Chapter 7: VP-Lab

1. Introduction

The final piece of empirically-based work in this thesis returns to the theme of principles for evaluation and design. Some mainly qualitative evidence will be presented in support of the author's view that real-world design decisions can be informed by underlying principles such as examined in this thesis. This takes the form of a study of the changes made between the early prototype versions of a web site for experimental psychologists called VP-Lab.

VP-Lab (Virtual Psychology Laboratory) provides online access to an archive of psychological experiments and research material, via an Internet site (VP-Lab 1999). Users of the web site may retrieve bibliographic information and download the contents of research papers, the latter including data sets and experimental software. VP-Lab is a collaborative project involving the universities of York, Cardiff and Essex. The archive itself is maintained at the university of Essex; the interface and search engine were developed at the universities of Cardiff and York respectively. (The author is not a member of the VP-Lab team.)

There have been three main versions of VP-Lab, the first two (V1 and V2) being developmental prototypes. This Chapter focuses on the changes proposed and made between versions V1 and V2, and those which were proposed for version 3. The aim was to illustrate how interface design principles such as those used in Experiments 1 to 3 have been put into practice in an actual developmental process. Reference will be made to the results of evaluations performed on the interfaces to the first two versions of VP-Lab, and to the principles which the between-version proposals might be said to embody. Though a direct causal influence for the principles material proved to be impractical to demonstrate, the implicit use of such principles is offered as an indication of their potential application.

2. Versions 1 and 2 Evaluations

Evaluations of the first two prototypes were carried out by members of the VP-Lab team (and therefore did not involve the author). The evaluation of V1 took place during June 1997, that of V2 in January 1998. Subjects consisted of psychology researchers, students and other academic personnel with experience of bibliography searches. 10 subjects (5 Cardiff, 5 York) were used in the V1 evaluation, 16 subjects (5 Cardiff, 5 York, 6 Essex) in that for V2. In each evaluation subjects were asked to work through a set of tasks designed to emulate typical search scenarios. User-system interactions were logged and user protocols recorded, subjects being asked to 'think aloud' while working through the tasks. Experimenters also elicited further comments and reactions. Subjects were asked to complete a post-test SUS (Standard Usability Score) questionnaire (DEC 1986). See Appendixes K and L for the task sets and SUS questionnaire.

2.1 Version 1

The interface to V1 is shown in Figure 7.1. It consisted of one or more sets of three search terms (input fields), each set occupying one row. The default was two rows. Clicking on the 'refine search' button added a new row; rows subsequent to the first could be removed via a row-specific 'delete' button. The current search specification was submitted via the 'submit' button. The results of each search appeared on a new (web site) page.

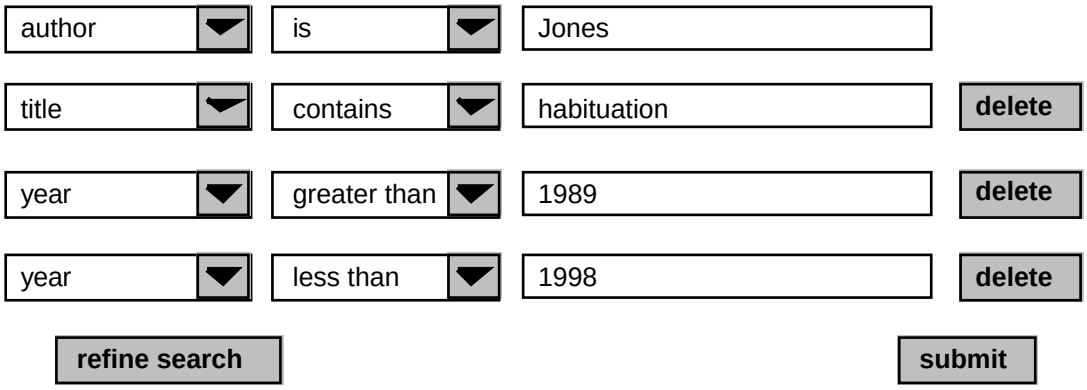

Figure 7.1. VP-Lab Version V1 search interface.

The contents of the first two fields in each row were selected from drop-down menus, the third requiring text input. The first field menu included {title, author, abstract, year}; the second included {is, contains, greater than, less than}. Search terms were ANDed together in a conjunction rather than a disjunction (thus retrieving only records matching all parts of the specification rather than all which match each part). Figure 7.1 thus would search the archive for contents matching the following:

{author is Jones} AND
{title contains habituation} AND
{year greater than 1989} AND
{year less than 1998}

thereby representing a search for papers whose author is Jones and whose title contains the word 'habituation' and which were published between 1990 and 1997 inclusive.

2.1.1 Version 1 Evaluation Results

Mean SUS ratings for version 1 were 67.75% (Standard Deviation 10.17, minimum 55.00, maximum 92.50).

Interface-specific changes which were proposed (by the VP-Lab team) as a result of the V1 evaluation included the following:

2.1 Add 'includes' to the second field of each row, so that (e.g.) author and title searches could be 'title includes xyz' and 'author includes zyx' rather than 'title is xyz' or 'author is zyx'.

2.2 Include additional operators (e.g. 'less than or equals', 'greater than or equals') in the second field, in order to extend the scope of searches by year.

2.3 Explicitly insert the label AND between each succeeding row.

2.4 Reproduce the search interface on each page of the results. This would avoid having to go back a page in order to submit a new or changed search.

2.5 Change the focus of search results from experiments to source papers. Thus rather than show experiments (from which source papers could be accessed), present papers with links to associated experiments. (Links from experiments back to source papers were to be retained).

2.6 Allow the user to undo a failed or unwanted search (and thereby return to previous search results). This was implemented in V2 as a 'previous search' button.

2.7 Include a 'clear all' button (enabling all fields to be cleared), but retain the 'delete' buttons (enabling individual rows to be removed).

2.8 Move the 'submit' button from bottom right to top right.

2.2 Version 2

The interface to V2 is shown in figure 7.2. It features the changes listed in the previous Section. It also introduced a 'back again' (redo) button, so that after use of 'previous search' users could retrieve the current search results without having to re-submit. ('Previous search' and 'back again' were thus complementary, acting as Back and Forward navigational moves.)

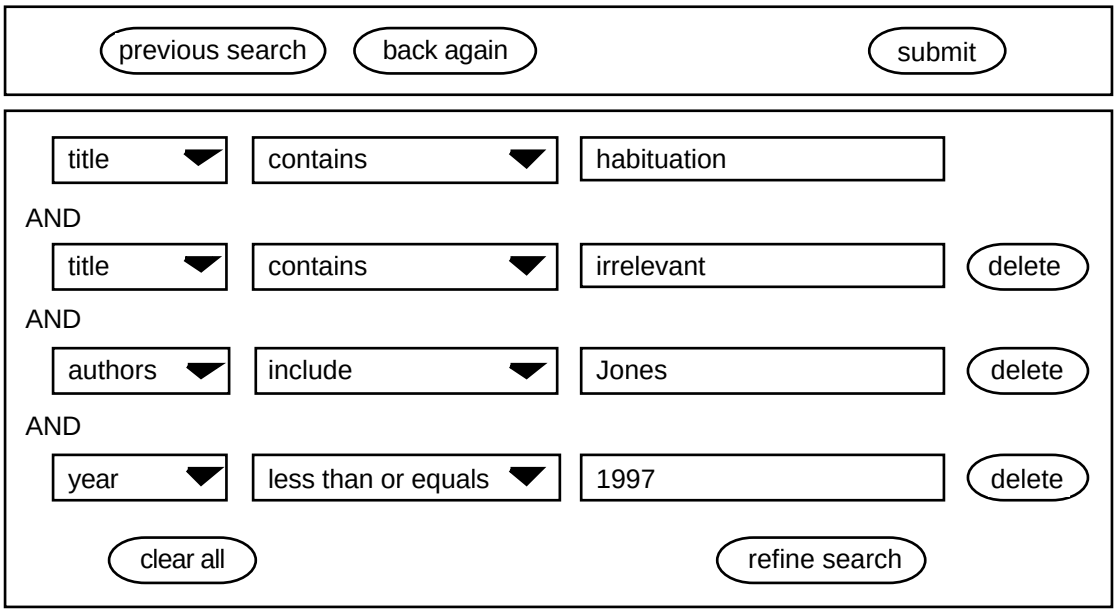

Figure 7.2. VP-Lab Version V2 search interface.

The first field menu now included {title, authors, abstract, topic, year}; the second field included {is, contains, include, greater than, less than, less than or equals, greater than or equals}. Figure 7.2 would thus search for the following:

{title contains habituation} AND
{title contains irrelevant} AND
{authors include Jones} AND
{year less than or equals 1997}

thus representing a search for papers by author(s) Jones and any other(s) whose title includes both the words 'habituation' and 'irrelevant' and which were published in or before 1997.

2.2.1 Version 2 Evaluation Results

Mean SUS ratings for version 2 were 65.78% (Standard Deviation 10.50, minimum 55.00, maximum 80.00).

Interface-specific changes considered for version 3 included the following:

3.1 Relocate the 'refine search' and 'submit' buttons (move 'submit' back to bottom right).

3.2 Remove the 'clear all' button (the row-specific 'delete' buttons were considered to be sufficient).

3.3 Explore alternatives for 'previous search', 'back again' navigational combinations (e.g. retaining only the former button, introducing a 'new search' button).

3.4 Rename the 'previous search' button (several possibilities considered).

3.5 Rename the 'submit' button (to 'Ok').

3.6 Rename the 'refine search' button (several possibilities considered).

3.7 Make the contents of the drop-down menu for the second field of each row contingent on the first field selection. For example, if the first selection is 'title', the second would offer only 'contains', 'is', etc.; if the first is 'authors', the second would offer only 'include'.

3.8 When the first field selection is 'topic', make the third field offer cascading (drop-down) menu items in line with the archive topic organisation.

3.9 Make it clear to the user that the word 'and' should not be included in the (third) text field, in order to AND together separate fields. For example, rather than {title contains 'habituation and irrelevant'}, use {title contains 'habituation'} AND {title contains 'irrelevant'}. (However, e.g. {title contains 'irrelevant sound effect'} would be acceptable.)

3.10 Put navigation links ('previous search', back again', 'new search') at the bottom of each page of the search results (in addition to those elsewhere).

3. Principles Usage

It proved to be possible to make at least partial (though post-hoc) matches between most of the changes made following the two evaluations and the evaluation materials featured in this thesis. This was explored in interviews with members of the VP-Lab project team in April 1998 (three months after the V2 evaluation), using the three-page (short) version of the principles set used in Experiment 1. (See Appendix B for a transcript.) Some matches with Nielsen's ten heuristics (Molich & Nielsen 1990), listed in Appendix A, were also possible. The ratio of the former to the latter is 11:7 (counting matches made more than once) or 9:5 (counting more than once only).

In the following discussion, matches with the principles and heuristics appear after each summary. Post-evaluation proposals which were implemented are shown in (brackets). Between-version changes can be grouped into five categories.

Navigation

Reduction in the number of movements between search inputs and results were achieved by reproducing the searcher interface (2.4) and navigation links (such as 'previous search' and 'new search') (3.10) on each results page. Allowing direct return to the previous results (2.6) enabled users to undo a failed or incomplete search. Consideration was also given to the naming of navigation buttons (3.4, 3.6) and the alternatives for backward navigation (3.3). The change of focus in the results from experiments to papers (2.5) facilitated the retrieval of experimental source material.

Principles: Navigational Effort, Location and Navigation
Heuristics: Provide Clearly Marked Exits

Mental Models

Attempt was made to design for the user's anticipated mental model rather than to force users to create complex interaction models (e.g. of navigation between results and search specifications) (interview with Will Stevenson, April 1988). To this end, it was considered preferable to support only a few between-state moves, such as 'previous search', 'new search' (2.6, 3.3), and to consider dropping the 'back again' button (3.3). The inclusion of the word AND between each row of the search terms (2.3), and the proposed clarification that 'and' should not be used in the third text field (3.9) are examples of making explicit the logical ANDing of search terms.

Principles: Navigational Effort, Memory load, Error Management
Heuristics: Simple and Natural Dialogue, Error Prevention

Consistency

It was considered important to maintain internal consistency by keeping the format of search results the same (2.4, 2.5), and to foster external consistency with other search engines, so that users would be able to anticipate the results of any particular combination of search parameters based on prior experience (but nevertheless to utilise a relatively simple user model of interaction behaviour: interview with Will Stevenson). External consistency was also to be assisted by changing the name of the 'submit' button to 'Ok' (3.5) and relocating this button at bottom right (3.1).

Principles: Consistency
Heuristics: Be Consistent

Accuracy

Improvements to the likely match between search inputs and results, and thus to the accuracy of results outputs, were to be made via changes to the field contents (2.1, 3.7, 3.8), the range of operators (2.2) and the clarification of the AND search model (2.3, 3.9).

Principles: Accuracy of Content, Appropriateness of Content
Heuristics: none

Redundancy

The improvements to field and operator contents (2.1, 2.2) and menu contingency (3.7, 3.8) were intended to reduce redundancy in matches between search inputs and archive contents. (This can also be considered as a form of error prevention.) Redundancy was also to be reduced by allowing users to delete unwanted entries (2.7, 3.2).

Principles: Error management, Choice Availability
Heuristics: Simple and Natural Dialogue, Error Prevention

Terminology

Changes proposed for button names (labels) included those of 'refine search' (3.4), 'submit (3.5)' and 'previous search' (3.6), as well as to input fields such as 'authors include' (2.1).

Principles: System - User Match
Heuristics: Speak the User's Language

4. Summary of Results

1. The changes made and considered following evaluations of version 1 (V1) and version 2 (V2) of the VP-Lab search interface were summarised and illustrated.

2. The changes were matched, where possible, to the contents of both the (short) principles set used in Experiment 1 and Nielsen's ten heuristics. There were found to be more candidate matches with the principles set than with the heuristics.

5. Discussion

The matches between the VP-Lab changes and the principles set are offered as illustration of the way in which at least implicit use can be made of such principles (and heuristics) in an iterative development process. All of the changes considered following the evaluations of versions 1 and 2 of the VP-Lab interface could be related to at least one of the principles (in the form of the short version used in Experiment 1). The fact that fewer matches were made with the heuristics (and the 11 : 7 ratio of principles : heuristics matches) may be indicative of the limited nature of the heuristics relative to the principles.

This Chapter represents mainly qualitative support for the view that evaluative principles such as presented in this thesis have a role to play in interface design. Though the iterative process which the VP-Lab interface went through does not demonstrate a direct causal influence for the principles set (and the lack of improvement in the SUS scores for version 2 were disappointing), it is nevertheless suggested that at least some of the principles were embodied in the between-version changes. The decisions made by the design team can be seen as an implicit representation of what has been proposed in this thesis, namely, that there exist general evaluative principles which can be applied to a wide range of interface domains, now including internet sites and search engines.

Perhaps unsurprisingly, few attempts have been made to establish a direct causal role for design or evaluative principles in product development. In order to do so, it would be necessary to demonstrate a difference between the outcomes of design processes both with and without principle-based input (as distinct from measuring improvements during an iterative design process). Only two examples are known to the author: first, the series of experiments reported by Andriole & Adelman (1995) and Gerhardt-Powals (1996), in which comparisons were made between versions of a prototype submarine warfare system engineered both with and without 'cognitive-design' principles; second, the comparison by Vaughan (1998) of different implementations of a HyperCard diary interface, designed to manipulate the requirements of two user-centred design principles. The case study reported by John & Marks (1997), which tracked the effectiveness of changes made to an authoring tool as a result of problems predicted by five UEMs, is unusual in comparing versions with and without design changes. (Of these three experiments, only the first

reported unequivocal support for such input.) The original intention of this thesis had been to demonstrate that iterative development principles such as those proposed by Gould & Lewis (1983, 1985) do, in fact, produce measurable improvements over interfaces produced without such principles. The recent widespread adoption of user-centred design and usability testing may mean that this has become un-necessary, but nevertheless, empirical demonstration of this sort appears conspicuous by its absence.

It may seem relatively simple to demonstrate some results for principles over heuristics, given the nature of the exercise carried out for VP-Lab. However, this thesis shows that a principles-heuristics effect turns out to be more elusive than first envisaged. Why this should be so will be taken up in full in the next and final Chapter.

Summary of Chapter 7

This Chapter presents some evidence for the implicit use of the usability principles such as are embodied in the principles set used in this thesis, in the form of the changes which followed two evaluations of an internet web site providing access to psychological experiments and research material.