Novice and Experts Strategies for Understanding Complex Big Data

Andreas REITER[3], Xianxu HOU[3], Genovefa KEFALIDOU[1], James GOULDING[2]

[1] *Human Factors Research Group, University of Nottingham*
[2] *Horizon Digital Economy Research, University of Nottingham*
[3] *Horizon Centre for Doctoral Training, University of Nottingham*

Abstract. Personal data is everywhere. Its complexity grows exponentially as more devices generate data. Understanding and making sense of complex data is fundamental as critical decisions may depend on its interpretation. In this lab-based observation study both novices and experts were exposed to complex medical information. The findings suggest that medical professionals employ different strategies from non-medics during sense-making and task completion. We discuss implications for designing new decision-making tools that support sense-making complex big data.

Keywords., Big Data, sense-making, decision-making, medical data

1. Introduction

Sense-making research has increased in recent years as new technologies and frameworks are introduced and explored across different domains. At the same time, web-based collaborative tools have changed the way we interact with each other and with information, leading to new dynamics in human-human, human-computer and human-data interactions while introducing new volumes of data to process (Haddadi et al., 2013). With the recent rise of the "quantified self" movement and the accompanying move of big data analysis into the home and organisations, the need for a deeper understanding of how people perceive their own personal data and how they perform data analysis that may impact our daily lives has become apparent. The medical sector is no exception, with medical professionals now seeking to make sense of a multitude of different Big Datasets on a regular basis, both for purposes of supporting health service provision and to improve care. In this paper we report findings of an investigation on how both novices i.e. non- medical professionals, and experts i.e. medical professionals, interact with specific volumes of health-related data. Our first aim was to identify the strategies and behaviours medical professionals and non-medical professionals exhibit when confronted with sense-making tasks that involve understanding medical data. Our second aim was to understand how we can inform the design of new tools that can support and enhance understanding of complex medical data to aid decision-making.

1.1 Big Data and Sense-making processes

According to Jacobs (2009), Big Data can be defined as "data whose size forces us to look beyond the tried-and-true methods that are prevalent at that time". It is estimated that 2.5 quintillion (2.5 billion, billion) bytes of data are being created every single day, increasing the volume, variety and complexity of data being generated. Big Data permeates across many domains as ubiquitous technologies continuously harvest data. For example, loyalty cards, mobile phone usage, health and census statistics are embedded into both workplace and leisure spaces capturing personal data. Sense-making methodology has been called a 'black box' that needs to be opened to uncover the information and processes involved in understanding it (Dervin, 1999). Dervin understood this 'gappiness' as a metaphor to express ill-structured problems, ambiguity and uncertainty. As such, sense-making implies the pre-condition of fuzziness and constant enquiry – situations that can indeed be triggered by interruptions. Since then, much research has studied sense-making both as a methodological tool (Savolainen, 2006) and as a research question on its own due to the applicability of its nature. Sense-making has been studied within different contexts ranging from firefighting (Dyrks, Denef & Ramirez, 2008) to notetaking and spreadsheet manipulation (Russell, Stefik, PirollI & Card, 1993). Weick (1995) considers sense-making as a strong social construct that aims to extract cues to generate an argument (or tell a story) in which plausibility over accuracy may emerge. Baber Attfield, Wong and Rooney (2013), followed recently an intelligence analysis approach to understand sense-making processes focusing on data, frames and narrative elements. Kefalidou and Houghton (2016) adapted this technique to identify co-constructions of meanings and collaborative approaches emerging from fabricated ambiguous data that fed into the design of a collaborative sense-making platform for crises incidents (Blum, Kefalidou, Houghton, Flintham, Arunachalam & Goulden, 2014).

1.2 Experts vs. Novices

'Expertise' in decision-making research often refers to people who have attained some form of external validation e.g. degrees, titles, certificates. However, in practical terms, experience often becomes a synonym for expertise within given contexts. As such, Shanteau (1988) redefines experts for his purposes as people who are highly regarded by their peers. Within his proposed pyramid framework, *Naives* sit at its base with very little or no experience in the subject matter, *Novices* in the middle and *Experts* at the top. The study of Naives is of little use here but observing Novices on the other hand does provide an opportunity to find out how humans - largely unbiased by prior experience - attempt to tackle data swamping. The on-going generation and control (or lack of) of personal information blurs the traditional boundaries of expert and novice roles in understanding generated data. Individuals become data hubs as personal devices emit and sync. Although research on expertise and expert systems has led to a rather negative view on expert decision-making due to their cognitive limitations - "Indeed, it can be difficult to find cited psychological studies which have anything positive to report about experts" (Christensen-Szalanski and Beach, 1984) they do exhibit certain characteristics that are desirable to laypeople, especially when it comes to trying to improve human well-being through effective medical human-data interaction.

The difference between experts and novices has been investigated in the domains of systems analysis, decision making and problem solving (Haddadi *et al.*, 2013; Schenk *et al.*, 1998). In a study similar to the one reported here, Schenk et al. (Schenk *et al.*, 1998) employed a verbal protocol method to analyse behavioural differences between expert and novice system analysts in approaching a sample requirements engineering task. One of their findings was that experts often refer to their episodic memory of past experiences and can therefore better distinguish relevant from irrelevant information.

2. Methods

The method adopted in this study was lab-based observation using think-aloud protocols (Ericsson and Simon, 1985). Participants were asked to perform a task, followed by a 15-minute semi-structured interview. The task required them to make sense of a lot of pre-generated data printed on A4 sheets of paper. The data for these graphs stemmed from a dataset used in medical data analysis, the General Practitioner Records Database (GPRD) (see Figure 1 for example).

2.1 Participants

Nine novice participants (M_{age}=26; 4 female, 5 male – all postgraduate students) and three experts (M_{age}=37; 1 female, 2 male) were recruited at the University of Nottingham. The three experts were recruited from the National Health Services (NHS) Nottinghamshire Trust Information Sciences department and were professionals that handle patient records on a daily basis. Recruitment was done through opportunity and snowball sampling, and no participant was supplied with any financial incentive or reimbursement.

2.2 Materials

The General Practitioner Records Database (GPRD) is a set of anonymised general practitioner records which was formerly provided by the Medicines and Healthcare Products Regulatory Agency (MHRA) and is now worked on jointly by NHS England through the Clinical Practice Research Datalink (CPRD). It provides anonymised information on administered tests, prescribed therapies and medication in participating practices. This dataset was reduced to 54 different cumulative statistics of patients in nine practices across the UK East Midlands Area. The timeframe of all clinical events in these reports was 1.10.2002 to 1.02.2012. These reports featured e.g., per-practice breakdowns of patients joining and leaving the practice or per-practice and per-consultation charts of the ten most common types of consultations within the entire GPRD data set limited to the nine case study practices. Figure 1 shows an example of the style and layout used in all of these reports. All graphs displayed time-series split by events that occurred in the selected time period and are representative medical data that medical professionals handle in their everyday tasks. Participants were also given an A4 paper sheet with a fictitious task scenario. According to the scenario, participants had just started to work as data analysts at a research institute – requiring experts and novices alike to profile themselves and leave a "strong impression" with their new employer. In this scenario they were given the case study GPRD dataset and told that they had to prepare themselves for a meeting in which they would present key understandings from the dataset to a project team of medical professionals, their line managers and policy makers. They were asked to:

"Predict the demands that are likely to be placed on practices over the next few years with respect to health services and how these will vary across practices."

This phrasing was used to preserve some level of freedom of interpretation by participants and to give them an incentive to develop "a feel" for the data and explore the dataset.

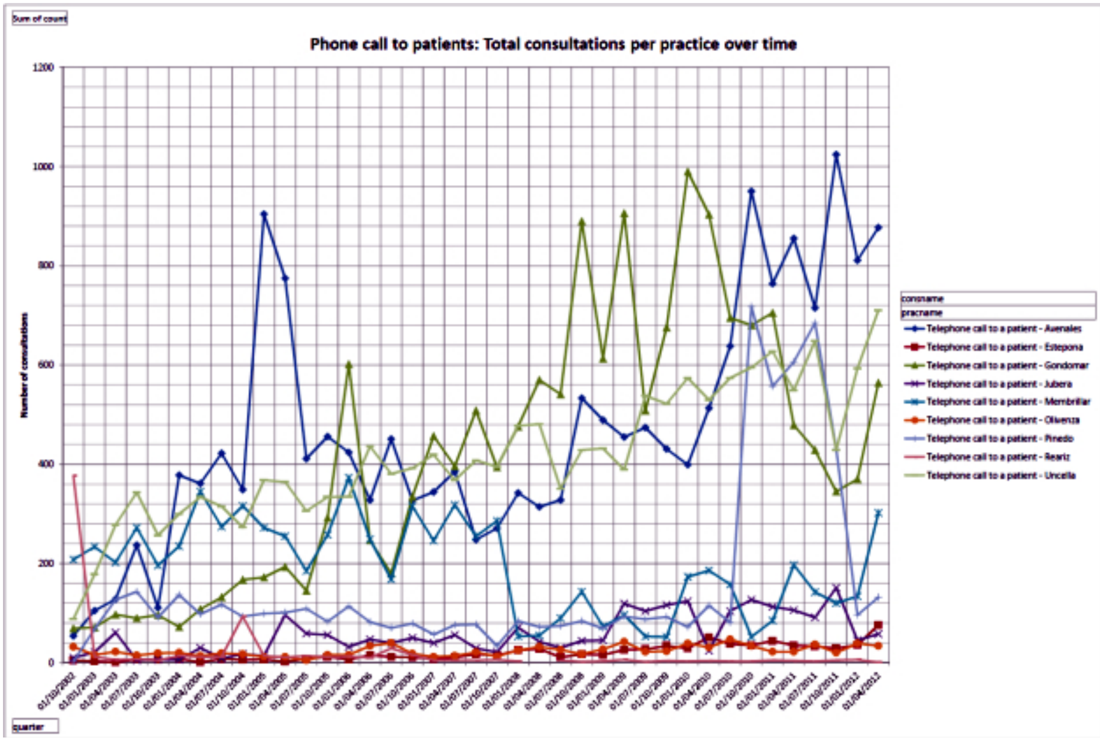

Figure 1. Example of stimuli graph

2.3 Procedure

Each participant was briefed on the experimental procedure and gave informed consent. they were each given the stimuli portfolio (the task dataset GPRD) and a separate sheet of A4 paper with the task scenario to complete. The participant sat on the left and the observer on the right, with the materials provided in two separate stacks initially covered by blank sheets of paper. One stack contained: database description documents, a graphical view of the tables and fields; then the database dictionary; and a Lookup table for the database codes. The second contained the graphs and reports and was initially ordered with general information on top, followed by increasingly specific groups of reports. Participants were instructed to concurrently think aloud while performing the task. The researchers observed the session throughout, recording field notes and auditory input using a voice recorder. Participants were told they could leave markings and notes on all materials and on separate sheets of A3 paper provided. They were also told that the data and reports they were working with were anonymised, but from real existing GP practices from the UK's East Midlands region. Finally, participants were told that they were not being tested on performance and validity of their predictions. The session involved one participant, lasted for about 1 hour and was audio-recorded.

3. Results

3.1 Novice vs. Experts behaviours

Thematic analysis was applied to all field notes and audio transcriptions**.** Two behavioural analysis approaches stood out as clearly identifiable throughout: the Scanning approach and the Sequential approach. Eight out of the 9 novice participants were observed switching between these, while all three experts exclusively chose the Sequential approach. These strategic differences are modelled in Figure 2.

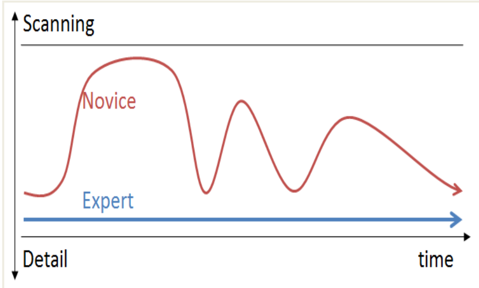

Figure 2. *Switching behaviour*

3.1.1 Scanning Approach

The scanning approach refers to all instances where participants spread out on the table all the information in a sub-group so as make them visible at-a-glance.

> "*well, when you've got so much data what I always do is separate it into themes*" (P02 – novice)

An example of this strategy is shown in Figure 3. Using this approach, participants could spot graphical features such as spikes or drops in time-series, as well as pick up and compare those graphs. It also led to them: losing their overview when they moved from one group of reports to another; and then searching for graphs they had seen in a previously-scanned group (but not knowing where they were within that group). Those who relied mainly on this strategy verbalised more interconnections of events within the data and their own pre-existing knowledge, possibly because they were visually aware of more linking points. They were also more prone to not noticing conflicts within the pre-made charts.

3.1.2 Sequential Approach

This approach involved going through the pile of graphs in the sequence as provided by the researchers. Participants either moved the charts from one pile to another or kept a few papers in their grasp until they decided to put them down with charts they have already read. Figure 4 shows how participants interacted with the stimuli in a sequential mode. This approach was mainly used by those who were closely investigating details of the graphs such as scaling or individual time series when more than one was present. The drawback of this strategy was that it proved to be very time consuming. Its advantage was that participants often switched back and forth between charts to investigate the more detailed hypotheses they had stated. Those using this method also wrote down more notes than those who did not.

> *"moving on from one to next, taking one at a time"* (P11 – expert)

3.2 Interview Themes

Several themes emerged from the follow-up semi-structured interviews. These are discussed below.

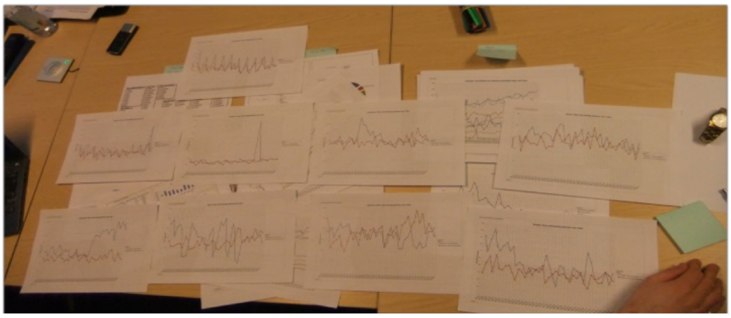

Figure 3. *Scanning over graph groups*

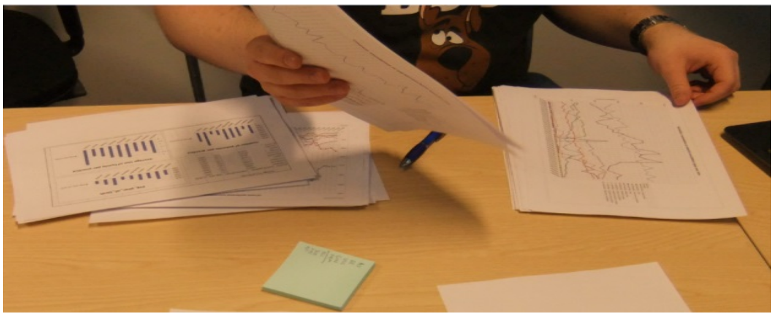

Figure 4. *Sequential approach*

3.2.1 Normalising data for easier processing

Four participants (all Novices) said they would have preferred to look at the data in percentage of total practice size, partly because they felt that would have made the task of comparing practices easier. All the experts also expressed this view very early in the task.

"I don't know, I think, I'd rather have to deal with percentages? I think it makes it easier to look and process" – P02 (Novice)

3.2.2 Too Big data

Six participants (5 Novices, 1 Expert) said they felt overwhelmed by both the number of graphs and the multitude of information presented within one graph, which then led to a loss of significance of a single time-series.

> "*I think I showed you one or two where there was maybe one surgery that there was importance with, but because it was overlapping with all the other surgeries it kind of lost it's meaning in a way*." - P01 (Novice)

3.2.3 Use of Tools

Eight participants (5 Novices, all 3 Experts) reported that additional tools would have helped with the interrogation task, indicating that some tools for human-data interaction already exist and are used by both experts and novices.

> "*Datasets, no. Tools, however: So, we've got graphs here, we can get the numbers out of oracle if I wanted to, if I had more time to construct, to work out what the queries are that I wanted to get the numbers out, but then need on top of that something like Excel to plot it.*" - P05 (Novice, when asked about additional datasets needed to help with the task)

3.2.4 Collaboration in Data Analysis

A common strategy of all experts (but none of the novices) was that they would consult with colleagues or subject area experts to gather either more missing medical or external background knowledge. Their reasoning for this was to either: find out what the actual need behind the "customer's" request was, to avoid possible follow-up requests; and to detail the request further, so that appropriate database queries could be formulated.

> "*I write the [SQL] code, try and get a good feel for the data, write the code, extract the data and then just play with it, trying to get a feel for it, erm, think about what I want to look at and just try to pull together a table and some charts and then perhaps then go and discuss it with people, "what do you think about this", and they might, because they've got more knowledge about the area, say "Oh, that's, the reason you've got that spike is because it's such and such" and that's something I don't know about.*" - P11 (Expert, talking about the process they employ when making sense of datasets)

3.2.5 Scanning and Grouping Strategies

When participants who used the scanning strategy were asked about the reasoning behind the different groups they had created on the table, they identified the three following strategies, presented in order of frequency of occurrence:

Theme: separate piles for illnesses, demographic information, and per-practice information

Visual characteristics: such as spikes or sudden steep drops (also preference for looking at numbers)

Medical / Temporal process: ordering from diagnosis-related information to test charts to therapy.

While some participants came up with very interesting predictions about the actual demands, these were not the main research objective due the lack of comprehensible criteria or definitions for what distinguishes good from bad predictions.

3. Conclusion

This case study presents first observations of how humans interact with over-abundant data, and how analysis approaches differ between expert medical data analysts and novices. The study outcomes span both behavioural and interactional observations across our two groups. For example, experts appear to cope with overwhelming aggregated data by iterating through each item in turn, while novices pass over graphs multiple times, using grouping. Novices used both the Sequential and the Scanning approach, while experts stuck to sequential analysis. The reason for this could be that experts sought in-depth knowledge about the dataset itself and its contents. However, the task itself could have been too broad and the time limit too short for experts to fully engage, although this is unlikely since participants received no reimbursement for the time spent doing the case study. Also, and like the heuristics applied in other domains e.g. optimisation problem solving (Kefalidou & Ormerod, 2014), novice participants appeared to use clustering behaviours to try to handle complex data. All the experts also tried to refine the original request and task. That this behaviour was only apparent in experts and not novices could be an indication of experts' prior, real-world job experience, as their "customers" very rarely can articulate and verbalise what it is exactly that they want when they request a certain analysis. Direct conversation with the customer and identification of their needs helps the expert, as it allows them to predict to some degree what other additional data the customer might need but has not included in the verbal formulation of the request. This saves the analyst both time and effort. The scanning strategy naturally comes with a higher visual exposure to possible linking points between graphs and enabled the development of a broad overview of the dataset. This provides further support to the notion that humans may employ visuospatial strategies when tackling complex data as indicated by previous research (Kefalidou & Ormerod, 2014). Collaborative features and toggling options between different process / visualisation modes appear to be critical for designing new tools to support data processing and decision-making. Based on the experiences and knowledge gathered in this case study in a laboratory environment using a fabricated task, a next step would be to extend our understanding of human-data interaction through more ethnomethodological approaches or across different domains. These contributions will then further enrich the growing body of knowledge in design for human-data interaction.

4. Acknowledgements

We gratefully acknowledge support from RCUK's Horizon Digital Economy Research Hub grant, EP/G065802/1.

5. References

Baber, C., Attfield, S., Wong, W. & Rooney, C. (2013). Exploring sense-making through an intelligence analysis exercise, In H. Chaudet, L. Pellegrin & N. Bonnardel (Eds.), *Proceedings of the 11th International Conference on Naturalistic Decision Making* (NDM 2013), 2013, Marseille, France.

Blum, J., Kefalidou, G., Houghton, R., Flintham, M., Arunachalam, U. & Goulden, M. (2014). Majority Report: Citizen empowerment through collaborative sense-making, In *Proceedings of 11th International ISCRAM 2014*, Pennsylvania, USA

Boyd, J. R. (1995). *The essence of winning and losing*. Retrieved 3/12/16: http://pogoarchives.org/m/dni/john_boyd_compendium/essence_of_winning_losing.pdf

Busby, J. S. and Hibberd, R. E. (2004). Artefacts, sense-making and catastrophic failure in railway systems. In *IEEE International Conference on Systems, Man and Cybernetics*, 7, 6198-6205.

Dyrks, T., Denef, S. & Ramirez, L. (2008). An empirical study of firefighting sense-making practices to inform the design of ubicomp technology. In *Sense-making workshop of the SIGCHI conference on Human Factors in Computing Systems* (CHI 2008).

Weick, K.A. (1995). *Sense-making in Organisations. London*, UK: Sage.

Haddadi, H., Mortier, R., McAuley, D. and Crowcroft, J. (2013), "Human-data interaction", *University of Cambridge, Computer Laboratory, Technical Report* UCAM-CL-TR-837.

Kefalidou, G. and Houghton, R. (2016). "How can a Horse be in two places at once? Group Sense-making using diverse and ambiguous information", In: Ergonomics and Human Factors 2016.

Kefalidou, G., Ormerod, T. (2014), "The Fast and the Not-So-Frugal: Human Heuristics for Optimization Problem Solving", *Proceedings of the Annual Meeting of Cognitive Science Society,23rd-26th July, Quebec, Canada*.

Schenk, K.D., Vitalari, N.P. and Davis, K.S. (1998), "Differences between novice and expert systems analysts: What do we know and what do we do?", *Journal of Management Information Systems*, Vol. 15 No. 1, pp. 9–50.

Shanteau, J. (1988), "Psychological characteristics and strategies of expert decision makers", *Acta Psychologica*, Vol. 68 No. 1, pp. 203–215