

## THE EVALUATION OF DESPERADO—A COMPUTERISED TOOL TO AID DESIGN REUSE

NICOLA. J. LAMBELL, LINDEN. J. BALL

*Institute of Behavioural Sciences, University of Derby, Mickleover, Derby, DE3 5GX, UK.*

AND

THOMAS. C. ORMEROD

*Department of Psychology, Lancaster University, Fylde College, Lancaster, LA1 4YF, UK.*

**Abstract.** The reuse of previous design knowledge when tackling a current problem is a potentially important way to improve design efficiency. However, unguided reuse is fraught with problems. Our current goal is to try and limit the problems associated with reuse by means of a computer-based support tool which guides the effective encoding of information for subsequent reuse. This paper presents an evaluation of such a tool that we have developed, which is known as Desperado. The evaluation involved observing academic designers working on a short design brief, either with or without Desperado. Our observational data are reported in conjunction with designers' subjective experiences of using Desperado. It was found that whilst the introduction of the system changed the nature of the users' design work, these changes were primarily associated with a better exploration and evaluation of possible design options.

### 1. Introduction

The reuse of previous design ideas is an attractive proposition. It can reduce development costs, maintain quality, improve productivity, avoid repetition of design effort and maintain upward product compatibility and consistency with company standards. The reuse of software code and specifications has been shown to benefit software engineers (Sutcliffe and Maiden, 1990) and work has been undertaken to produce appropriate tools to support this process (see Tracz, 1988, for an overview of past research, and Pena-Mora and Vadhavkar, 1997, as an example of more recent research). Supporting reuse should not, however, be limited to

software designing. Reuse is applicable in all areas of design and with the appropriate tools could actively enhance innovation in these areas. The notion of *innovative design reuse* sounds at first like an oxymoron. However, providing access to previous design options enables the maintenance of innovation over time. When designers know that ideas, if not implemented immediately, may be useful for future projects, they may explore them more creatively.

### 1.1 ISSUES IN DESIGN REUSE

In spite of the potential benefits associated with reuse, unguided reuse is fraught with problems which tend to reduce designers' willingness to reuse and the effectiveness of their actual reuse activity. Ball, Lambell, Ormerod, Slavin and Mariani (in press) present an in-depth discussion of such problems which we summarise here. First, designers appear to demonstrate difficulties in recognising potentially reusable information in previous designs (Woodfield, Embley and Scott, 1990), a problem which is exacerbated the further into the past the previous design recedes (Busby 1998). Sutcliffe and Maiden (1990) similarly found that inexperienced designers tended to reuse specifications based on surface features rather than deep-level analogies. This resulted in the erroneous transfer of domain knowledge.

Second, designers are inclined to underestimate the amount of modification required on those occasions when they do recognise relevant candidate designs (Busby, 1998). In addition, when modifying designs, because designers tend to rely heavily on their memory of previous information they are also likely to replicate the faults of the original design (Ball, Maskill and Ormerod, 1998; Busby, 1997, 1998). This problem is especially problematic given the absence of efficient indexing of previous design information.

Third, the invocation of previous design ideas may restrict the range of options that are explored by designers (Ormerod, Mariani, Ball and Lambell, 1999). Even expert designers engage in minimal solution-search behaviours and reveal severe *satisficing* tendencies – often becoming fixated upon single solutions rather than exploring alternatives in order to improve choices (see Ball, Maskill and Ormerod, 1998, for a review of evidence and explanations of such tendencies in the reuse of design solutions).

Finally, as Banker, Kauffman and Zweig (1993) have noted, at an organisational level design companies may provide little incentive for reusing previous designs, whilst at an individual level designers may be reluctant due to problems of quality assurance (Busby, 1998). However,

our own experience (Ormerod et al, 1999) suggests that many design companies and individual designers are currently becoming more proactive in their attempt to capitalise upon the potential benefits of design reuse.

Despite the problems which confront designers reusing past solutions, we believe that there is considerable potential for cost-effective design reuse in innovative design practice if such limiting factors can be eliminated or ameliorated. The task remains to develop a support system that addresses these problems.

## 1.2 USING DESIGN RATIONALE TO SUPPORT DESIGN REUSE

Our approach to supporting design reuse has been guided by the notion of 'design rationale' (DR). This is the documenting of the underlying reasoning behind the design of an artifact and describes the actual design process which culminated in a set of particular design features (see Moran and Carroll, 1996). The documenting of DR itself can lead to a better understanding of the issues involved, aid maintenance and redesign, improve critical reflection and reasoning, and enable more efficient communication between designers. Of primary importance to our aims is the notion that documenting of DR can aid reuse and limit some of the problems discussed previously.

Lee (1997) suggests that DR can serve as an index of past knowledge, thereby improving the ability of designers to spot relevant designs. Pena-Mora and Vadhavkar (1997) used DR to help index software code to promote reuse. DR itself is also potentially reusable. It would additionally improve the understanding of previous designs as it provides a common language of understanding, enabling designers to assess the implications of attempting to modify a previous design (McKerlie and Maclean, 1994) and providing a means of checking the integrity of previous designs to overcome quality assurance problems. Thus the current problems of reuse to do with understanding, adaptation, indexing and relevance could all be reduced by reusing DR and by using DR to index previous solutions. However, the benefits of DR for reuse depend heavily on the technique used to capture and represent DR.

Shipman and McCall (1997) distinguish between three different perspectives on DR: argumentation, documentation and communication. The Argumentation approach focuses on capturing the thinking itself in an orderly way. Documentation focuses on recording decisions and the actual thinking is fairly implicit. Communication captures information in the form of communications such as emails, faxes and design meeting notes and much of the DR is implicit within these communications.

Shipman and McCall argue that the communication approach is the most effective method of capturing all the DR involved in a design but it is difficult to index and therefore retrieve relevant DR. The argumentation approach, on the other hand, has more problems associated with effective capture of DR but is very useful for retrieval of relevant DR. Thus the argumentation approach would seem to be the most suitable for supporting reuse.

There are several different argumentation techniques, the two most widely considered being the Issue Based Information Systems (IBIS) approach developed by Rittel (e.g. Rittel, 1984; Rittel and Weber, 1983) and the Questions, Options and Criteria, or QOC approach pioneered by Maclean, Young, Bellotti and Moran (1991). Both of these representations provide structural components to enable the development of a graphical network of nodes and links which captures the design rationale.

Proponents of the issue-based approach tend to consider DR to be a record or a history of design deliberations. IBIS seeks to capture the key design issues that are articulated in the course of the conversation between stakeholders, along with the various positions that are raised in response to these issues and the pros and cons of these various positions. In contrast, Maclean et al do not see DR as being a record but a co-product of design which is constructed in addition to the artifact itself. QOC seeks to represent and explain the relationship between an artifact and its alternative design possibilities (Options) using Questions and Criteria, to guide the construction and evaluation of these design alternatives. Whilst both approaches would be suitable candidates for our computer support tool, the QOC approach was chosen for several reasons.

First, the QOC notation is more consistent with recent empirical observations that design activity, in both individual and cooperative contexts, seems to be naturally segmented in terms of questions and their associated options and criteria (Ball and Ormerod, in press). The IBIS approach, on the other hand, appears to suffer from the lack of an explicit criterion space (see Lee and Lai, 1991), and thus seems less reflective of natural design activity than the QOC formalism (see also Maclean et al., 1991). Second, the evaluation of options in relation to plausible alternatives is fundamental to QOC, and is especially useful for reuse as it specifies not only the reasoning behind an artifact but also offers a host of plausible alternatives which can be reused, thereby challenging satisficing behaviours, design fixation and confirmation biases (cf. McKerlie and Maclean, 1994). Third, the assessment of options against criteria is useful for design modifications since the criteria emphasise what was considered important when questions and options

were first examined. As such, the negative implications of a change in the design are readily apparent since certain criteria will no longer be satisfied. Questions also help to shape ideas better thus making it much easier to detect relevant rationales. Finally, it is noteworthy that proponents of the QOC approach emphasise its value for design reuse (e.g., McKerlie and Maclean, 1994), whereas advocates of IBIS-based models explicitly state that reusability is a secondary concern (e.g., Conklin and Burgess-Yakemovic, 1991).

Our approach to encoding DR is based on the notion of Questions, Options and Criteria but does not fully implement the complete philosophy associated with this approach. For example it does not support a hierarchical representation. The next section describes the system and the encoding of questions, options and criteria to aid effective reuse of design ideas.

## **2. System Description**

### 2.1 OVERVIEW OF THE SYSTEM

The development of Desperado involved phases of needs analysis, requirements specification, conceptual and detailed design, and implementation. The development was also informed by frequent discussions with end user companies and from ethnographic data collected at the beginning of this project (Ball and Ormerod, in press). Desperado is now in its second phase of development (see Ball et al. in press for a detailed discussion of the implementation of Desperado I and subsequent development of Desperado II). The current version of Desperado has a web-browser front end written in HTML and JavaScript to a relational database using Filemaker Pro<sup>TM</sup>.

Desperado is based around the notion of a 'design episode'. This is the unit of encoding which is the basis of the indexing of reusable information. An episode is focussed around the notion of Questions, Options and Criteria. It begins with the pursuit of a question and the subsequent encoding of relevant indices (such as user and client information) pertaining to that question. It then enables the exploration of options and criteria pertaining to that question and the storage of other relevant documents. Distinctions between the end of one episode and the beginning of a subsequent episode are determined by shifts from one question to another. Thus Desperado supports structured, goal-oriented shifts in activity which are thought to be the hallmark of design

expertise (Ball and Ormerod, 1995; Ball, Evans, Dennis and Ormerod, 1997) as well as allowing for more opportunistic shifts to emergent design issues (e.g. Sen, 1997) to facilitate crucial aspects of design flexibility.

Desperado was designed to provide an environment for the simultaneous encoding and reuse of design information during ongoing design work. Relevant episodes from the database are made available for retrieval during encoding. The system additionally provides pop-up menus with information from previous design episodes. These menus promote reuse of key terminology whilst also ensuring naming consistency and reducing the need for manual data entry.

The prompting of relevant previous episodes and prioritisation of information in the pop-up menus are performed by an interpreter. This interpreter uses prioritisation data to rank order previous episodes and labels. There are five sources of prioritisation data: time of episode, user selected defaults, key word matching, frequency of retrieval, and weighting derived during the process of rating options against criteria and weightings based on whether an option was adopted or not. The interpreter contains an algorithm which evaluates the value of the data from each design episode and then it presents the most appropriate episodes and labels at the top of the retrieval menu and pop-up menus.

## 2.2 GUIDED ENCODING AND RETRIEVAL

The system provides a sequenced procedural dialogue which elicits information from users. This method of information elicitation makes the process of encoding, reusing and designing interactive. The four phases are:

Phase 1: *Data-oriented encoding*. The first phase in the encoding of an episode involves specifying the design question to be pursued and the encoding of episode indices. These indices consist of information pertaining to: the user, the component function, the stage of design (e.g., requirements vs. conceptual vs. detailed design), the scope of the design activity (e.g., project-specific, organizational) and episode type (e.g., notes, meeting). During encoding the system provides pop-up menus containing previous episode labels and relevant episodes for reuse.

Phase 2: *QOC encoding*. This involves the elicitation of the options and criteria that relate to the question specified in Phase 1, and the subsequent cross-rating of options and criteria together with associated decision-making. Again, relevant episodes are made available in the

retrieval window. Additionally, users can access previous option and criteria names directly via the pop-up menus.

Phase 3: *Indexing of supplementary documentation*. This involves the indexing of relevant design documents such as CAD files and requirements specifications.

Phase 4: *Specification of 'consequents'*. This involves the specification of questions that emerge as a result of the current episode. The notion of consequents resembles the construct of dependencies (e.g., McCall, 1991) where answering a design question may depend upon how other questions are answered.

There are also five modes of system use: browsing, creating a new episode, editing an existing episode, searching, and retrieving. These distinctions prevent the erroneous loss of information from the database (you cannot edit an episode unless you are specifically in edit mode) and enable user-initiated reuse (via searching and retrieving of existing information) as well as system-initiated prompts.

## **2. The evaluation study**

### 3.1 AIMS

The purpose of this study was to evaluate the usability and effectiveness of the Desperado system (at an interface level and at the more fundamental level of encoding QOC during ongoing design). Direct observations of participants using Desperado whilst designing were complemented by data from a post experimental questionnaire. These evaluation techniques enabled objective and subjective insights into participants' use of the tool and how using Desperado affected participants' design behaviour.

### 3.2 DESIGN

For this study we developed three design briefs in consultation with managers at end-user companies. The first brief related to the design of an automated car-rental facility, the second to the design of an automated cloakroom, and the third to the design of an automated short-loan library facility. All three briefs shared analogous deep and surface level features - for example, they all involved automation of non-

automatic facilities, they all focussed on product conceptualisation, and all involved inputs and outputs to the system. The briefs were designed to be non-routine since Buckingham Shum, MacLean, Bellotti, and Hammond, (1997) found that participants had some difficulties implementing QOC analysis on routine design tasks. The briefs were, therefore, based on design concepts that participants would not be familiar with, but for which they would possess the necessary skills in their repertoire to tackle them effectively.

The experiment involved three sessions and each session was uniquely associated with one of the three design briefs. This arrangement enabled control of the information available to the designers for reuse during each session. In Session 1, all participants worked on the car-rental brief, using pen and paper only. In Session 2 (the cloakroom brief) the Control group used pen and paper only while the Desperado group used Desperado with encoding mode (i.e. without retrieval). In Session 3 (the short-loan library brief) the Control group used pen and paper only while the Desperado group used Desperado with encoding and retrieval. At this stage Desperado was seeded with design work from the previous session which was available for retrieval.

This experimental design allowed the effects of encoding to be partialled out from those of retrieval. It also allowed comparisons to be made within-participant (Session 1 vs. 2 vs. 3) and within design brief (Control vs. Desperado Sessions 1, 2 and 3). The inclusion of a control group also allowed us to consider the natural occurrence of reuse.

Eighteen MSc students from Lancaster University were paid £15 per session and 15 participants completed all three sessions (seven Control and eight Desperado). These students were all trained in mechanical and/or electronic engineering and had an average of six years of academic and company-based experience. All participants had attended an introductory course on the concepts of documentation and DR. After Session 1, participants were matched according to their verbalisation ability and the clarity of their transcripts. One person from the pair was then randomly assigned to the Desperado condition and the other remained a Control.

Design sessions were separated by intervals of about two months. During each session each participant spent approximately one hour developing a conceptual design to meet the brief. Participants made notes and sketches, and, where appropriate encoded and retrieved design episodes using Desperado. All participants were instructed to produce concurrent think-aloud verbalisations during the experiment and were told that we were interested in capturing the initial hour of their normal design process.

All of the participants in the Desperado condition received an overview and training exercise in their first encounter with Desperado and a top-up training exercise in their second encounter. The training and exercises were based on a 'carton-seal' design brief which they were familiar with from their Master's studies. The training was deliberately brief to allow a conservative evaluation of Desperado.

The data were analysed to examine the nature of reuse and the usability of the Desperado system. Additionally, participants transcripts were coded in terms of Questions, Options and Criteria. This allowed comparisons to be made of the nature of their design work, with and without Desperado.

### 3.2 DESIGN OF USABILITY QUESTIONNAIRE

Henderson, Smight, Podd, and Varela-Alvarez (1995) found that direct observation tended to fail to produce any useful insights into how to improve computer programs whereas questionnaire data lacked the ability to directly pinpoint the nature and location of possible problems with a system. Thus it was decided to combine the two techniques in order to maximise the informativeness of our usability data. This two-pronged approach to usability analysis was considered to be the most suitable for enabling a comparison of participants' behaviour and their subjective experience of the system and for providing clear insight into possible areas of improvement. Subjective experience was considered to be important because it was felt that demonstrating the system's usefulness was not sufficient in itself to ensure that people would use the system in the future. It was hypothesised that people would not use the system unless they explicitly felt that they would gain something from it.

The first two sections of the questionnaire required participants to rate statements about the Desperado system and their use of QOC analysis on a five-point Likert scale. Half of the statements were phrased positively and half were phrased negatively. Additional space was provided below each statement to enable justifications of their ratings and a number of open ended questions. The design was informed by Henderson et al's research which established that feedback would enable greater insight into participants' perceptions than just simple ratings. The statements employed and the questions asked were also informed by their questionnaire. These statements and questions are discussed individually in the results section. The questionnaire was given to participants at the end of Session 3.

## 4. Results

### 4.1 OBSERVATIONAL DATA

Full analysis of our extensive data-set is ongoing and the present analysis is therefore based on a subset of data (four participants from each condition). Initial analysis of the transcripts was centred around identifying specific episodes that were present (for the Desperado groups, these episodes were in the main directly retrieved from the databases for each participant). Episodes were then coded according to elements comprising the episode-specific question, options and criteria (in many instances, the questions in the Control group transcripts were inferred from the presence of option/criteria clusters). A composite hierarchy of these elements was produced for each transcript to provide an overview of participants' design work and to enable consideration of the design levels traversed. Finally, the 'design view' of the question and option elements of the episode were identified. For example, questions such as "How does the machine collect keys on return?" were coded as being functionally-oriented whereas questions such as "When should the keys be dispensed?" were considered to be procedurally-oriented. Table 1 presents the results of the episode and QOC analysis.

While caution should be exercised in the interpretation of the data in Table 1 owing to the relatively small number of designers being considered, it seems clear that Desperado is having an impact upon design activity. The effects can be interpreted both as negative and positive. The Control group are generating many more questions, and addressing the design hierarchy in apparently greater depth and breadth than the Desperado group. This suggests that the Desperado group achieved less coverage of the design brief. This is only to be expected: not only do the Desperado group carry the burden of interacting with a relatively unfamiliar system, but they are also required to undertake the additional requirement of documenting their design rationale (in the form of option and criterion encoding and rating). The latter point is particularly important, since over two thirds of interactions that participants made with Desperado were to rate options under criteria. In contrast, there was scant evidence in the protocols of the Control group for participants undertaking the systematic evaluation of options under all the criteria that they generated. It is also worth noting a decrease in time taken per question between sessions 2 and 3 for the Desperado group, which suggests a strong system learning effect.

TABLE 1. Mean Design Activity Scores

Session Characteristics	Control			Desperado		
	Session 1	Session 2	Session 3	Session 1	Session 2	Session 3
Design levels traversed	6.8 (0.96)	7.3 (0.58)	7 (1)	6.3 (0.58)	4.3 (0.58)	4.3 (0.58)
Basic design functions	11 (1.63)	9 (1)	10.7 (0.58)	10.7 (1.15)	3 (2)	4.7 (1.15)
Mins. taken per Question	1.12	1.3	0.88	0.76	13.86	6.67
Total/Unique Questions	53/33 (9.1/ 2.1)	51/30 (11.6/ 5.5)	70/40 (19.1/ 14.9)	88/55 (18.5/ 9.2)	4/4 (1.2/ 1.2)	9/9 (1.7/ 1.7)
Options per question (Total/Unique)	1.1/0.9	1.5/1.2	1.1/0.9	1.2/1.1	3.8/3.0	2.6/2.0
Criteria per question (Total/Unique)	0.6/0.5	0.8/0.6	0.5/0.4	0.6/0.5	3.5/2.8	3.2/1.4

**Note:** Standard deviations are presented in parenthesis. Design levels traversed provides a measure of the depth to which design work was carried out, while Basic design function provides a measure of breadth of key design aspects developed in each session. Total / unique scores for each question refer to the total number considered in each session, and the number of questions that were unique.

Positive evidence for Desperado comes from inspection of the mean number of options and criteria considered per question. The Desperado group considered up to three times as many options per question as the Control group, and up to six times as many criteria. This is important, since the pursuit of multiple options is precisely the kind of innovative design behaviour that Desperado was intended to encourage as a way of overcoming satisficing in design. We are also encouraged by the increased number of criteria considered for each question, since this suggests that Desperado is encouraging reflective design.

There are clearly some problems with Desperado II, that require attention in future versions. For example, it is apparent from Table 2 that there is a complete absence of any procedure-orientated QOC views. This suggests that the system is restricting the kinds of view that designers naturally adopt.

TABLE 2. QOC Characteristics

	Control			Desperado		
	Session 1	Session 2	Session 3	Session 1	Session 2	Session 3
% Function-oriented	35.8 (13.6)	31.9 (8.9)	41.0 (7.1)	36.4 (3.4)	29.6 (10.8)	41.0 (10.8)
% Structure-oriented	49.1 (16.3)	56.9 (4.2)	35.9 (12)	46.8 (10.6)	69.2 (10.4)	57.4 (9.3)
% Procedure-oriented	12.7 (10.9)	7.1 (7.4)	22.9 (13.9)	16.1 (9.8)	0	0
% Optimisation-oriented	2.4 (3.24)	3.7 (5.0)	0.2 (0.3)	0.7 (1.1)	1.2 (2.1)	1.6 (2.7)

**Note:** Standard deviations are presented in parenthesis. The QOC characteristic scores refer to the percentage of episode elements in each stage for which participants held each design view.

The data allow a number of other observations to be made. The performance of the Control group across the three sessions provides some evidence for the formation of spontaneous analogies to previous and existing designs. There was also evidence for the recognition of similarities across stages, but little evidence for reuse of functional design concepts and criteria. Examination of the protocols from the Desperado group show the frequent reuse of criteria in Session 3 generated in Session 2 that were made available through the prioritised pop-up menus.

Comparisons of Desperado Session 2 versus Session 3 show that there were few occasions where participants actively retrieved and browsed one of the previous episodes prompted by the prioritisation interpreter. Protocol analysis suggests two reasons for this. First, the relatively short time (six weeks, on average) between sessions 2 and 3 meant that participants recognised the episodes. Secondly, the pop-up menus of reuse class names and QOC elements allowed them to access relevant information without needing to retrieve whole episodes. In one sense this is disappointing, since by not retrieving whole episodes, participants were not accessing the detailed design rationale information (e.g., the option/criterion ratings). However, we are encouraged by evidence that there are reusable elements beyond solution options. We suspect that the failure to browse previous episodes may be an artifact of the small scale of the study. With a larger database, longer time intervals between

sessions, and episodes from multiple users, we believe that prompted episode retrieval will play a greater role.

#### 4.2 QUESTIONNAIRE DATA

Our descriptive analysis of the questionnaire data were based on the full data-set. Participants responses were assigned numerical values, whereby 1= strongly disagree and 5= strongly agree. All questions were converted to produce positive scores. The key outcomes from these analyses are presented below.

##### *4.2.1: Section 1 of Questionnaire—Ratings and Comments About QOC*

a) *QOC was easy to understand* (M=4.13, Sd=0.35). On the whole participants thought that QOC was very easy to understand and made little comment to support this rating

b) *QOC was expressive enough* (M=3.25, Sd=0.89). QOC was not quite expressive enough, some felt thinking in QOC terms interrupted their thought processes and forced oversimplification. However, others felt that the ability to add notes overcame this as it allowed them to expand on the QOC representation

c) *QOC helped understanding of design deliberations* (M=3.75, Sd=0.71). Generally QOC helped participants to make sense of their design deliberations. It helped them to clarify individual points and was flexible enough for them to adapt it to what they wanted. One person felt that it would be more useful to implement on a larger scale project

d) *QOC aided design work* (M=2.75, Sd= 1.16). Participants were generally undecided as to whether QOC aided their design work. Perhaps because some felt that it slowed them down so much. However, many of the participants felt that to some extent QOC provided them with a convenient break in which to think.

e) *Would use QOC when designing in the future* (M=3.88, 0.35). Participants felt that they would probably use QOC again or a similar design rationale approach. A few mentioned that they would use QOC again but possibly in their heads or on paper.

#### 4.2.2 Section 2 of Questionnaire—Ratings and Comments About the Desperado System in General

a) *Desperado aided their design work* (M=3.25, Sd= 0.89). They were undecided as to whether the system as a whole aided their design work. Some commented on the fact that it slowed them down so that they produced less. Others felt that it regulated their work and allowed them to keep track of their choices.

b) *Making previous design work accessible aided their design work* (M=4, Sd=0.58). Participants felt that making their previous design work accessible did aid their designing. They felt that it jogged their memory and prompted alternative solutions. One person, however, felt that it stopped them thinking.

c) *Easy to navigate through the system* (M=2.5, Sd=1.31). Participants did not think it was particularly easy to navigate through the system. They found that it was difficult to establish which mode they were working in.

d) *Would use Desperado in the future* (M=3.56, Sd= 0.53) . Participants tended to suggest that they would use Desperado again in the future. It was considered to have potential and be particularly useful when working as part of a team to provide structure and traceability. One participant felt that they would use Desperado after they had sketched their ideas out on paper as a means of documentation.

e) *The retrieval part of the system aided their design work* (M=4, Sd= 0.58). Participants thought that the ability to retrieve previous designs was advantageous. They felt it was useful to check similar previous design work and to browse through to recap on previous ideas. One participant felt the choice to browse or ignore it was worth commenting on, however, one participant voiced concern as to scaling up issues when there are lots of episodes in the database.

f) *The system was easy to use* (M=3.56, Sd= 0.5). Participants though it was fairly easy to use but voiced concerns over the awkwardness of the interface and how it improved once you got used to it. This is reflected in question (g) below.

g) *The system was easier to use the second time* (M=4.6, Sd= 0.52). Participants felt that the system was much easier to use once they became more familiar with it.

#### 4.2.3 Section 3 of the Questionnaire—Open Question Section on the System as a Whole

a) *What do you understand to be the purpose of the Desperado system.* References were made to aiding design, recording the design process, tracking current design, structuring design and supporting the derivation of criteria and functionality.

b) *What was most liked about the system.* References were made to the ability to reselect criteria and identify possibly missed options, being able to use ideas from previous projects, referencing similar past designs. One participant commented on the ability to format everything the same to allow comparison of ideas.

c) *What they most disliked about the system.* References were made to the slowness of the system, refilling of boxes from pop-up menus was thought to be tedious, the fact that the boxes provided required everything to be so brief, in particular in expanding on criteria. One participant suggested it did not allow functional decomposition.

d) *How could Desperado be improved.* The most common response was the introduction of a hierarchical facility which enabled participants to view the history of the project and to provide an overview of the questions considered throughout the whole project.

e) *Whether they thought that their design work differed when Desperado was introduced and if yes, to explain how if they could.* Four participants thought that the system changed the nature of their designing: some positively and some negatively. Some participants made reference to it slowing the whole process down whilst others commented on their work being more structured. One participant felt that he concentrated more on deciding between safe options than innovative solutions thus encouraging him to produce a safe design.

f) *Participants were asked to define what they thought was meant by the terms, Questions, Options, and Criteria.* All participants responses were rated by an independent rater who was not involved in the experiment proper. It was found that 96% of all responses were rated as being within the definitions of the terms as defined by Maclean et al. Thus participants demonstrated common understanding of QOC terminology.

g) *Any other comments.* One participant expressed the opinion that using the system was good fun, another added that he thought his usual design approach was quite random and Desperado helped to organise his work. One participant requested the inclusion of more examples during the learning phase of the experiment.

## 5. Conclusions

The observational data were encouraging. Despite showing less coverage of the design brief than control participants, participants using Desperado demonstrated a more comprehensive evaluation of criteria against options and a greater production of options per question. Spontaneous reuse of analogous designs was observed but there was little evidence that explicitly retrieving prompted episodes aided design. Instead, participants tended to reuse criteria and option names through the pop-up menus. The absence of explicit reuse of complete episodes is partially a product of seeding the database with each participant's own design work. Experiments are currently in progress to test Desperado using professional designers and a database of 50 episodes collected from our end-user companies.

The usability data suggested that people could think in QOC terms and provided a useful structure for their design deliberations. The difficulty seemed to be that the QOC formalism was not expressive enough to capture all aspects of our participants' design work. This is supported by the complete absence of any coding of procedural information in the database. The Desperado system as a whole was also viewed fairly favourably and despite a lack of participants using explicitly retrieved design information, participants felt that this information provided a useful memory aid. They also reported that reusing criteria and options from the pop-up menus was useful during ongoing design. The major concerns voiced were interface difficulties - especially issues of navigation and limitations of the current pop-up menus. Perhaps a more fundamental problem was the inability to provide participants with a hierarchical overview of their design deliberations. This hierarchical overview is associated with the traditional approach to QOC analysis and is an important omission in our system. These problems are currently being addressed.

It appears that using Desperado did change the nature of participants' designing, and this was explicitly reported in the questionnaire. On the whole these changes were fairly positive, suggesting that Desperado improved participants' reasoning about options and criteria. However, one major concern is the finding that using Desperado precluded a complete coverage of the design brief. We suggest that a number of factors must be borne in mind in assessing the negative impact of Desperado use. The first is that participants received minimal training in the use of Desperado. It can take many sessions for designers to gain sufficient familiarity with any system that it does not impact upon their design productivity, and this is undoubtedly the case with Desperado. For example, we suspect that there is a natural level of encoding detail that

participants will adopt when using the system, such that they become selective in the kinds of design episode that they encode. The evaluation study may have created a demand characteristic among participants to encode episodes continuously rather than selectively as one might expect in natural design environments. Second, there is clearly a trade-off between designing in breadth, as seen with control participants, and designing in depth, as seen with Desperado users. Clearly in a realistic context, there are situations where a rapid and superficial exploration of a design problem is valuable. However, reflective and evaluative design practice as seen with the Desperado group is also clearly advantageous. In the short term, the control group may have made more progress, but it is ephemeral. In the longer term, the presence of a detailed record of previous design work means that Desperado users have a broader set of design information with which to begin a project.

### Acknowledgements

This research was supported by an ESRC Cognitive Engineering Programme grant (L127251027). We are grateful to the managers and designers in our end user companies for their extensive assistance. We also thank John Mariani, Simon Slavin and Simon Lock for their assistance with the programming.

### References

- Ball, L.J., Evans, J.St.B.T., Dennis, I. and Ormerod, T.C.: 1997, Problem solving strategies and expertise in engineering design, *Thinking and Reasoning*, **3**, 247-270.
- Ball, L.J., Lambell, N.J., Ormerod, T.C., Slavin, S. and Mariani, J.: in press, Representing design rationale to support innovative design reuse: A minimalist approach, to appear in *Automation in Construction*.
- Ball, L.J., Maskill, L. and Ormerod, T.C.: 1998, Satisficing in engineering design: Causes, consequences and implications for design support, *Automation in Construction*, **7**, 213-227.
- Ball, L.J. and Ormerod, T.C.: 1995, Structured and opportunistic processing in design: A critical discussion, *International Journal of Human-Computer Studies*, **43**, 131-151.
- Ball, L. J. and Ormerod, T. C.: in press, Applying ethnography in the analysis and support of expertise in engineering design, to appear in *Design Studies*.
- Banker, R.D., Kauffman, R.J. and Zweig, D.: 1993, Repository evaluation of software reuse, *IEEE Transactions on Software Engineering*, **19**, 379-389.
- Buckingham Shum, S.J.B., MacLean, A. Bellotti, V.M.E. and Hammond, N.V.: 1997, Graphical argumentation and design cognition, *Human- Computer Interaction*, **12** (3), 267-300.
- Busby, J.S.: 1997, Why engineering designers don't learn from feedback, in the *Proceedings of the International Conference on Engineering Design (ICED)*, Tampere (August 18-22, 1997), 105-110.

- Busby, J.S.: 1998, Causal Explanations of the Absence of Reuse in Engineering Design Organisations, in the *Proceeding of the Engineering Design Conference, EDC'98*, (June 23-25, 1998), 475-482.
- Conklin, E.J. and Burgess-Yakemovic, K.C.: 1991, A process-oriented approach to design rationale, *Human-Computer Interaction*, **6**, 357-391. Also in T.P Moran and J.M. Carroll (Eds.), *Design rationale: Concepts, techniques and use*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Henderson, R.R., Smight, M.C., Podd, J. and Varela-Alvarez, H.: 1995, A comparison of the four prominent user-based methods for evaluating the usability of computer software, *Ergonomics*, **38**(10), 2030-2044.
- Lee, J.T.: 1997, Design rationale systems: Understanding the issues, *IEEE Expert Intelligent Systems and Their Applications*, **12**(3), 78-85.
- Lee, J. and Lai, K-Y.: 1991, What's in design rationale ? *Human-Computer Interaction*, **6**, 251-280. Also in T.P Moran and J.M. Carroll (Eds.), *Design rationale: Concepts, techniques and use*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Maclean, A., Young, R.M., Bellotti, V.M.E., and Moran, T.P.: 1991, Questions, options and criteria: Elements of design space analysis, *Human-Computer Interaction*, **6**, 201-250. Also in T.P Moran and J.M. Carroll (Eds.), *Design rationale: Concepts, techniques and use*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Maher, M. and Garza, A.: 1997, Case-based reasoning in design, *IEEE Expert*, **12**(2), Mar/Apr, 34-41.
- McCall, R.: 1991, PHI: A conceptual foundation for design hypermedia. *Design Studies*, **12**(1), 30-41.
- McKerlie, D. and MacLean, A.: 1994, Reasoning with design rationale: Practical experience with Design Space Analysis, *Design Studies*, **15**, 214-226.
- Moran, T.P and Carroll, J.M. (Eds.): 1996 *Design rationale: Concepts, techniques and use*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ormerod, T.C., Mariani, J.A., Ball, L.J. and Lambell, N.J.: 1999, Desperado: Three-in-one indexing for innovative design, in M.A. Sasse and C. Johnson (Eds.), *Proceedings of the Seventh IFIP Conference on Human-Computer Interaction—INTERACT '99*, London: IOS Press.
- Pena-Mora, F. and Vadhavkar, S.: 1997, Augmenting design patterns with design rationale, *AI-EDAM—Artificial Intelligence for Engineering Design Analysis and Manufacturing*, **11**(2), 93-108.
- Rittel, H.W.J.: 1984, Second generation design methods, in N. Cross (Ed.), *Developments in design methodology* (pp. 317-327).
- Rittel, H.W.J. and Weber, M.M.: 1973, Dilemmas in a general theory of planning, *Policy Sciences*, **4**, 155-169. (Reprinted in *Developments in design methodology*, 1984, pp.135-144).
- Sen, A.: 1997, The role of opportunism in the software design reuse process, *IEEE Transactions on Software Engineering*, **23**(7), 418-436.
- Shipman, F.M. and McCall, R.J.: 1997, Integrating different perspectives on design rationale: Supporting the emergence of design rationale from design communication, *AI-EDAM, Artificial Intelligence for Engineering Design Analysis and Manufacturing*, **11**(2), 141-154.
- Sutcliffe, A.G. and Maiden, N.A.M.: 1990, Software reusability; delivery productivity gains or short cuts, in D. Diaper, D. Gilmore, G. Cockton and B. Shakel (eds) *Proceedings Interact '90* (North- Holland, Amsterdam), 895-901.
- Tracz, W.: 1988, *Software Reuse: Emerging Technology*, IEEE Press, New York.
- Woodfield, S.N., Embley, D.W. and Scott, D.T.: 1987, Can programmers reuse software? *IEEE Software*, July, 52-60.