

An ecologically valid study of categorisation by designers

Thomas C. Ormerod, Ryan Rummer
Lancaster University

Linden J. Ball
University of Derby

Abstract

This paper reports the use of ethnographic and experimental methodologies to investigate the conceptual representation of design expertise. An ethnographic study was first conducted in a commercial software development group. This study yielded a wide range of data, from which thirty design episodes were selected. Members of the design team then undertook a card sort task, in which they sorted the episodes into natural categories. Differences in the sorts produced by designers reflect their specialist roles within the design team. The same design information is evaluated under different metrics across roles. This corroborates previous studies that show how conceptual design representations differ in individual and team contexts. The results also show how combining the outcomes of ethnography with a traditional cognitive method enable us to verify objectively the observations made in ethnographic studies.

Introduction

The use of ethnography in design research has proliferated in recent years (e.g., Bucciarelli, 1988). Often ethnographies are carried out to inform the development of design tools and methods. While ethnographic methods provide rich and naturalistic data, they face the criticism that observations are open to subjectivity. We have argued (Ball & Ormerod, in press) that where the goals of an ethnography are purposive, then there is a requirement for objectivity. One approach to enhancing objectivity is to use experimental methods of cognitive psychology. However, examples of the successful use of experimentation to explore design are few. This is hardly surprising given the complexity of the domain, and the consequent difficulty and validity of imposing experimental controls upon expert behaviours in pursuit of objective validation of hypotheses. In this paper, we argue that ethnography can play a role in *enhancing* the ecological validity of experimentation in design research. We report a study in which the ecological validity of an experimental method is enhanced by the provision of realistic materials from an ethnographic analysis.

A basic tenet of the ethnographic approach is the adoption of a non-interventionist research tradition. Most cognitive psychology, on the other hand, is of an interventionist nature, gaining its strength from the systematic manipulation and control of variables that might affect performance. Although ethnographic methods provide rich and naturalistic data, observations are open to subjective sampling and interpretation. Conversely, attempts to use experimental methods to explore design immediately confront problems of *ecological validity*. Notwithstanding these problems, we argue that both approaches have a role to play in design research. The role of controlled experimentation is demonstrated in usability testing of products and prototypes where objectivity in observation and interpretation is essential. Similar requirements for objective assurance face the researcher in developing theories of design expertise, accounts of optimal design practice, or tools and methods for supporting design. We suggest that ethnographic and experimental methods can be used to triangulate observations across methodologies.

In the remainder of this paper, we report an empirical study which combines ethnography with the use of a sort method for investigating the categories designers use to classify information. Participants are given sets of cards that contain domain descriptions and are required to sort these into categories significant to them. The method originates in research into memory and the mental representation of concepts. The rationale is that, by examining the nature of sorts, one can infer participants' mental representations of conceptual knowledge. In applied research, the method is used to study expert/novice differences (e.g., Davies, Gilmore and Green, 1995). It has been suggested (e.g., Burton, Shadbolt, Rugg, & Hedgecock, 1990) that the sort method provides an informative but cost-effective method for eliciting expertise in the development of expert system. However, the sort method is not without its problems. In early psychological research into conceptual categories, there was an assumption that what was being elicited through a sort method reflected relatively static conceptual representations. However, this view has been challenged by Barsalou's (1985) description of goal-directed categories. In this view, categories are not fixed, but are determined by the task faced by the individual at any one time. If this view is correct, then the assumption that a sort method elicits static representations of concepts is unfounded.

A particular problem facing the use of the sort method concerns the nature of the items to be sorted. In traditional sort tasks, the items consist of object descriptions, often varying in familiarity, representativeness, or prototypicality. In applied research, they may contain descriptions of domain objects or terminology (e.g., programs - Davies et al, 1995; archaeological shards - Burton et al, 1985). However, it is in the selection of sort items that the method can lose ecological validity. If concepts are goal-directed, then the sorts that experts produce are likely to be a function of the items they are given. One way around this problem is to use as sort materials items produced as part of expert performance. To date,

we are unaware of any studies employing the sort method that have used items produced by the study participants themselves. This is not surprising given the difficulty of collecting such items. We suggest that ethnography can be used to elicit design episodes that provide ecologically valid materials for a study employing the sort method to explore design expertise.

The aim of the study reported here was to investigate whether natural categories can be identified that are used by expert designers to classify design information. The study is part of a wider project to develop a computer based indexing system for supporting design reuse. Thus a further aim was to identify natural categories to provide a notation for encoding design information within the system. The study took place at a major manufacturer of personal computers (PCs), and involved two phases, ethnographic data collection and the experimental study.

Phase 1: The ethnographic study

In the first phase, our researcher spent two weeks conducting an applied ethnographic study of design practices within the company's software development team. A full report of the first phase is available elsewhere (Ball & Ormerod, in press). The general remit of the team was to develop software networking solution. The team had twelve full-time members. As well as the team leader, two members were principally involved in managerial and high-level planning or commercial liaison activities. Four members were specialist designers, each with their own area of expertise (e.g., local area networks, hardware systems). The remaining members were programmers. The team had an hierarchical management structure, in which one or two programmers and one specialist designer worked under the remit of one of the three senior team managers. At the beginning of each week, the whole team had a two hour meeting in which current projects were reviewed and team goals for the week were shared. Otherwise, team members worked mixture of single and shared office spaces, with numerous informal conversations and meetings taking place but otherwise few formal exchanges.

The researcher followed a standard ethnographic approach, collecting a mixture of field notes, taped interviews and meetings, texts and other company documentation, whilst noting personal observations and commentary. At the end of this period, we analysed the resulting data with a view to establishing a set of design materials from the data that could form the basis of a sort study. Previous work has suggested that the 'design episode' forms a useful unit for describing design activity. As a metric for determining the scale of a design episode, we chose the Questions, Options and Criteria approach of (MacLean, Young, Bellotti, & Moran, 1991). The identification of episodes was made on the basis of judgements by two researchers as to the extent to which a segment of data reflected the pursuit of a single design question. Thirty episodes were chosen to reflect the types of design information that were evident in the ethnographic data.

Phase 2: The experimental study

The aim of the experimental study was to investigate the natural categories of design episode information produced by the designers who had been the focus of our observations in the ethnographic phase.

Method

Nine members of the design team took part in this phase (three managers, three specialist designers and three programmers). Materials consisted of a thirty episodes, each printed on an A5 size card (examples as in Table 1). Participants carried out the sorts individually. Each participant reviewed a instructions that included an example sort. They were then given 30 cards, each displaying one episode. Participants were asked to sort the episodes into natural categories segregated by a dimension of their own choosing, and that captured a view of the episodes that was deemed important or relevant to their own design activities.

They were told to use as many or as few categories as they found necessary to represent each grouping. Each participant completed two sorts, and the dimensions used to differentiate categories and the cards assigned to each category were recorded.

Table 1
Example episodes

(audiotaped conversation): “Actually, ‘Jim’ said to me yesterday, or Friday last week maybe, that ‘John’ wanted to talk about a way of formalising the process from Lotus Notes to NBP. You know, how do we link these and have an actual cut off that says...here’s the point where we stop talking about this stuff and say that’s the decision. Now put it to a Word Pro doc and up to the library, and make that a design point.”

(email): TARDIS Web Page & Logic Diagram

As the Web page idea has crystallised and grown!, so has the need to firm up the design. I think the logic diagram I put on the [...] Web site [..address..] will describe where we are headed. You will see how the Web page will navigate the user to the right download site ... It will also give him the right code that will automatically route his support request to the correct Helpcenter. The response has been no less than enthusiastic from all and we want to keep up the momentum.

(field notes): 'Bill' describes to 'Mike' during an informal morning meeting in Mike's office how The GUI and ‘engine’ of TARDIS v0.1 are distinct but closely coupled - always running on the same system. TARDIS v0.2 removes this restriction by introducing a new GUI in the form of a combination of HTML web pages and Java applets to allow [..... deleted.....]. Bill is seeking clarification of whether he should continue to work on the new GUI. 'Mike' queries whether he is the right person to give this clarification.

Results

Data analysis focuses upon the category assignments for each sort and category labels provided by designers. The average number of categories produced in each sort was 5.8 (standard deviation = 2.3). Much of the variability comes from a single designer, who produced two sorts of 10 and 11 categories. Example category labels produced are shown in Table 2. From these, dimensions that govern distinctions between category labels were inferred, mainly from concurrent verbal reports of the designers.

The data were subjected to multi-dimensional scaling (MDS) and hierarchical cluster analysis methods using Statistica™ using two data matrices. The first was a similarity matrix scoring the number of designers making each possible episode pairing. The second was a dissimilarity matrix in which pairing of episodes was shown for each participant. The first matrix allowed for analyses of episode clustering across dimensions, whilst the second matrix allowed for the analysis of participant groupings and outliers. Figure 1 shows a mapping produced by an MDS run from the similarity matrix. Note that the dimensions are those determined by MDS, not those inferred from the category labels. However, these dimensions fit reasonably closely to the dimensions of Product (with a scale from 'general' to 'specific') and Role (with scale from 'managerial' to 'implementor'). The boundaries drawn around episode groups were identified as a result of conducting an hierarchical cluster analysis on the matrix data (using a complete linkage algorithm). Similar mappings were produced for the dissimilarity matrix data. However, space limitations preclude their inclusion.

Table 2
Dimensions of category labels produced by designers

Dimension / labels

Designers (d1-d9)

Relevant to specific group/role (e.g., "technical group; systems; bureaucratic; test group; methods; status meeting")	d1, d2, d5, d7, d9
Design process/stage (e.g., "process; marketing; function; logistics; program development")	d2, d3, d6, d9
Episode content type (e.g., "obvious procedures; press release; discussion of specification; general discussion; specification details; general narratives; long-winded info of unknown quality; status notes")	d1, d4, d6
Product (e.g., "TARDIS v0.1; POLAXv0.2; not product-specific")	d3, d8 (sorts 1 and 2)
Information utility (e.g., "trivial items; info relevant only to participants; sensible suggestions; important work items; hi-level design descriptions; detailed design info that does not affect designers")	d5
Decision stage (e.g., "Things already done; defined options requiring decision; early design discussions; oddballs")	d7

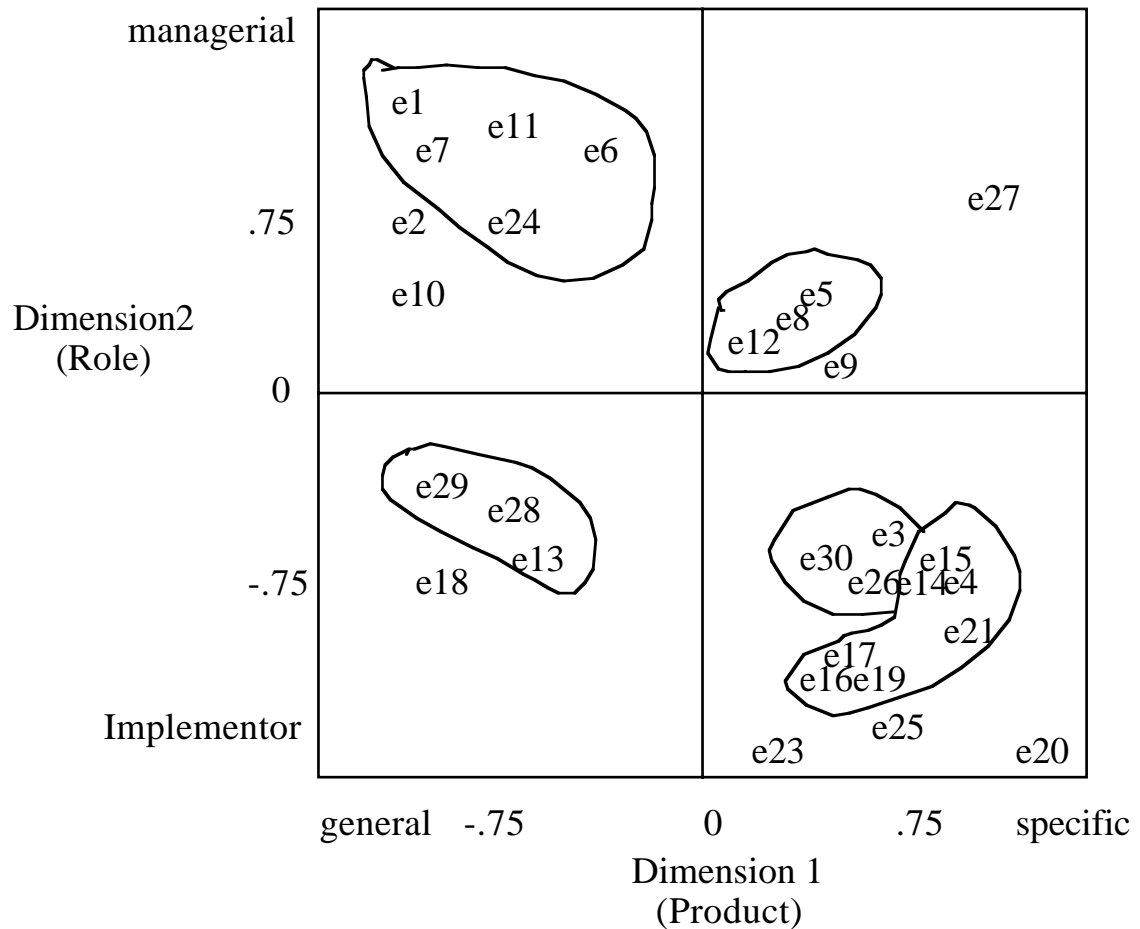


Figure 1. Two dimensional scaling solution for the episode similarity matrix ('e' refers to episode).

Discussion

It is notable that, of the 17 sorts generated (one designer was only prepared to undertake a single sort), 15 are captured by just four dimensions. What is interesting is how some of these dimensions, and more specifically the categories generated within them, corroborate the observations made by Bucciarelli (1988). Notably, 'relevant to specific group/role' corresponds with his notion of object worlds, and 'design process/stage' corresponds with his notions of specification and constraint. Thus, the sort study is able to verify the observations that emerge from this ethnographic study. In the limited space available, it is inappropriate to comment upon the clustering of cards in each category across designers and consistency of designer's sorts beyond a few broad observations. Designers were highly consistent in their category assignments under the dimensions of 'relevant to specific group/role' and 'product', but less so under dimensions such as 'episode content type' and 'design process/stage'. The role played by each designer in the software development team had a strong influence on their categorisation. For instance, programmers produced 'critical' evaluative categories (e.g., "trivial information"; "excellent program structure"), while managers produced open evaluative categories (e.g., "detailed decision-making"; "external influences"; "common building blocks"). Differences in sort performance between the various specialist roles within the design team, and the implication that the same design information is evaluated under different metrics across roles, corroborate our observations from the ethnographic study that design knowledge representations differ in individual and team contexts.

Acknowledgements

This research was supported by an ESRC Cognitive Engineering Programme grant (No: L127251027). We are grateful to our grant co-holder (John Mariani) and to Gary Spiers for their contributions and the company and design team members for their co-operation. A synopsis of the study appears as part of Ball and Ormerod (in press).

References

- Ball, L.J., & Ormerod, T.C. (in press) 'Applying ethnography in the analysis and support of expertise in engineering design'. *Design Studies*.
- Barsalou, L. W. (1985) 'Ideas, central tendency, and frequency of instantiation as determinants of graded structure in categories'. *Journal of Experimental Psychology: LM&C*, 11, pp.629-654.
- Bucciarelli, L. L. (1988). 'An ethnographic perspective on engineering design'. *Design Studies*, 9, pp. 159-168.
- Burton, A. M., Shadbolt, N. R., Rugg, G., & Hedgecock, A. P. (1990) 'The efficacy of knowledge acquisition techniques: A comparison across domains and levels of expertise'. *Knowledge Acquisition*, 2, pp167-178.
- Davies, S. P., Gilmore, D. J., & Green, T. R. G. (1995) 'Are objects that important? Effects of expertise and familiarity on classification of object-oriented code'. *Human-Computer Interaction*, 10, pp. 227-248.
- MacLean, A., Young, R. M., Bellotti, V. M. E., & Moran, T. P. (1991) 'Questions, options and criteria: Elements of design space analysis'. *Human Computer Interaction*, 6, pp. 201-250.

