing House

February 2001

Summary

We report a study of expert-novice differences in case analysis, a management and marketing training technique widely established in business schools and industry. While there is evidence for the technique's effectiveness, little is known about the cognitive processes that underlie case analysis, a lacuna this study set out to remedy. A secondary motivation was to explore expertise in tackling problems that are essentially undefined, and where problem discovery and description is a key activity.

The study consisted of the collection and analysis of verbal and written protocols of expert and novice marketers working individually in analysing up to four short marketing cases (approximately 1 hour per case). The experts were 11 members of academic Marketing departments, each of whom had extensive experience in teaching with the case method. The novices were 22 third year Marketing students, whose performance was tested prior to and after a one semester module on case analysis. The resulting protocols (over 100 hours of audio and videotaped transcripts) were subjected to two kinds of analysis. First, protocols were transcribed, chunked and coded using a scheme reported in Ormerod et al (1999), in which each transcript segment was coded according to cognitive act (e.g., evaluate, monitor, generate, paraphrase) and referent (e.g., case statement, external knowledge sources, prior knowledge, concerns and recommendations). Second, qualitative analyses of case outcomes were undertaken. From these analyses, a number of expert/novice and before/after training comparisons were made.

While a large proportion of the data await quantitative analysis, a number of initial results can be highlighted from the study:

- Although the time spent on analyses did not differ reliably between experts and novices, large qualitative differences were found between the groups. Experts generated more alternative recommendations, identified more critical issues and used more evaluative criteria than novices. The outcomes of their analyses were generally qualitatively better than those of novices, and were more likely to bring in issues not specifically referred to in the case statement. Novices also tended to reach a firm viewpoint or recommendation early (often during the first reading of the case statement), while some experts deferred reaching a recommendation until later in the analysis, were more likely to change their stance during the analysis, and in some cases did not reach a specific recommendation at all. Novice analyses focussed more upon outcome while expert analyses were more likely to focus upon process issues.
- Novice analyses tended to be disappointingly shallow, and constrained by the content and order of the case statement. Perhaps the most important general finding is that the final output of the analysis often seems to be a poor summary of the richness of the process that has gone before. This contrasted with expert analyses, which became more focussed yet did not lose the richer issues generated early in the analysis.
- Comparisons between novice analyses before and after training revealed evidence, both of an improvement in the quality of analyses, and changes in the processes of analysis. In particular, the depth of analyses increased, and solution development tended to be deferred until later in the analysis.
- A number of differences between case analysis and other creative problem-solving domains were identified. While phases of problem understanding, solution development, evaluation and review are common across domains, there was relatively little evidence in the expert protocols of the control activities (specifically scheduling

and monitoring) that dominate protocols in domains such as design problem-solving. We putatively suggest ‘snowballing’ as a component of expert case analysis skills, that is, the online assembly and collation of key issues and recommendations that carry through an analysis in the absence of explicit control activities. There was also surprising little overt use of analogy or remembered exemplars in either expert or novice protocols. This contrasts strongly with evidence from studies (both within the current study and in an associated project) of group processes in case analysis. Another surprising outcome is that experts tended to show more variability than novices. This contrasts with the majority of expertise domains, where increasing expertise typically leads to a convergence of process and outcome across individuals. Finally, the use of external documentation by experts was subject to great variability – some experts claimed deliberately to avoid early note-taking, as ‘reification of bad ideas’. There is some evidence that early documentation by some novices inhibited the quality of their subsequent analyses, since they tended to collate the documentation in place of a properly structured analysis output.

- The kinds of outcomes generated by both expert and novice groups are, at least in part, a function of the nature of the case itself. Where case statements represented the analysis as a choice between alternatives, this highly constrained the nature of the analysis, while cases that were open-ended typically elicited richer analyses, especially with the expert and post-training novice protocols. In the pre-training novice protocols, however, there is evidence that open-ended cases gave rise to large individual differences: while some participants produced richer analyses than with more constrained cases, others produced highly superficial and truncated analyses.

Analysing cases is a clearly a complex process with complex outcomes. As a result it is crucial to understand that while there may be recurrent cognitive and social processes common to all case analyses, these are modified by particular contingent factors. It is therefore important to realise the impact of the analyst, the setting and the case in trying to offer advice to case analysts and instructors. It would appear that experience and training help but the most effective ways are not entirely obvious, especially since initial learning effectiveness seems to be a function of individual differences in response to case type. Despite the complexity of, and variability in, the results, a number of provisional recommendations and suggestions for further work can be drawn from the study. First, it would be helpful to find ways to capture the richness of early analyses, to prevent the losses of issues, recommendations and evaluative criteria that characterise case outcomes in novice protocols. We propose to explore mechanisms (both methodological and technological) that offer a form of ‘inverse scaffolding’ across case sessions. Second, despite the apparent gains made by novices after training, the expert protocols were much richer and more compelling. It may be valuable to develop case method training materials that provide, not only a case statement, but also transcripts of expert analyses as target exemplars of best practice. The effectiveness of this kind of ‘vicarious learning’ from analysis episodes, rather than simply from accompanying teaching notes, presents an interesting topic for further research. Third, it is clear that different things happen when students analyse cases as individuals and in groups. While it is fair to say that the majority of case method use focuses upon group activities, it is important to recognise both the differences between individual and group settings, and also the contribution that individual activities make to group processes. To do this requires a set of data elicited from groups that is directly comparative with the data reported here from individuals (notably before versus after training data from group analyses of new cases). Fourth, we

focussed in this study on short case statements, as a way of establishing a baseline of performance in as simple a context as possible. Notwithstanding this deliberate simplicity, we have uncovered a richness of issues that might not have been anticipated from such simple cases (notably the between-case effects). However, studies of simple cases preclude the observation of many of the skills associated with more realistic cases, such as data handling, simulation, and possibly the kinds of control activity that are not apparent (and even not necessary) in small scale analyses.

Introduction

"Every curriculum design is a hypothesis about learning". Stenhouse (1979)

Implicit in the choice of any learning experience, in whatever educational context, is the presupposition that it will promote learning of some kind. But what kind of learning and whether this is the best way to acquire it are questions educators would dearly love to know the answers to. The case method presents a particularly difficult problem in this regard since it is, itself, a complex learning experience compared, for example, to a lecture. In addition, it can employed a variety of different ways in an infinity of contexts. The key question which we addressed in the current study was as follows: What, in its simplest form (single session analyses of short case statements by individuals) comprise the stages, processes and outcomes of case analysis? In addressing this question we aim to contribute to the design of optimised case learning experiences so that students acquire problem solving skills most effectively.

We start with the belief that the case method helps students to acquire problem-solving skills (this contrasts with an alternative view that the case method simply delivers pre-packaged experience of real management and marketing issues to inexperienced students). What makes case analysis unusual as a problem-solving domain is that there is often no statement of a 'problem' or goal (and where there is, it is often incomplete, incorrect or peripheral). Instead, the analyst is presented with information pertaining to company, individual, product or current issue, which may include historical, statistical, personal, corporate, and other contextual data. The analyst may specify recommendations for action or simply provide a structured commentary. This sets up a second key question that we addressed in the study: How does case analysis differ from problem-solving in other domains? In addressing this question we aim to situate an understanding of the case method within a broader learning and training context, to allow its future evaluation against related training approaches in different domains and against different training approaches (e.g. simulations) within the same domain.

An obvious approach to answering these research questions is to measure respondents' initial case analysis performance, put them through a series of case method learning experiences and measure their performance again. However there are research strategy issues which result from (a) the complexity of the behaviours involved in case analysis and (b) the influence of contextual and content-based variables (e.g. type of student / instructor, type of material, media etc) on the outcomes. Both pose problems in terms of generalising research results.

There are, broadly, two research strategies that could be adopted. The first is the broad, shallow, output approach (BSO); this recognises the importance of context and seeks to discover whether, for example, learning varies if the analysts are experienced or tyro, the length of the case study differs or the subject is medicine or management. As a result a large number of data points are required to cover different contexts and so the measurement of learning behaviour is confined to the output ignoring the process. So we learn how various factors affect case study performance but the measures of such performance have to be simple, robust and as a result are invariably somewhat superficial. But in the case method there is a problem concerning the judgement of the quality of the output. A distinction is often made between cases and exercises. While the general form of the two types of learning experience may be quite similar they are in practice quite different. An exercise has a single answer which can relatively easily assessed; a case will

usually have a number of more or less defensible answers. Indeed one of the objectives of the case method is to persuade students that, as in real life, there are usually alternative solutions. As a result, judging the content quality of case solutions becomes somewhat problematic. Clearly it is done in educational institutions throughout the world as part of student assessment. But the reliability of such measures is generally regarded as being suspect. As will be discussed later, our approach to this problem is to measure performance indirectly by comparing the behaviour of novices to experts, a technique widely used in cognitive psychology.

The pros and cons of the BSO approach are clearly illustrated in the case of Problem Based Learning research. The case method can be considered as one of a number of different learning experiences which may be used to promote what has come to be called Problem Based Learning (Boud, 1981, 1998; Schmidt et al, 1995; Bouhuijs et al, 1993). PBL is defined as adopting an "...emphasis on active learning using a problem as the stimulus... as a starting point" (Bouhuijs et al, 1993) and originated in the applied (social) sciences including law, medicine social work, architecture, and education. A large amount of research has been carried out under the banner of PBL but, not surprisingly, most is of the BSO kind. The emphasis has been on contextual issues such as curriculum design, planned implantation of PBL, problems of assessment and the role of tutors, and a great deal has been learned about the role of such variables. Such research has been possible since the performance criteria (output) have largely been confined to knowledge acquisition and application rather than the development of higher level cognitive skills. This focus has made it possible to check for learning relatively easily, often by means of objective testing. The results suggest that PBL produces "better" learning (output) than more conventional learning experiences, but very little psychological research has been carried out to discover why this occurs (for an exception see Brown, Collins and Duguid, 1989).

The second approach is the one we propose; observations of deep, narrow processes (DNP). Depth comes from examining, in detail, what it is that subjects are doing (process) when they analyse case studies. However, in order to do so one sacrifices breadth since each performance measure necessarily involves a great deal of analysis. As a result only a narrow range of contextual variables can be investigated. But a major advantage that the DNP approach has over BSO is that it can help researchers design better learning experiences, since it is concerned to understand the processes of problem solving and not just the outcomes. For example, poor analysis may result from students deciding on a solution too quickly. It is not difficult to see how a case learning experience could be designed to correct this behaviour.

While neither approach is superior per se, we would argue that the DNP approach has the additional advantage that it can be based upon research-validated theories of the cognitive processes of individuals. From this research, which is reviewed below, there is evidence of consistency of behaviour among participants who are carrying out problem solving tasks. We, therefore, sought to discover whether case analysis, as a particular form of problem solving, also exhibits similar uniformity. To the extent that it does, we would claim to be able to generalise or contrast across domains.

A key conceptual assumption of this research design is that case analysis represents a particular form of problem solving. Even in its most primitive usage, case analysts are required to "analyse the case and come up with solutions". In other words, the case contains problems and the analyst is required to discover and define what they are and suggest means by which they can be solved. A problem orientation for case analysis is particularly appropriate for management since it can be argued that managers, and their

organisations, are driven by problems (or opportunities). One definition of a problem is the difference between what is and what should be (Easton, 1992). Such gaps, it could be argued, are precisely what drive managers to action. They are salient and motivating and so it behoves management teachers and trainers to help apprentice managers to gain problem solving skills.

Rational problem solving prescriptions abound and can provide a template against which to judge actual behaviour. Easton (1992) uses one example to structure a book designed to help students analyse cases. It comprises a simple step by step approach. The steps include understanding the situation, diagnosing problem areas, generating alternative solutions, predicting outcomes and evaluating alternatives. However, research in cognitive psychology suggests that such approaches to problem solving are rarely met in actual behaviours especially among novice problem solvers.

Initially, it is helpful to distinguish between performance and learning. On the one hand, one can identify the process of case analysis itself as being a complex problem-solving activity whose performance involves problem comprehension, idea generation and solution development. On the other hand, the use of case analysis in providing exemplars for understanding new problems invokes the issue of learning.

The most established account of problem solving performance is Newell & Simon's (1972) GPS (General Problem Solver). It is worth a brief description here since it is upon this theory that almost every recent cognitive theory of problem solving and learning has built. Part of GPS is a description of the mental representation of a problem, which Newell and Simon term the problem space. The problem space is determined by the nature of the problem itself (e.g., the initial state and the desired goal state, the 'operators' or methods available to assist in reaching that goal, and constraints or rules for undertaking the problem), the task environment (e.g., whether the problem is couched in abstract or familiar thematic terms) and the information-processing characteristics of the solver (notably, the proposal that humans have a limited capacity working memory system which constrains the amount of information that can be consciously accessible at any one time). The other key component to GPS is an account of how the problem space is searched using strategies that reduce the size of the problem space to that likely to produce viable solutions. For example, means-ends analysis is a strategy in which the distance over the problem space between the current and goal states is estimated, and if a problem solving method cannot be applied directly to reduce this difference, the problem is decomposed into sub-goals that can be solved directly. Strategies for searching the problem space are one of the key determinants of expert-novice differences. For example, experts in geometry theorem proving adopt a strategy of working forwards from problem lemmas to required proof, whilst novices adopt a strategy of working backwards from proof to lemmas.

The issue of decomposition is critical in problem-solving domains such as case analysis. A case statement presents a problem that is ill-defined in terms of the precise goal state, constraints, and operators that might be applied in analysing it. It typically involves a complex series of problem-solving steps that are not obvious from the outset and that generate too great a memory load to be solved as a whole. Thus, a major part of the case analyst's task is to restructure the problem by decomposing it into manageable units that can be tackled in isolation whilst contributing to an overall analysis. From studies of other complex and ill-defined problem-solving domains such as design (e.g., Simon, 1981; Ormerod & Ball, 1993; Ball & Ormerod, 1995) we know that the effective use of appropriate decomposition strategies is one of the principal factors underlying expertise. One of the skills of an expert designer is knowing when to switch between prescriptively

optimal breadth-first decomposition (which minimises premature commitment to particular solution types before a problem is fully understood) and cognitively economical depth-first decomposition (which minimises memory loads during problem solving). ‘Control’ strategies such as these (i.e, strategies for controlling the order and depth of ongoing problem-solving activities and foci) enable the expert designer to explore alternative design solutions without overloading working memory, while still making progress towards meeting their goal. One might expect to see similar strategic switches in expert case analyses.

Another key factor underlying problem-solving expertise is the retrieval and use of relevant prior knowledge, which brings us to the issue of learning from cases. In its most general form, case analysis can be seen as a kind of analogical reasoning, in which knowledge of previous exemplars is used in solving a target problem. Analogy plays a fundamental role in most current cognitive theories of learning and performance. For example, in Anderson's (1993) ACT-R theory, the acquisition of procedural skills (knowledge about how to carry out a task) is based upon a process of making structural analogies (i.e. where the structure of the problem is similar though the content might be quite different) in order to access known procedures for solving similar tasks. In the same vein, Gick & Holyoak (1983) propose that learning be based upon the induction, over a series of analogical problem-solving episodes, of general 'schemas' (i.e., knowledge that provides relatively abstract solutions for a wider class of problems).

Another theory, similar in name as well as apparent cognitive foundations to case analysis, is the case-based reasoning account of human cognition proposed by Riesbeck & Schank (1989; see also Schank & Leake, 1989). They argue that problem solving and learning are based upon the retrieval of similar cases held in memory. Cases are organised in terms of knowledge structures called memory organisation packets (MOPS), which characterise a sequence of expected events. MOPs share lower level components called scenes, which describe discrete events. For example, there might be distinct MOPs describing a visit to the doctor and a visit to a dentist, but they share the scene of an office visit to a professional. The importance of MOPs is in providing a theory of how humans develop flexible explanations for new events upon the basis of previous case (or MOP) knowledge. Despite sharing a common foundation in analogy, these theories have important differences. For example, whilst ACT-R proposes an account of expertise which is based upon the refinement of task-specific procedures, schema induction and MOPs are accounts of the development of abstract and generalisable knowledge structures. The issue of generalisability of case knowledge underlies an untested assumption that case-analytic education leads to knowledge that is both abstracted from specific cases and generalisable to novel situations. Also, whilst both ACT-R and schema induction invoke learning mechanisms in response to successful problem-solving episodes, learning in case-based reasoning is invoked only where retrieval of a case fails to provide the expected solution to a problem. In terms of case analysis, the issue might be summarised as follows: Is the value of learning from cases in providing a rule or in highlighting the exceptions? Whether experts who reason from case analyses rely upon identifying the matches between known and new cases or on exploring the mismatches is another area that requires empirical evaluation.

To summarise the project, it entailed the following key activities:

- i. Preparation of case materials, piloting of sessions, development of domain-specific coding schemes;
- ii. Collection, transcription and analysis of data from experienced participants;

- iii. Collection, transcription and analysis of data from individual novice participants in pre-training situation;
- iv. Comparison between experienced and novice data sets;
- v. Collection, transcription and analysis of data from individual novice participants in post-training situation;
- vi. Collection, transcription and analysis of data from group discussions in pre treatment situation;
- vii. Comparison between novice data sets before and after training;
- viii. Initiation of data collection from other expert samples;
- ix. Production of final report.

The current status of the work is that all the empirical data collection that we believe feasible and necessary from studies of individuals has been carried out (though certain changes to the initial proposals were made in response to emerging findings and contingencies – these are detailed in footnotes). The project generated over 100 hours of video and audio-taped data. Given that it takes in the order of three working days to produce a complete analysis of one hour of case analysis session, our detailed examination of the data is limited to a sample (approximately 25%). This was always our intention, though we now have the opportunity, in future analyses, to confirm quantitatively many of the hypotheses and qualitative observations generated in our analyses to date. Thus, the results reported in this report are based mainly upon the 25% data sample.

Research design

We conducted an empirical study of case analysis to address the two key questions discussed above, namely:

1. What, in their simplest form (single session analyses of short case statements by individuals), comprise the stages, processes and outcomes of case analysis?
2. How does case analysis differ from problem-solving in other domains?

We set out to address these questions by targeting four kinds of case analysis context, described below:

Expert-novice comparisons

We compared the performance and outcomes of a group of ‘experts’, comprising experienced academic teachers of the case method, with a group of ‘novices’, comprising students entering their third year of a Marketing degree course at Lancaster University. A second (repeated measures) factor was the type of case: cases were directive or non-directive¹. A directive case statement presents an instruction to the analyst that dictates both their role (e.g., “Imagine you are an external marketing consultant”) and their goal (e.g., “...making recommendations to the board about the planned merger”). A non-directive case statement, in contrast, leaves the assignment of role(s) and goal(s) to the analyst. Four cases were used in the study, two directive and two non-directive. Because of the relatively small sample of experts available and constraints upon the time available, the expert/novice comparisons focus upon just two of the cases, Metropool Zoo (directive) and Swatchmobile (non-directive).

Effects of case method training on novice performance

The second context focuses upon the novice participants, and manipulates the factor of training, examining participants' analysis performance at the beginning and end of an undergraduate course on case analysis. Participants were recruited from a 3rd year marketing strategy course in the Management School. The first two case analysis sessions were carried out at the beginning of course, with a further two sessions toward the end of

¹ ‘Case type’ replaced the factor of ‘case study domain’ suggested in the original proposal, in which we had proposed to give participants cases from within-domain (marketing) and outside-domain (non-marketing). This decision was taken as a result of pilot work suggesting that the degree of direction offered to the analyst by the case statement was likely to be an important factor in determining subsequent performance. Given that many textbook cases for novices are directive, for example in establishing role of point of view of the analyst, we felt it was important to focus upon this issue. However, we suggest that a future study might usefully examine the domain specificity of case analysis skills. In principle, expertise in case analysis might be attributable to any of three sources: domain-specific knowledge (i.e. marketing knowledge), skills specific to case analysis, and general problem-solving skills. We can discriminate between skills specific to case analysis and general problem-solving skills by comparison of performance in this study with findings in the literature. However, in order to differentiate fully between the roles played by domain knowledge and case analysis skills, we need to examine participants' performance in reasoning about cases that lie within the target domain and those that lie outside it. In the case of experienced analysts, we might expect to see some degree of transfer of skills from within-domain to outside-domain cases. In the case of novice analysts, we might expect shift from reliance upon (albeit limited) domain knowledge and general problem-solving skills prior to taking a course on case analysis, to use of skills specific to case analysis.

the course. This course is usually taught by Professor Easton, but was taken by another member of staff during the data collection period, allowing us to run the study independent of any potential conflicts between assessment and research priorities. It is important to note that our participants were 'novices' rather than being entirely naive to the domain², that is, they had some familiarity with the terminology and nature of cases and the marketing domain in which the cases were based. The students had had experience of case studies and case analysis in previous courses, though these were relatively few in number and had been used as just one of a number of different learning experiences. A few of the students also had one year's industrial experience.

A rationale for choosing students on this course as our target novice sample is that undergraduates with little or no work experience are likely to be strongly influenced by a case course and therefore show the largest changes in behaviour. We also chose this particular course because it comprises a group of specialist marketing majors who have comparable educational experience. However, in order to reduce further the variability and increase understanding of the results, each student was interviewed prior to undertaking the study to assess the degree of their relevant life experiences (e.g., part time work).

The marketing strategy course lasted two academic terms. Every fortnight there was a two-hour case presentation session, in which groups presented their case analyses to the whole class, who also had to analyse the case, were questioned and received a mark. In addition, one whole session was given over to presenting a case analysis framework, in effect a summary of the contents of "Learning from Case Studies" by Geoff Easton. Trained novices would therefore have analysed 10 cases, seen them presented and discussed the presentation as well as presented one case in depth themselves. Part of the objectives of the study was to discover whether this training programme had any effect.

In addition to the Training factor, cases were also manipulated systematically across phases, such that equivalent numbers of participants tackled each case before and after training. In addition, of the two cases tackled at each phase, one was directive and one was non-directive.

Group case discussions

Many teaching formats adopted for delivering the case method involve group interaction. In order to provide some preliminary data concerning group processes in case analysis, we carried out a very limited study that examined the nature of social interaction in case learning situations. Two groups of five participants were selected at random from the individual student participants³. Each group met, shortly after having analysed the last of the cases as individuals, for an hour long videotaped case session. In this session, the aim was for the group to reach a group consensus regarding the case. The case that they were given was one that they had already analysed as individuals. The sessions were videotaped using multiple cameras.

The case session was initiated by the case tutor (Prof. Easton). He then left students for the majority of the session to work in an undirected fashion, before returning to the group for a review, where he operated in a non-directive mode. In this way, we aimed to come as close to the traditional case discussion situation as possible.

² A risk of employing naive participants in expert/novice comparisons is that the large group differences that emerge can simply be explained as gross effects of domain comprehension, which are of little psychological or pedagogical interest.

³ We had hoped to undertake the same kind of group session with experts. However, it did not prove to be possible to arrange such a session due to teaching and other commitments of the participants.

Other "expert" categories⁴

It became obvious during the course of the research project that the notion of an expert was less straightforward than had first been believed. In order further to study the nature of expertise, it was decided to carry out one case session with non-marketeers. A further participant was recruited from the staff of the Psychology Department at Lancaster, a senior lecturer in Social Psychology, to see how maturity and general intellectual expertise affects the performance of case analysis. At the same time, an opportunity arose through the PhD work of Lloyd Davies of the Management School, Lancaster University, to begin the process of collecting comparable data from practising senior managers in industry. This work is ongoing.

Method

Participants

The novice group comprised 22 3rd year student volunteers from Lancaster University who were paid to participate in the project (approximately £10 per session plus a £10 bonus on completion of all four sessions). Of the initial 22 students who completed two case analysis sessions prior to training, 18 went on to complete two further cases after the training courses.

The expert group comprised 11 lecturers, senior lecturers and professors who currently teach marketing in three of the leading business schools in the UK: 4 from Lancaster, 3 from Warwick and 4 from Strathclyde. They took part on an unpaid volunteer basis. Experts were defined as people with a minimum of 10 years experience gained in both academic and commercial settings, who are currently involved in teaching Marketing at University level.

Cases

The form of case study used in the tests and re-tests represents another key design variable. A short case study format was used for three main reasons. First, it represents the kinds of cases found in introductory level textbooks on the case method. Second, it enabled a snapshot of the whole process of case analysis (with the exception of presentation), moving the participants beyond reading and comprehension and on to the complex problem solving behaviours within a single session. Third, it ameliorated potential problems of availability, motivation and fatigue among participants.

Each case was subjected to a three-part a priori description, first by subject, second, by structure, and third by degree of constraint.

Subject

Four cases were used. These are shown in Appendix 1.

- i. Metropol Zoo (a case presented as a fictional setting, though almost certainly London Zoo) involved a decision about whether to replace some animals with virtual reality systems in order to increase visitors. This case was directive, in that it contained an explicit question "If you were acting

⁴ This context was not part of the original proposal, but allowed us to take advantage of opportunities that arose during the project period.

- as an external business consultant brought in to help the board of the zoo come to a decision, what would you recommend they do and why?”
- ii. The Swatchmobile case (1277 words) implicitly asked the analyst to assess the viability of a possible joint venture between Swatch and Mercedes to produce a new fashionable microcar. It was non-directive, though it did list three issues at the end of the case statement that might have focussed the analysis for some participants.
 - iii. The Aldi case (759 words) described the retail food market in the UK and Aldi's increasing share. The issues were not set out but might have involved an assessment of the likely success of their existing strategy. This case was also directive, in that it contained the explicit question “As a consultant to Aldi, what do you think their marketing strategy should be over the next few years?”
 - iv. In the final case, the increasingly competitive situation facing BUPA (762 words) in the complex private health care market is portrayed. It was a non-directive case.

Each of these cases varied in a number of dimensions. The topic of any case (e.g. marketing, HRM) is mostly determined by the context in which it is provided and the views of the case instructor. The topic determines whether the participants would regard themselves as an expert or someone who is interested or indeed uninterested. This in turn is likely to affect both their competence and motivation. The context helps determine the amount of knowledge that a participants brings to the analysis and the number and type of assumptions that they have to make. For example, almost everyone has been to a Zoo but fewer people know about the complex structure of the private medical insurance industry (BUPA). As it turned out, while most of the experts had varying degrees of knowledge about the Swatch case, none of the novices had any relevant background knowledge, and few participants had anything but cursory knowledge about Aldi.

Structure

The cases were subjected to a hierarchical structure analysis, in which the key content points were organised according to natural groupings. This allowed us to determine the order in which participants pursued issues raised in the case statement. It also enabled us to judge both the scope of the case a priori and the relative length of time participants spent on each issue. The structure hierarchies for each case are shown in Appendix 2. The way in which these structures were developed was as follows. Semantically related sentences from each case statement were arranged together, and then common semantic elements were extracted that enabled a minimal hierarchy (in terms of breadth and depth) to be obtained without compromising semantic relatedness. Once the hierarchy was obtained, the individual sentences were deleted. For reasons of brevity, the hierarchies are shown decomposed to six levels, though for most cases there was also a seventh level of decomposition in some areas of the hierarchy. The resulting structures do not necessarily reflect text order since case statements often raised semantically related points in non-linear order within the text.

The structures shown in Appendix 2 immediately reveal differences between the cases that might impact upon participants' performance. For example, even when the length in words is approximately the same, the number of distinct semantic propositions identified can differ radically. This is illustrated by the Aldi and BUPA case statements, which are of

similar length, but there are 28 semantic propositions in the Aldi case, compared with 54 in the BUPA case. This, of course ignores the fact that the Aldi case contains a table of figures which was not broken down into semantically independent units, since the semantic content of the table is inferred by the analyst. As it turned out with this particular case, no participant spent longer than 1 minute looking at the table, and did not derive any semantic units from it that are not already captured in the hierarchy. Thus, some cases were informationally impoverished relative to others. This, however, does not take into account the nature of the semantic content. For example, it might be argued that large parts of the Swatch case statement (e.g., the search for a partner) were tangential to the core of the case. However, this is of course a subjective and even loaded judgement, since it might be precisely this tangential information that determines the success or otherwise of the case as a pedagogic tool.

Degree of constraint

Each of the cases was subjected to an examination as to the degree to which the case statement is likely to constrain the processes undertaken by participants. This focussed upon issues of problem identification, solution generation and evaluation.

In terms of identifying the problem, some cases like BUPA and Aldi do not have obvious problems to solve or issues to resolve, and one might expect that it would be more difficult for novices to analyse. Metropol Zoo, in particular, had at its centre two apparently clear-cut alternatives for the case analyst to resolve. Swatch also appeared to be more straightforward, involving a go - no go decision and included some rhetorical questions that novices used to structure their answers. In terms of creating solutions, some cases like Zoo seemed to disallow creative thinking since the options are seemingly laid out for the analyst. All the other cases seem to leave reasonable scope for creativity. In terms of evaluating solutions, Metropol Zoo was the easiest case to evaluate since it gave a lot of internal detail about not only the options but also the values of the various protagonists. In the other cases, the absence of corporate values (or in the Swatch case, apparent contradictions about the values of Hayek – as designer, entrepreneur, marketer and financier) and lack of detail meant that option evaluation was much more difficult and many more assumptions had to be made.

Procedure

The majority of case analysis sessions took place in a laboratory in the Psychology Department at Lancaster University. Sessions with experts at Warwick took place in the participant's office, while sessions with experts at Strathclyde took place in an empty meeting room. Participants were seated at a table, and were provided with pens of different colours, paper (lined and plain) and the case statement, copied from a textbook and presented on 1 or 2 sheets of A4 paper. The experimenter was present at each session (at Lancaster, in an adjoining video control room), but did not directly engage in dialogue with the participant. This was done because it is recognised that conversational styles of verbalisation often lead to different verbal accounts (notably more retrospective and hypothetical verbalisations) than concurrent verbal articulation alone. Participants were given a minimal set of verbal instructions to orientate them to the task. These included the following points:

- i. They would find a short case statement on the A4 sheet, which they should read aloud when they referred to it (while all the novices complied with this requirement, the majority of experts insisted on reading the case silently at the start of the session).
- ii. They should spend as long on the case as they deemed appropriate. In essence, they should do the analysis "as they normally would if they were asked to conduct an analysis for some other purpose". Thus, the sessions were self-terminating. [Some experts commented that they 'normally would not do a case analysis at all', though all subsequently commented at the end of the session that the endeavour they had undertaken was a fair reflection of how they thought a case analysis should be done. Whether it is possible to select cases for teaching without undertaking some degree of analysis is open to dispute, and indeed most experts agreed post-hoc that the extent to which they had analysed the current cases was approximately equivalent to the analysis they would give the cases in choosing them as teaching materials].
- iii. They were told that they could make any written notes they wished, but they were not compelled to do so.⁵
- iv. They were allowed to summon the experimenter to ask questions at any stage of the analysis, though participant rarely did so.
- v. Critically, participants were not told what stages or processes to undertake, or what the final required outcome was.

During each individual session, participants were asked to verbalise concurrently whilst conducting the analysis. The concurrent verbalisation approach was adopted for a number

⁵ This represents a change from the procedure specified in the proposal. There are both pros and cons for the idea of having the equivalent of a written output. One argument against it is that it makes further demands on the participants. The advantages are that subjects have to commit to paper and it also captures possible differences between the way that subjects oral and written thoughts are expressed. Also, one of the most common expert / novice differences found in other domains is in the ways that they vary problem solving strategy between phases of problem solution and implementation. However, we felt that it was better to leave participants free to choose whether to make notes, and specifically whether to produce a written 'output', since pilot studies suggested that some experts did not wish to commit to a written output after a relatively short analysis.

of reasons (see Ericsson & Simon, 1992, for a review).⁶ Key among these were the richness of resulting data, and the ecological validity of observing ongoing analysis performance in a relatively realistic context.⁷ Concurrent rather than post-hoc verbalisation was used because of the need to avoid rationalisations interfering with the stream of ongoing cognitive activities and foci of interest at any particular stage. To facilitate the verbalisation process with novices, a familiarisation task consisting of an anagram solution task was given at the beginning of each session, to offer practice to those unfamiliar to the process of concurrent verbalisation. At the end of each session, post-hoc reflections on the analysis were obtained in a taped post mortem. All expert and novice sessions conducted at Lancaster University were videotaped and audiotaped. Sessions at Warwick and Strathclyde were audiotaped only, though participants' notes were retained and the experimenter made written notes of any items that needed recording in visual format.

⁶ The use of verbal protocols has been the subject of a number of critical reviews (e.g., Nisbett & Wilson, 1977), concerning potential problems of incompleteness, inaccuracy, subjective interpretation and interference. In our view, the benefits of the approach outweigh the problems. However, of these criticisms, we believe the possibility of interference is most directly relevant and potentially damaging to our research (though we took steps to minimise the impact of the other potential inadequacies). In particular, it has been demonstrated that verbalising concurrently can actively impair some forms of insightful or creative problem-solving (Schooler, Ohlsson & Brooks, 1993). Thus, unlike most studies that employ concurrent verbalisation, we did not insist upon participants maintaining a constant stream of verbalisations, but allowed them to fall silent when they wished. We are currently conducting a comparative analysis of verbalisation rates (in general, experts verbalised more than novices, though only by 25% or so), but believe that tolerating pauses in verbalisation did not greatly increase incompleteness while it reduced the risk of interference.

⁷ Purists might argue that a single-participant and laboratory-based setting is far from natural. We would agree that the individual nature of the sessions, combined with their snapshot nature (e.g., no periods of down-time or seeking out of additional resources were allowed) made the sessions somewhat artificial, though they offer a valuable baseline against which to examine the role of situational constraints and opportunities in future studies.

Analysis

A sample of participants' protocols were analysed according to two main approaches. The first set out to investigate the processes underlying cases, specifically coding protocols in detail according to psychological activity and textual or other referent at each point in the transcript. The second set out to examine the outcomes of the analyses qualitatively, focussing upon the type and quality of proposed solutions. To date, 12 expert and 12 novice protocols have been analysed in detail. The remainder have been subjected to broader qualitative analyses. Two of these protocols, one novice before training and one Lancaster expert, both tackling the Metropol Zoo case, are shown in the Appendices as examples of each aspect of the process analysis.

Process

Subsequent coding and interpretation of session transcripts and comparison was undertaken to reveal the strategies employed by participants, as well as revealing impasses or change episodes in their behaviour.

Audiotapes of each session were transcribed by a research assistant⁸. Samples of uncoded transcripts are shown in Appendix 3. Once transcribed, the resulting protocols were segmented according to clear changes in topic or pauses of more than three seconds, each segment typically comprising approximately 5-30 words. The grain size of each segment is nicely illustrated by three segments distributed across Tables 2 and 3 that relate to the transcript of Participant NS12 (Novice, before training, conducting an analysis of the Swatchmobile case). Over segments 14-16 of the transcript, the participant reads an item, paraphrases it, and schedules an activity to carry out later in the session.

The segmented protocols were then loaded into a database written by Dr. Ormerod using FileMaker Pro software that aided in the process of coding and subsequently interpreting the protocols (e.g., through automated counts of items across time referred to by participants). Figure 1 shows a screen capture that illustrates the database system. A particular feature of the software was that it facilitated the counting of cognitive acts and referents across time quartiles (i.e., equal quarterly units of time over the protocol). This an assessment of how cognitive acts and referents change over time.

Each segment was then coded using a scheme developed by Ball, Evans, Dennis & Ormerod (1997) for identifying two key components in each protocol item: the cognitive act being undertaken, and the referent (that is the focus of each segment). This scheme was chosen because it has been applied to other domains such as engineering design (Ball et al, 1997) and instructional task development (e.g., Ormerod & Ridgway, 1999). This allows direct comparison of case analysis with other domains of expertise.⁹

⁸ Transcription by a researcher from outside the domain was not without its problems. For example, in one transcript, a number of references to 'mackerel organisational theory' required editing!

⁹ We had intended to draw up problem behaviour graphs (e.g., Byrne, 1977) and count activity transitions (e.g., Ormerod & Ball, 1993). However, neither analysis was deemed relevant in the light of the very low occurrence of control items among the cognitive acts.

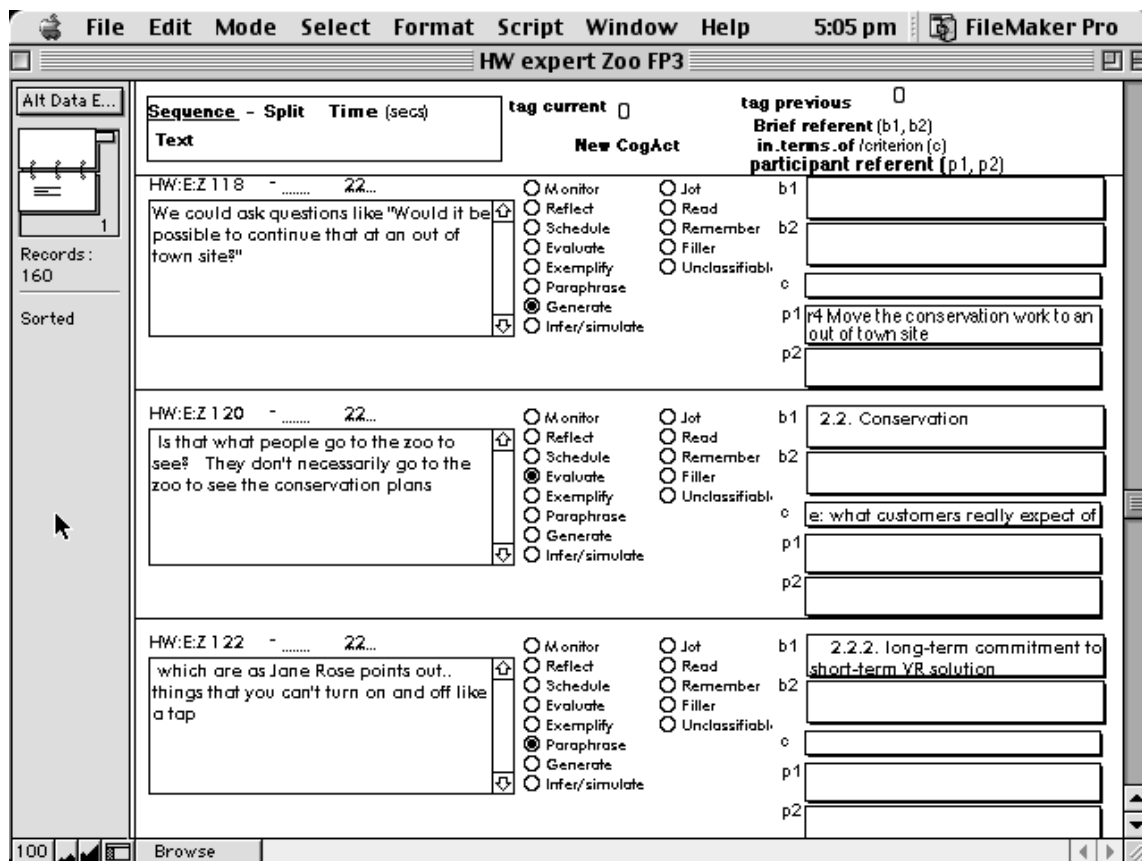


Figure 1. A screen capture from the Filemaker Pro™ coding software

Tables 1-4 show the cognitive act classification scheme, along with explanatory notes and examples. The scheme can be split into four main components. Table 1 shows cognitive acts that are associated with evaluative activities that occur once a referent has been selected, read, paraphrased or generated. Table 2 shows cognitive acts that are related to understanding the case statement. Paraphrases are of two types; first, ones that involve a restatement in the participant's own words, and second, ones that invoke the use of jargon or some kind of phrase that seems to go beyond the issue being comprehended. Of the two examples of paraphrases given in Table 2, the NS12 example paraphrases within the content of the case referent, while the LA example, goes beyond the text he is summarising to introduce the notion of visitors as customers. Table 3 shows cognitive acts that are associated with maintaining control of problem-solving activity (choosing what to tackle next, reviewing what has happened, discussing how the participant 'normally' undertakes such activities, etc.). Table 4 shows cognitive acts that are associated with generating referents that are not given directly or unambiguously in the case statement. Acts of remembrance are distinguished from acts that do not have a temporally related source.

Table 1. Cognitive acts involving evaluation

Cognitive Act	Explanation	Example
Evaluate(E)	Assessing merits/de-merits of a referent, drawing out qualities or issues associated with referent (with or without positive or negative weighting)	NS14: Novice after: Aldi, lines 119-120: “Promoting price as a benefit as well as.. I mean that doesn’t really go towards solving the scale issues which in terms of actually getting sites is out of the aspect of the marketing strategy.”
Exemplify (E)	Illustrating in detail the operation/nature of a referent {Note that the example’s first line consists of a generative act, which is then exemplified}	SA:Expert, Lancs: :Zoo, lines 103- 104 “..but one would want interactive possibilities to support the live exhibits .. so that if you've got some giraffes in the park you'd also have some kind of game or something that involved giraffes in the jungle.”
Infer/simulate (E)	Working through consequences / operation of a referent, beyond the information/ideas given in the statement	SB: Expert, non-Lancs: Swatch: Lines 184-185: “so its cost SF 15,000-21,000.. seem to remember he thought this was going to be low cost and I might be desperately trying to do a conversion in my head about whether that is (writes ‘low cost’) and not assuming that its low cost and I would put a (writes ‘question mark’) and I would (writes ‘check the exchange rate’) figure out exactly how much that is.”

Two other cognitive act categories were available, though their eventual use accounted for less than 1% of segments across all coded transcripts. The first concerned ‘Filler’ acts, which contained items not relevant to the ongoing analysis (e.g., requests for water, toilet breaks, comments pertaining to other activities that the participant was undertaking that day). The second concerned ‘Unclassifiable’ acts, that is, items whose classification was ambiguous or was relevant to case analysis but was not captured by the remaining scheme.

Table 2. Cognitive acts involving understanding

Read (U)	Reads verbatim from case statement , notes or other written source.	NS12: Novice before:Swatch, line 014: “..In 1992 SMH’s profit grew 64 per cent to SFr 413m. And was set to grow 23 per cent in 1993. But after years of steady growth its share price was levelling out...”
Paraphrase (U)	Summarising or rephrasing an item from the case statement, written notes or verbalisation.	NS12: Novice before:Swatch, line 015: “OK so it was on a (writes ‘plateau..’)” LA:Expert, Lancs:Zoo, line 038: “So that's kind of another view of the world in a sense.. is kind of.. it's taking the zoo as this kind of sanctuary.. this kind of well protected area which is kind of unique in some way and putting it right on the plane.. seen from the point of view of the customer.”

Table 3. Cognitive acts involving control

Schedule (C)	Selection, de-selection, planning for future selection of a task/item.	NS12:Novice before:Swatch, line 016: “maybe we could look later at a bit of a growth.. business product life cycle”
Monitor (C)	Reviewing progress made, outlining tasks/items to be tackled.	NS14: Novice after: Aldi, line 024: “I’m just trying to reformulate these problems I’ve got into general issues that the marketing strategy needs to address”.
Reflect (C)	Consideration of meta-issues concerned with analysis process, usually reflections upon self as analyst, teacher, student, etc.	HW:Expert, Lancs:Swatch, line 012: “[I] usually start looking at a case from aspects of marketing theory and practice.. which we know to be sort of ..good practice”

Table 4: Cognitive acts involving generation

Generate (G)	Developing a new referent (idea, critique, etc)	NS19: Novice after: Bupa, line 78: “The threat could be working with the NHS (writes ‘NHS partnerships – good idea?’)” NS20: Novice after: Bupa, lines 56-58: “Bupa also should (writes ‘target specific audiences but differentiate itself’) .. so its really a multi-stage strategy.. and continue to market also.. if a new service or product is needed they should be pioneers in offering that”
Jot (G)	Externalising (either through diagram, note, or written text) – NB, all segments containing externalisations were dual coded	NS19: Novice after: Bupa, line 002: “Just underline that (underlines ‘Bupa is the leader in the provision of private health care’)”
Remember (G)	Recall event or example from past experience (NB. This does not include recall of theory or techniques, unless tied to a particular episode of use).	SB: Expert, non-Lancs: Swatch: Lines 132-135: “I would note VW because again VW suggests all kinds of different things to me... I immediately jump up to the fact that I supervised a Ph.D. that was for somebody that worked for VW.. I have taught previous cases that are about VW, Skoda, Audi and particularly about eastern Europe and I’m looking at VW and I’m looking at China and thinking emerging markets other links with the types of analysis that I have done on other cases”

Three kinds of referent were identified and coded in the transcripts. The first concerned references to parts of the case statement. These were coded according to the item in the case structure hierarchy (see section 4.2.2) that was the focus of the referent in each segment. Appendix 4 shows summaries of case statement referents, summarised in terms of the referents that were explicitly referred to over four time quartiles in the protocols of the two example participants.

The second kind of referent concerned criterial referents, that is, references to beliefs or assumptions, evaluative criteria or specific rephrasings that add meaning to paraphrased items. Four of the evaluative criteria were common to all participants. These were agree/support, disagree/reject, key issue(s) and personal preference. These were used to code evaluations where participants did not supply rationales beyond general positive or

negative weightings. These were coded according to a summary of the precise words used by the participant. Appendix 5 shows summaries of criterial referents for example participants.

The third kind of referent concerns participant-generated referents, that is, any other reference that the participant made that was not either part of the case statement or a criterial referent. Participant-generated referents could be of a number of different types, depending upon whether they pertained to the case problem, solution, the analyst, and so on. A generic category scheme used to categorise self-referents is shown in Table 5. Appendix 6 shows summaries of participant-generated referents for example participants.

Outcome

Each case analysis was assessed in terms of the final outcome of the process, that is, the recommendations that each participant made with respect to the particular organisation that was the subject of the case study. This form of analysis was not mentioned in the original proposal since it was assumed that the analysis would concentrate on process rather than content. It was added because it was felt that some measure of the relative quality of the analyses, however subjective, should be made since the ultimate goal of any research programme of this kind would be to improve the case analysis. It should therefore be regarded as a pilot study that sought to discover what the problems associated with outcome assessment might be. Geoff Easton rated the recommendations in terms of quantity, detail and overall quality, each on a 10 point scale.

Coverage was conceived of as a measure of the extent to which a particular set of recommendations provided possible solutions to the all problems involved in the case. Detail was meant to measure the extent to which the solutions were articulated. Quality was the overall measure of how well the solution(s) were evaluated as solving the problems identified. The basis for assessment was a complete analysis of each case completed by Geoff Easton. These analyses were marginally changed after reading through the novice and expert case analyses responding to aspects of the case, particularly novel solutions, not included in the original analysis. Different solutions and combinations of solutions were scored and used to mark the S's analyses. In addition, higher marks were given if the recommendations were suitably justified. An attempt was made to move away from the usual university standards, where marks over 8 are never awarded, to one where a performance as good as the marker could achieve would be given a 10 rather than expecting total perfection. In addition, an attempt was made to look at the content of the recommendations. Each set of recommendations was summarised in a few sentences in order to be able to make comparisons.

Table 5. Categories and examples of participant-generated referents.

Referent type	Explanation	Example
Analogies	An analogy to illustrate or generate concept	CS: Expert non-Lancs: Zoo, lines 009-010: “So you’ve got a social kind of political thing.. but you’ve got the fact that other attractions are recruiting visitors to them so.. and no money from the (writes ‘government [funding]’) sounds like a University...”
Concerns/ questions	An issue referred to beyond case that analyst believes significant	DW: Expert, non-Lancs: Swatch, line 73: “so how on earth do you (w innovate) uh consistently in mature industries?”
Recommendations	A recommendation for action/outcome/focus	RW: expert, non-Lancs: Swatch, line 092: “so um one might say that one of the key issues is going to be.. we’re going to have to make a choice at some stage.. and it can’t really be a a Swatch Mercedes or a Mercedes Swatch because those really do seem to be rather different images ..”
Meta- recommendations	A broader kind of recommendation	NS14: Novice after: Aldi, line 95: “In the light of these problems I mean the (writes ‘strategy’) strategy essentially is to come up with solutions for them in terms of I guess the marketing mix.”
Rememberings	Anything explicitly retrieved from past experience other than an analogy	N10:Before:Zoo: line 113 “ I know this is fictional and things but (w Windsor Safari Park) shut down because it just wasn't pulling in the people any more because of all these things and they had diminishing turnover and diminishing margins and they just couldn't afford to keep the animals any more.”
Self references	A reference to self’s practices	LA:Expert, Lancs:Zoo: line 111: “As a consultant it is extremely difficult to actually say in a situation like this.. to say essentially I think x y z..”
Techniques	References to marketing theory or methods	NS10:Before:Zoo: line 064: “Just trying to look for some strengths to stick into the SWOT.”

Results

Process

Processes underlying individual analyses

In all, 24 cases were subjected to coding and quantitative analysis of process information. The mean analysis time was 32 minutes per case (not including start up or subsequent discussion with the experimenter) and the mean number of segments in each transcript was 108. A summary of mean analysis times (in minutes) and mean number of segments per analysis is shown in Table 6. Analyses of variance were conducted on both these data sets, between the factors of Expertise (expert versus novice before versus novice after) and Case (Zoo, Aldi, Bupa and Swatch). However, no significant effects were found in either analysis. It seems likely that any underlying differences on these measures, between either levels of expertise, case or the interaction between these factors, are masked by the large individual differences in analysis times in all three participant groups. However, given the complexity of the design, it may be that we simply have not analysed sufficient numbers of participants yet, and that significant differences might be found once the complete analysis is undertaken.

Table 6. mean analysis times (in minutes) and mean number of segments per analysis for each participant group across the four cases.

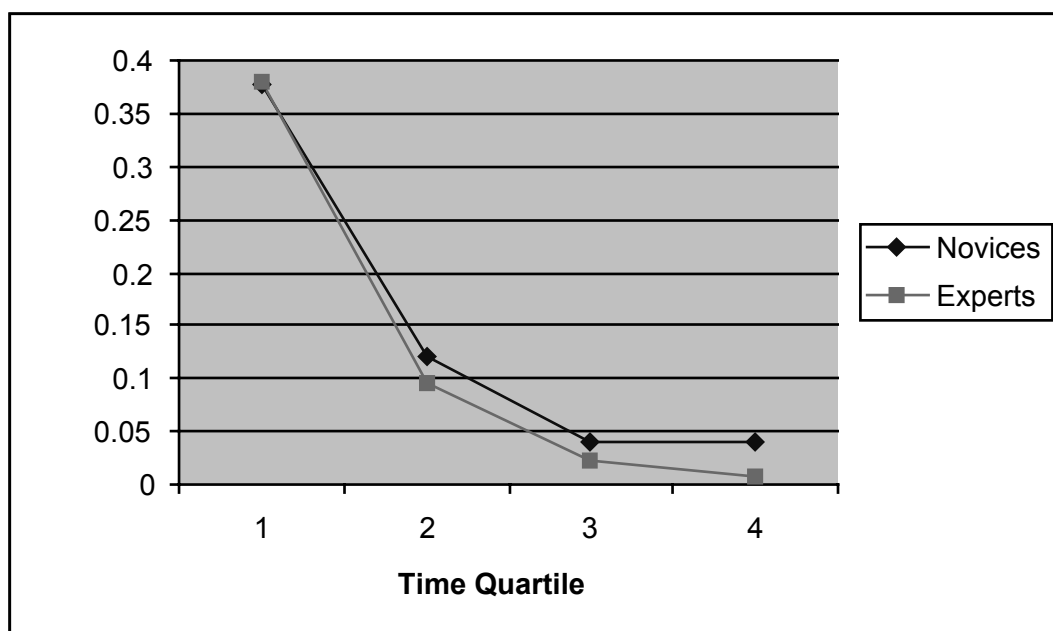
		Zoo		Swatch		Aldi		Bupa	
		Time (mins)	Segs	Time (mins)	Segs	Time (mins)	Segs	Time (mins)	Segs
Experts	Mean	37.8	193.0	40.5	206.8	---	---	---	---
	sd	11.1	30.7	25.9	96.1	---	---	---	---
	Min	30.0	148.0	12.0	123.0	---	---	---	---
	Max	54.0	215.0	90.0	316.0	---	---	---	---
Novices before	Mean	32.3	141.5	23.8	99.3	19.1	78.3	24.4	122.0
	sd	15.5	74.1	12.6	48.2	8.0	34.6	9.9	67.9
	Min	16.0	46.0	15.0	49.0	12.2	55.0	9.8	61
	Max	52.0	227.0	42.0	159.0	35.0	299.0	41.0	331.0
Novices after	Mean	26.3	132.5	33.1	185.3	23.3	148.5	23.6	172.0
	sd	6.2	55.6	9.9	38.4	9.9	54.1	11.7	78.5
	Min	13.0	56.0	14.0	49.0	12.2	55.0	9.8	61
	Max	55.0	244.0	42.0	299.0	44.0	309.0	48.0	271.0

While there were no overall significant differences, a number of interesting observations emerge from the above data. First, expert analyses were generally longer than those of either novice group. However, there are large individual differences among the experts. For example, Appendix 3 contains two radically different analyses, one of approximately 12 minutes (CS, non-Lancs, Metropol Zoo) and one that was terminated by the

experimenter after 90 minutes. The difference between these two protocols is not a function of location (e.g., one non-Lancs participant to over 50 minutes for both analyses) but is more likely one of personal style. Second, the effect of training does not make novice analyses necessarily longer – indeed, on two of the cases (Metropol Zoo and Bupa) novice analyses after training were on average shorter than those before training. Third, there is great variability among cases, the longer cases (Metropol Zoo and Swatch), not surprisingly taking longer on average to analyse than the shorter cases. However, even this pattern is not reliable, with novices before training analysing the Swatch case quicker than the Bupa and Metropol Zoo cases.

One aspect of behaviour that was also dominated by individual differences concerns the use of externalisations (jotted notes, written output, diagrams, case statement highlighting, and so on). Figure 2 shows the proportion of transcript segments that were accompanied by some form of externalisation. At first sight, the graph shows a remarkable equivalence between experts and novices (on the before training group are shown) in the amount of externalisation that they engaged in. However, this masks the facts, a) that some experts did not externalise at all while two took over 9 pages of notes for each analysis they undertook, and b) that some novices simply underlined parts of the case statement while others wrote out parts of the statement in long hand.

Figure 2. Proportion of transcript segments accompanied by externalisations



Cognitive acts

A count of the percentage of each type of cognitive act as a proportion of the total number of cognitive acts in each time quartile was conducted, averaged across cases, for experts and novices (before training only: an analysis of cognitive acts of novices after training remains to be completed). Table 7 shows the percentages of cognitive acts associated with control of problem-activity. Table 8 shows the percentages of cognitive acts associated with all other cognitive act types that accounted for more than 5% of acts in any quartile. All other cognitive acts (exemplify, infer/simulate, remember, filler) accounted for fewer

than 4% of cognitive acts in any quartile. Also shown in these tables are comparative data from the study of instructional task design expertise, conducted by Ormerod & Ridgway (1999).

Table 7. The percentage of cognitive acts associated with control

Domain	Act	Experts					Novices				
		1 st	2 nd	3 rd	4 th	All	1 st	2 nd	3 rd	4 th	All
Case Analysis	Monitor	2	1	2	4	2	1	8	13	19	10
	Schedule	5	3	3	1	3	7	10	6	2	5
	Reflect	4	2	6	14	7	5	6	4	4	4
Task design	Monitor	7	11	8	6	9	3	6	8	14	9
	Schedule	8	12	13	15	10	12	10	8	2	10
	Reflect	8	6	7	8	7	2	4	3	6	4

An important finding to emerge from these data concerns the relative absence of control activities among the experts conducting case analyses. Only 2% of their transcript segments contained acts of monitoring, compared with 9% by experts in the domain of instructional task design. Similarly, only 3% contained scheduling acts, compared with 10% in task design. The figure of 7% of reflective acts is confounded by the data from one expert from Lancaster, whose transcript contained over 30% of segments that were coded as reflections upon self (i.e. “how I usually do it” or “how I think it should be done” kind of statements).

Interestingly, the proportion of control acts was higher in novice transcripts. In particular, there was a marked increase across time quartile in the amount of monitoring that novices undertook. This might be explained as reflecting the need for novices to devote efforts to reviewing and assembling ideas, checking on the state of completion and coverage, and explicit search for contradictions, as the analysis proceeds. There was also more evidence of scheduling acts in the middle phases of the analysis. This might be explained as an attempt by novices to control the problem space as the set of concepts increases in size beyond their immediate cognitive capacity.

Table 8 also reveals some interesting differences between experts and novices, and differences across domains. First, experts carried out many more generative and evaluative acts than novices, in all phases of analysis. 21% of expert transcript segments contained generative acts (where they came up with concerns or recommendations that might contribute to the analysis output) compared with 12% for novices. Similarly, 31% of expert transcript segments contained evaluative acts, compared with 24% for novices. Novices, on the other hand produced more paraphrase segments (25%) than experts (19%). The greater predominance of evaluations in expert than novice transcripts is common with other domains of expertise. However, both groups produced relatively more evaluative segments than comparable groups conducting instructional task design.

Table 8. The percentage of cognitive acts of generation, understanding and evaluation

Domain	Act	Experts					Novices				
		1	2	4	4	All	1	2	3	4	All
Case Analysis	Generate	8	22	27	19	21	4	9	21	13	12
	Evaluate	25	38	37	26	31	15	28	25	25	24
	Paraphrase	14	18	15	24	19	30	26	21	25	25
	Read	38	12	4	4	12	38	9	2	1	13
Task design	Generate	6	28	31	44	35	7	56	43	55	46
	Evaluate	22	23	11	14	18	10	13	12	13	12
	Paraphrase	6	6	4	3	5	7	8	6	7	7
	Read	15	14	14	12	13	56	10	5	5	22

References to the case statement

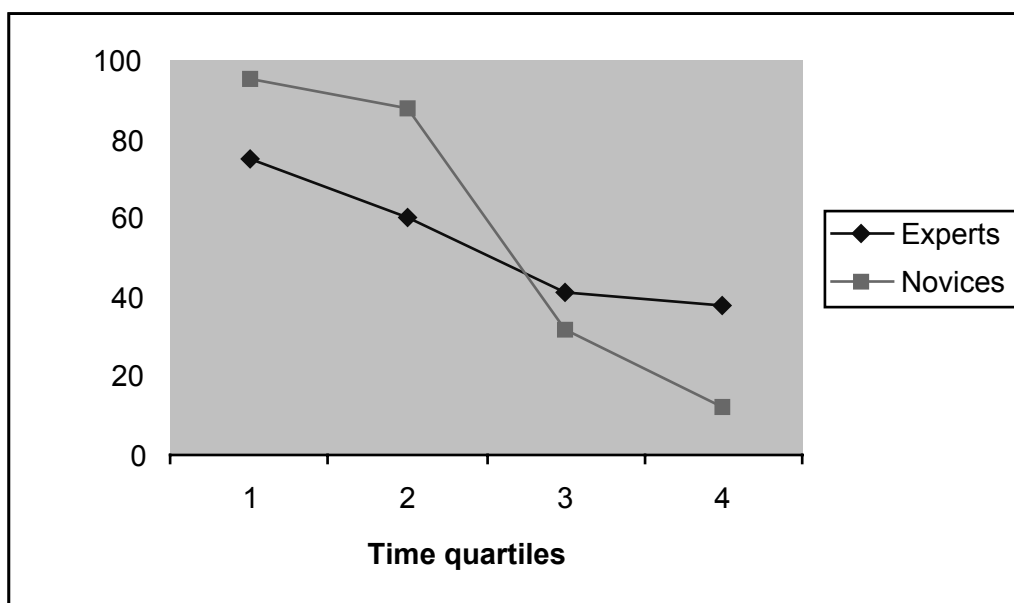
A number of differences between experts and novices in their use of the case statement are apparent in the transcripts. Figure 3 shows the percentage of segments in which experts and novices made reference to items explicitly mentioned in the case statement over time quartile. From this figure it is clear that, while novice transcript segments are dominated in early stages by references to the case statement, later stages show only a small proportion of such references. Experts show the opposite pattern.

Inspection of the transcripts suggests that this pattern of results can be accounted for by two separate factors. First, novices were generally more rigorous in their early reading of the case statement, and tended to focus upon its contents to the exclusion of any other referent (e.g. self-generated recommendations). The recommendations that novices generated were typically those offered as solution alternatives in the case statement itself. Experts often read the case statement in a more cursory fashion (a point neatly illustrated by the sample transcript of participant CS, Appendix 3). However, they were more likely to generate their own referents, based on knowledge or assumptions external to the specific case. Second, experts were more likely to return to the case statement in later phases than novices. In part, this appeared to be a strategic decision (e.g., three experts stated explicitly that they liked to review the case at the completion of the analysis, “to check if it all makes sense..”, or to “see if I’ve covered everything that might be relevant”). Novices, particularly those who had made extensive notes, tended not to make explicit reference to the case statement in later phases.

Another expert/novice difference emerges from an analysis over time quartile of the focus and level of detail in the case statement of participants’ references. Appendix 4 shows two examples of participants conducting an analysis of the Metropol Zoo case. The tables included show how many references were made to each part of the case structure hierarchy in each time quartile. Two key findings are demonstrated in these examples. First, the expert tended to ignore the details and focus on the broader issues of the case, while the novice tended to visit the detailed points as often as the broader issues. This is shown by the larger number of case statement items mentioned in all phases by the novice, and the fact that they tend to be at a lower hierarchical level than the experts. For example, the expert ignores all of the aspects of the case to do with ‘resolution issues’ (the role of the chairman, the ‘customer is always right’ clichés of the salesman, etc), while these are the focus of some attention for the novice. Second, the novice appears to process the case

in a more linear fashion than the expert. Note, for example, that the novice does not refer explicitly to the description of ‘problems’ offered in the brief after the first time quartile.

Figure 3. Percentage of segments containing reference to the case statement over time quartile.



Criterion-based referents

As discussed in Section 5, three kinds of criterion-based reference were made by participants: evaluative criteria, beliefs or assumptions, and semantically-rich paraphrases. Again, there was considerable variation among participants, which is illustrated by the examples provided in Appendix 5. The most typical expert/novice contrast is illustrated by comparing the novice NS16 (26 minutes analysis, producing 13 evaluative criteria, 3 beliefs and 4 paraphrases) with the expert SA (39 minutes analysis, producing 34 evaluative criteria, 18 beliefs and 10 paraphrases). While SA spent a bit longer analysing the case than NS16, she produced considerably more criteria proportionally for making judgements and decisions. Inspection of the evaluative criteria reveals that, while most of the novice’s were directly derived from, or relevant to, the case statement, those of the expert came from a wider range of sources or pertained to issues outside the brief. For example:

NS16: “VR is exciting for children” (a positive evaluation)

SA: “Business mission and vision would be destroyed by VR concept”

The ideas of business vision and mission are not explicit within the case (though there is a discussion of the zoo’s purpose, as conservation). Similarly, SA evaluates the case itself:

“Using anecdote to prove theory”; “ ‘its what the customer wants’ is an assertion”; “adoption of innovation factors is of interest in the case”.

There are no such general case evaluations among the novices criteria. As another example, consider the difference between the following criteria:

NS16: “customer survey only covers zoo visitors”

SA: “customer surveys are useless because visitor numbers are falling”

While they both point to the inadequacies of the customer surveys, one might argue that the expert’s criterion is ‘deeper’, in that it relates the critique to the core problem rather than superficially linking it to a property of the critiqued item itself.

Participant-generated referents

Appendix 6 shows four examples of the self-generated referents. In all four examples, the recommendation (a compromise solution, in which VR technology is installed incrementally to fund the conservation of the zoo), it is clear that the experts generated far more referents of their own devising than the novices. Novice protocols reflect a centrality and salience of a few ideas while ignoring many others. Experts on the other hand, seem to tackle a wider range of solution ideas and concerns

Interestingly, experts generally pursued their solution recommendations in more conceptual depth than novices. This occurred, in the main, because experts were attempting to link the outcomes of each recommendation into the whole solution. For example, expert SA developed the idea (r4.3.1) that interactive exhibits should be tied specifically in to live animal exhibits. This led to the idea (r4.3.1.2.3) that the environmentalists mentioned as antagonists in the case statement might be involved in the design of such exhibits, thus diffusing that source of difficulty. NS12 produced a similar solution, but did not expand it to any depth. Instead, she was side-tracked by drawing a layout map of the re-designed zoo. Interestingly, this led to the serendipitous discovery of an evaluative criterion, that walking past the animals to get to the VR attraction would feed back into enhancing the publicity of the live animal part of the zoo.

It is interesting how, in both cases, a plunge into detail generated important conceptual issues that fed into support for the recommendations. This contrasts with best practice recommendations in domains such as design, where plunge into detail is generally seen as premature commitment. In essence, the best solution recommendations for both expert and novice participants seemed to reflect a process of coming to a solution through the process of creating and rejecting options.

Group processes

These group sessions were very long and complex and only a preliminary aggregate analysis has been carried out. What is very obvious from the tapes is that the process differs very much indeed from that of individual case analyses. There is a wealth and variety in all aspects of the case analysis that is all too evident. This richness means that the case is better understood, more problem areas are elicited, a greater number of more creative solutions are proposed and each solution is discussed in greater depth. However there is clearly a process problem since this richness does not carry through to the final outcome; the summary of the recommendations. This phenomenon has been observed many times during case tutorials that are given to the presenting groups in the 3rd year marketing strategy course that all the novice Ss took part in.

The addition of a social dynamic to case analysis provides complexity as well as richness to the whole process. There was, for example, the taking of roles.

One traditional way of analysing such roles is to distinguish content and process roles. The former are largely concerned with creating output or meeting the task objectives of the group. The latter are concerned with facilitating the tasks. For example, in each group there were scribes, who wrote things down, summarisers, who every so often summarised the content of a particular or indeed the whole discussion, and controllers, who suggested what the group should do next. Scribes tended to be scribes for the duration of the session but other roles could swap between members of the group. Similarly content providers included members who knew something about the topic of the case, had useful experiences, were particularly creative or were good at criticising and evaluating. It was notable that the scribes were rarely asked to read back what they had written and as a result many useful insights were lost to the group memory.

Perhaps as a result of the "controllers" there was more overall structure to the discussion than was evident in the individual case analyses though there was never any evidence of an overall agenda. Indeed new topics could be raised at any time and were tolerated. Within a particular topic there was usually a fairly free form conversation leaping from one idea to another with often quite loose connections. When there was a hiatus, the usual routine was to return to the case material for stimulation.

The group seemed to be particularly influenced by the real world experience of one of its members and this seemed to take precedence over what was written in the case. However some offerings were met by complete silence and ignored for no apparent reason. Others were taken up and developed, often in a creative rather than logical way, with everyone contributing. On occasion personal feelings were brought up and the groups treated these as legitimate.

Both groups were almost entirely conflict free since they knew or were acquainted with each other and were not being assessed either individually or as a group. However there were clearly mild political and social pressures in the groups. One of the most important of these seemed to be wishing to promote ones own agenda without being seen to do so overtly. This wish to avoid being "pushy" meant that the groups had difficulty in both foreclosing, and summarising and there was little attempt by individuals to reintroduce interesting ideas that they had previously put forward. This reserve was demonstrated by the fact that statements were often put to the group in the form of questions. Humour was used throughout to help create group cohesiveness, to make the task more pleasant and to ease any tensions that might have arisen.

6.2 Outcomes

6.2.1 Overall

It was, as might be expected, difficult to assign numbers to the outcomes of the participants' case analyses. This occurred partly because we did not ask the participants to provide written output. This was a deliberate decision that involved a trade-off. On the one hand if we had asked Ss to provide written output this would have seemed a more daunting task and might have resulted in a rather impoverished summary of the analysis. We would, in any case, have had to take into account the oral as well as written evidence. On the other hand, writing might have forced participants into clarifying their recommendations and provided a stronger basis for deciding what the outcomes actually were.

The lowest quality mark in the whole set of analyses was 2 and the highest was 10 although the average was around 6.5. These results are hardly surprising given the

competencies and abilities of the groups of participants. Similar results were obtained for coverage and detail though both were marginally higher than quality. There was a tendency for participants to choose solutions that were not necessarily the best ones or particularly well justified but which were detailed and broad in scope and this probably accounted for the result. In general, the three measures correlated pretty well within each participant's analysis with the exceptions showing no obvious pattern.

In terms of content a great deal of interpretation was required since it was not always obvious that the same phrase meant the same thing in a different analysis context. It was also the case that there seemed to be a general tendency to avoid wholehearted commitment to a particular course of action. It was therefore often necessary to interpret the output from amongst the analysis being described at all stages of the analysis and not just at the end.

6.2.2 By contexts

6.2.2.1 Novices before training

While novices performed reasonably well gaining quality marks of around 6, their recommendations were flawed in a number of obvious ways by comparison with the experts. In particular knowledge of the world and the way it works in general, and in particular for any case, is an important variable for novices. The quality of novice S solutions were often marred because they had misunderstood what particular technical words meant (e.g. broker), how particular processes worked (e.g. store location decisions) or the nature of particular constraints (e.g. financial or managerial). In addition, since all cases require the making of assumptions, it was also apparent that novices made unrealistic assumptions. For example they often assumed that something would happen if it were simply suggested, ignoring the organisational and political problems associated with one solution or another. There was also a lack of creativity especially in the untrained novices with solutions being often simply chosen from those either offered in the case or minor variations on what was already being done

6.2.2.2 Trained novices

Happily in terms of quality of outcome novice participants improved after training! (mean 5.8 to 6.8). We begin to see the emergence of pursuing more than a single solution alternative. For example:

" Another solution to the problem of this over capacity and the fact that the streets are over congested with cars and the pollution is that this car is very small and environmentally friendly, and friendly to the environment" (NS14, Swatch, line 71).

At least part of this improvement could be attributed to better structured recommendations. Several participants actually mentioned elements of the case analysis framework they were exposed to during the course and seemed to be using it. In some cases, it was, essentially, by rote recall:

"Right before I do anything I'm going to write down what the bloke said. He said (writes 'understand situation, diagnose problem areas, generate alternative solutions, predict outcomes, evaluate alternatives and then communicate results').. (NS10, Bupa, line 001).

Notwithstanding the somewhat dogmatic use of the framework, the participant then went on to apply it in a fairly vigorous fashion, for example:

“ OK the competition, so for this the structural issue one... Solutions - reorganise the company or do nothing. ... If they do nothing they can just function as they do but they're not doing too bad at the moment....They're market leaders but there's a bit of confusion going on so if they reorganise the company they can have like the brand core at the centre then like 3 separate businesses so to speak.... Outcomes of that - there'd be no confusion going on (NS10, Bupa, lines 108-111).

Many other examples of increased structuring can be found in the protocols. Some of this is evident in planning phases. For example:

"So, we'll first identify what are, identify, some of the problems in this case" (NS16, Metropol Zoo, line 12).

In other cases, structuring of output is increased. For example:

"I'll just write down some problem areas that I highlighted as I went through the case" (NS17, Metropol Zoo, line 71).

6.2.2.3 Experts

It came as no surprise that the marketing experts gained much higher "quality" marks than the novices (mean around 8). In addition their scores for coverage and detail were very much better and somewhat higher than the quality scores. The quality marks would have been higher except for a pronounced tendency on the part of the experts to avoid coming down solidly in favour of one solution rather than another. The highest marks were obtained by those Ss who were not only subject but also process experts. In other words they had a great deal of experience of teaching with case studies.

6.2.2.4 Non marketing expert

The subject performed as one might have expected; the quality mark was higher than that for a trained novice but lower than the average for the marketing experts. What was particularly interesting was that the analysis made use of a number of well known marketing ideas (e.g. Unique Selling Proposition, differentiation) and suggested a solution that no one else had put forward. The concept of expertise will be discussed in more depth below.

6.2.2.5 Groups

The quality scores for the group were again higher than for individual trained novices but slightly lower than for experts. The group processes led to a greater understanding of the complexities of the case situation and more creative solutions as evidenced by the fact that the coverage and detail marks were the highest in all the categories of participants. This difference was largely due to the inability of the groups to summarise with enough clarity what their joint recommendations actually were. In terms of content the groups produced rather more subtle variations on the usual theme discussed in great detail.

6.2.2.6 Cases

It is possible that the clustering of the quality scores were a result of the relative shortness of the cases and the ease of their analysis. The rank order of the case scores was as follows Aldi, Metropol Zoo, Swatch, and BUPA. This is rather a surprising result in one aspect since Metropol Zoo has a clear cut decision to take, a known context and a relatively

simple structure and set of data. However, the other cases rank as one might expect taking into account familiarity of context, clarity of task and complexity of problem situation. There did, however, seem to be a strong interaction in some cases between the participant and the case study. One novice scored a 7 after the training while scoring 9 before. It may be just how he or she felt on the day or it may be that something about the case, its topic, its length, the context, the task that either reduced motivation or meant that the participant didn't have the capability of tackling it. However it was noted that content in some cases, particularly Metropol Zoo, elicited rather similar recommendations in contrast to, for example, Swatch. This point will be discussed in more detail later.

Discussion

The data we report above can offer only a flavour of the richness of analytic behaviours we observed in our transcripts. In part, this is because we sampled only 24 analyses for detailed investigation, and we clearly need to complete a detailed analysis of all 74 novice and 24 expert analyses that we recorded. We are also aware of the restrictive nature of the case analysis sessions, and the questions that this raises as to the generalisability of our findings. Nonetheless, it is clear that there are expert/novice differences, that there are effects of training, that groups do differ from individuals in the way they approach analyses, that many of the behaviours we would expect to find were present, and that some we expected to dominate the analyses were peripheral or absent.

First, there is no simple relationship between case size and length of analysis. Instead, the time spent on analysing individual cases varies more by participant and by case structure than any other factor. There may, of course, be some concerns about participant's motivation. Many seem to be content to do quite limited analyses and stop. Others didn't seem to know when to stop or may even have been conditioned by the statement that we had said it would take about an hour. Nonetheless, all participants reported that the analyses they had completed represented the kind of first pass that they would normally expect to carry out, prior to any further group work, teaching use, or presentation. The question remains then as to how participants decide when to close. This leads to the important of time pressure and the acceptable solution. In terms of simulating the "real" world this is clearly an important issue. We cannot answer this question here, beyond suggesting that case materials need to be tolerant of different individual approaches and styles. Consider the analyst who stops initially after 8 minutes saying:

“well when you say have I analysed it I've got what I consider to be the big issues out of it ... what I would expect from students , and that's because I mean I started drawing important things here” (CS, expert non-Lancs, line 58).

Is this participant necessarily doing a worse job with this case than an analyst who labours on with a case for over an hour, before reaching essentially the same broad conclusions? In a training context, maybe – a key aim, after all is to learn the reflective and communicative skills of problem-solving. However, in a managerial context, maybe not – rapid decision-making can be a positive attribute (though in a 'consultant' context, one might argue that the appearance of detailed analysis is at least as important to client satisfaction and the products of the analysis).

Another aspect of performance that does not yield to simple description concerns the use of externalisations, such as written notes, diagrams, underlining of parts of the case statement, and so on. These data are illustrated in Figure 2. Perhaps not surprisingly, more externalisations were made early in the analyses than late, when participants were reading through the case statement in the first phase of problem understanding. However, it is perhaps disappointing that we saw little attempt to structure the outputs of analyses in written form (though this, of course, was not a strict requirement of the participants). There are two aspects to this observation that we believe are important.

First, our expert sample do not seem to reveal a formalised or systematic approach to documenting their analysis outputs. This, may simply be that they judged it to be expedient in the context of data collection to verbalise their output rather than write it down. However, it is interesting to contrast the use of externalisations by experts here with evidence from other domains. For example, Davies (1992) found that expert programmers, working in similarly contrived laboratory studies to the present study, externalised far

more than novices. He argues that experts develop a strategy of externalisation, in part to reduce memory loads, but also because external repositories of ideas can be re-organised and parsed in creative ways that are not possible with mental representations of the same information. In the present study, it is interesting to see that the same strategic use of externalisations does not occur. As well as the explanation of expedience, it is possible that experts are aware of a potential danger of externalisation, that of ‘reifying’ early ideas that are later dismissed after subsequent evaluation. Indeed, two experts, when questioned post-hoc as to why they chose not to make any notes, claimed that they never made notes on a first analysis precisely because of the dangers of fixing poor quality ideas in their documentation. On the other hand, some participants outlined a specific pedagogic role for externalisation. For example:

“if you can’t draw it you can’t have simplified it enough” (CS: Expert, non-Lancs, line 001).

Second, while some novices made copious notes early on, their subsequent use of these was sporadic. In some cases, participants did not review their notes at all. In other cases, it is clear that participants simply collated their early observations and recommendations as the output of their analyses. One clear relationship is between the time taken to analyse cases, and the use of externalisations by novices. Essentially, the more the novice externalised, the slower their analysis was. Sadly, however, there is no clear relationship between the amount of externalisation and the quality of the output. If anything, there is some evidence that the most diligent note-takers tended to produce the least creative solutions. This raises some interesting questions regarding best practice. As we will argue below, there is clearly a need for some kind of support for the documentation of analyses. Yet, these findings seem to suggest that simply encouraging learners to be more diligent throughout the analysis process in their documentation might be counter-productive.

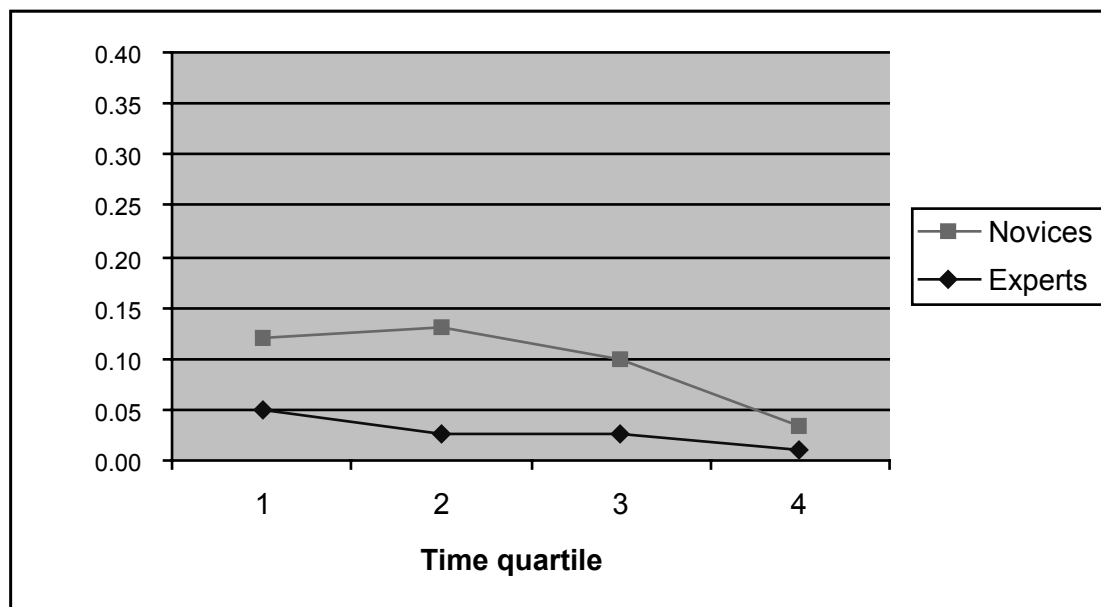
The analysis of cognitive acts revealed a number of interesting, and in some cases surprising, differences between experts and novices. These specifically concern differences in control activity, evaluation and generation.

The relative absence of control activity among experts in case analysis contrasts with all other expertise domains that we are aware of. For example, in the domain of engineering design (Adelson & Soloway, 1985; Ball, Evans, Dennis & Ormerod, 1997), experts differ from novice in being increasingly breadth-first in their structured decomposition of complex design problems, a process they maintain by extensive scheduling of design problems and monitoring to check extent of completion. This is usually regarded as good practice, since it prevents premature commitment to solution details while ensuring that the design problem is pursued systematically and completely. Even in the domain of designing instructional tasks for teaching English as a Second Language, we have observed that experts are highly structured (Ormerod & Ridgway, 1999). This is despite the fact that a search for creative and novel solutions is perhaps the key performance requirement in task design. Figure 4 shows the proportion of scheduling activities by experts and novices over time quartile. From this figure, it is apparent that novices seem to undertake more scheduling, an activity usually associated with expert behaviour in other domains.

Why then, do we see so little evidence for structured decomposition in case analysis? It is not, we are confident, because our experts were not producing high quality case analyses. The outcome measures reported in Section 6 suggest that expert analyses were qualitatively better in all respects than those of novices, even those after training.

How, then do experts manage such a complex and apparently unstructured problem-solving activity without explicit control acts?

Figure 4. Proportion of control acts made by experts and novices over time quartile



There are two possible explanations, and it is likely that both play a role. First, experts will have stored knowledge concerning the typical structure of cases. They will have expectations about what the case is likely to contain, and how and when particular knowledge types become relevant. At the same time, they will have richer domain knowledge, both about cases in general and about the specific case domain. These knowledge structures may provide the kinds of organisation that allow experts to evaluate and generate on-line, without having to formulate plans, review work conducted to date, and so on. Second, experts may have strategic knowledge about the value (or relative lack of it) of planning and reviewing. That is, they may deliberately tackle a case in a data-driven fashion, avoiding the creation of later sub-goals in order to be fully influenced by thoughts and issues as they arise. Indeed, as one expert commented in a post-hoc debriefing “Its best to let a case wash over you at first... the chances of spotting anything interesting decline the more you process the information as you read it”.

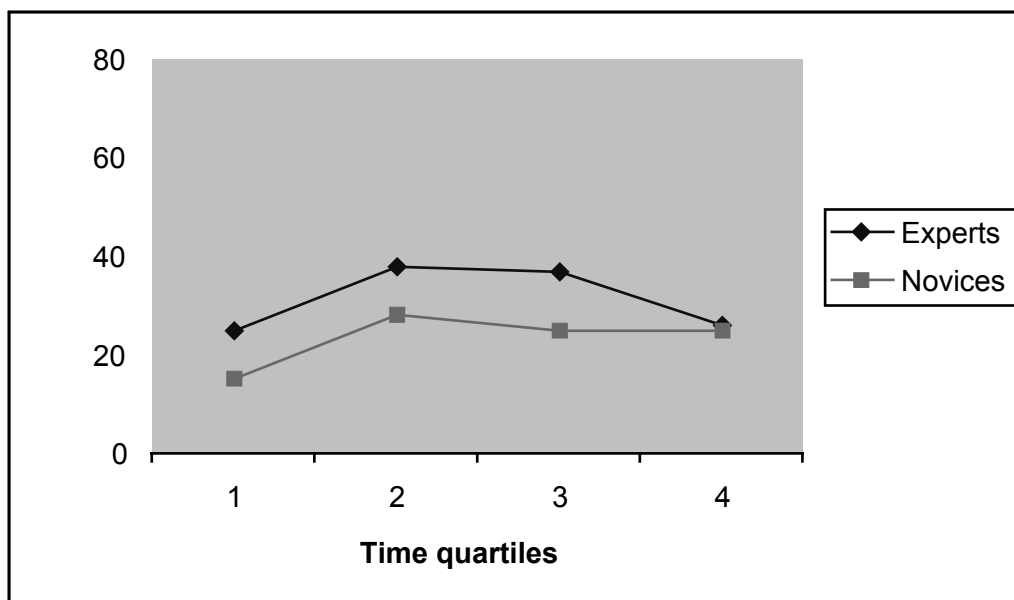
One possibility that cannot be ignored is that the cases were of necessity short cases since longer cases would have resulted in longer, more complex and more tedious experiences for the participants. As it was, novices saw a two-page case as more taxing and difficult than a one page one. It is possible that long cases are tackled differently than short ones. There is more information and presumably less need for assumptions. However it would require organisational skills and would probably lead to Ss ignoring more material. Thus, one might expect to see a greater preponderance of control activities among experts with longer cases.

Why, then, do novices show signs of control activity? It seems likely that acts of monitoring and scheduling are ‘coping strategies’. They provide ways of dealing with a large assembly of apparently relevant issues and ideas when there are no knowledge structures that promote on-line evaluation. In essence, when a novice encounters an issue,

either self-generated or more frequently from the case statement itself, they have no way of knowing until later in the analysis whether it will turn out to be important.

If experts are not scheduling or monitoring, then what are they doing? The answer appears to be that they are generating and evaluating. Expert analysts evaluated more and generated less compared expert instructional task designers (Ormerod & Ridgway, 1999). This is perhaps not surprising, since one might argue that evaluation is the key goal of the case analyst, while production is the key goal of the task designer. Perhaps more significantly, expert analysts generated more referents (concerns, recommendations) than novices, and they also evaluated more throughout than novices. Figure 5 shows the proportion of evaluative acts over time quartile for each group. It is apparent that experts evaluate roughly twice as much as novices. Clearly, the amount of evaluation is an important indicator of expertise. This is common to all domains of expertise, such as engineering design (ball et al, 1997) and mathematical problem-solving (Shoenfeld, 1990).

Figure 5. The percentage of evaluative acts by experts and novices over time quartile



Experts did not simply evaluate more often than novices, they did so with more and better evaluative criteria. The examples given in Appendix 5 demonstrate both the wider range and greater conceptual depth and abstraction of the evaluative criteria used by the experts. These expert/novice differences in the role of evaluation are consistent with the ways in which participants made use of the case statements. The data suggest that novices' early phases of analysis were dominated almost exclusively by attention to explicit case statement items, while experts were more likely to generate their own issues and ideas. In later phases, however, experts were more likely to return to the case statement, in part to evaluate the products of their analysis against the given statement. In the early stages of the analysis novices seem to just read and make simple evaluations. Experts were more likely to question the material in the cases and actually ignore a point if they thought it was incorrect. They were also more likely to pick out value judgements and separate them from facts. Novices were much less critical.

In essence, novices' recommendations generally arose simply from reading the case statement. Given that their recommendations were so closely tied to the case statement,

there might appear little need to the novice to re-visit the case statement. It is interesting to note that experts continued to generate solution recommendations throughout the analysis. Sometimes these recommendations emerged after a re-reading of the case statement. It appears then that creative ideas spring, not from simply reading the statement, but often from re-reading the case material.

The circumstances under which experts go back to reading the case may offer some suggestions for refining the case analysis process as it is delivered to novices. For example, re-reading seems to serve a number of roles for the expert analyst: for inspiration, to confirm a point, and as part of a systematic process of review. It is possible that a prescription for novices to read the case again every 15 minutes or so might improve their resulting analyses. However, it is important that such a methodological change stresses to the novice that their recommendations should be tested, not only for fit with the given case statement, but also for differences that might signal creative intuitions that are worth retaining.

There does seem to be a phenomenon one could call a flash of inspiration. These can be either in terms of understanding or a creative idea. How do they occur in the sense that there seem to be no precursors? One possibility is that they are serendipitous, in that the 'best' ideas often stem directly from attempts to predict the outcomes of a putative idea. The examples of expert SA (her solution concept for Metropol Zoo involving the direct involvement in exhibit design of the environmental protagonists) and novice NS12 (her idea to force a walk past the live animals to get to the VR attraction as a boost to conservation publicity) have this property. In both cases, the 'bright idea' was a by-product of an otherwise fairly pedestrian initial solution concept. It is interesting to note that this suggests a different role for the "Predict outcomes" phase of analysis suggested by Easton (1992) in his case method account. As well as providing an evaluation of the current idea, this phase seems to be a key source of truly innovative ideas.

One surprising aspect of the coding of self-generated referents is the relative absence of analogies or remembered events or examples in both expert and novice transcripts. Typically, expert transcripts would reveal a single analogy that played no further role in the analysis beyond the segment in which it was mentioned. For example, one of the 'better' expert solutions (SA. Metropol Zoo) revealed only one self-reference to an analogy, that of Sinclair's C5 car. This appears on a single line, and is not referenced again. Its source is likely to be the previous case that the participant tackled, the Swatchmobile case, where explicit reference was made to Sinclair's 'black watch' design. Some analogies seemed to crop up repeatedly. For example, no less than four experts came up with the same analogy to the Metropol Zoo case ("declining public funds and falling visitor numbers... sounds like a university"). There are a number of interesting issues that this repeated occurrence of the same analogy generates. For example, the fact that it springs to mind with the current expert cohort may not be surprising given that they are university academics. However, their use of it, and the absence of any evidence that it changes the subsequent analysis that they conduct, suggests that it serves a conversational rather than generative or analytic role. Furthermore, a moment's thought reveals it to be a rather poor analogy! Government funds may be declining for universities, but (perhaps unfortunately) 'visitor numbers' in terms of students wanting to attend, are certainly not. Novice transcripts had more remembered examples ("the last time I went to the zoo", "my sister's experience with the Ford Ka", etc), but rarely more than three, and typically they were restricted to a single segment. Nonetheless, novices' personal experience seemed to weigh more heavily in making judgements about how the real world works and what is possible. This is perhaps not surprising in view of the lack of relevant case and domain

knowledge that a novice has relative to an expert. Novices seem to use their own experience directly and openly while experts seem to have abstracted it and use such knowledge more covertly.

Critically, from the point of view of promoting the case method, we gave uncovered evidence that it works, at least to the extent that a one-term course can change novice analysts' behaviours. The increased depth of analysis (despite equivalent case analysis times) and better solution quality are important indicators of the approaches success. What remains to be seen, however, is precisely how much of the improvement came from the case method, and how much from other sources? A critic might reasonably argue that we did not have a control group, that is, a group who did a course in Marketing based on some other technique, or simply upon examples delivery. It would be interesting to see whether similar improvements can be achieved by other interventions. However, the evidence of rote reproduction of Easton's (1992) approach suggests that the method must be having at least some specific impact.

Finally, it is clear that groups produce even richer analyses than individual analysts but seem to be no better at capturing that in the final product. Our analyses are incomplete, but they show some interesting patterns, such as the spontaneous adoption of roles, and the more complete pursuit of multiple alternative recommendations. The study, however, raises more questions than it answers. For example, does the individual learn to rely on others in the group and become lazy or a specialist or do they learn by copying? Do individuals who perform in groups then perform differently as individuals in subsequent analyses? Questions such as these require further study.

Implications

While a large proportion of the data await quantitative analysis, a number of initial results can be highlighted from the study:

- Although the time spent on analyses did not differ reliably between experts and novices, large qualitative differences were found between the groups. Experts generated more alternative recommendations, identified more critical issues and used more evaluative criteria than novices. The outcomes of their analyses were generally qualitatively better than those of novices, and were more likely to bring in issues not specifically referred to in the case statement. Novices also tended to reach a firm viewpoint or recommendation early (often during the first reading of the case statement), while some experts deferred reaching a recommendation until later in the analysis, were more likely to change their stance during the analysis, and in some cases did not reach a specific recommendation at all. Novice analyses focussed more upon outcome while expert analyses were more likely to focus upon process issues.
- Novice analyses tended to be disappointingly shallow, and constrained by the content and order of the case statement. Perhaps the most important general finding is that the final output of the analysis often seems to be a poor summary of the richness of the process that has gone before. This contrasted with expert analyses, which became more focussed yet did not lose the richer issues generated early in the analysis.
- Comparisons between novice analyses before and after training revealed evidence, both of an improvement in the quality of analyses, and changes in the processes of analysis. In particular, the depth of analyses increased, and solution development tended to be deferred until later in the analysis.
- A number of differences between case analysis and other creative problem-solving domains were identified. While phases of problem understanding, solution development, evaluation and review are common across domains, there was relatively little evidence in the expert protocols of the control activities (specifically scheduling and monitoring) that dominate protocols in domains such as design problem-solving. There was also surprising little overt use of analogy or remembered exemplars in either expert or novice protocols. This contrasts strongly with evidence from studies (both within the current study and in an associated project) of group processes in case analysis. Another surprising outcome is that experts tended to show more variability than novices. This contrasts with the majority of expertise domains, where increasing expertise typically leads to a convergence of process and outcome across individuals. Finally, the use of external documentation by experts was subject to great variability – some experts claimed deliberately to avoid early note-taking, as ‘reification of bad ideas’. There is some evidence that early documentation by some novices inhibited the quality of their subsequent analyses, since they tended to collate the documentation in place of a properly structured analysis output.
- The kinds of outcomes generated by both expert and novice groups are, at least in part, a function of the nature of the case itself. Where case statements represented the analysis as a choice between alternatives, this highly constrained the nature of the analysis, while cases that were open-ended typically elicited richer analyses, especially with the expert and post-training novice protocols. In the pre-training novice protocols, however, there is evidence that open-ended cases gave rise to large individual

differences: while some participants produced richer analyses than with more constrained cases, others produced highly superficial and truncated analyses.

Analysing cases is a clearly a complex process with complex outcomes. As a result it is crucial to understand that while there may be recurrent cognitive and social processes common to all case analyses, these are modified by particular contingent factors. It is therefore important to realise the impact of the analyst, the setting and the case in trying to offer advice to case analysts and instructors. It would appear that experience and training help but the most effective ways are not entirely obvious, especially since initial learning effectiveness seems to be a function of individual differences in response to case type. We putatively suggest ‘snowballing’ as a component of expert case analysis skills, that is, the online assembly and collation of key issues and recommendations that carry through an analysis in the absence of explicit control activities. Support for this notion comes from the fact that, despite showing little overt scheduling and monitoring, experts were able to conclude their analysis with a rich and complete set of issues and recommendations that focussed upon key issues and showed significant gains over the given case statement.

Despite the complexity of, and variability in, the results, a number of provisional recommendations and suggestions for further work can be drawn from the study.

First, it would be helpful to find ways to capture the richness of early analyses for both individuals and groups, to prevent the losses of issues, recommendations and evaluative criteria that characterise case outcomes in novice protocols. We propose to explore mechanisms (both methodological and technological) that offer a form of ‘inverse scaffolding’ across case sessions. By inverse scaffolding, we mean support for the process of documenting the analysis process incrementally, but not exhaustively. This contrasts with the educational research notion of scaffolding in learning, where scaffolding (i.e. learner support through examples, instruction, feedback, and so on) is given in full at the start and then gradually withdrawn as the learner achieves competence. We suggest that novice case analysis sessions require the opposite. Too much documentation support early on would make the novices reliant on the initial ideas and critiques that seem to stem from their first reading of the case statement. Thus, it is important to increase documentation support (e.g., through prompts, summary sheets, etc.) as an analysis session proceeds. Hence, the scaffolding is added as the learner’s analysis grows in sophistication, with the target of supporting a structured and articulate outcome.

Second, despite the apparent gains made by novices after training, the expert protocols were much richer and more compelling. It may be valuable to develop case method training materials that provide, not only a case statement, but also transcripts of expert analyses as target exemplars of best practice. The effectiveness of this kind of ‘vicarious learning’ from analysis episodes, rather than simply from accompanying teaching notes, presents an interesting topic for further research.

Third, it is clear that different things happen when students analyse cases as individuals and in groups. While it is fair to say that the majority of case method use focuses upon group activities, it is important to recognise both the differences between individual and group settings, and also the contribution that individual activities make to group processes. To do this requires a set of data elicited from groups that is directly comparative with the data reported here from individuals (notably before versus after training data from group analyses of new cases). Fourth, we focussed in this study on short case statements, as a way of establishing a baseline of performance in as simple a context as possible.

Notwithstanding this deliberate simplicity, we have uncovered a richness of issues that might not have been anticipated from such simple cases (notably the between-case effects). However, studies of simple cases preclude the observation of many of the skills associated with more realistic cases, such as data handling, simulation, and possibly the kinds of control activity that are not apparent (and even not necessary) in small scale analyses.

Dissemination

There are a number of routes to further dissemination of our results, and we are hoping to engage actively in these, under the guidance of the ECCH. First, we would be happy to engage in seminar under the auspices of ECCH in line with their existing seminar series, offering a form of research based case teaching. Second, we wish to write research papers, for both psychological and management/marketing audiences. The target journals include the Journal of Marketing Education, Instructional Science, the Journal of Management Studies, Cognitive Science and Applied Cognitive Psychology. In addition at least one paper will be targeting at the Academy of Marketing Conference in 2001 and the Cognitive Sciences Society in 2002.

At the level of particular findings, the main outputs will be recommendation to case teachers about how cases should be designed and taught. These recommendations will be neither comprehensive nor definitive but will hopefully provide both help and stimulus for further thought and research. It is suggested that the main medium for the presentation of this, and other research funded by the ECCH Research Panel, be a national seminar on Research on Case Teaching in Management, to be held in Cranfield sometime during 2002. The seminar should, of course, aim to be self-funding but would require organising. Invited speakers could also be included and the emphasis would be on research, rather than anecdotally, based prescriptions. Management teachers and management developers outside the academic world would also be part of the target audience.

Failing this alternative we would propose to hold either a regional seminar at Lancaster under the joint auspices of the British Academy of Management and the Academy of Marketing with presentations of this research and perhaps one or two other invited researchers. It is also suggested that all researchers be routinely invited to make presentations at the ECCH Teaching with Case Studies Workshops.

One of the most important outputs will be continuation of the research. Should the results prove sufficiently interesting it is proposed to seek further funding to extend the research to cover MBA and post experience managers, to vary the form of the case studies included and to increase the overall size of the sample, choosing subjects based on the individual differences identified as a result of the research proposed here.