

INTERNATIONAL JOURNAL OF CONCEPTUAL STRUCTURES AND SMART APPLICATIONS

January-June 2013, Vol. 1, No. 1

Table of Contents

EDITORIAL PREFACE

- iv *Simon Polovina, Conceptual Structures Research Group, Sheffield Hallam University, Yorkshire, UK*
Simon Andrews, Conceptual Structures Research Group, Sheffield Hallam University, Yorkshire, UK

RESEARCH ARTICLES

- 1 **Eliciting People's Conceptual Models of Activities and Systems**
Ann Blandford, UCLIC, University College London, London, UK
- 18 **Towards Scalingless Generation of Formal Contexts from an Ontology in a Triple Store**
Frithjof Dau, SAP AG, Walldorf, Germany
- 39 **From Existential Graphs to Conceptual Graphs**
John F. Sowa, VivoMind Research, LLC, Rockville, MD, USA
- 73 **Advances in FCA-based Applications for Social Networks Analysis**
Marie-Aude Aufaure, Ecole Centrale Paris, Paris, France
Bénédicte Le Grand, Université Paris 1 Panthéon-Sorbonne, Paris, France
- 90 **CGs to FCA Including Peirce's Cuts**
Simon Polovina, Conceptual Structures Research Group, Sheffield Hallam University, Sheffield, UK
Simon Andrews, Conceptual Structures Research Group, Sheffield Hallam University, Sheffield, UK

Copyright

The **International Journal of Conceptual Structures and Smart Applications (IJCSSA)** (ISSN 2166-7292; eISSN 2166-7306), Copyright © 2013 IGI Global. All rights, including translation into other languages reserved by the publisher. No part of this journal may be reproduced or used in any form or by any means without written permission from the publisher, except for noncommercial, educational use including classroom teaching purposes. Product or company names used in this journal are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark. The views expressed in this journal are those of the authors but not necessarily of IGI Global

Eliciting People's Conceptual Models of Activities and Systems

Ann Blandford, UCLIC, University College London, London, UK

ABSTRACT

People using computer systems are required to work with the concepts implemented by system developers. If there is a poor fit between system concepts and users' pre-existing conceptualisation of domain and task, this places a high workload on the user as they translate between their own conceptualisation and that imposed by the system. The focus of this paper is on how to identify users' conceptualisations of a domain – ideally, prior to system implementation. For this, it is necessary to gather verbal data from people that allows them to articulate their conceptual models in ways that are not overly constrained by existing devices but allows them to articulate taken-for-granted knowledge. Possible study types include semi-structured interviews, contextual inquiry interviews and think-aloud protocols. The authors discuss how to design a study, covering choosing between different kinds of study, detailed planning of questions and tasks, data gathering, and preliminary data analysis.

Keywords: Computer Systems, Conceptual Models of Activities and Systems, Conceptual Structures, Task Structures, Verbal Data

INTRODUCTION

If there is poor conceptual fit between users and the systems that they work with, users are forced to “translate” between concepts that may only be meaningful to system developers and ones that are meaningful to themselves. This imposes additional workload on the user (Payne et al., 1990). The focus of this paper is on techniques for eliciting people's conceptual models as a stage in the process of designing or evaluating interactive systems.

Cooke (1999) defines a conceptual structure as comprising “domain-related concepts

and their interrelations”, while Johnson and Henderson (2011) define a conceptual model as comprising the target task domain; the concepts the application presents to users; the relationships between those concepts; and the mapping between task-domain concepts and application concepts. This definition makes it explicit that task-domain concepts and application concepts may not always be identical. In this paper, we focus on what Cooke calls “domain-related concepts”, and Johnson and Henderson refer to as “task-domain concepts” – i.e., the concepts users are working with when performing a real-world activity.

DOI: 10.4018/ijcssa.2013010101

These concepts may include what Johnson and Henderson (2011) refer to as “application concepts”, and that we call “system concepts”, because the way users think about their activity is often shaped by the systems they use, as discussed in more detail below. Sometimes, this is unavoidable, and presents few problems to users. However, when people have to expend effort on “translating” between domain concepts and system concepts, we refer to these as “misfits” between the user’s model and that implemented in the system. Misfits typically indicate usability difficulties, and may also represent new design opportunities, highlighting possibilities for systems that better fit the users’ needs (Blandford et al., 2008a). This approach of focusing on concepts complements most traditional methods for designing and evaluating interactive systems, which typically focus on task structures and processes rather than concepts (Blandford et al., 2008b).

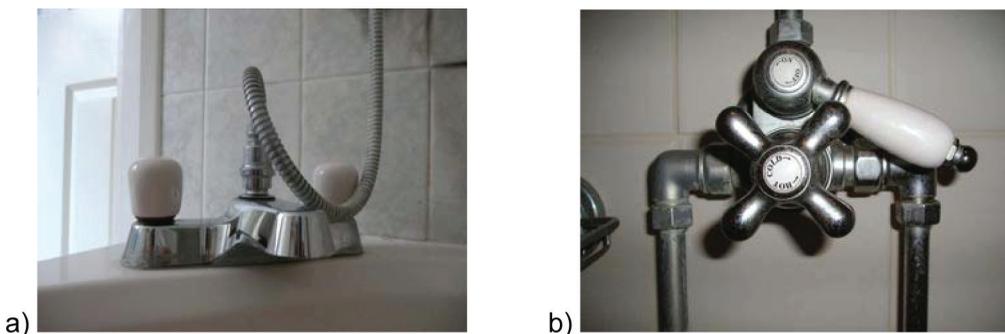
People’s conceptual models that they bring to an interaction are based on their prior experience, both of performing the activity (“doing work”, or the leisure equivalent) and of using analogous systems. To take a very simple example: someone using a shower will typically think of their requirements in terms of the temperature (too hot / too cold / just right) and the pressure (too forceful / too feeble / just right) of the water. Earlier generations of taps

and showers forced the user to work separately with the force of hot water and of cold water, engaging directly with the underlying “system model” of separate hot and cold water supplies being delivered and mixed together (Figure 1a). Prior experience with such taps means that people can quickly work out how to use new showers they encounter that are based on this model. However, such showers can be difficult to control well. More modern shower controls that allow the user to control temperature and pressure independently are typically easier to work with: this interaction better matches the user’s conceptual model, even when, as illustrated in Figure 1b, the controls have an old-fashioned appearance.

BACKGROUND

Many traditional approaches to the design and evaluation of interactive systems have focused on task structures – i.e., on understanding how people structure tasks in terms of sub-tasks and procedures (e.g. Card, Moran & Newell, 1983) and on the design of task structures for interactive systems (e.g. Hackos & Redish, 1998). Task structures have an important role to play in the design and evaluation of interactive systems, but can be difficult to manage when tasks are ill-structured or highly complex. Even a task as superficially simple as running a shower of

Figure 1. Shower taps that (a) directly reflect the underlying system model and (b) better match the user’s conceptual model



the desired temperature and pressure can be surprisingly challenging to describe in terms of task structure. For example, using hot and cold taps (Figure 1a), the task structure might be:

1. Turn cold tap 30 degrees anticlockwise;
2. Visually check that water is flowing;
3. Turn hot tap through 90 degrees anticlockwise;
4. Wait for a minute, or until you expect hot water to be running hot;
5. Visually check that water doesn't look scaldingly hot:
 - a. If it does then turn cold water anticlockwise further (how far?!?);
6. Manually check temperature of water:
 - a. If too hot then turn cold tap anticlockwise; wait; check;
 - b. If too cold then turn hot tap anticlockwise or turn cold tap clockwise; check;
7. When temperature as desired, step into shower.

This task description is, even now, incomplete: it says nothing about checking the pressure; how to choose whether to turn the cold water pressure up or the hot water pressure down; assessing how far to turn each tap; how to adjust the task description if the tap in fact opens clockwise (rather than anticlockwise, as assumed here); or how to respond when someone elsewhere in the building turns on a tap and the temperature fluctuates wildly. An every-day activity has a surprisingly complex task structure even though it is conceptually fairly simple.

From Task Structures to Conceptual Structures

Compared to the attention that has been paid to understanding task structures, little attention has been paid to understanding people's conceptual structures. Norman (1986) argues that an important responsibility of the design team is to communicate the designers' conceptual model to the user through the interface. However,

this downplays the importance of any existing conceptual models that users may come with.

Norman (1986) emphasises the importance of the interface and interaction design because users construct their own "mental models" (Johnson-Laird, 1983; Gentner & Stevens, 1983) of systems through their interactions with them. These mental models may be more or less complete, and more or less accurate. They are built up through experience of working with a system, and through gathering other information about the system – e.g., through training or reading manuals, or by analogy with related systems. Someone's mental model of shower controls (e.g. that turning the tap controls the pressure of water, or that other people running taps in the building affects the pressure) helps them work out how to control the shower without needing detailed task instructions every time. Mental models are typically neither complete nor accurate, and evolve over time through interacting with a system. Mental models are of the system being interacted with, whereas conceptual models are of the domain of activity, which may include some system concepts that shape that understanding. A full conceptual model of showering might include concepts relating to shampoo, soap and towels as well as those relating to water temperature and pressure. For the purpose of misfit analysis, one would focus on the concepts relating to just the part of the activity supported by the system of interest. In this case, this would focus on managing water. Attention might be paid to the broader activity if considering novel design options; for example, one might consider designing a showering system that manages shampoo and soap as well as water, and maybe even provides a drying feature!

Payne et al (1990) present an analysis of text editors in terms of "Yoked State Spaces", arguing that, as well as conceptualising the domain in terms of the goals that matter to them, the user has to understand and track how the device represents the task domain. Further, the user needs to be able to map between these two. The activity of interacting with a computer system in order to achieve domain goals

requires the user to continually map between these two state spaces: their domain space and the device space. Implicitly, if this mapping is complex then the system will be cognitively demanding to use.

Green (1989; 1990) developed a set of “tools for thought” to reason about some of the mappings between the user’s conceptual model and the device model that they have to work with which he termed “Cognitive Dimensions”: ways in which a system might be easy or difficult to work with. For example, “viscosity” in a system means that something that is conceptually simple is, in practice, difficult to achieve; this is usually because a concept that is significant to the user is not directly represented in the system model. To give a simple example: while writing this paper, if I chose to insert a new Figure between two existing Figures, I would have to manually update the numbers of all the subsequent Figures and all cross-references to them. This is because I have not invested effort in creating cross-references to Figure numbers or defined Figure numbers explicitly in the Figure captions; consequently, a task that is conceptually simple (insert a new Figure and update all Figure numbers) is in practice slow and tedious to perform.

Building on these ideas of there being sources of misfit between user and system, Blandford et al (2008a) present CASSM (Concept-based Analysis of Surface and Structural Misfits) as an approach to evaluating an existing or proposed system in terms of the quality of fit between the user’s conceptual model and that encapsulated in a design. CASSM fills a gap in the space of HCI evaluation methods by focusing attention on concepts rather than processes. It addresses evaluation questions that are missed by more traditional evaluation techniques, most of which are based on task structure and processes, or rely on heuristics (Blandford et al., 2008b). This focus on concepts rather than processes also makes it possible to evaluate complex systems where there may be many possible ways for users to achieve the same outcomes, because the space of concepts

does not explode in the way that the space of possible procedures does.

Shifting attention from evaluation to design, Johnson and Henderson (2011) propose a method for designing to support conceptual models. As noted above, their definition of a conceptual model is that which is implemented within the application rather than the user’s conceptual model. Although they recognise and discuss the importance of choosing appropriate conceptual models for design, they do not address the question of how to identify or support users’ existing conceptual models.

CASSM in a Nutshell

Whether gathering data to identify user requirements or to evaluate an existing system, the way the data is gathered and analysed will depend strongly on how it is to subsequently be used. In this paper, we are focusing on data gathering as an early stage of conceptual analysis. Our experience of this is based on our work on CASSM, so we present a brief overview of CASSM to set the scene for the subsequent discussion of how to plan and conduct data gathering.

A CASSM analysis involves identifying user concepts, interface concepts, and underlying system concepts and comparing them. Initially, the comparison is typically at the level of concepts. Concepts may be:

- **Present:** readily accessed);
- **Difficult:** hidden, disguised or hard to work with in some other way; or
- **Absent:** unavailable to user or system.

Later, the analysis will split concepts into entities and attributes, and consider the effects of actions on the system state, as described in more detail by Blandford et al (2008a). For the shower example presented above, focusing on the version with separate hot and cold taps (Figure 1a), these might be as shown in Table 1. There are other possible representations, but this one is a good first approximation for our purposes in this paper.

Table 1. An outline CASSM analysis for the taps controlling shower output

Concept	User	Interface	System
Temperature	Present	Difficult	Absent
Pressure	Present	Difficult	Absent
Hot water pressure	Difficult	Present	Present
Cold water pressure	Difficult	Present	Present

What is shown in Table 1 is that concepts that matter to, and are meaningful to, the user (temperature and pressure) are indirectly represented at the interface; i.e., the user can feel the output that is the result of their actions. However, they are absent from the underlying system model. In other words: temperature and pressure are present for the user, difficult at the interface and absent from the underlying system model.

Conversely, the system model has a direct representation of hot and cold water pressure, through the positions of the taps, and consequently of the valves being controlled. The user cannot access these directly, and they have little significance for the user. So these concepts are difficult for the user, but present at the interface and in the underlying system.

This short example captures the key ideas of a CASSM analysis: that the analysis involves identifying all concepts that either are directly relevant to and of concern to the user, or that the user needs to be aware of and work with in order to interact effectively with the system. Misfit analysis involves systematically laying out the space of concepts and working through them to identify which are difficult or absent, and what the consequences of those difficulties are likely to be for effective user interactions with the system. Identification of these misfits should highlight possible redesigns that reduce the misfits between user and system. This process is presented in more detail by Blandford et al (2008a). In this example, the misfits are addressed by more modern shower tap designs that give the user direct control of temperature and pressure (e.g, Figure 1b).

CASSM analyses have been presented for a variety of interactive systems. For example, Connell, Blandford and Green (2004) present an analysis of a ticket machine while Nishino (2011) presents an evaluation of a music programming language. Focusing on team interactions, and how a system supports people performing different roles, Blandford, Wong, Connell and Green (2002) present an analysis of an ambulance dispatch system from the perspectives of three different kinds of user: radio operator, telephone dispatcher, and allocator. The allocator is the individual responsible for allocating ambulances to incidents, who also serves as team leader delegating work to radio operators and telephone dispatchers.

Blandford, et al (2008a) present the method for conducting a CASSM analysis in detail. They illustrate the approach with analyses of three systems: a robotic arm for use by people with limited arm or hand movement; a digital library; and a drawing tool.

In summary, CASSM analyses have been conducted for a wide variety of systems, but all presentations have focused on the method and outcomes of analysis without discussing approaches to data gathering in detail. A focus on data gathering is also missing from the method presented by Johnson and Henderson (2011).

That is the focus of the current paper: how to design and conduct a study to gather user data as a basis for identifying users' conceptual models to inform design and evaluation. This is the outcome of reflections on extensive experience of conducting CASSM analyses, and of teaching this approach to students on a Masters programme in HCI, including both insightful and impoverished analyses.

PRINCIPLES OF DATA GATHERING

Any study design must be open to participants' various ways of conceptualising and articulating their activities, and must also be sensitive to taken-for-granted knowledge that is not naturally articulated. While there are individual differences between people, those differences are typically small compared to the differences between people who have different roles when interacting with a system. For example, in our study of ambulance despatch, everyone taking the "allocator" role had very similar conceptual models, although these were significantly different from the models of radio operators. Allocators had much richer models of the despatch task, including the external world of ambulance stations, hospitals, etc., while radio operators had much richer models of communication, e.g. different types of communication with ambulance crews, reflecting their different roles and responsibilities.

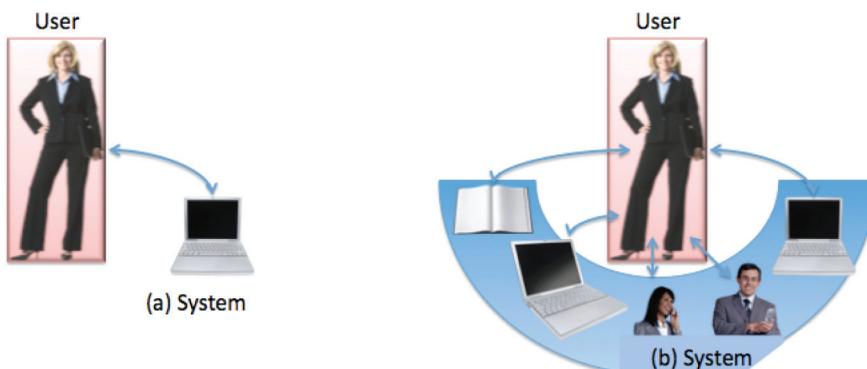
In contrast to approaches such as Activity Theory (Nardi, 1995) or Distributed Cognition (Hutchins, 1995; Hollan et al., 2000), which model the entire system including people and artefacts at equivalent levels of concern, not discriminating between one person and another, a CASSM analysis focuses on *conceptual misfits* between one individual and the system with which that individual engages. Thus, a

"system" might be a computer system or a particular software application running on a computer (Figure 2a), or it might be an aggregate of tools and other people with whom the individual interacts to achieve work (Figure 2b). The focus of data gathering for the user model is on the individual at the centre of the analysis. In Figures 2a and 2b this role is taken by the woman labeled "user".

As is centrally recognised and represented in Activity Theory, people's conceptual models are shaped by the tools that they use to achieve their goals. Sometimes, this is unavoidable: the tool creates and defines the activity. For example, an alarm clock defines the activity of setting and responding to an alarm. Without an "alarm" feature, people would be unlikely to talk about alarms or about setting them to go off at a particular time. In other cases, interactive systems support activities that have an independent, real-world significance for people. For example, an electronic calendar supports people in managing their time, particularly focused around appointments. In a study of time management tools, Blandford and Green (2001) found important disparities between the ways that people manage their time and the ways that a particular electronic calendaring system worked, including the following:

- People rarely thought about the durations or end time) of meetings unless they had very

Figure 2. CASSM focuses on individual, interface and system, where that system might be (a) a single computer application or (b) a collection of tools and other people



busy diaries, whereas electronic calendars demand that every meeting has a duration;

- People rarely made explicit travelling times or times for preparatory activities in their diaries, but omitting these considerations from an electronic diary could lead to difficulties, because time that is not marked as “busy” is implicitly assumed to be “free”, when in fact it might be needed for solo activities such as preparation or travelling to a meeting.

In conducting a CASSM analysis, or for designing new systems that better support users than existing systems do, one of the most difficult challenges is to elicit users’ conceptual models that are, as far as possible, independent of the particular tools they use.

PRACTICES OF DATA GATHERING FROM USERS AND POTENTIAL USERS

To construct a model of users’ concepts, it is essential to have some form of verbal data from current or intended users of a system. What is possible depends on what resources are available, particularly in terms of potential participants and prototype systems. The most common forms of data gathering are semi-structured interviews, contextual inquiry interviews (Beyer & Holtzblatt, 1998; Holtzblatt & Beyer, 2013) and think-aloud protocols (Charters, 2003). Each has strengths and limitations as discussed below. Other forms of data gathering, such as focus groups and those used for knowledge elicitation for the development of Expert Systems (e.g. Burton et al., 1990; Cooke, 1999) can also be adapted and applied, following the principles outlined here. Whichever data gathering approaches are used, the detailed study design needs to be focused on eliciting users’ conceptual models.

Practicalities of Conducting a Study: Who and Where

Early considerations for any study are likely to include how to recruit participants and where to conduct the study (Cooke, 1999). These considerations often shape what is possible in terms of data gathering techniques.

For studies that focus on work activities and the design of systems to support work, the most obvious participants to recruit are people who perform the role in question, and the most obvious place to conduct the study is in the work-place. For such studies, it is usually possible, and appropriate, to recruit participants through their managers, particularly if the study is to take place within their working time, for which management approval is likely to be needed. In this case, data gathering techniques are often chosen to minimise the disruption to ongoing work. So, for example, contextual inquiry interviews, which take place within the context of work may be preferable to semi-structured interviews, which take the participant away from their daily work. Conversely, there may be situations (e.g. high security or safety-critical contexts) where it would be too disruptive to conduct contextual inquiry interviews; then, it is essential to find another place, such as a staff room or coffee bar, for interviewing.

For studies that focus on home or leisure activities, there are many possible ways to recruit participants:

- Direct contact (e.g. Of friends and acquaintances); this is likely to result in an “accidental sampling” or “convenience sampling” approach;
- Indirect contact (following leads to people in the desired participant population); this is often described as a “snowball sampling” approach;
- Advertising in physical locations (e.g. On noticeboards) where people with relevant interests or activities are likely to encounter the invitation to participate; or

- Advertising online, through popular lists (e.g. Gumtree or craigslist) or social media.

Depending on the recruitment method, it may be more or less appropriate to meet face-to-face, in a public space or at the participant's home. The choice of meeting place will depend on many factors including minimising the need to travel great distances; ensuring the safety of both researcher and participant; putting the participant at ease; absence of distractions or excessive noise; and availability of relevant artefacts to support the data gathering.

We discuss three possible approaches to data gathering that are well suited to eliciting people's conceptual models: semi-structured interviews; contextual inquiry interviews; and think-aloud protocols. All of these deliver audio data and we have audio-recorded data in all of our studies. However, we do have one exception (Blandford & Green, 2001), where we only kept interview notes; we soon learned from this mistake: it proved very difficult to identify the more nuanced users' concepts reliably when the only data we had was notes, however thoroughly we had tried to gather them. Conceptual modelling from user data is most easily done with rich audio records that can be reviewed repeatedly; we have chosen to transcribe most audio data to make it easier to revisit the data as needed.

In the following sections, we outline each of the three data gathering approaches, focusing particularly on aspects relevant to identifying users' conceptual models. Where possible, we provide examples of resources (interview script, task instructions) prepared prior to data gathering. We also present a brief extract from early data analysis, to illustrate important aspects of that analysis. Fuller details on analysis are provided by Blandford et al (2008a).

Semi-Structured Interviews

Because people's conceptual models are subjective rather than objective, interviews are well suited to eliciting them. This contrasts with

self-reports of procedures or of past events, which tend to be unreliable because people only remember, and report, items that are salient to them.

Interviews may be more or less structured. A structured interview involves asking identical questions of all participants; an unstructured interview is more like a conversation with a purpose. Gill et al (2008) compare these different kinds of interviews. For eliciting conceptual models, semi-structured interviews give flexibility for adapting questions in the light of participants' responses, while also ensuring some consistency across interviews and coverage of important points. Preparation for a semi-structured interview involves planning key phases of the interview and possible forms for questions, but then being alert to exploring avenues of investigation that might not have been anticipated ahead of time. For a CASSM analysis, the interviewing needs to explore people's understanding of the domain in an open way.

Examples are powerful in interviewing: if invited to describe activities or structures in an abstract way, people will tend to provide generalities or approximations. To elicit details, it is usually much more effective to focus in on particular examples; the most effective examples are either recent or ongoing ones or highly memorable ones. The critical incident technique (Flanagan, 1954) has been widely used for eliciting people's understanding of past incidents, drawing out user requirements in terms of both concepts and functionality.

To illustrate the semi-structured interview approach, we draw on a study with lawyers, investigating how they think about their information resources and information management. Figure 3 illustrates a semi-structured interview prepared for some of the interviews. This shows the key phases of the interview: introduction; context; detail; closure. It also exemplifies some of the kinds of questions that might be asked, including detailed questions about a recent example.

Figure 3. Example semi-structured interview script (some details omitted)

Introduction:
 Thank you for agreeing to take part in this study.
 What we're interested in is understanding how you use information within the context of your legal work, how you organise it, how you get it, how you interact with other people about it, what you might like that you don't have: those kind of questions.
 We're interested in how you've obtained and used information, both the formal material that you would automatically know about, and also informal ways of gathering information: the chats in corridors, random telephone calls, and the like. When you gather information: just in time or when you can plan for it. How you gather information, how you reuse it. And any requirements you might have for working remotely.
 We'll be interviewing you for about an hour about your work and asking you to describe some particular examples in detail. If at any point you need to break off to answer the phone or to deal with anything else, please do. If you need to send us out the room while you deal with something confidential that's fine.

Gathering context:
 Can you describe your role in the organisation, and your main responsibilities?
 Are you largely based in your office or do you spend a lot of time out with clients or in Court or elsewhere?
 How is your team structured?
 How many projects are you typically working on at one time?

Into detail:
 Can you think of a recent project and talk through how you decided which parts of it you were dealing with yourself and which parts you were delegating?
 What were the key documents that you needed to work with?
 [For each]: find out how it was accessed, modified, stored, used...
 What were the key resources you accessed, and why?
 Was there any point in that project where you really needed information that you couldn't find easily? [Talk through the instance]
 Information reliability is important. How do you assess it?
 [...]

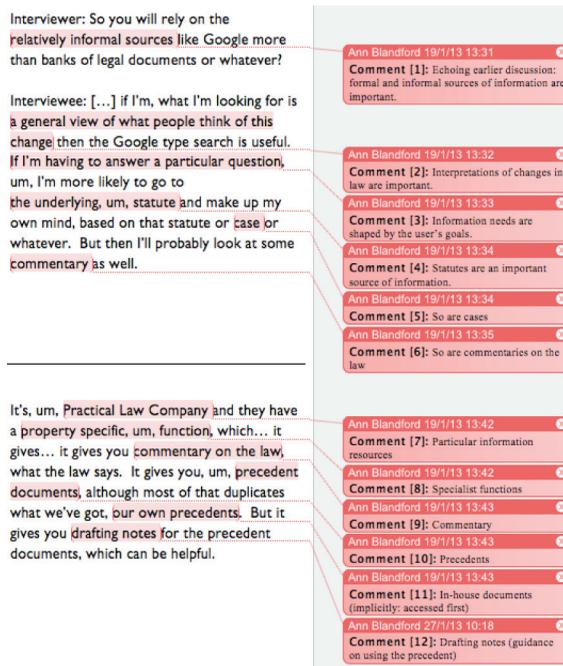
Closure:
 Is there anything else you'd like to add - any other questions you think I should have asked?
 That's been really helpful. Thank you!

As noted above, semi-structured interviews allow the interviewer to follow up on unanticipated points made by interviewees. For example, in one interview with a property lawyer, it emerged early on that a major aspect of her role was viewing properties. This was followed up in the interview, both in terms of how she managed data while out on site and also in terms of what this meant for the kinds of information she needed to access. One consequence of her site visits was that she could perceive value in being able to access information while on site via a mobile device, even though this was not currently part of her standard practice. Another consequence was that maps and plans were

important information resources for her, and needed to be managed alongside information resources that would be used by lawyers across a range of specialisations. This insight – that lawyers need to manage specialist resources alongside legal documents – could then be explored further in subsequent interviews.

As well as gathering interview data, it is often possible, and valuable, to take photographs or copies of important artefacts relating to the work to support the analysis. For example, photographs of documents that the lawyer had annotated provided an additional record of the kinds of annotations made, their role in her

Figure 4. Two annotated extracts from a semi-structured interview with a property lawyer



information work, and how she manages and communicates information to others.

Following each interview, it is valuable to review the interview and findings from it, and in particular to consider whether the interview script should be adapted—e.g. to revise questions that did not work well, or to explore emerging lines of inquiry in more detail.

Figure 4 shows preliminary annotations of extracts from the interview with the property lawyer, highlighting concepts that she articulated as being important to her. It quickly became apparent that, as well as obviously legal sources of information (statutes, cases), others' interpretations of that law (commentaries) were important, and so were generalist information sources such as maps, company information or location-based information. Precedent documents ("boiler plate text" for re-use in future legal documents) and information to make those usable (drafting notes) were also highlighted.

Interviews with further lawyers made it possible to determine which concepts:

- Are of widespread importance in their information work, and should therefore be implemented in future designs of systems to help lawyers access and manage information resources;
- Are important to subsets of lawyers (e.g. those specialising in property law), and might be implemented in more specialist products; and
- Are unique to an individual.

Contextual Inquiry Interviews

Although interviews may usefully take place within the working context, or any other place where the individual generally performs the activity in question, the standard semi-structured interview takes place outside the normal work activity. This enables people to reflect on the activity and how they think about it, but does not engage them directly in it. Contextual Inquiry (Beyer & Holtzblatt, 1998) is an approach to interviewing that explicitly builds in observation

of the work, and the work artefacts, to construct understanding of the work, and of the conceptual models people are working with. As described by Beyer and Holtzblatt, Contextual Inquiry involves putting oneself in the position of an inquiring outsider, partnering with and valuing the expertise of the people whose work is being studied. The analyst observes individuals performing work, and asks timely questions to probe the participants' understanding of their work, covering both the tools they are working with and the broader domain concepts. This does not focus just on domain and device concepts, but also builds a broader understanding of the work. As with semi-structured interviews, it can be challenging to focus attention on the concepts that are of primary interest, because there is more to work (processes, values, cultures, etc.) than concepts. Questions are likely to include who, what, when, where, how and why questions. However, unlike semi-structured interviews, it is difficult to prepare questions in advance, because the conversation is grounded in the work setting, and the themes that emerge are shaped by the work. Holtzblatt and Beyer (2013) describe this as "active inquiry".

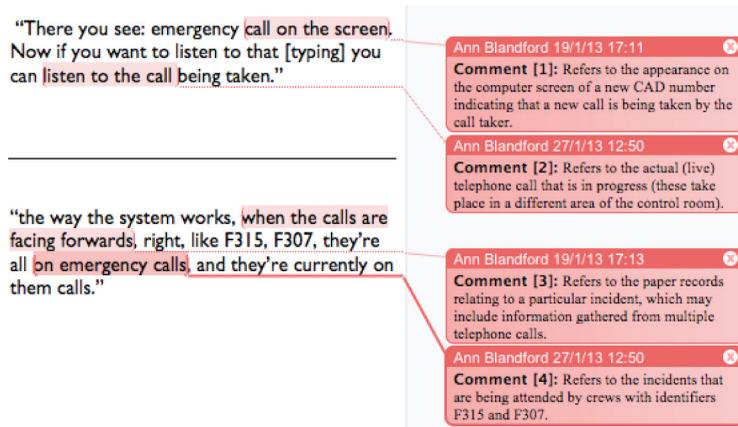
Contextual inquiry was a suitable approach to study the work of ambulance controllers (Blandford et al., 2002), to understand how they used the ambulance dispatch system, and how to make the information display better match their conceptual models. Because the control room has a fixed layout, and people's work is largely desk-based, it was possible to simply pull up a chair next to a participant and watch them working, with their permission. Given the nature of ambulance control work, which is very demanding while dealing with incidents but with "down time" between calls, it was important to engage in questioning at appropriate times, and to accept that the conversation would be suspended when the next call came in. This made the early stages of data gathering very difficult, because it was hard to make sense of the work and of the concepts that they were working with. At a superficial level, it is easy to understand that ambulances are being sent

out in response to emergency calls to specified locations. It was easy to understand that every call was allocated a Computer Aided Dispatch (CAD) number as soon as the call-taker accepted the call, and that this was the primary way they referred to calls after that. However, the nuances of the work were difficult to grasp, and understanding was built up over several days of observation and contextual inquiry interviews. It became apparent quite early on that controllers thought of incidents principally in terms of their features, such as time, location and type (fall, cardiac arrest, road traffic accident, etc.), and not in terms of the CAD number that was assigned to the call.

This was identified as a misfit, because the information screens they worked with at the time of the study facilitated access to call details via the CAD number, and the home screen listed just these CAD numbers. As a result of the observations and analysis, a simple design recommendation to present more details of each call on the home screen was made; this change was implemented, and was found to make it easier for controllers to select a call of interest if they needed to refer back to it after dispatching an ambulance to it.

The study of ambulance control exemplified the challenge of identifying concepts. This is illustrated by two extracts from the transcript of one ambulance allocator talking about his work. In the extracts annotated in Figure 5, the word "call" is being used to mean four different things. The first and third references are to device objects (digital and paper) that provide information gathered from one or more telephone calls about the incident in question. The second reference is to the ongoing telephone call that is providing information about an incident, while the fourth is to the incidents being attended. At one time, there was approximately a one-to-one mapping between calls and incidents, but with the growing use of mobile phones, there may now be several calls to report one incident in a public place. But the ambulance controllers we worked with never use the word "incident": they always referred to them as "calls". Of more

Figure 5. Different concepts being described using the same word (“call”)



importance, as their systems were gradually being transferred from being paper-based to digital, was that they were using the paper-based system to collate information about incidents from multiple calls, but their computerised system at the time of the study required them to process each call separately, and did not support the grouping of information from multiple calls about an incident. This made it difficult for them to manage separate incidents, particularly when there were multiple calls about each.

An analysts, it took us several days of observing and discussing the work of ambulance controllers before we recognised that the word “call” was being used to mean significantly different things that had important implications for design.

Think-Aloud Protocols

A third approach that can be valuable for gathering data about users’ conceptual models is the use of think-aloud protocols (Charters, 2003). A think-aloud protocol is a recording of the participant thinking aloud while performing a task, typically with existing computer-based tools that support that task. Think-aloud was popularised in the study of human problem solving (Ericsson & Simon, 1984) and is widely used in the evaluation of interactive systems (e.g. Nørgaard & Hornbæk, 2006).

The challenge in gathering data using think-aloud for identifying user concepts is to create activities that allow participants to articulate their understanding of the domain, and not just of the device with which they are working at the time. Unlike Contextual Inquiry, for which participants are generally expected to carry on with their work (albeit being observed and questioned), think-aloud usually involves participants performing tasks defined by the analyst. For gathering user concepts, those tasks should be as realistic as possible, including being appropriate to the participants’ knowledge, motivations, context and expertise. For example, in studies of information interactions, we have sometimes been able to invite participants to define their own information tasks (finding information that was relevant to them at that time) (e.g. Makri et al., 2007; Blandford et al., 2008a). Even when tasks are artificial, they should have real-world relevance; for example, in a brief study of flight booking sites, we asked participants, all from overseas, to plan a journey home, rather than booking a flight between specified airports, and to talk about their thoughts and decision making processes, rather than focusing on interface details. See Figure 6 for example task instructions.

Figure 6. Example think-aloud instructions

"I'm studying how people plan journeys and book flights. Imagine that you are planning a journey home to visit your relatives for a holiday. Use internet resources to find a suitable flight. Think aloud while you're looking. For example: How are you choosing dates and flights? What other factors do you need to take into account in choosing tickets? How easy is it to find the next thing to do? Are you getting closer to your goal? Stop when you've found a suitable flight or flights. Thank you for taking part in this study."

Figure 7. Annotated, anonymised interview transcript

I'm going to try to get to my town in [country], it's a small airport called [local airport]. OK, so I'm going from London to [local airport]. I want to go on Christmas Day because it's cheaper to fly that day, so it's 25 of December. If I'm flying all the way to [country] I need to be there at least two weeks. So I'll be coming back on the 15th? Let's see the calendar: so 1, 2, 3, I'll be coming back on the 17th. Round trip. From London, which is LON, any airport in London, and going to [local airport]. It doesn't like the date. Oh, I'll be coming back in 2007. So, one ticket, economy and, searching in all those places [long pause] it's taking longer than I expected. OK, so, 650 pounds to go there for three weeks [...]

I'll take this one, 374 pounds London Heathrow to [capital city] and I won't worry about how I'm going to get home.

- Ann Blandford 19/1/13 18:07
Comment [1]: Destination
- Ann Blandford 19/1/13 18:08
Comment [2]: Airport (may be large or small)
- Ann Blandford 19/1/13 18:08
Comment [3]: Starting city
- Ann Blandford 19/1/13 18:10
Comment [4]: Outgoing date
- Ann Blandford 19/1/13 18:10
Comment [5]: Cost of flight
- Ann Blandford 19/1/13 18:09
Comment [6]: Visit duration
- Ann Blandford 19/1/13 18:09
Comment [7]: Return date
- Ann Blandford 19/1/13 18:10
Comment [8]: Ticket / journey type
- Ann Blandford 19/1/13 18:11
Comment [9]: Recognises that a city or region may have multiple airports
- Ann Blandford 19/1/13 18:12
Comment [10]: The system being used permitted the user to enter return date that was before outgoing date
- Ann Blandford 19/1/13 18:12
Comment [11]: Ticket price
- Ann Blandford 19/1/13 18:33
Comment [12]: Highlights the important difference between destination airport and final destination

A short extract from one of the resulting transcripts is shown in Figure 7, with the preliminary annotations highlighting user concepts.

Whereas in the previous example we highlighted the challenge of recognising that one word might have multiple meanings, representing multiple concepts, in this example, we see that the participant names concepts in terms of instances (“London”, “Christmas day”, etc.) rather than the classes those instances belong to (“starting city”, “departure date”, etc.). In this case, it is up to the analyst to name the categories, or to draw on data from other participants to identify suitable labels for the

concepts. More broadly, participants may not all use the same terms for the same concepts.

In this case, the analysis highlighted design implications including:

- Some which should be simple to implement, such as making the locations of stop-overs on non-direct flights easier to discover;
- Some requiring larger design changes, such as making additional information about flights (e.g., entertainment and meal options) easily available; and

- Some that would offer significantly different functionality, such as helping people to plan entire journeys, including booking train or bus tickets, airport parking and car hire.

NEXT STEPS

Depending on the purpose of the study, and the stage of development, the next step is likely to involve either proceeding to a misfit analysis of an existing system, or directly to conceptual design for a new product.

Conducting a Misfit Analysis of an Existing System Design

If the purpose of the project is to complete a misfit analysis, typically prior to re-design, then it is necessary to identify system concepts at the same level of abstraction as the user concepts, for which the process is described above. System concepts are identified from system descriptions, such as a specification or other system documentation, and other existing system representations such as a running system or prototype. The data is analysed in whatever ways are possible given the data sources. For example, system documentation might be analysed in a similar way to a user transcript, whereas for a running system it might be necessary to do a systematic review of the concepts that the user has to interact with. These would normally be implemented as entities and attributes, taking a traditional entity-relationship modelling approach to system development (Chen, 1976).

In identifying system concepts, there may be a temptation to include all interface widgets; this should be avoided except where the interface concepts are likely to cause user difficulties. For example, to return to the shower example above, the focus was on temperature and pressure, rather than on the taps themselves and how to control the taps.

The detailed steps of misfit analysis are presented by Blandford et al (2008a).

Moving to Conceptual Design

If the next step is proceeding to conceptual design then the user concepts that have been derived from user data need to be reviewed and structured into a conceptual model. There may be user concepts that are very difficult to implement directly as system concepts; for example, Faisal et al (2006) identified *ideas* as being important to users of a digital library. Ideas are not directly represented in the documents in a library, but are embodied in readers' interpretations of those documents. Consequently, the concept of an *idea* cannot be directly implemented without substantial additional effort. Developers might decide that this additional effort is not justified.

Alternatively, they might decide to extend the system significantly. For example, they might explore the possibility of adding a social tagging feature to the library, to enable users of the library to tag documents with their *ideas* and to review the tags of others. This would take the library into a new design space; whether or not to do this is a major design decision that is beyond the scope of this paper.

The development and use of conceptual models are discussed in detail by Johnson and Henderson (2011).

CONCLUSION AND DISCUSSION

We have presented, and exemplified, three approaches to data gathering for conceptual modelling, as an early stage to inform the design and evaluation of interactive systems. This focus on concepts complements other established approaches to design and evaluation, which typically focus more on processes and functions than concepts. Even other approaches that put concepts at the centre (e.g. Norman, 1986; Johnson & Henderson, 2011) focus on how to communicate system concepts to the user rather than how to base the design around user concepts. For information systems, techniques

such as card sorting (Hudson, 2013) elicit user data to identify structure in a set of concepts, e.g. to design intuitive navigation through an information system. However, they do not address how to identify the concepts in the first place. Our focus here has been on how to gather data that can be analysed to construct users' conceptual models.

Not every user conceptualises a domain in the same way. Greater diversity of the participant group might result in a broader set of concepts being elicited, but in our experience the set of concepts stabilises (thus reaching "saturation") within a small number of participants—typically fewer than ten—provided that the participants are focusing on the same activity. Where greater divergence in concepts has been found, this generally means that participants do not have the same goals, or have significantly different ways of thinking about the activity. This often indicates that they would work more easily with different systems that support their differing goals, or different ways of thinking about the activity. Cooper (1999) would argue that different groups of participants correspond to different *personas*, needing different interfaces or underlying conceptual structures.

The most effective data gathering often involves using multiple methods and triangulating findings (Mackay & Fayard, 1997), or tailoring methods that are shaped to fit the context and questions. For example, interviews can be used to find out generally what people do in, and think about, a particular activity. This can be used to design suitable tasks for a think-aloud study that elicits more information about how people

perceive a particular system for achieving those tasks, assuming that an implementation already exists. In general, more sources of data yield more information, but this has to be balanced against any need for speed, efficiency or the practicalities of accessing different sources.

Individual methods can be adapted to the demands of a situation; for example, somewhere between the semi-structured interview and the contextual inquiry interview lies an approach that involves explicitly valuing the expertise of the participants and shaping the interview as a teaching session in which the domain expert teaches the analysts about their world of work, focusing on the activity of interest. Approaches such as this which place the participant in a position of authority, with a responsibility to explain their understanding effectively, go by various names such as teach-back and peer tutoring.

As highlighted above, different approaches to data gathering have different strengths and weaknesses, as summarised in Table 2 for the three approaches discussed here.

There is, of course, more to system design than mirroring existing user concepts: innovation is achieved through creating new possibilities that extend beyond what is currently available. However, every new concept that is introduced into a system has to be learned by users. And every misfit between the user's conceptual structure and that implemented in a system forces the user to focus on the system rather than the domain goal at hand, and increases the demands on the user.

Table 2. Summary of features of each data-gathering approach

	Semi-Structured Interview	Contextual Inquiry	Think-Aloud
Easy to probe past experience?	√		
Direct focus on work	√	√	
Concurrent engagement in activity?		√	√

ACKNOWLEDGMENT

The original research developing CASSM was conducted with Thomas Green and Iain Connell. My understanding of the challenges of data gathering and analysis has been greatly improved by interactions with Anne Adams, Simon Attfield, Sarah Faisal, Dominic Furniss, Amir Kamsin, Stephann Makri, and many MSC students (too many to name individually); it's been a pleasure to work with them all. Particular thanks to Jo Iacovides, Katie Parker and Sheila Pontis, who provided constructive feedback on a draft of this paper.

REFERENCES

- Beyer, H., & Holtzblatt, K. (1998). *Contextual Design*. San Francisco, CA: Morgan Kaufmann.
- Blandford, A., & Green, T. R. G. (2001). Group and individual time management tools: What you get is not what you need. *Personal and Ubiquitous Computing*, 5(4), 213–230. doi:10.1007/PL00000020.
- Blandford, A., Green, T. R. G., Furniss, D., & Makri, S. (2008a). Evaluating system utility and conceptual fit using CASSM. *International Journal of Human-Computer Studies*, 66, 393–409. doi:10.1016/j.ijhcs.2007.11.005.
- Blandford, A., Hyde, J. K., Green, T. R. G., & Connell, I. (2008b). Scoping usability evaluation methods: A case study. *Human Computer Interaction Journal*, 23(3), 278–327. doi:10.1080/07370020802278254.
- Blandford, A. E., Wong, B. L. W., Connell, I., & Green, T. R. G. (2002). Multiple viewpoints on computer supported team work: a case study on ambulance dispatch. In X. Faulkner, J. Finlay, & F. Détienne (Eds.), *Proceedings of the People and Computers XVI (HCI'02)* (pp. 139-156). Springer.
- Burton, A. M., Shadbolt, N. R., Rugg, G., & Hedgecock, A. P. (1990). The efficacy of knowledge elicitation techniques: a comparison across domains and levels of expertise. *Knowledge Acquisition*, 2(2), 167–178. doi:10.1016/S1042-8143(05)80010-X.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human computer interaction*. Hillsdale, MI: Lawrence Erlbaum.
- Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal OLD*, 12(2).
- Chen, P. P. S. (1976). The entity-relationship model—toward a unified view of data. *ACM Transactions on Database Systems*, 1(1), 9–36. doi:10.1145/320434.320440.
- Connell, I., Blandford, A., & Green, T. R. G. (2004). CASSM and cognitive walkthrough: Usability issues with ticket vending machines. *Behaviour & Information Technology*, 23(5), 307–320. doi:10.1080/01449290410001773463.
- Cooke, N. J. (1999). Knowledge elicitation. *Handbook of applied cognition* (pp. 479-510).
- Cooper, A. (1999). *The inmates are running the asylum*. Indianapolis, IN: Sams Publishing.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Faisal, S., Cairns, P., & Blandford, A. (2006). Developing user requirements for visualizations of literature knowledge domains. In *Proceedings 10th International Conference on Information Visualization* (pp. 264-269). London, UK: IEEE.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327. doi:10.1037/h0061470 PMID:13177800.
- S. Gentner, & A. L. Stevens (Eds.). (1983). *Mental models*. Hillsdale, NJ: Erlbaum.
- Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Methods of data collection in qualitative research: Interviews and focus groups. *British Dental Journal*, 204(6), 291–295. doi:10.1038/bdj.2008.192 PMID:18356873.
- Green, T. R. G. (1989). Cognitive dimensions of notations. In A. Sutcliffe, & L. Macaulay (Eds.), *People and computers V* (pp. 443–460). Cambridge, UK: Cambridge University Press.
- Green, T. R. G. (1990). The cognitive dimension of viscosity: A sticky problem for HCI. In D. Diaper, D. Gilmore, G. Cockton, & B. Shackel (Eds.), *Proceedings of the Human-Computer Interaction (INTERACT '90)* (pp. 79-86). Elsevier.
- Hackos, J., & Redish, J. (1998). *User and task analysis for interface design*. Wiley.

- Hollan, J. D., Hutchins, E. L., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on CHI*, 7(2), 174–196.
- Holtzblatt, K., & Beyer, H. R. (2013). Contextual design. In Soegaard, Mads and Dam, Rikke Friis (Eds.), *The encyclopedia of human-computer interaction* (2nd ed.). Aarhus, Denmark: The Interaction Design Foundation. Retrieved from http://www.interaction-design.org/encyclopedia/contextual_design.html
- Hudson, W. (2013). Card sorting. In Soegaard, Mads and Dam, Rikke Friis (Eds.), *The encyclopedia of human-computer interaction* (2nd ed.). Aarhus, Denmark: The Interaction Design Foundation. Retrieved from http://www.interaction-design.org/encyclopedia/card_sorting.html
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Johnson, J., & Henderson, A. (2011). *Conceptual models: Core to good design*. Morgan Claypool Synthesis Lectures on Human-Centered Informatics.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.
- Kirwan, B., & Ainsworth, L. K. (1992). *A guide to task analysis*. London, UK: Taylor and Francis.
- Mackay, W. E., & Fayard, A. L. (1997). HCI, natural science and design: A framework for triangulation across disciplines. In *Proceedings of the 2nd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (pp. 223-234). ACM.
- Makri, S., Blandford, A., Gow, J., Rimmer, J., Warwick, C., & Buchanan, G. (2007). A library or just another information resource? A case study of users' mental models of traditional and digital libraries. *Journal of the American Society for Information Science and Technology*, 58(3), 433–445. doi:10.1002/asi.20510.
- B. A. Nardi (Ed.). (1995). *Context and consciousness: Activity theory and human-computer interaction*. MIT Press.
- Nishino, H. (2011). On conceptual misfits in computer music programming. *Journal of the Japanese Society for Sonic Arts*, 1(1), 17–21.
- Nørgaard, M., & Hornbæk, K. (2006). What do usability evaluators do in practice? An explorative study of think-aloud testing. In *Proceedings 6th ACM Conference on Designing Interactive Systems* (pp. 209-218). New York, NY: ACM Press.
- Norman, D. A. (1986). Cognitive engineering. In D. A. Norman, & S. W. Draper (Eds.), *User centered system design* (pp. 31–62). Hillsdale, NJ: Lawrence Erlbaum.
- Payne, S. J., Squibb, H. R., & Howes, A. (1990). The nature of device models: The yoked state space hypothesis, and some experiments with text editors. *Human-Computer Interaction*, 5, 415–444. doi:10.1207/s15327051hci0504_3.

