DOING ETHNOGRAPHY AND EXPERIMENTS TOGETHER TO EXPLORE COLLABORATIVE PHOTOGRAPH HANDLING

Tom Ormerod, Nicola Morley, John Mariani
Lancaster University, LA14YF, UK
{t.ormerod;nicki.morley;j,mariani}@lancaster.ac.uk

Kevin Lewis, Graham Hitch
York University, YO10 5DD, UK
{k.lewis;g.hitch}@psych.york.ac.uk

James Mathrick, Tom Rodden
Nottingham University, NG8 1BB, UK
{jym: tar}@cs.notts.ac.uk

ABSTRACT

We describe the application of ethnography and experimentation to the specification and evaluation of a browser to support collaborative sharing of photographs in family contexts. In an extension of this mixed-method approach, we adopted an ethnographic stance to the post-hoc analysis of couples' interactions recorded during the experimental studies. The experiments and post-experiment analyses yielded usability enhancements and unexpected interaction patterns among couples in categorising and retrieving photographs that led to browser redesign.

Keywords

Experimentation, Ethnography, Digital photography.

1. INTRODUCTION

As interactive systems move out of the workplace and into the home and the wider environment, it is increasingly clear that no single methodological approach can deliver optimal design specifications or provide unequivocal evaluative data. Champions of qualitative approaches point to the poverty of contextual understanding, as well as the expense and invasiveness, of quantitative approaches such as experimentation. Proponents of quantitative methods point to problems inherent in qualitative methods, such as ethnography of subjective interpretation, and the lack of measures of frequency or criticality to systems use. The project reported in this paper was set up to challenge an apparent qualitative vs. quantitative dichotomy. Experimental psychologists, HCI researchers, systems developers and sociologists worked together to investigate collaborative remembering of digital photographs.

1.1 Ethnography and Systems Development

The use of ethnography in the design and evaluation of interactive systems is now common in HCI [1,2,3]. Ethnography is the rich, situated, longitudinal and immersive observation and account of activities in their place. Ethnography allows observations to be made of the real, complete and unmodified activities of an actual or potential user group in a non-invasive manner. As well as revealing regular occurrences, it captures exceptional or extreme behaviours and events that can provide exemplars of best and worst practice or trigger novel design ideas.

Ball & Ormerod [4] characterise ten features of a pure ethnographic stance that are necessary to fully understand and articulate a community of practice (see Figure 1). However, they suggest that ethnography for systems design must balance these feature against three other requirements: i) observations must be *verified* prior to design commitment, ii) ethnographers must be *purposive* in seeking to observe activities and events that surround the proposed point of technological intervention, and iii) observations must balance richness with *specificity* to meet timescales and budgets.

Figure 1. Features of a pure ethnographic stance [4].

1. Situatedness - Data are collected by a participant observer located within a community of practitioners.
2. Richness - The observer studies behaviour in all forms (conversations, documents, non-verbal interactions etc).
3. Autonomy - The observees need not comply with rigid, pre-determined study arrangements.
4. Openness - The observer remains open to unexpected issues that come to light as a study progresses.
5. Personalisation - The observer notes their own feelings in relation to situations encountered.
6. Reflexivity - The observer takes account of, rather than striving to eliminate, their effects upon observees.
7. Self-reflection - The observers interpretations are influenced by the tradition to which they belong.
8. Intensity - Observations are intensive and long-term, the observer immersed in the observee's environment.
9. Independence - The observer must not be constrained by pre-determined goal-set, mind-set or theory.
10. Historicism - The observer connects observations to a backdrop of historical and cultural contingencies.

1.2 Ethnography and Experimentation

The experimental method has long been part of HCI. Experiments offer a reductionist approach to understanding human-system interaction, controlling or holding constant the complexities of real-word environments to assess the influence of specific factors. However, limitations have been recognised, both by psychologists who seek ecological validity in research [5], and by HCI practitioners who seek practical and effective methods for evaluating human-system interfaces [6]. Indeed, the rise of ethnography as a systems design method reflects in part the

inadequacies of experimental research methods. Yet, the rejection of experiments in favour of qualitative approaches is not without cost. For example, hypothesis-driven experiments can reveal implicit influences on behaviour that individuals are not always able to report [7].

One remedy to this problem is to do both. A mixed-method approach might generate hypotheses from ethnography that are then tested in experiments. This approach has its place, but it has three potential limitations. First, it is expensive in time and participant/researcher involvement. Second, not all observations from ethnographic studies are amenable to experimental testing. Third, it limits the outcome of experimental methods to a confirmatory role. Another approach is to extend the notion of self-reflection (see Figure 1) to include reflection upon the use of experimental methods. In particular, we set out to conduct analyses of observations of participants undertaking experiments that investigated collaboration around a digital photograph browser. In terms of self-reflection, we were seeking to understand how the act of taking part in the experiments revealed interactions among participants that might inform our understanding of the experiments themselves, the future development of the browser, and the use of cognitive ethnography in further ethnographic studies of the situated use of a digital photograph browser.

2. DIGITAL PHOTOGRAPH HANDLING

Reduction in cost combined with technological advances has led digital photography rapidly becoming the medium of choice. The number of photographs held digitally by families has increased, and with it a need to store, organize and retrieve photographs from large databases. Commercial systems that help individuals manage their photograph collections (e.g, iPhoto, Adobe Photoshop) typically use browsing as an interaction style for exploring collections, with keyword annotation to guide the retrieval of targets and related images. Recent research, however, suggests that keyword systems lead to low levels of cataloguing and retrieval performance [8]. Exploratory technologies to support collaborative handling of photographs have emerged, though to date there is little empirical data. For example, the Personal Digital Historian [9] is a table-based environment around which users share and construct stories using images. Our aim was to explore how, within an *familial* context, people handle digital images, and how practices change because of technological interventions.

The handling of digital photographs presents an interesting test case for a mixed-methods approach to HCI. Photographs are inherently social artefacts: the reasons for taking pictures, the uses we put them to, and the ways in which we handle, store and reveal them reflect their contexts of use. To specify technologies without ethnographic observation risks missing social activities that surround image handling. Yet, processes invoked in categorising, remembering and recalling photographs also determine the success of digital photograph technologies. Like many aspects of human cognition, these processes are not easy to observe with methods like ethnography.

2.1 The TW3 browser

We developed a prototype browser (*The Way We Were,* or TW3) as an environment for empirical studies of partners managing collections of digital photographs. An initial ethnographic study of family interactions around collections of physical photographs informed the browser's initial design [10]. To the extent that family members categorized photographs at all, categories consisted of seemingly haphazard sets stored in different locations. Wedding pictures were in formal albums; pictures of children over the years were in less formal displays; photographs remained in developer's packets in a box, ongoing projects were in a bag under a sofa, and so on.

The ethnography prompted us to adopt a 'Who, What, Where and When' category scheme for the TW3 browser. It captures key episodic dimensions associated with the event portrayed by sets of images, which studies of autobiographical memory play a key role in remembering [11]. These dimensions are available within in existing photograph management software but typically as part of an unrestricted keyword category scheme. The TW3 browser limits users to the dimensions of who, what, where and when, allowing only six sub-categories under each dimension. The browser also imposes a procedure for cataloguing photographs: users sort photographs under one dimension then re-sort them under other dimensions in turn. This procedure aims to increase commonality in how family members organise collections, to reduce effects of *collaborative inhibition* found in psychological studies of shared recall [12]. Collaborative inhibition arises because individuals differ in their subjective organisation of material, so the cues used by one person to remember interfere with another person's ability to retrieve information. To retrieve a photograph, users click on a sub-category under each of the four dimensions (see Fig. 2). Should they fail to find a photograph, they drop one dimension at a time, increasing the set of images returned.

Figure 2. The retrieval interface of the TW3 browser.

We conducted an experiment comparing the TW3 browser with the commercial Adobe™ Jasc browser, in which individuals catalogued and then retrieved photographs of

the UK royal family [13]. When individuals retrieved photographs using their own categories, Jasc gave better first-attempt retrieval than TW3, though there was no difference in overall retrieval. When individuals retrieved using someone else's categories, the TW3 browser gave much better performance. Users of Jasc provided large numbers of detailed keywords in cataloguing photographs. If an attempt to retrieve failed, they tended to switch keywords or add a further keyword to the search. With TW3, users were restricted in the category labels they could provide, which increased the degree of commonality among participants' cataloguing schemes, and the dropping of dimensions proved a better approach to recovering lost photographs than switching or adding keywords.

3. EXPERIMENTS WITH THE TW3 BROWSER

A comparative evaluation demonstrated the feasibility of TW3 as a practical browser [13], but our real interest is in using the browser as a research tool to explore collaboration around digital collections. We conducted two experiments in which couples worked separately (Expt. 1) or together (Expt. 2) in cataloguing and retrieving from a database of photographs taken from their own collection. A full report of these studies is in preparation, but since Expt. 2 served as the basis for our meta-application of ethnography to experiments, we give a brief description.

Participants were heterosexual couples (12 in each experiment) who had been together a minimum of 2 years with an age range of 21-57. In Experiment 1, each partner completed four phases individually; selection of photographs, cataloguing photographs, retrieval using their own codes and retrieval using codes of their partner. In Experiment 2, the selection phase was the same; thereafter, couples worked together in the cataloguing phase and either together or alone in the retrieval phase (a third experiment in which couples catalogue separately and retrieve together, as a test of collaborative inhibition, is in progress). Intervals between selection, encoding and retrieval phases were 2 weeks, each phase lasting approximately 1-2 hours.

Prior to the experiment, each partner selected 100 photographs that they considered representative of their life or a period of their life. The experimenter put the photographs for each couple into the TW3 browser. In the encoding phase, participants were free to review thumbnails of all 200 photographs for up to 10 minutes. Then they sorted the photographs under each dimension, the order of dimensions counterbalanced. Participants moved a photograph from the pile to one of 6 locations by clicking the mouse arrow on that box. Participants were free to re-sort within each dimension and review photographs in each pile. They then labelled each of the 6 sub-categories before proceeding to the next dimension. Targets in the retrieval phase consisted of 40 photographs, 20 from each partner. The retrieval interface displayed the dimensions and subcategories for each. The target was presented for 5 seconds. Then participants selected sub-categories for the target and clicked on a Found Photo button. This brought up thumbnails of photographs that matched the selected subcategories. If the target was not found, participants dropped each dimension in turn.

Figure 3 summarises successful retrievals (using 4- and 3-dimension searches). The results of Experiment 1 show that participants more accurate using their own coding scheme than their partners, though this effect was restricted to females. By way of comparison, a group of control participants (12 single females with no shared knowledge of photographs) achieved 40% correct. Retrieval was higher in Experiment 2, showing an advantage of developing a category collaboratively on both shared and individual retrieval. This result is interesting in suggesting an effect of collaborative facilitation during cataloguing upon individual retrieval. In both experiments, participants made more errors with When and What dimensions (approx. 25%) than with Who and Where (approx. 10%).

Figure 3: % target photographs retrieved in Expts 1 and 2.

Expt. 1	75 (own codes)	64 (partners codes)
Exp.t 2	94 (retrieve together)	85 (retrieve alone)

4. POST-EXPERIMENT ANALYSIS

The experiments reported above reveal interesting and important data regarding differences in cataloguing and retrieval performance between individuals and couples. However, as we ran the first experiment it became clear that the experiments themselves were an interesting and potentially important object of empirical study. Even though participants were taking part individually, they referred continually to their partner's knowledge, disposition, personality, and so forth, to the extent that one researcher commented that the browser might offer a valuable marriage-guidance tool! Thus, in Experiment 2 we video-recorded couples undertaking each phase and their subsequent discussion with the experimenter, and adopted a modified ethnographic stance to analysing these video recordings. In analysing the five or so hours of video-tapes for each couple, we lost three of the features of pure ethnography described in Figure 1: situatedness, autonomy and intensity. However, we retained four features: openness, personalisation, reflexivity and self-reflection. The status of the remaining features was variable: for example, the data were rich in comparison with experimental data but their specificity and purposiveness limited their scope considerably. The ethnographer set out to maintain independence, but was informed about psychological theory on collaborative and individual remembering and also was likely (at least implicitly) to seek validating evidence for the browser.

To conduct the analysis, the researcher first transcribed the videotapes, and then coded segments according to themes of interest. The role of the coding was to allow estimates to be made of the frequencies with events and activities occurred. At the same time, she collected notes on her observations and thoughts regarding the couples' interactions around the TW3 browser. A key role for these notes was to document the things (unusual or exceptional events or behaviours) that coding the transcripts would ignore. Finally, she produced a written report documenting recommendations concerning browser usability enhancements and psycho-social aspects of couples'

interactions around the browser. For reasons of space, we focus here on interactions during cataloguing.

Two main usability issues arose from the analyses. First, the TW3 browser was severely limiting cataloguing by poor image contrast and lack of zoom facilities. Second, couples had difficulty with the post-sort labelling of categories: they made frequent reference to wanting to document the rationale for categories, and on a number of occasions were unable to reconstruct the reasons for category boundaries. We have added annotation features to a revised browser.

Couples adopted one of two kinds of cataloguing strategy: bottom-up in which category piles emerged without planning, and top-down in which after the thumbnail preview they plan category boundaries. This has led us to look again at the data to see if strategy predicts retrieval (currently in progress). Categorisation by exclusion also emerged as a common theme. For example, couple NC had a category "recent" and a category representing where they live "shireshead". During cataloguing, N identified that the categories overlap and reiterated to C that 'shireshead' does not include the last twelve months. Also common was the cataloguing of hard-to-categorise photographs relative to major events represented by other categories (e.g, a photograph location identified through being taken not long after graduation). One couple assigned a photograph by superficial similarity to images on the top of piles. After this observation, the researcher noted how discussions traded category consistency against accuracy.

Many observations reveal the participants acting on expectations of the experiment. For example, most couples indicated the need to have equal pile sizes. Although not a task requirement, it may be a product of the requirement to use six piles: moreover, it may make sense psychologically to aim for category boundaries that distribute photographs evenly across a dimension (again, a hypothesis requiring investigation). Couples were conscious of producing pile names that were meaningful to the researcher: often they catalogued in terms of life eras (pre-wedding, post-kids), but their final category names might be produced in years in order to "make more sense to the experimenter". The generation of hypotheses (possibly erroneous) about how people without shared knowledge of a collection cope with the category scheme is an avenue for future research.

5. CONCLUSIONS

The mixing of ethnography and experimentation has allowed us to collect and analyse a rich set of data, and to generate observations and design hypotheses that a single method would not allow. The concept of using one method to investigate the outputs of the other presents a form of meta-analysis that allows the immersive strengths of ethnography to apply to data sets collected under (relatively) controlled experimental conditions, and vice versa. We suggest that, far from being dichotomous methods, ethnography and experimentation are usefully practiced together in HCI and *on* each other.

ACKNOWLEDGEMENTS

REFERENCES

[1] Suchman, L. (1987). *Plans and situated actions: The problem of human-machine comm^n*. Cambridge: CUP.

[2] Viller, S., & Sommerville, I. (2000). Coherence… Ethnographic analyses in systems design. *IJHCS*.53.

[3] Hughes, J., et al. (1997). Ethnography, communication and support for design. C. Heath & P. Luff (Eds.), *Workplace Studies*.

[4] Ball, L. J., & Ormerod, T. C. (2000). Putting ethnography to work. *IJHCS*, 53, 147-168.

[5] Neisser, U. (1976). *Cognition and reality*. San Fransisco: Freeman.

[6] Jeffries, R. et al (1991). User interface evaluation in the real world *Proc. CHI*, 11-18, Seattle, ACM Press.

[7] Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84, 231-259.

[8] Rodden K., Wood, K. (2003) How do People Manage Their Digital Photographs? *CHI 2003,* 409- 416.

[9] Shen, C. et al (2002). Sharing and building digital group histories. *Proc CSCW*, New Orleans, 324-333.

[10] Crabtree A. et al (2002) Probing for Information. *Proc. 2nd EQUATOR Conference*.

[11] Wagenaar, W. A. (1986). My Memory: A study of autobiographical memory over six years. *Cognitive Psychology*, 18, 225-252

[12] Finlay, F. Hitch, G. & Meudell, P. (2000). Mutual Inhibition in collaborative recall. *J.Exp. Psy:LM&C,*.

[13] Ormerod, T.C. et al (2004). Mixing research methods in HCI. Proc. ECHI 04, Hamburg, Germany.