Adaptation of Method-resources Between Projects: A Case Study From a Dynamic and Complex Work Domain

Jessica Wardlaw
University College London
Gower Street, London, WC1E 6BT, UK
j.wardlaw@ucl.ac.uk

Anna L Cox
University College London
Gower Street, London, WC1E 6BT, UK
anna.cox@ucl.ac.uk

Muki Haklay
University College London
Gower Street, London, WC1E 6BT, UK
m.haklay@ucl.ac.uk

ABSTRACT
In this case study we describe how method-resources were reconfigured across three design and evaluation projects conducted by an in-house design team within the same company during a six-year action research collaboration with academics from the field of Human-Computer Interaction (HCI). This case study specifically focuses on the reconfigurations that occurred in participant recruitment, task selection, reporting format and problem identification between the three projects. The underlying contextual factors behind the reconfigurations, in particular the application domain, organisational factors and project constraints, will be discussed to give unique insights into the realities of design work from within a single organisation over the six-year collaboration. This case study demonstrates the complexity of comparing methods across projects, particularly within dynamic and complex work domains, and that existing attempts may be too simplistic because they fail to account for these factors.

Author Keywords
Action research; design and evaluation methods; resources.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Design; Human Factors.

'OLD' METHOD-RESOURCES IN NEW CONTEXTS

In this case study we describe how the same method-resources were applied differently across three projects conducted by an in-house design team within the same company during an action research (AR) collaboration. AR encompasses methods and approaches for collaborative research with partners towards addressing problems they experience; this is done through cycles of planning, action and reflection, which offer HCI the opportunity to address gaps between theory and practice [4]. It is interesting to compare the projects since the business environment significantly changed during the collaboration, imposing very different constraints on the projects despite their strategic importance to the company. The resources that this case study will focus on are: participant recruitment (finding the right type and number of participants), task selection (specifying tasks for inspection or user testing), reporting format (communicating problems and solutions for subsequent analysis, evaluation auditing, iteration and customer communication) and problem identification (tools and approaches for identifying/discovering problems) [16].

Dr Foster Intelligence (DFI) is a public-private partnership in the United Kingdom (UK) health informatics sector that provides independent health and social care information to healthcare managers and clinicians for the improvement of clinical effectiveness and efficiency. DFI was formed in 2006 as a partnership between the National Health Service (NHS) Information Centre and Dr Foster Ltd. DFI has produced a range of web-based data analysis tools which give NHS managers access to the Hospital Episodes Statistics (HES) database that contains admitted patient care data from 1989 onwards and outpatient attendance data from 2003 onwards. Whilst live access to a database of 825 million hospital records presents many challenges, particularly with users that vary greatly in requirements and geographic location, it also presents great opportunities that are unavailable in any other country's health system.

In 2010, however, after a change in Government, the Department of Health bought the NHS Information Centre's shareholding and announced a strategic review of the future of DFI [9]. Following this review, and the Government's Spending Review, urging Departments to maximise value from assets that do not need to be held in the public sector, it was announced that DFI would be marketed for sale [11]. This sale is still being negotiated.

The company therefore faces many challenges to maintain their position as a leading provider of health informatics in the UK. Since the company began, the market has become more competitive and many trusts will develop internal solutions to save money. Financial constraints have resulted in customers having more complex and changing needs and demanding more choice. The usability of health informatics tools, and how well they meet users' requirements, is thus an increasingly important factor when health organisations are deciding whether or not they will invest in them.

Project 1: Obtaining user requirements for and evaluating Population Health Manager

Primary healthcare services in England (e.g. doctors, dentists, opticians and pharmacists) are managed by local Primary Care Trusts (PCTs). PCTs control 80% of the NHS

budget [2] to determine and provide the health services local communities need, including hospitals. Many PCTs use DFI's Population Health Manager (PHM) tool for this work, which provides PCTs with the information to:

- Understand the local population and develop segmentation models of their health needs
- Identify and analyze local health inequalities to target unmet needs or gaps in care
- Monitor admission trends, forecast population health needs and predict future health trends.

PHM offers various datasets, which are regularly updated, along with the facility for users to upload their own local datasets. Maps of PHM data can identify spatial inequalities in the provision of health outcomes to inform the location of appropriate services and interventions, in addition to understanding patient referral patterns (Figure 1).

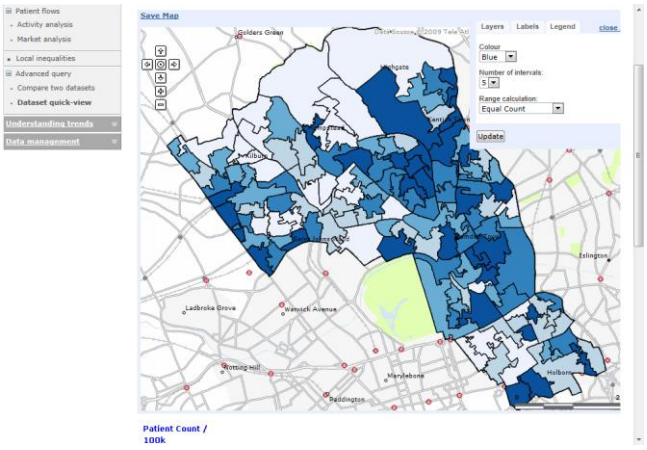

Figure 1. The mapping interface for PHM

During the PHM project, an online survey was used to establish potential users' preferences for the cartographic presentation of the data and thus inform the design of a new mapping interface. Questions presented two or three maps covering various cartographic aspects of the maps including data classification, number of ranges, colour schemes, the representation of point data, raster or vector data for the background map and mapping multiple datasets. For each aspect participants were asked to answer a question about the data that required interpretation of the map and to identify which map enabled them to answer it more easily and which option they preferred. Survey results were then incorporated into the software design. The final interface was then inspected using Heuristic Evaluation [5] and Cognitive Walkthrough [8].

Project 2: Developing company Personas

Following the PHM project a need was recognised to improve the developers' understanding of the end users of DFI products, since they did not have regular opportunities to meet the end users to understand why the tools were used and the development team personnel was often changing. To achieve this, a specific project was initiated to create personas of the key DFI users in which a variety of methods were combined according to the resources that were available. Fourteen semi-structured interviews with key DFI users were supplemented with information from database server log files (reflecting usage of the entire user population) to confirm which job titles represented the most frequent users and which parts of the tools were used the most. In addition, a user-generated screenshot survey required users to take a screenshot of their entire computer screen whilst using a DFI tool which revealed rich information on users' working environments and taskflow.

These methods were specifically chosen due to the wide geographical distribution of users and the relevance and richness of the information they could provide with limited resources. These personas were used, in part, to inform the redesign of DFI's flagship product, which was to be relaunched under the name Quality Investigator (QI). For this product user testing was also carried out.

Project 3: User Testing of Quality Investigator (QI)

QI is a web-based tool that monitors quality outcomes and patient safety by assessing clinical, process and coding factors. This was developed three years after PHM. Its user interface comprises tabs for Mortality, Length of Stay and Readmissions, all key indicators of clinical quality and efficiency (Figure 2). A dashboard highlights a hospital's 'CUSUM alerts' for diagnosis and procedure groups; negative CUSUM alerts (indicated by red bells) are given when indicators diverge sufficiently from expectations to suggest a systematic problem. 'Relative Risk' also provides the observed cases as a percentage of the risk-adjusted expected (reflecting case mix and national average). This permits analysis of patients by diagnosis or procedure group and comparison of clinical performance. The five diagnosis and procedure groups with the highest 'observed' (number of cases within the selected dataset) exceeding 'expected' (expected cases given the case mix) and crude rate (observed cases as a percentage of volume) are also shown.

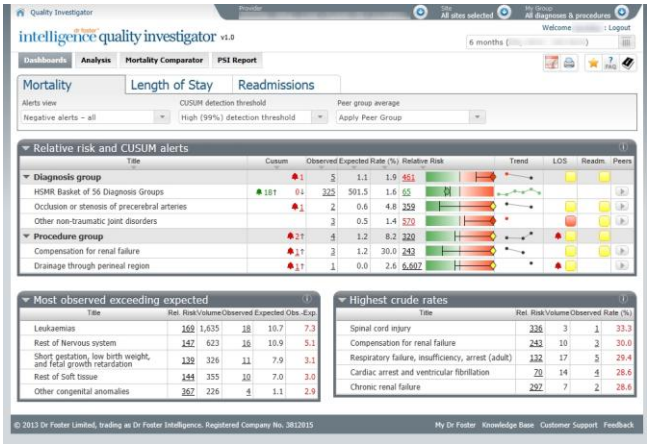

Figure 2. Quality Investigator dashboard

Before these studies' resources are compared, Table 1 presents the acronyms introduced thus far for reference.

Acronym	Full Term
DFI	Dr Foster Intelligence
PHM	Population Health Manager
QI	Quality Investigator

Table 1. Acronyms.

Stories of Transfer: Triumph or Tragedy or Both?

In this paper we examine the reusable HCI method-resources across the projects described to identify where they have proven to remain useful and compatible or require substantial modification when applied across the usage contexts of the projects.

Participant recruitment:

Each of the projects required recruitment of participants.

At the time of the PHM project, there were very few resources for usability work and usability work had little status within the organisation. Furthermore, it was a new product in a new market for DFI, so there was only a limited network from which to recruit participants. Survey participants were therefore recruited by emailing users that were suggested to the researcher by colleagues.

Recruitment of key users for interviews to inform the personas was negotiated and managed with the Customer Service Managers, who regularly meet with users, so that they could approach any potential participants initially.

In contrast, participant recruitment for the QI project was assisted. This was in recognition, by the organisation, that users must be engaged in the design process because contracts were at risk of non-renewal.

It can be concluded that the facilitation and effectiveness of participant recruitment has improved across the three projects. Whilst Amazon vouchers were offered for participation in the maps and user-generated screenshot surveys, and the suitability of some interviewees for the personas can be questioned because of their very limited experience of the tools, participants in the user testing were motivated to take part by the opportunity to be involved in the development and direction of the new tool and to try it ahead of its launch. The low number of users who tested QI, however, may have impacted upon the reliability of the results given the complexity and breadth of users of the application domain [15].

Task selection:

Two of the projects required the creation of tasks to evaluate the tools being developed. Data collected to create personas also included identification of real users' tasks.

Tasks were not well defined for the PHM project because it was a new product in a largely untested market for DFI. In order to design a task for the Cognitive Walkthrough it was necessary to look at job descriptions of the target audience, in an attempt to understand the type of work that they did. A DFI colleague who supported a particular local health organisation to carry out tasks similar to those PHM was designed to support was also available to consult on the types of reports that his client wrote.

To create personas, we used Contextual Inquiry [14] consisting of semi-structured interviews and unstructured observation sessions during which participants demonstrated a task that they commonly performed using a DFI tool. This produced a range of closed and open tasks so arguably provided a much more realistic picture of the tasks users aim to accomplish with the tools.

For the user testing of QI, tasks were designed based on information provided by customers as to the tasks they would like to be able to perform with the new tool, which included new functionality, to which the project team gave priority and improvements on existing functionality. Some tasks could not be completed during the initial user tests because of the development stage of the tool; however these tasks were implemented for subsequent tests in addition to some quick fixes to issues that were identified during the initial tests. There was also a second phase of user testing in order to test a more completed version of the tool.

The ability to select realistic and appropriate tasks to carry out the methods has gone from tragedy to triumph during the three projects through a growing understanding of the application domain and users' requirements. However, there has been no opportunity to reuse the tasks generated as each project has focused on a different tool that supports a different part of the users' work.

Problem identification:

Each project resulted in the identification of numerous usability problems, however they varied dramatically in terms of their success with this. The Heuristic Evaluation of PHM exploited a structured report of 296 heuristics available online [5] loosely grouped according to Nielsen's ten usability heuristics [7]. Each usability problem identified was then assigned a severity rating according to the classification in [10]. Similarly the Cognitive Walkthrough followed the format outlined by [8], with the addition of a fifth evaluation question that asked what the system provided beyond the normal method by which users would carry out the task. These methods together identified 32 usability problems; 12 rated as irritants, ten rated as moderate, nine rated severe and one as unusable.

To create personas, interviews were conducted in which users revealed their frustrations with existing tools. Additionally, participants were asked to demonstrate a typical task they perform on the website, which revealed some additional problems; this was recorded using video

capture software. Although activities conducted to create personas are not necessarily designed to elicit usability problems with existing software, this was a serendipitous outcome of this project. This demonstrates extension of the textbook scope of Contextual Inquiry and reproduces some of the view developed by [6].

The QI methods included a user study, which identified a large number of problems. It was a relatively easy job to prioritise the problems: for example, feedback on results was requested for a Monday morning scrum meeting and results were quickly compiled after Friday's final user test. Standard usability evaluation methods are known to vary in terms of the number and severity of problems identified. The expertise of the evaluator is also known to influence these outcomes. It is likely that all of these factors impacted our projects. An unexpected triumph was that interviews conducted for the purpose of persona generation also revealed usability problems that could be addressed by DFI.

Reporting format:

The reporting format also varied between the projects.

The personas were first compiled into a PDF file comprising of a page for each persona that included details of their goals, working environment, typical behaviour, attitudes and skills, in addition to a photo and some personal information to bring the personas to life. These were then printed as A3 posters for the walls of the office. Subsequently a more detailed report of the main themes that arose in the interviews was written to extend the scope of information conveyed by the personas; this comprised of a summary of tool usage patterns, who the users are (their job titles, roles and responsibilities and main motivations for the tools) in addition to problems highlighted with the data in the interviews (transparency of data source, data quality, timeliness of data, unclassified data, data complexity, analysis and information presentation) and recommendations that interviewees gave for improving the tool functionality. The results were communicated to the rest of the team through a presentation of the personas and summary of the report (with a focus on the issues that interviewees identified with using the tools) at the Product Development team monthly meeting; both the personas and report were shared with the team after this meeting as well as the interview transcripts, so that developers could develop an understanding of the language used by the users.

In the new business context of QI, since the user testing was formally part of the development process a much more concise report was provided. In addition to a written report, highlights of the user testing sessions were communicated at a project meeting through a presentation that included a summary video. This video was designed to show examples of both unsuccessful and successful task completion so as not to discourage the developers.

The reporting format has largely been configured according to its audience and therefore been successful in the transfer of redesign proposals through to development. For PHM the audience was primarily the designers working with the developers, for the personas the Product Development team, and the results of the user testing were communicated to the project leads before being prioritised and put through to development at the weekly scrums. However, there was no opportunity to reuse existing formats for reporting. Development of a standard reporting format to be used within DFI would facilitate reporting in future projects.

ANALYSIS OF CONTEXTUAL FACTORS

Over the time course of the three projects a number of key contextual factors changed. We now discuss their influence on method-resources and on the success of applying the design and evaluation methods in practice. To do this we use the classes of resources defined in [16].

Axiological resource types

Axiological resources refer to the values that motivate an approach, for example clients' needs and expectations from a method and corporate culture and values [16]. Across the project presented, the action research approach taken necessitates consideration of the clients of DFI (i.e. the end users of DFI's systems) and the client within the collaboration (i.e. DFI); both relationships require careful management of expectations at the individual level and clients will place different values on the methods [3]. This was evidenced in the participant recruitment and test protocol for the user testing of QI; DFI were keen not to give the participants the impression that any suggestions they made for improvement would be implemented. In addition, the nature of the collaboration requires that there is a mutual understanding that any one method would not provide the 'silver bullet' that DFI might hope for. The reporting formats were also sensitive to how long each project allowed for analysis and reporting of results. The end users' needs and expectations for PHM were not clear since there was no formal requirements gathering process; moreover, the organisation's expectations of the final mapping interface had to be managed according to what functionality could be implemented given limitations arising from the underlying architecture of the product.

At the team level, the development team for QI was much smaller than the PHM team, which has facilitated more effective and efficient communication of evaluation results. More fundamentally, changes in development team personnel resulted in a change in developers' skills between the two projects. Agile developers were recruited for the QI project, which meant that they were much more accustomed to an iterative design process. As has been noted, a user experience expert was also recruited for the QI project who brought with them their own knowledge and expertise in designing and running user testing sessions.

At the organisational level, there has been a change in the perception to user testing towards employing it as a tool to build and maintain client relationships. Client relationships

are largely protected, as would be expected for any commercial development organisation with busy users with whom they often have to consult due to the complexity of the domain; this impacted upon participant recruitment particularly during the first stages of the collaboration. This was especially important since persona interviews were pitched as an exercise for the Engineering Doctorate that forms the basis for the collaboration, despite the benefits for DFI being explained in full before users consented to their participation. Participants for user testing were recruited from users whose contracts were about to end and the organisation had reason to be concerned that they might not renew. The aim was to encourage these participants to renew by exposing them to upcoming developments. This in turn has increased management support for usability work; the visions and values of key stakeholders can be an important influence on how other resources are assembled and configured in design work [13]. This approach to participant recruitment is partly necessary due to the complexity of the work domain but in stark contrast to the approach of many design consultancies that are able to send screening questionnaires to many potential participants.

It is important to note that whilst there was not a shortage of budget for the PHM project overall, due to lack of awareness around usability, usability work itself was not allocated a separate budget. In this case study more project resources were allocated to usability when the company was performing less well financially than when it was performing well financially. The need for improvement in the quality of design work largely stemmed from the increased market pressure described, and for improvements to be effected there had to be a change in the status of usability work across the organisation and recognition of its importance, particularly from senior management. This resulted in the recruitment of knowledge and expertise in usability, integration of design and evaluation methods into the development process and more successful use of resources and methods.

Expressive resource types

Expressive resources are those that communicate evaluation findings [16]. In this case study, the format and medium for reporting the results of design and evaluation methods required adaptation with the shift from a Waterfall development approach to an Agile approach, which demands a faster and more concise reporting format. This demonstrates that design and evaluation methods can be more effective in dynamic contexts if they support the rapid analysis and feedback of results. As Sy (2007) describes, for the Agile development process results were reported through the weekly scrums, whereas the Waterfall development approach enabled the writing of much more detailed reports [12]; the usability process was much more informal for PHM compared to how it was integrated within the development process for QI.

Knowledge resource types

Knowledge resources refer to knowledge of the system under evaluation, users and their abilities and tasks, and the application domain [16]. Such resources can impact upon all other method-resources. Growth in knowledge resources during the collaboration has directly resulted in more successful transfer of resources between projects.

Over the course of the collaboration the primary researcher developed her own experience of using design and evaluation methods, through guidance from academic experts, Masters courses and various workshops; this will have unavoidably coloured the way in which the methods have been applied and the projects have been planned. For example, the list of heuristics used and the format of the Cognitive Walkthrough were deliberately selected to be highly structured to provide additional support for problem identification. One particular consequence of this was that the heuristics used were partly adequate but also partly inadequate for the complexity of the interface inspected; many heuristics were assigned ‘not applicable’ and therefore the support provided by the heuristics used for problem identification in this case is questionable.

It should also be noted that between the PHM and QI projects three years had passed, during which a user experience expert was recruited by DFI who brought his own expertise in managing client relations when conducting user testing. His experience was particularly useful in the design of the testing protocol, for example letting the clients talk about their general experience of using the original DFI tool before introducing them to the new design. This was in addition to the growth in knowledge resources within the product development team through training, experience and arrival of new staff.

Finally, for information systems such as the ones described, knowledge of how the users think about the data is manifested in the database and software architecture. In this case study, easy modification of the user interface was found to depend on this knowledge much more than its separation from the software architecture [1]. At the time of PHM the architecture of the underlying database made it fundamentally very difficult to implement some redesign suggestions but in the three years between PHM and QI the architecture was reconfigured with an explicit aim of making changes much easier and more stable to implement. This can be attributed to the complexity of the application domain and had a big impact on the resulting design decisions made and whether redesign suggestions were put through to development. This improved flexibility has furthered receptiveness of usability work within DFI.

CONCLUSION

Methods and resources were employed across design and evaluation projects: a survey to inform the design of a mapping interface and its subsequent inspection using Heuristic Evaluation and Cognitive Walkthrough; the

creation of company personas using database server log files, a screenshot survey and interviews; and finally the user testing of an interface whose redesigned was informed by these personas. Over this period, a number of significant changes took place within and outside of the organisation. The influence of these changes was that participant recruitment, problem identification, the reporting and dissemination of results and task selection have had to be reconfigured according to local resources over the course of the collaboration, with increasing success as reported.

The reality of design work illustrated by the changes in business environment described in this case study is that designers consider the "ingredients" available to them before deciding which "recipe" to follow [16]. Recipes can come from their own recipe book, or the "HCI" recipe book. The ingredients, and importantly the cooking methods/utensils, available are determined by the application domain, organisational factors and project constraints. This case study demonstrates that local resources can be more influential than those indicated by the textbook versions of methods and more important than any financial cost of the ingredients and cooking utensils: development context (especially the process being used and how methods fit within this), organisational culture (supported from the highest levels of the organisation), knowledge resources (the expertise currently available) and the clients' expectations and needs were especially important. Studies that compare methods used in different contexts frequently ignore such factors, which case studies such as this suggest is to their detriment.

We conclude that this action research project has been of great benefit to DFI in raising the awareness and status of usability at the organisation and integrating methods into the design and development process; this has included making the method-resources available for the high quality usability work required in this complex domain.

ACKNOWLEDGEMENTS

This work was undertaken for an Engineering Doctorate supported by the Engineering and Physical Sciences Research Council (EPSRC) and DFI, in addition to an Industrial Fellowship from the Royal Commission for the Exhibition of 1851. We further thank all participants in the studies described and the three reviewers for their feedback.

REFERENCES

[1] Bass, L. and John, B.E. 2003. Linking usability to software architecture patterns through general scenarios. *Journal of Systems and Software*, 66(3), 187–197.

[2] Department of Health 2006. *Health reform in England: update and commissioning framework.*

[3] Furniss, D. 2008. *Beyond Problem Identification: Valuing methods in a "system of usability practice"*. University College London.

[4] Hayes, G.R. 2011. The relationship of action research to human-computer interaction. *ACM Transactions on Computer-Human Interaction*, 18(3), 1–20.

[5] Heuristic Evaluation – A System Checklist: 1995. *http://www.stcsig.org/usability/topics/articles/he-checklist.html*. Accessed: 2013-01-18.

[6] McDonald, S., Monahan, K. and Cockton, G. 2006. Modified contextual design as a field evaluation method. *Proceedings of the 4th Nordic conference on Human-computer interaction changing roles - NordiCHI '06* (New York, USA, Oct. 2006), 437–440.

[7] Nielsen, J. 1994. Heuristic Evaluation. *Usability Inspection Methods*. J. Nielsen and R.L. Mack, eds. John Wiley & Sons, Inc. 25–62.

[8] Polson, P., Lewis, C., Rieman, J. and Wharton, C. 1992. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36(5), 741–773.

[9] Review of future of Dr Foster Intelligence: 2010. *http://www.dh.gov.uk/en/MediaCentre/Pressreleases/DH_118452*. Accessed: 2012-04-24.

[10] Rubin, J. and Chisnell, D. 2008. *Handbook of usability testing : how to plan, design, and conduct effective tests*. Wiley Pub.

[11] Sale of Dr Foster Intelligence: 2010. *http://www.dh.gov.uk/en/MediaCentre/Pressreleases/DH_121537*. Accessed: 2012-04-24.

[12] Sy, D. 2007. Adapting Usability Investigations for Agile User-centered Design. *Journal of Usability Studies*. 2, 3 (2007), 112–132.

[13] Uldall-Espersen, T. 2008. Visions in software development: Achieving value in organizations. *Proceedings of CHI '08 workshop: Values, value and worth* (Florence, Italy, 2008).

[14] Wixon, D., Holtzblatt, K. and Knox, S. 1990. Contextual design: an emergent view of system design. *Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people* (1990), 329–336.

[15] Woolrych, A. and Cockton, G. 2001. Why and When Five Test Users Aren't Enough. *Proceedings of IHM-HCI 2001 Conference (Volume 2)* (Toulouse, France, 2001), 105–108.

[16] Woolrych, A., Hornbæk, K., Frøkjær, E. and Cockton, G. 2011. Ingredients and Meals Rather Than Recipes: A Proposal for Research That Does Not Treat Usability Evaluation Methods As Indivisible Wholes. *International Journal of Human-Computer Interaction*, 27(10), 940–970.