

Augmenting the Business Intelligence Lifecycle Model with Usability: Using eye-Tracking to Discover the *why* of Usability Problems

T.R. Beelders and J.E. Kotzé

University of the Free State, Bloemfontein, South Africa

beelderstr@ufs.ac.za

kotzeje@ufs.ac.za

DOI: 10.34190/EJISE.20.23.1.007

Abstract: Humans and machines generate data at an alarming rate but still struggle to analyse and process this data fast enough. In order to meet this rising demand, business intelligence (BI) software firms are constantly incorporating new reporting or analytical functionalities in their software. At the same time, in order to be used optimally, they must ensure these features do not overwhelm users. There has however recently been cases where organisations, even after significant monetary investments, fail to realise the full potential of BI tools, sometimes reporting less than 20% utilisation. The purpose of this study was to investigate possible causes of the low utilisation and test the inclusion of a BI usability process. The study uses an off-the-shelf tool as a matter of convenience but proposes that including an empirical evaluation process as part of the BI design and development could increase BI adoption and usage. A small feasibility study was conducted to determine whether eye-tracking is viable as a means of determining why the usability problems exist. Findings suggest that analysis of eye-tracking data, together with traditional usability measures offer insight into how users solve problems and use the BI tool. Therefore, it will be a valuable aid in the design and testing of BI tools before release to enhance the usability therefore and potentially increase adoption and usage rates and should be included as a phase in the design and development lifecycle.

Keywords: business intelligence; usability; BI tool; dashboard; eye-tracking

1. Introduction

Good decision making is paramount to the success of a business, particularly in the fast-paced modern world. This is of course dependent on the analyses and processing of data relevant to the business (Lim, Chen, and Chen, 2013). Humankind has long since identified the importance of gathering such data and has developed processes, techniques and tools for collecting and analysing intelligence to support decision making (Gilad and Gilad, 1986; Kinsinger, 2007). Decision support systems (DSS) are computerised systems which are often used to aid managers in making informed decisions (Arnott and Dodson, 2008), for example data warehouses (DW) and business intelligence (BI) solutions. These data-driven decision support systems (BI/analytics) are considered to be an integral part of the business world and were, according to Gartner (2018), the highest priority technology items for chief information officers (CIO's) in 2018.

According to Rouhani, et al. (2018), much research has been conducted on the adoption of BI solutions, with a focus on IT/IS adoption theories such as the Technology Acceptance Model (Davis, 1989) and the Technology-Organisation-Environment (TOE) perspective (Tornatzky and Klein, 1982). Even though earlier work based on these models played an important role in investigating the adoption of BI, the theoretical lens of these models did not take into consideration alternative approaches to ensure a successful BI system implementation. Yeoh and Koronios (2010) stated that “understanding critical success factors for BI is key for successful implementation of a BI system”. Their research found that “those organisations which address the CSFs from a business orientation approach will be more likely to achieve better results.” However, other than the absence of critical success factors leading to a high failure rate, the managerial usage rates of deployed BI tools has also been low (Arnott, 2010). This phenomenon, dubbed the “utilisation problem”, indicates that only 20% of users actually make use of BI technologies, although industry research reports that the real usage level is actually much lower – possibly approaching 11%. Despite ongoing investments in BI and their growing importance, not all firms are successful in developing their BI capabilities (Audzeyeva, and Hudson, 2016; Kulkarni, Robles-Flores and Popovič, 2017). There are also claims that BI usability is not well researched and needs more attention (Chang, Hsu, and Shiau, 2013). These factors form the basis of this paper which will show that usability testing is applicable to BI but that it only shows the level of effectiveness, efficiency and learnability and not why these levels are achieved. Eye-tracking is suggested as a means of determining the why and a small feasibility test is

conducted to determine whether the metrics will indeed shed light on why users might struggle with using a BI tool.

This paper starts by discussing some of the available BI and usability literature, highlighting the intersection of these two fields. Thereafter, the usability of business intelligence and information regarding the theory of eye tracking is discussed. The paper will conclude with the results of a feasibility study and finally propose a new model for the design of BI tools.

2. Background

2.1 Business Intelligence

For the purposes of this study the following definition of BI is accepted:

“BI can be defined as the process whereby data are gathered, stored and transformed into information through analysis (Shollo and Kautz, 2010), thereby providing decision makers with valuable information and knowledge by leveraging a variety of sources of data as well as structured and unstructured information” (Sabherwal and Becerra-Fernandez, 2011, 6).

While BI is a generic term, a BI process refers to the process through which an organisation obtains, analyses and distributes information and knowledge (Sabherwal and Becerra-Fernandez, 2011). The BI process will typically be supported by a BI solution, which in turn will use certain BI tools (Sabherwal and Becerra-Fernandez, 2011). BI solutions are deployed within an organisation and will utilise BI tools acquired by the organisation. The BI solution will draw upon the vast pools of data from existing data warehouses, transactional processing systems, structured as well as unstructured information to provide information to managers (Sabherwal and Becerra-Fernandez, 2011) using BI tools. BI tools are essentially a collection of decision support technologies for the enterprise aimed at enabling knowledge workers such as executives, managers and analysts to quickly make better decisions (Chaudhuri, Dayal, and Narasayya, 2011) based on empirical information (Grossmann and Rinderle-Ma, 2015). These tools should allow its users to leverage the best use of its data and provide summarisation and aggregation to determine (a) what happened, (b) why it happened, (c) what is likely to happen and (d) what just happened (Eckerson, 2003). BI tools are developed by BI vendors such as Tableau®, Microstrategy®, Oracle®, Microsoft® or IBM®. Based on the combination of BI processes, BI solutions and BI tools, business intelligence offers a potential means of harnessing the vast amount of data at the disposal of businesses to improve decision making and understanding of the environment.

An effective BI solution should be able to foster innovation in your organisation and simplify decision making, which in turn will drive business success (Olszak, 2016). In addition, the purpose of a BI solution is far more than a presentation layer on top of your data (Naudé as cited in ITWeb 2014). Managers must be able to make real-time decisions quickly and without having to read through extensive reports or study numerous screens. Therefore, the BI tool must provide quickly accessible actionable information in order to be effective and if it is difficult to use, this will not be achievable (Sabherwal and Becerra-Fernandez, 2011). Today, it is very difficult to find a successful enterprise that has not leveraged BI technology in their business (Llave, 2017). But not all firms are successful in developing their BI capabilities (Kulkarni, Robles-Flores and Popovič, 2017) and it stands to reason that the problem could lie with the use of the tool once it has been implemented into an organisation. The presentation capability of a BI tool is of particular interest since it is the point of contact between the BI tool and the user and therefore possibly the source of the low usage rate. Here, usability plays a critical role and since poor usability cannot be addressed with training (Sabherwal and Becerra-Fernandez, 2011), it must be corrected to make the tool more usable to increase the usage.

2.2 Human-computer Interaction and Usability

Human-computer interaction (HCI) focuses on the interaction of humans with computing technology (MacKenzie, 2013). The user interface is the conduit for the interaction (Dix, 2009) and it is vital that any new user interface be rigorously evaluated (Lazar, Feng and Hochheiser, 2017). The interface can be considered the overriding factor which determines the success or failure of an application (Abran et al., 2003). As previously mentioned, adoption of BI within organisations is high, yet there are high failure rates and low usage levels. This could be a consequence of the interfaces.

The International Standards Organisation (ISO) provides two usability standards, namely, “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” and “the capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions” (ISO, 1998, ISO/IEC, 2001).

From these, three roles can be identified that play an integral part in usability, namely users, tasks or goals, and the measurements of efficiency, effectiveness and satisfaction. The importance of learnability, as identified in the latter definition, must not be underestimated as it too plays an important role in the acceptance of software and will therefore also be considered a measurement for the purposes of this study.

User experience encompasses usability but extends to a broader, more encompassing concept, including aspects such as thoughts, feelings and perceptions and can include behavioural and physiological metrics (Tullis and Albert, 2013). This study will however stay within the field of usability as user experience is beyond the scope of the study as the study.

These measurements are seen as applicable within the domain of BI and can be used to address the shortcomings and/or problems that have been identified. The straightforward idea that usability is “being able to do the things you want to, not the things you have to” (Cato, 2001) may provide an answer as to why the end use of BI tools is lower than expected.

2.2.1 Measuring usability

The concept of usability, with the appropriate measurements, can be evaluated in order to determine how successful the interface is at its primary goal, namely facilitating the desired/intended communication. Effectiveness measures how well the user is able to achieve tasks using the system (ISO, 1998) as indicated by accuracy and completeness (Cato, 2001; ISO, 1998). The amount of resources used to complete a task is a measure of efficiency (ISO, 1998), and includes time, mental effort (Bevan and Macleod, 1994), the number of actions required to complete the task, “lostness” or a combination of task success and time (Tullis and Albert, 2013). Efficiency and effectiveness are objective measures, while satisfaction measures the subjective opinion of the user towards the system (ISO, 1998). Learnability measures both the time a user requires to become familiarised with the system and the extent to which that user can remember system functionality (Cato, 2001). Learnability is determined by periodically measuring the same efficiency and effectiveness metrics (Tullis and Albert, 2013) with the user repeating the same tasks.

2.2.2 Business Intelligence usability

BI projects are prone to high failure rates and of those that are successful only 10-20% of the intended users actually use the system (Arnott, 2010). Therefore, it is suggested that greater attention be given to the effective development and implementation of a BI system (Arnott, and Pervan, 2008) as a means to potentially minimise or at least lower the failure rate. Additionally, it is essential that the process of using these tools is better understood to leverage the potential of these tools and increase usage. Shollo and Kautz (2010) argue that too little work has been done on how the BI tools were used and how intelligence is acted upon in decision-making processes. What is known is that, from a managerial perspective, BI tools are often too complex or time-consuming to use and are cumbersome to customise for individual use (Sabherwal and Becerra-Fernandez, 2010). This study will attempt to bridge this problem and propose guidelines that will enable BI vendors to design BI tools that are more user-friendly, and provide information that is in a more usable format to ultimately facilitate better planning for the future.

Since this study will concentrate on the usability of BI tools, it falls within the scope of human-computer interaction research. As far back as 2011 Mayer, Winter and Mohr (2011) proposed six guidelines to improve user acceptance of a management support systems. However, these guidelines are not relevant to the current study and therefore there could still exist a gap in the evaluation of BI tool usability. Design guidelines for the development of dashboards to assist novice users in mapping data to visual constructs have been proposed (Smuts, Scholtz and Calitz, 2015). These include improved navigation and help features (Smuts, Scholtz and Calitz, 2015). A sustainability BI tool for higher education institutes included a usability step at the end of the process, once the product was already released (Scholtz, Calitz and Haupt, 2018). Furthermore, some studies have investigated the usability of business intelligence tools. Using a standard usability questionnaire, Wu, et al. (2019) could identify some usability issues with an interactive dashboard. While standard questionnaire may be applicable, business intelligence tools and dashboards are very different to other desktop applications. Hence,

a usability questionnaire specifically designed to evaluate a business intelligence application has been designed and tested for validity (Mousa, Raheem and Mohammed, 2019). This questionnaire will be applied to a finished product to evaluate the usability thereof and is thus applicable once the software is already in use.

For the purposes of this study the BI tool developed by Tableau® will be used to evaluate several performance dashboards.

2.3 BI Usability Guidelines

As an extension to standard usability guidelines, the unique nature and complexity of BI has led to the proposal of BI usability guidelines. These guidelines include the need for flexibility, in the sense that a user should feel in control of the application and the application should be customisable (Jooste, van Biljon and Mentz, 2014). Additionally, the system should be learnable to limit the load placed on the user's memory and to ensure that infrequent use does not prohibit effective and efficient use (Jooste, van Biljon and Mentz, 2014). Above all, the application should be effective, efficient, learnable and helpful, and the user should be in control of the use thereof (Jooste, van Biljon and Mentz, 2014). These are identical to standard usability metrics and therefore these metrics and testing techniques are applicable in the BI field. It has already been established that the uptake and usage rates of BI tools are low, hence there must be a reason behind this. Usability testing, which is applicable to BI tools, can be used to determine hindrances to the effective and efficient use of these tools.

2.4 Eye-Tracking

While usability metrics can serve to quantify how usable an application is, there is no means by which it can be determined why the application was not as usable as intended. If it can be established why a particular task could not be completed successfully or determine how a user approached a certain task, the usability can potentially be increased.

In particular, it suggests that eye gaze be evaluated in order to determine how the user approached a particular task. In other words, where they looked, how long they looked and what path they followed in order to complete the task. Analysis of the eye gaze will assist in determining whether the users saw the required widget in the application required to complete the task, but neglected to make use of it. In this case, one could extrapolate that the user saw the correct path or action but was unsuccessful in connecting the action and the task, thus necessitating a change in the interface. Conversely, the user may not have seen the required action at all which also necessitates a change in the interface, albeit of a different kind. Therefore, eye gaze will assist in explaining why the observed usability quantification is perhaps low.

2.4.1 Eye movements

In order to see an object, the eye must be moved to a position so that the fovea is positioned over the object (Rayner, 1998). This rapid, ballistic movement, during which visual sensitivity is suppressed (Rayner, 1998), is called a saccade (Gregory, 1996). Between saccades, the eyes exhibit periods of relative stability when a person focuses attention on an object. These fixations typically last between 200 and 300 milliseconds (Rayner, 1998).

2.4.2 Measuring eye gaze

In order to study eye movements, there must be a way to capture these movements. An eye-tracker is a device capable of capturing eye movements (Duchowski, 2007). In eye-tracking, the fixations are of importance as they give an indication of the intention and behaviour of the participant or user. In order to analyse the gaze behaviour, there are two qualitative visualisations which can be generated from the fixations. A heatmap is a colour overlay on any stimulus which was used during the study or test. Fixations are aggregated for all users for any particular stimulus. The warmer the colour of the overlay, the more (and longer) fixations occurred in that area and vice versa.

The second qualitative visualisation is a gaze plot. Since eye gaze alternates between fixations and saccades, a scan path can be constructed for each participant or user. This scan path shows saccades as straight lines between fixations, which are shown as circles or dots - the larger the circles, the longer the fixation. The number of the fixation is also shown which allows the order of fixations, in other words the order in which elements were looked at, to be extrapolated. The saccade between successive fixations gives the direction the eye travelled in, from the lower indexed fixation to the higher indexed fixation.

For more quantitative analysis, fixation and visit metrics can be analysed. An area of interest (AOI) is any area which is deemed to be of importance and is marked as an area on the stimulus. A visit is the total amount of time the gaze is present within the bounds of a particular AOI and can consist of multiple fixations. Once the gaze leaves the AOI bounds, the visit ends and the return of the gaze to that AOI is regarded as a subsequent, separate visit. Number of fixations and visits and fixation and visit duration are typical metrics which are analysed.

2.5 Usability and eye-tracking

Tracking and evaluating eye movements can assist in the evaluation of usability. Previous studies show that providing a task to the user changes the way they approach their visual search in terms of where they look, how long they may look and many other factors (cf. Castelano, Mack and Henderson, 2009; Borji and Itti, 2014). User testing by its very nature requires that representative tasks be given to the user to complete. Therefore, including eye gaze poses the inherent danger of changing the interaction behaviour, in terms of eye gaze. However, since the tasks are representative it could be argued that the tasks are in fact the actual use of the application.

Previous studies have evaluated the use of eye-tracking in usability studies (cf. Szymański, et al., 2015; Wang, et al., 2019) and have advocated the use thereof. For instance, the usability, in terms of experience and performance, of web pages was compared (Akpınar and Yesilada, 2015). Other studies include using eye-tracking in combination with other metrics to establish preferences regarding web page design (Goh, et al., 2013).

According to Ehmke and Wilson (2007), researchers often make use of eye-tracking in usability studies in an anecdotal way which is subject to the interpretation of the researcher. At that time there was no model available which could be used to extract meaning from eye-tracking data. In 2006, Armitage as stated in Ehmke and Wilson (2007) proposed a number of guidelines for the interpretation of eye-tracking data. These include that long fixations indicate interest or confusion, back-track saccades indicate possibly confusion, back and forth movements between elements can indicate confusion or distraction, the first fixation indicates an attention-attracting feature. While these are more concrete guidelines to interpretation and the studies discussed appear to use anecdotal interpretations, they were in fact applying these principles without explicitly basing them within a framework or model.

Ehmke and Wilson (2007) identified correlations between specific usability problems and eye-tracking patterns. For example, many short fixations across a particular page indicate that the user expects to find some sort of information there which is not present.

In a study (Calitz, Bosire and Lane, 2012) similar to the current study, eye-tracking was used to conduct a usability study on a BI dashboard. The study showed that eye-tracking can identify usability issues. Thus it is a valuable asset in BI usability. However, that study only used eye-tracking to assist in identifying usability issues, whereas the current study proposes to go further and uncover not only the issues but how and why they occur and eventually propose an addition to the BI development process pre-release to the software.

This study will advocate the use of eye-tracking as a supplementary tool to usability testing as a means of determining both how and why the BI tools have low adoption. Therefore, a small feasibility study, using eye tracking, will be conducted on an off-the shelf BI tool. For the scope of the feasibility study, heatmaps and gaze plots will be analysed and some statistical analysis of the metrics will be reported on. The inspection of the scan paths will assist in determining what the usability problems are and will also assist in determining whether cognitive compatibility is achieved with the tool. More in-depth analysis of the scan paths and the behaviour as it relates to the afore-mentioned guidelines are beyond the scope of the paper.

3. Feasibility study

3.1 Dashboards

In order to determine the feasibility of the proposed model, a pilot usability study was conducted using 3 dashboards. The dashboards were samples available in Tableau – a powerful end-to-end analytics platform. The three dashboards, while fairly simple, did increase in complexity. Next, four questions were compiled for each dashboard, these were based on questions managers would typically want to answer concerning their business in order to make informed decisions. Additionally, the questions ranged from basic graph reading to requiring a

small amount of navigation such as drill-down or a similar dashboard related activity. Following this, there were some usability questions aimed at determining the subjective opinion of the participants regarding the dashboards.

The order in which the dashboards were presented to the participants was counterbalanced.

3.2 Participants

Only 10 participants were tested as the process of collecting eye-tracking data is time consuming as it must be conducted on an individual basis. A convenience sample was used by requesting post-graduate students who were currently enrolled in a BI module to participate in the study. Therefore, the participants had background in business intelligence theory and had limited experience in working on dashboards within the controlled environment of the module.

3.3 Eye-tracker

A Tobii T120 eye-tracker was used to capture data. This eye-tracker captures data at a rate of 120Hz, which translates into a data point being collected every 8.3 seconds. In order to draw heatmaps and gazeplots, Tobii Studio was used as the analysis tool. Basic statistical data as well as raw data was exported from Tobii Studio.

3.4 Number of correct answers

Table 1 summarises the number of correct answers that were given for each dashboard.

Table 1: Number of correct answers for each dashboard

Number of correct answers				
	Question 1	Question 2	Question 3	Question 4
Dashboard 1 Coffee shop	10	8	10	10
Dashboard 2 Home sales	8	9	9	10
Dashboard 3 Oil reserves	5	6	5	0

As can clearly be seen, the majority of the participants answered all the questions correctly for the first and second dashboard. The final dashboard presented some challenges as 5 of the participants were unable to answer a single question correctly. Apart from the correctness of the task, other usability measures can be analysed to determine how usable the tools are as it has already been established via literature that these are applicable. For this study, only the proposed eye-tracking part of the usability tests will be analysed and discussed in order to highlight the usefulness of these metrics in gaining more insight into the execution of the task. In reality, the eye-tracking should be complemented with the standard usability measures.

3.5 Eye-tracking analysis

3.5.1 Dashboard 1

One of the tasks could be completed by simply looking at the product performance graph and correlating the colour of the required graph with the provided legend. All participants fixated on the correct area during the task, but only 70% looked at the legend and 60% at the legend text. Through observation, it appeared as though most participants simply calculated whether sales exceeded the projected budget by using the information which was displayed when hovering over the graph. This was later confirmed by their answers in the questionnaire as the majority referred to the sales exceeding the budget in terms of the figures given. During the task, one of the participants started calculating, saw the legend and immediately said "Oh that makes it so much easier".

A gaze plot of two participants (Figure 1) shows a vast difference in the number and spread of fixations. One participant had very few fixations while the other participant clearly experienced difficulty (looking at the

spread, direction and number of fixations) in locating the correct information and was perhaps overwhelmed by the amount of information presented.

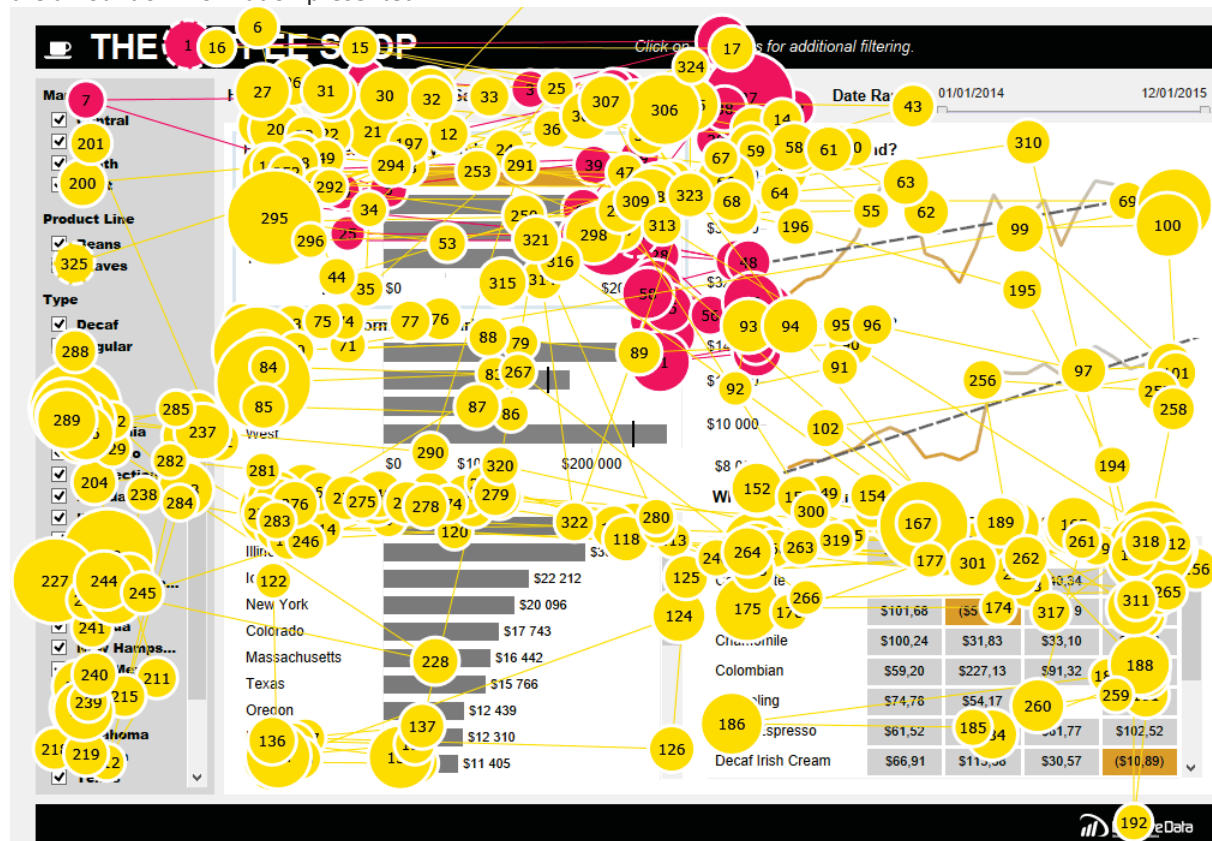


Figure 1: Dashboard 1 gaze plot for representative participants who did see legend

Two tasks required participants to filter data in order to show data for a specific state. While all participants were able to complete the task successfully, only two participants noticed the shortcut provided to quickly filter data while the remainder filtered on an individual basis which is a time consuming process. However for the next task all participants used the shortcut method for filtering, indicating learning. Average visit counts on available categories was down from 3.1 in the first filtering task to 2.5 in the second task and average fixation duration was also marginally less (0.35 versus 0.4). Therefore it appears that, in the short-term, the dashboard is learnable.

3.5.2 Dashboard 2

The first task on this dashboard could be answered by inspecting a legend below a graph. Similar to the first dashboard, the legend was looked at fairly quickly (after 3.89 seconds) but visited an average of 12.3 times which means the participants had to look at it repeatedly before realising what information was given there. The graph area was generally looked at first (on average within 1.79 seconds) which indicates that this is intuitively where participants expected to see the information required. Through observation, many of them hovered over the highest point on the graph, expecting this to give them the answer. When this failed, they continued to look for the answer elsewhere. The average fixation duration (as well as the average visit duration) was also the longest on the legend at 0.3 seconds which is not disproportionately high for an area which requires reading.

Two extreme behaviours were identified for this task (Figure 2), namely participants who quickly saw the information in the legend and those who experienced difficulty in determining how the information was being presented.

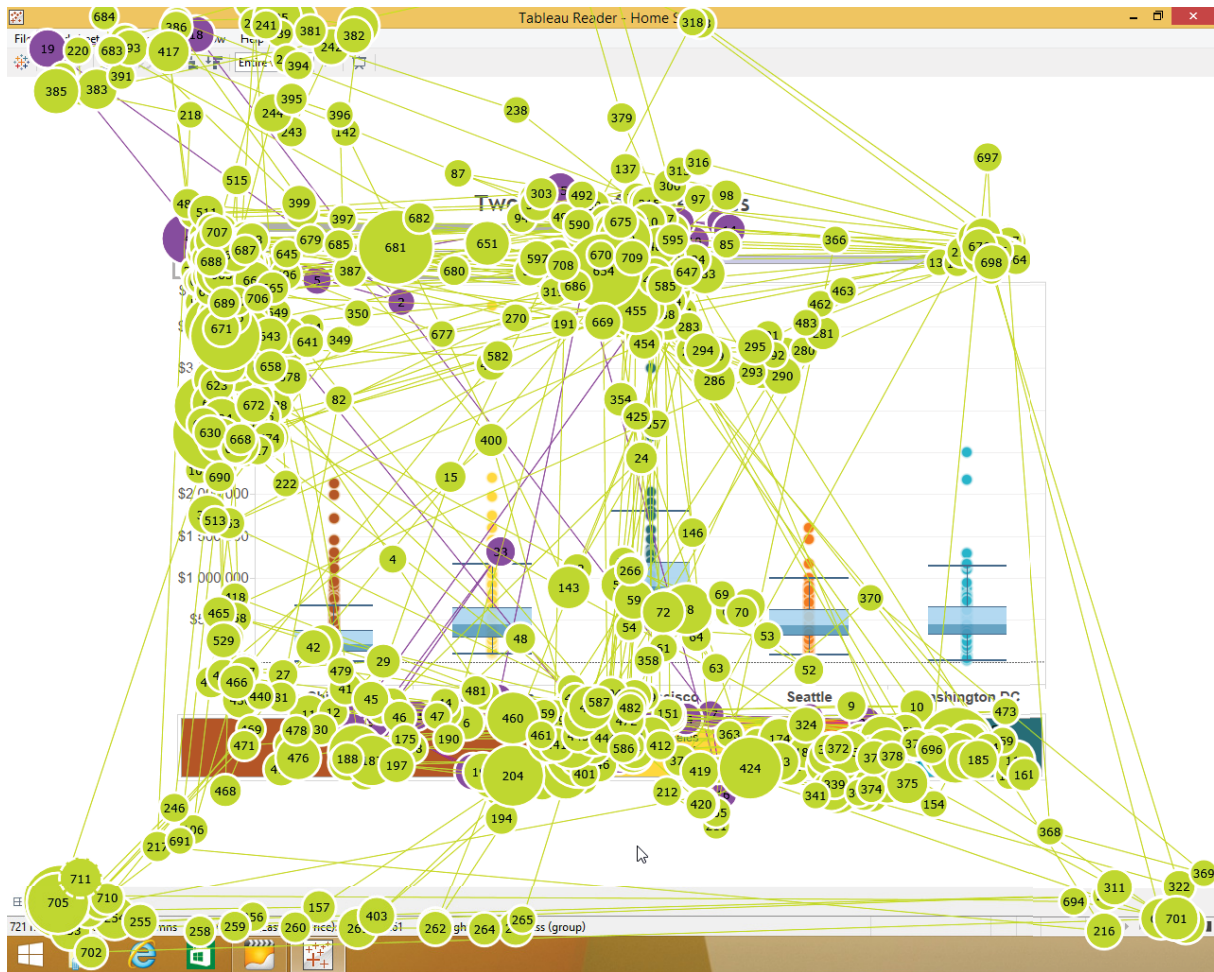


Figure 2: Gaze plot for single task on dashboard 2

The second task required participants to filter the data to show only that of San Francisco. There are numerous ways in which to do this, including double-clicking on the graph and clicking in the legend. Once again, the average time to first fixation was lowest on the graph (0.43 seconds) and 30% of participants clicked on the graph. Sixty percent of participants clicked on the legend and the remaining participant was unable to complete the task. The average number of visits was highest for the graph (9.9), followed by the legend (6.9) which indicates participants kept returning there – a good sign as these were the two areas which could be used to complete the task.

The next task could be completed by hovering over the highest point on the graph for a specific city. On average, the time to first fixation was lowest for the graph during this task but the visit duration was the longest and the number of visits was the highest. This once again indicates that the participants had to keep returning to the information in order to determine how to complete the task.

The pink gaze plot (with the fewer fixations) in Figure 3 is representative of nine of the participants since these participants all had few fixations and all concentrated on the relevant areas of the dashboard. Clearly, in this instance the information is obtainable exactly where the participants expect it to be and the answer can be obtained intuitively. In the case of the remaining participant, whose gaze is plotted below and who was the only participant who was unable to answer the question correctly, there was obviously a mismatch between what that participant was expecting to find and where the information was. While this is an isolated case, this type of gaze behaviour is very informative in terms of determining what the problem might be with an interface, if such a problem exists.

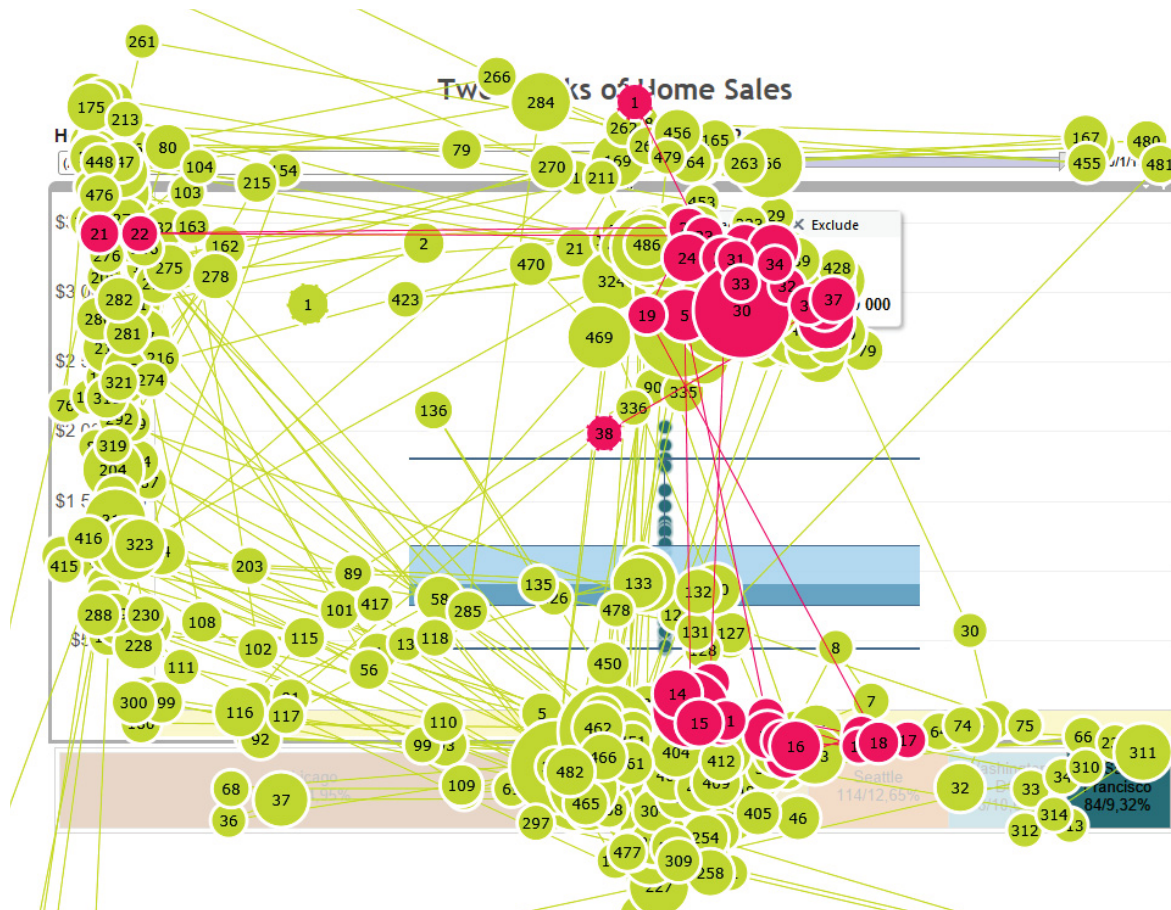


Figure 3: Gaze plot for single task of dashboard 2

3.5.3 Home sales – Task 5

The final task could be completed using knowledge obtained during a previous task on how to use the dashboard. All participants could answer this question correctly, including the participant who could not complete the previous task correctly. Again, this indicates that short-term learnability is sufficient for the dashboards.

3.5.4 Dashboard 3

The first two tasks of this dashboard will be discussed using more detail than in the previous dashboard discussion. The first task asked participants to identify the fourth largest oil importer for 2009. The data visualised included both importers and exporters of oil, which were distinguished from one another by shades of blue. The question therefore required participants to pay close attention to the fact that there were two shades of blue on the graph and either use the legend or the popup to determine what the difference was.

The majority of participants simply listed the fourth entry on the graph, with only half of the participants noticing that the graph contained a variety of data even though 70% of participants looked at the required legend during the course of the task. Three participants who looked at the legend still answered incorrectly and oddly enough one participant who did not look at the legend at all still answered the question correctly. It's possible this participant hovered over the bars in the graph in order to determine whether each was an importer or exporter and possibly also used the map to determine the import and export quantities of each.

The gaze plot below shows one participant (yellow) who completed the task incorrectly. In this case it can clearly be seen that the participant did not look at the legend (for illustration purposes the legend has been enclosed with a red rectangle) and stopped looking at the graph by the time they reached the fourth entry on the graph. The other gaze plot (blue) belongs to a participant who answered correctly and in this instance the participant very clearly looked at the legend. Interestingly it appears as though that participant read the graph until they

realised there were two shades in the graph and then located the legend which allowed them to jump directly to the fourth importer on the graph.

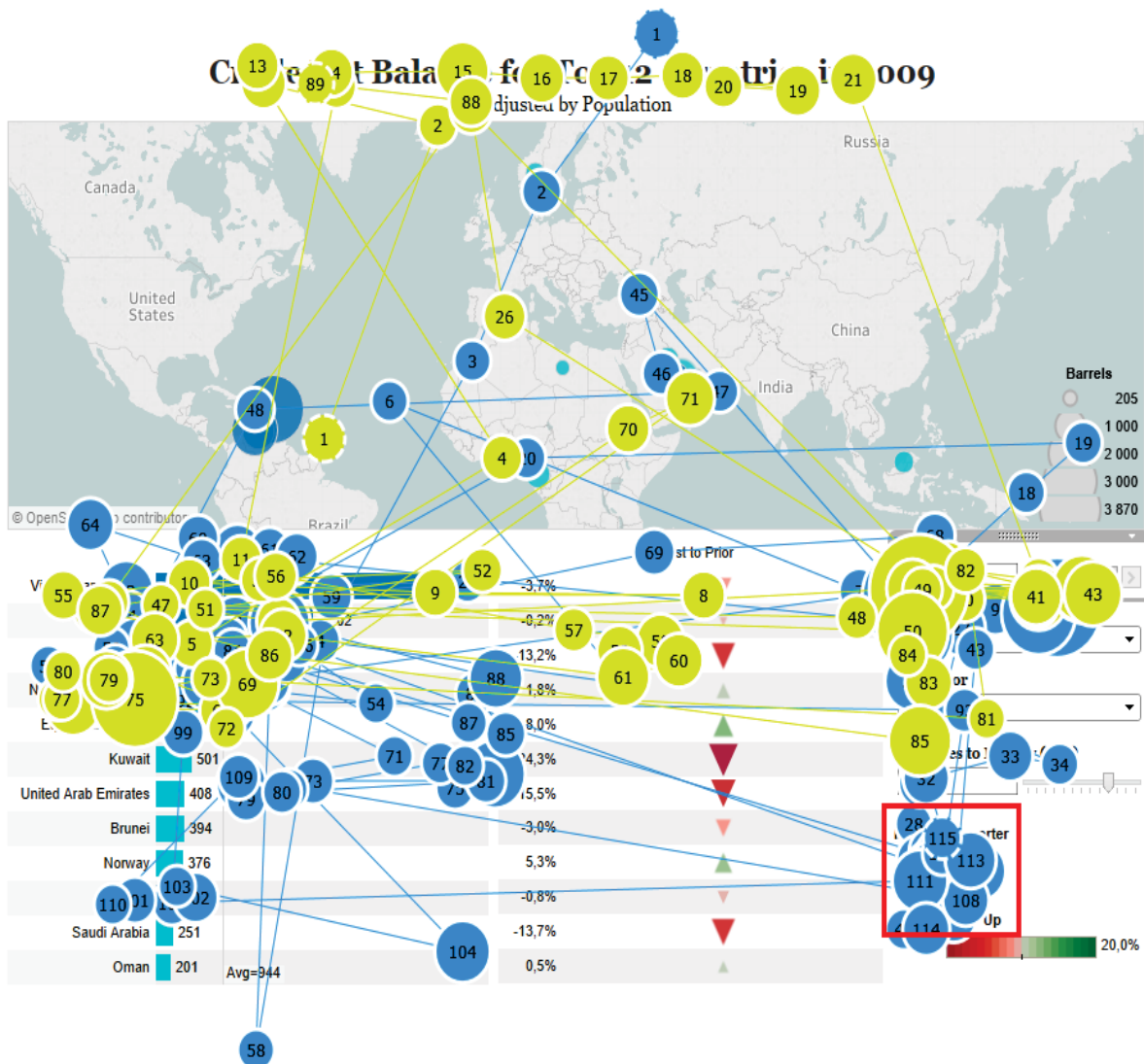


Figure 4: Gaze plot for single task of dashboard 3

The second task asked participants to determine who the largest net exporter of oil was. Immediately upon reading this question, some participants realised they had made a mistake in the previous question. Thereafter they paid much more attention to the graph and seemingly the legend as well. Eye-tracking metrics are similar between the two tasks for these AOIs although through observation it was clear the participants were now reading the legend and correlating it back to the graph – which was not evident with the first task. One participant commented that the shades of blue used were too similar and thus too difficult to distinguish between to the degree that one would realise it reflected different data.

Figure 5 shows the same 2 participants as with the previous task, both of whom now looked at the legend. The third gaze plot belongs to a participant who answered both questions incorrectly and clearly did not notice the different shades nor see the legend. It could be concluded that the shades are too similar and that the legend is situated too far from the graph, hence many participants simply did not see the legend until the follow up question alerted them to this fact.

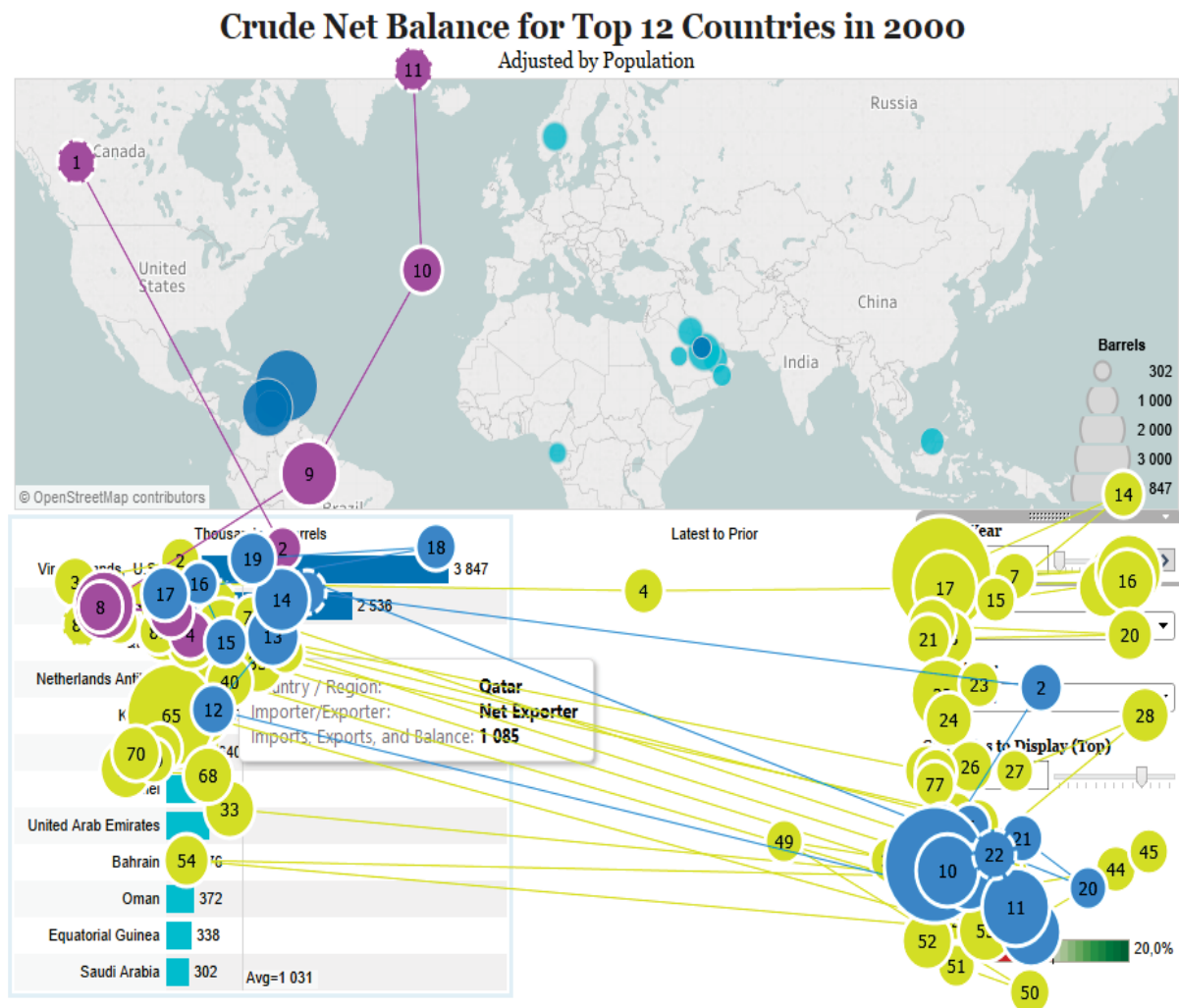


Figure 5: Gaze plot for task on dashboard 3

The easiest and quickest way to answer the next task was to filter the data using relevant dropdowns and legends. In this instance since a particular region had to be inspected, it might be expected that participants would click on the world map expecting it to function similarly to Google maps. However, only 20% of participants did so. On average, it took some time for participants to find the correct dropdown, within 5.82 seconds and after an average of 20.8 fixations. There were numerous visits to the region AOI, as to the other AOIs which were pertinent to the task including the world map. From the heatmap below it is clear that the largest proportion of the attention is below the map instead of on the map.

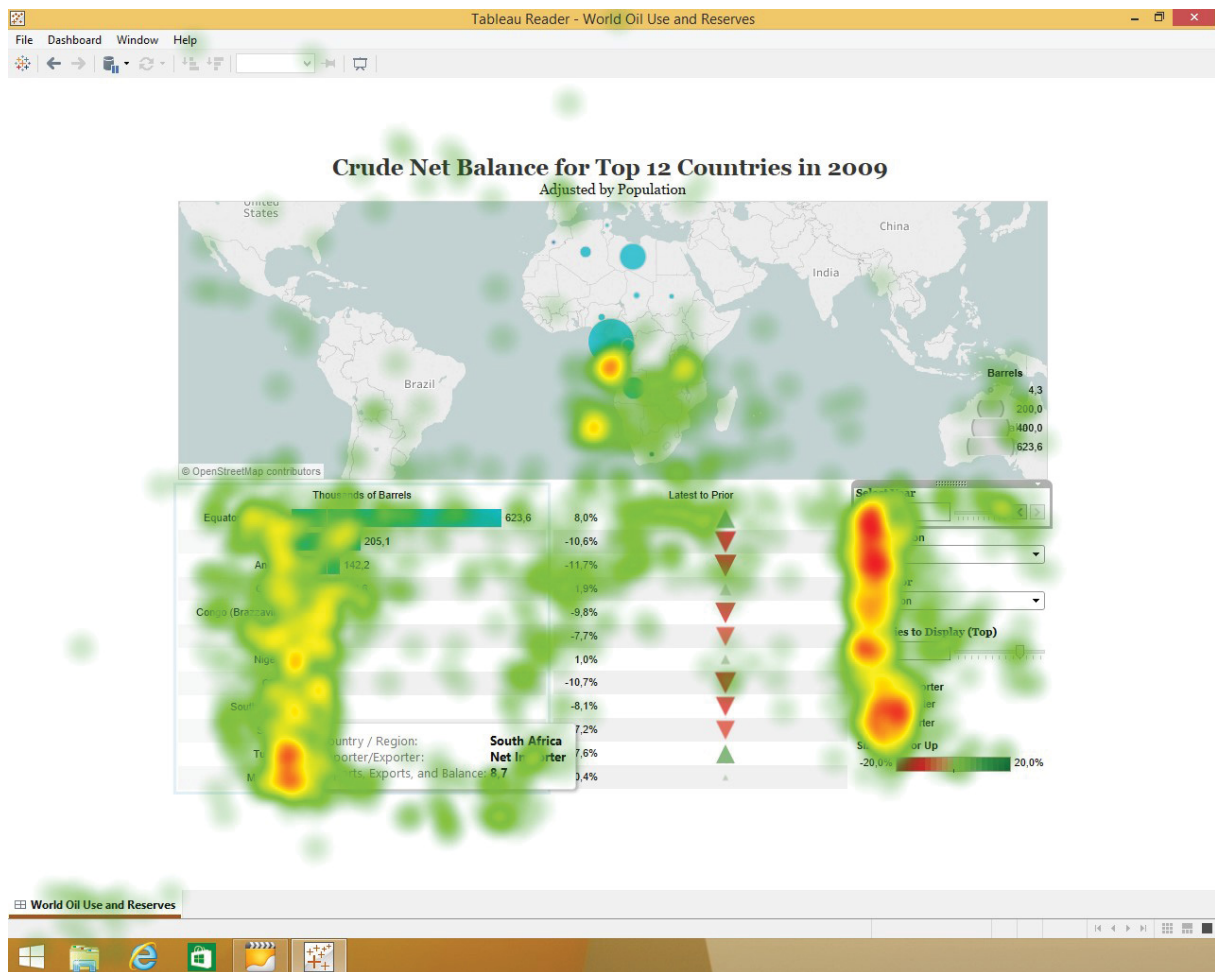


Figure 6: Heatmap for task on dashboard 3

As with the previous dashboards, the final task required a similar solution path to the previous task. In this instance, the region dropdown was, on average, the AOI that was looked at the fastest (within 2.75 milliseconds) and was also visited a number of times (average of 6.11 visits). Therefore, participants retained the information from the previous task. This is however, retention over a short period and it would be interesting to see if they retain the information over a longer time, particularly a time in which they do not actively work with the dashboard.

3.6 Gaze metrics across tasks

3.6.1 Dashboard 1

The average fixation duration of 8 participants increased from task 1 to task 2, 4 from task 2 to task 3 and only 2 from task 3 to task 4. Since task 3 and 4 were closely related and required the same actions, it could be expected that there would be a sharp decline in the fixation duration between these two tasks. Between task 1 and task 4, there were 5 participants who had a higher average fixation duration during the latter task. The increase in fixation duration between task 1 and 2 was not significant at an α -level of 0.05 ($t = 2.7$, $df = 9$), but that between task 2 and 3 was significant ($t = 0.1$, $df = 9$) with, on average, the fixations being longer in task 3 than in task 2. The difference between task 3 and 4 was also not significant at an α -level of 0.05 ($t = 2.1$, $df = 9$), but is at a less stricter significance level of 0.1.

In terms of number of fixations, 6 participants had more fixations during task 2 than during task 1, but all had less during task 3 than task 2. Only a single participant had more fixations during task 3 than during task 1. All but one participant had an increase in the number of fixations from task 3 to task 4, which is to be expected since the task required more reading and analysis than did task 3 (although the initial steps were identical). There were significantly more fixations during task 2 than task 1 ($t = 0.6$, $df = 9$) as well as in task 4 when compared with task 3 ($t = 1.8$, $df = 9$).

3.6.2 Dashboard 2

The total average fixation duration was lower during task 2 than task 1, with only 3 participants exhibiting an increase in fixation duration during task 2. The fixations were significantly less during task 2 ($t = 0.1$, $df = 9$). Six participants had longer average fixation durations during task 2 than task 3, a difference which is significant at an α -level of 0.05 ($t = 0.8$, $df = 9$). Fixations counts were also, respectively, less and more between tasks 1 and 2, and 2 and 3. The differences were also significant for both ($t = 1.6$, $df=9$) and ($t = 0.2$, $df = 9$).

3.6.3 Dashboard 3

The majority ($n = 6$) average fixation lengths were significantly longer in task 2 than in task 1 ($t = 1.3$, $df = 9$). Between task 2 and 3, the mean fixation length dropped but not significantly so. However, the increase in fixation lengths between task 3 and 4 was significant ($t = 0.4$, $df = 9$).

Fixation counts dropped between task 1 and 2, but not significantly. There was a significant increase in the number of fixations between task 2 and task 3 ($t = 1.5$, $df = 9$), with 8 of the participants exhibiting an increase in the number of fixations. Similarly there was an increase in the number of fixations between task 3 and 4, but in this case it was not significant ($t = 1.6$, $df = 9$).

3.7 Summary of analysis

Qualitative analysis of the eye-tracking data indicates that participants were able to learn to the dashboards and apply the knowledge to further tasks. However, participants often overlooked the valuable information given in legends as a means to assist them, preferring to search and sometimes calculate the answer themselves. This is contrary to the meaning of a BI tool which is meant to provide information at your fingertips without placing the burden of the analysis of the user. The high concentration of information could be a possible reason that the participants sometimes struggled to relate legends to given information or find the right information without having to refer back to the region a few times. Since these were fairly simple dashboards, the problem would be magnified on large and complex dashboards and this would be overwhelming to the user who cannot fathom how to interpret and use the data and eventually gives up using the tool. In isolated instances, the participants also neglected to see the shorter routes and screen widgets which would simplify the task, instead using a longer route to get to the answer. While this is not desirable, it is encouraging that they could eventually get the answer to all the questions. With extended use and if they become power users, they will undoubtedly make use of the "short cuts" provided for convenience. Subtle differences, such as different shades of blue were also largely not noticed by the participants and while they increase the aesthetic appeal to keep them similar it may decrease the visibility of the difference in data.

In general, both fixation length and fixation count increased as the exposure to the dashboard continued, oftentimes significantly so. This is indicative of an increase in cognitive processing. All tasks required some measure of reading, although in some instances it was more than others. This could therefore, in isolated instances, account for the difference in gaze behaviour.

The analysis shows that eye movements are a viable means of detected why users might be unable to complete a task or why they take long to complete a task. For example, since many participants did not notice some information they resorted to calculating it themselves. The consequence thereof may be that users feel using a BI tool is unnecessary since they must analyse and calculate results themselves regardless.

4. Proposed BI usability model

The fact remains that the adoption and usage rates of BI tools remains low despite the obvious advantages the tool could provide within the fast-paced business world. The complexity of BI tools may very well result in a violation of the mental model of the user. Additionally, the terminology used is analytical in nature and may not correspond to what a business person is expecting. Prior studies determined that usability studies are applicable within the scope of BI tools and can be used to determine whether users are able to make use of a tool effectively and efficiently. However, simply knowing that they are not usable does not solve the problem as to why they are not usable. Thus this study attempted to identify a way in which it could be determined why BI tools are not optimally used once deployed. For this, eye tracking was shown to be a useful means of determining how a user goes about solving a problem and how the BI tool is used. It can also be used as a measure of cognitive load. Thus, it can be considered to be a way to determine why the usability of a tool is low.

While the study performed the analysis on an off-the-shelf, commercially available tool, it would obviously be more advisable to perform these types of tests before the tool is released on the market. Therefore, the intervention must occur in the design phase of the tool. In so doing, the usability could be increased, thereby increasing the adoption rate and the eventual usage rate. If the user is able to perform their tasks effectively and efficiently and learn the software easily then the use thereof will not be so easily abandoned.

Therefore, a modified life cycle as depicted in Figure 7 is proposed, inserting a “BI Application Usability Evaluation” process prior the Deployment process. This process would feedback inputs from the Usability Evaluation back into the BI Application Design process.

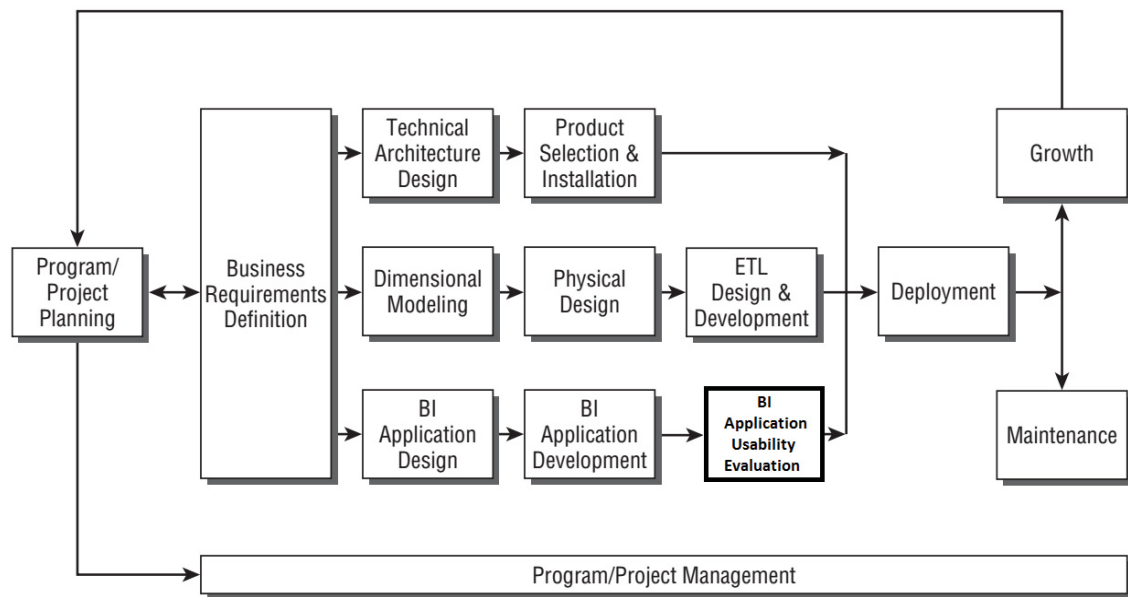


Figure 7: Business Intelligence Development Lifecycle

Source: Kimball and Ross (2002)

5. Conclusion

This paper recognised that BI tools, while widely acknowledged as powerful analytical tools, suffer from low usage rates after adoption, often leading to BI project failures. Since these problems are specifically encountered after adoption of a BI tool, the paper suggests that problem may be routed in the user interface, which is the contact between the end user and the BI tool. In order to test this argument, a feasibility study was conducted to evaluate the inclusion of such a process with limited results. From a technical perspective, an eye-tracking usability study was conducted to assist in explaining if usability problems were encountered and if so, why. The results were discussed as both qualitative and simple statistical analysis but it is clear from this small feasibility study that eye-tracking can play a vital role in evaluating the user interface (usability) of a BI tool. More detailed analysis, such as scanpath analysis top-down versus bottom-up gaze patterns, which is beyond the scope of this paper, can be used to give even greater insight into the usability of the tool and how users approach the use of the tool in order to complete a task. By examining the traditional BI developmental cycle, we suggest that a BI evaluation process be included as part of the traditional BI developmental cycle. This would lead to a direct intervention in determining whether BI usability is an issue in the use and adoption of BI tool. Moreover, the proposed life cycle model could be adopted by industry and the adoption and usage rates thereafter could be monitored to determine if there is any improvement. In future it could thus be investigated whether increasing the usage rates of the BI tool itself will decrease the failure rates of BI projects. Therefore the proposed inclusion of the usability testing, specifically including eye tracking to allow for more in-depth analysis of “why” there are any problems is considered to be a viable and a workable solution. In conclusion, it is therefore suggested that the testing occur before the BI tool is released in the market to improve the design and hopefully increase adoption and usage rates.

References

- Abran, A., Khelifi, A., Suryn, W., and Seffah, A. 2003. Usability meanings and interpretations in ISO standards. *Software Quality Journal*, 11, pp.325-38.
- Akpınar, E. and Yesilada, Y. 2015. "Old habits die hard!": eyetracking based experiential transcoding: a study with mobile users. In *Proceedings of the 12th Web for All Conference*, Florence, Italy, May 2015.
- Arnott, D. 2010. Senior executive information behaviors and decision support. *Journal of Decision Systems*, 19(4), pp. 465-80.
- Arnott, D. and Dodson, G. 2008. Decision support systems failure. *Handbook on Decision Support Systems* 1. Springer, Berlin, Heidelberg. pp. 763-90.
- Arnott, D. and Pervan, G. 2008. Eight key issues for the decision support systems discipline. *Decision Support Systems*, 44(3), pp. 657-72.
- Audzeyeva, A., and Hudson, R. 2016. How to get the most from a business intelligence application during the post implementation phase? Deep structure transformation at a U.K. retail bank. *European Journal of Information Systems*, 25(1), pp. 29-46.
- Bevan, N. and Macleod, M. 1994. Usability measurement in context. *Behaviour and Information Technology*, 13(1-2), pp. 132-45.
- Borji, A. and Itti, L. 2014. Defending Yarbus: Eye movements reveal observers' task. *Journal of Vision*, 14(3):29, pp.1-22.
- Calitz, A.P., Bosire, S. and Lane, E. 2012. Usability evaluations of ERP business intelligence dashboards. In *Proceedings of International Business Conference 2012*, Kenya.
- Castelhano, M.S., Mack, M.L. and Henderson, J.M. 2009. Viewing task influences eye movement control during active scene perception. *Journal of Vision*, 9(3):6, pp. 1-15.
- Cato, J. 2001. *User-centered web design*. London: Addison-Wesley.
- Chang, Y.-W., Hsu, P.-Y. and Shiau, W.-L.. 2013. An empirical study of managers' usage intention in BI. *Cognition, Technology and Work*, 16(2), pp. 247-58.
- Chaudhuri, S., Dayal, U. and Narasayya, V. 2011. An overview of BI technology. *Communications of the ACM*, 54(8), pp. 88-98.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly*, 13(3), pp. 319-40.
- Dix, A. 2009. Human-Computer Interaction. In L. Liu and M. Tamer Özsu, eds. *Encyclopedia of Database Systems*. Boston: Springer, pp. 1327-31.
- Duchowski, A.T. 2007. *Eye tracking methodology: Theory and practice*. 2nd ed. London: Springer-Verlag.
- Eckerson, W. 2003. Smart companies in the 21st century: The secrets of creating successful business intelligence solutions. [online] TDWI Report Series. Available at <http://download.101com.com/tdwi/research_report/2003BIReport_v7.pdf>. [Accessed 24 October 2013].
- Ehmke, C. and Wilson, S. 2007. Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers*. Lancaster, United Kingdom, 3-7 September 2007, pp. 119-128.
- Gartner. 2018. Mastering the new business executive job of the CIO. Available at <https://www.gartner.com/imagesrv/cio-trends/pdf/cio_agenda_2018.pdf>. [Accessed 23 October 23, 2018].
- Gilad, B. and Gilad, T. 1986. SMR Forum: Business intelligence, the quiet revolution, *Sloan Management Review*, 27(4), pp. 53-61.
- Goh, K.N., Chen, Y.Y., Lai, F.-W., Daud, S.C., Sivaji, A. and Soo, S-T. 2013. A comparison of usability testing methods for an e-commerce website: A case study on a Malaysia online gift shop. In *Proceedings of 10th International Conference on Information Technology: New Generations*, Las Vegas, United States of America, 15-17 April 2013.
- Gregory, R.L. 1966. *The Eye and the Brain: The Psychology of Seeing*. London: World University Library.
- ISO9241. 1998. *ISO 9241-11: Ergonomic requirements for office work with visual display terminals*. Berlin: Beuth.
- ISO/IEC 9126-1 2001. *Software Engineering, Product Quality, Part 1: Quality Model*, Geneva, International Organization for Standardization. Geneva. Available at: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=22749. [Accessed: 22 November 2012].
- ITWeb. 2014. BI landscape constantly evolving. Available at <http://www.itweb.co.za/index.php?option=com_contentandview=articleandid=135211:BI-landscape-constantly-evolvingandcatid=72>. [Accessed 10 June 2014].
- Jooste, C., van Biljon, J., and Mentz, J. 2014. Usability evaluation for business intelligence applications: A user support perspective. *South African Computer Journal*, 53, pp. 32-44.
- Kimball, R., and Ross, M. 2002. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modelling*. 2nd ed. New York: John Wiley and Sons, Inc.
- Kinsinger, P.C. 2007. The business intelligence challenge in the context of regional risk. *Thunderbird International Business Review*, 49(4), pp. 535-41.
- Kulkarni, U. R., Robles-Flores, J. A., & Popovič, A. 2017. Business intelligence capability: The effect of top management and the mediating roles of user participation and analytical decision making orientation. *Journal of the Association for Information Systems*, 18(7), pp. 516-41.

- Lazar, J., Feng, J. and Hochheiser, H. 2017. *Research Methods in Human-Computer Interaction*. 2nd ed. Burlington: Morgan Kaufmann.
- Lim, E.-P., Chen, H. and Chen, G. 2013. Business intelligence and analytics: Research directions. *ACM Transactions on Management Information Systems (TMIS)*, 3(4), pp. 17.
- Llave, M. R. (2017). Business intelligence and analytics in small and medium-sized enterprises: A systematic literature review. *Procedia Computer Science*, 121, 194–205.
- MacKenzie, I.S. 2013. *Human-Computer Interaction: An Empirical Research Perspective*. Burlington: Morgan Kaufmann.
- Mayer J.H., Winter, R. and Mohr, T. 2011. Utilizing user-group characteristics to improve acceptance of management support systems - State of the art and six design guidelines. In: *Proceedings of 6th International Conference, DESRIST: Lecture Notes in Computer Science 6629*. Milwaukee, 5-6 May 2011. Berlin: Springer, pp. 291-305.
- Mousa, A.H., Raheem, H.A. and Mohammed, N.T. 2019. An evaluation instrument (Q-U) for measuring the usability of business intelligence application. *International Journal of Engineering and Technology*, 8(1.5), pp. 456-60.
- Olszak, C. M. 2016. Toward better understanding and use of business intelligence in organizations. *Information Systems Management*, 33(2), pp. 105–23.
- Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), pp. 372-422.
- Rouhani, S., Ashrafi, A., Ravasan, A. Z., and Afshari, S. 2018. Business intelligence systems adoption model. *Journal of Organizational and End User Computing*, 30(2), pp. 43–70.
- Sabherwal, R., and Baccara-Fernandez, I. 2010. *Business Intelligence*. Hoboken: Wiley.
- Smuts, M., Scholtz, B. and Calitz, A.P. 2015. Information visualisation tools for novice users. In J. Steyn, J.P. Van Belle. eds. 2015. *Beyond Development. Time for a New ICT4D Paradigm? Proceedings of the 9th IDIA conference, IDIA2015*. Nungwi, Zanzibar, 8-10 November 2015. pp. 148-162
- Scholtz, B., Calitz, A.P. Haupt, R. 2018. A business intelligence framework for sustainability information management in higher education. *International Journal of Sustainability in Higher Education*, 19(2), pp. 266-90.
- Shollo, A., and Kautz, K. 2010. Towards an understanding of business intelligence. In *Proceedings of Australasian Conference on Information Systems: ACIS 2010*. Brisbane, Australia, 1-3 December 2010.
- Szymański, J., Sobiecki, J., Chynal, P. and Anisiewicz, J. 2015. Eye tracking in gesture based user interfaces usability testing. In N. Nguyen, B. Trawiński, R. Kosala. Eds. *Intelligent Information and Database Systems. ACIIDS 2015. Lecture Notes in Computer Science 9012*. Bali, Indonesia, 23-25 March 2015. Berlin: Springer.
- Tornatzky, L. G., and Klein, K. J. 1982. Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings. *IEEE Transactions on Engineering Management*, EM-29(1), pp. 28-45.
- Tullis, T., and Albert, B. 2013. *Measuring the user experience*. 2nd ed. Burlington: Morgan Kaufmann.
- Wang, J., Antonenko, P., Celepkolu, M., Jimenez, Y., Fieldman, E. and Fieldman, A. 2019. Exploring relationships between eye tracking and traditional usability testing data. *International Journal of Human-Computer Interaction*, 35(6), pp. 483-94.
- Wu, D.T.Y., Vennemeyer, S., Brown, K., Revalee, J., Murdock, P., Salomone, S., France, A., Clarke-Myers, K and Hankem S.P. 2019. Usability testing of an interactive dashboard for surgical quality improvement in a large congenital heart center. *Applied Clinical Informatics*, 10(05), pp. 859-69.
- Yeoh, W., and Koronios, A. 2010. Critical success factors for business intelligence systems. *Journal of Computer Information Systems*, 50(3), pp. 23–32.