1 METHODOLOGY FOR DATA COLLECTION OF THE CURRENT STATE OF BI ADOPTION

1.1 Introduction

A good way to measure the effectiveness of (the implementation of) a Business Intelligence (BI) tool has been a highly sought after topic in various journals and publications. People try to use surveys and analyze them with the balanced scorecard method or try a more direct approach with server logs and hard data analysis.

The thing is that, in general, there is no perfect method of measuring the implementation of BI in an organization. Granted, business will see costs drop over time due to easier reporting and a more efficient workflow, however when it comes to actually measuring the performance and adoption of BI, like for instance measuring the performance of a machine, things get more difficult due to personal affinity of the end user with the tool, the willingness to learn or just the personal work tempo.

It is in this setting this internship will try to find an objective method to analyze and increase the adoption of BI tools within a test group at ASML.

1.2 Available Methods for current BI adoption measurement

1.2.1 Cost Analysis

As mentioned before, one way to measure the performance of BI in an enterprise is to look at the costs it takes to make a report. In general the costs it takes to make a report in a program like Excel, using spreadsheets, will be higher due to the effort it takes to access data sources, format and display the requested results. Furthermore the annual cost to update the same report (e.g. add a new year or month column, get a new subset from the original data source, etc.) will be substantial as well.

For reports built in a BI tool the initial build cost may be higher, although this will be very unlikely. However since reports made in a BI tool are dynamically linked to the original data source, any changes in said source will automatically translate in a change in the report, making the maintenance cost much lower to nearly inexistent (Eckerson, 2013).

The only issue with the cost analysis approach is that, these measurements can only be made for reports that are used on a recurring timeframe, like for instance the monthly production overview or annual sales reports. These kinds of reports used to be made in a spreadsheet program like Excel within a certain timeframe. With the transference to a BI tool the increase or reduction in time spent can be measured, given there is some kind of time recording system in place.

This approach however does not work for one-time or occasionally used reports. These reports generally vary a lot in content making the time it takes for them to be finished hard to predict. Furthermore there is usually no predecessor made with another tool, so there is no baseline for reference.

1.2.2   BI portal usage

A second way to measure BI adoption in an organization is to examine the usage of the BI portals (if there are any). On these portals, people can find readymade BI reports, provided to them by the IT department or a colleague.

Analysis of the time accessed of these reports, combined with a click or navigation count may indicate how intensively these reports are used by the different end users (Eckerson, 2013).

This analysis however does not take into account that end users can copy reports from the BI portal and continue to work with them locally. An indication for this practice could be that an end user opened a report once for a very brief time and stopped using it afterwards. There is however no way to discriminate between actual copy actions and accidental clicks on wrong reports, which means this one time click analysis can be skewed and can result in the under- or overestimation of the amount of copied reports.

The portal usage analysis however can be combined with an analysis of the connection times to various data sources/data traffic from these sources in relation to the up time of BI tools (mentioned below). Normally a high uptime of a data source and a high uptime of a BI tool would imply that the tool is using the data source and reports are being made using this tool. If however the data source is queried intensely but no BI tool is up and running, this could imply that the report is being made using another program like Excel.

1.2.3   BI questions from the end users

A third way to measure the adoption rate of BI in a given department is to measure the amount of questions coming from the end users to the helpdesk or other IT divisions. Getting questions from the end users means the department is working on or has implemented BI in its workflows.

When the implementation date of BI for a certain department is known, the amount of questions or feedback can be an indication of the level of adoption for said department.

On the graph below, the different steps of transition are depicted. No questions or feedback usually means the end users are in phases 1 to 3. People are still in shock from the change in their work routine and there is no general interest in the new tool. Starting from phase 4 and onwards BI will be accepted and further

developed in the department and with these stages of transition, more and more questions will be asked at the key users and at the IT department

The only caveat with this kind of measurement is that, when the end-users reach transition level 7, they have acquired enough expertise to be self-sufficient in their problem solving, which evidently reduces the amount of questions and help requests as well. The content of the few questions that remain will however be very specific and involve niche topics, which a regular user cannot solve on his/her own.
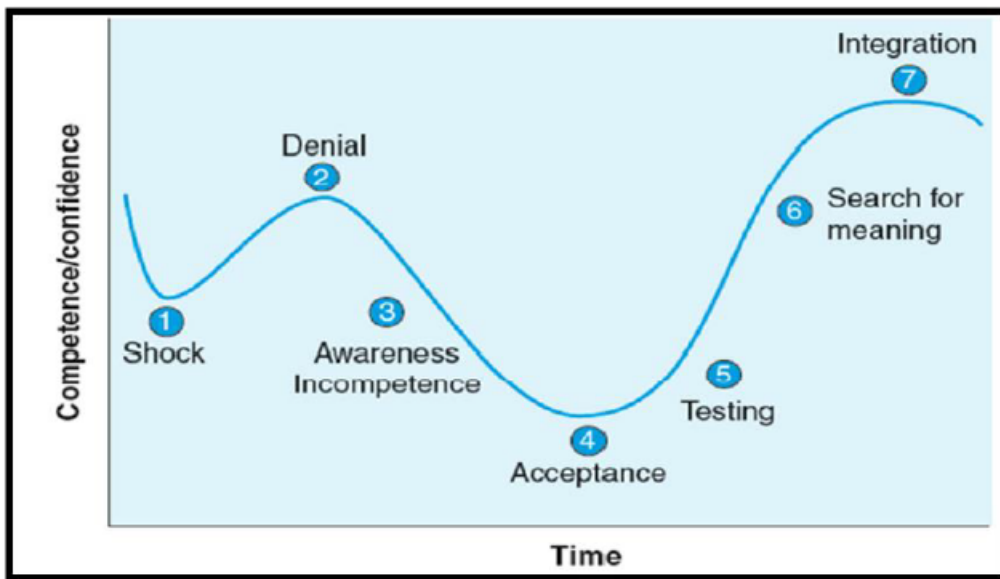

Figure 6-1: Transition curve - Adaptation of Adam et al. (1976) (Qlik, 2015, pp. 10)

1.2.4   Questionnaire

The next method to investigate BI adoption for a given group is through using the classic questionnaire. Specific questions are asked to discover if end users are susceptible to and are adopting the change required to adopt BI.

If filled in correct and truthfully, a questionnaire can render great insight in the psyche of the end user and into his willingness to change and experiment.

As with the other methods, this one too has some possible issues.

First the questions need to be really well written and very clear, so there is no interpretation possible (Owusu, 2017, pp. 9). If the interviewee has even the slightest doubt because more than one answer is possible, the question should not be used for further analysis. Therefore it's better to only use open ended questions with text fields for feedback on the questionnaire itself, but not for input data. Furthermore input text fields should be kept to a minimum as well and only be used

for number values. Text choices should be presented using a dropdown list with possible answers (Vannette, 2015).

Level of Agreement
- 1 – Strongly disagree
- 2 – Disagree
- 3 – Somewhat disagree
- 4 – Neither agree or disagree
- 5 – Somewhat agree
- 6 – Agree
- 7 – Strongly agree

Level of Agreement
- 1 – Strongly disagree
- 2 – Disagree
- 3 – Neither agree or disagree
- 4 – Agree
- 5 – Strongly agree

Figure 6-2: Seven vs. Five point Likert scale (Likert, 1932)

For most questions however a five or seven value answer, based on the Likert scale (Likert, 1932), should be used, ranging from strongly disagree to strongly agree. This type of answer will give the best results, given that the questions will be posed from the end user point of view (e.g. I use BI tools on a daily basis) (Vannette, 2015).

A possible issue could be that end users choose not to answer a question or cannot answer a question. For instance an employee on the sales department may have an account on a software tool due to his function level or job description, but it could be that among the different colleagues jobs are divided asymmetrically (e.g. someone does the report making, someone the text writing, etc. but they all work on the finance department). This means that one employee could be very proficient at working with a BI tool and therefore can fill in the questionnaire without any problems, whereas another employee has to stop after the second or third question (Lloyd, 2013).

A solution to counter this should be the addition of a Not Applicable (N/A) choice to the Likert scale answers.

The biggest caveat with this method is that the end users have to fill in the questionnaire truthfully. Needless to say that, if employees don't fill in the survey in a proper way, the results will be useless. Furthermore, there is no way of actually evaluating the gathered responses on validity. Also people can make mistakes while filling in the questions. In general an error margin of 10% is deemed acceptable for research purposes.

As a model for the questionnaire, the example from Chung-Kuang Hou (Hou, 2014, p. 587-588) is proposed. In this article three psychology methods are tested concerning user acceptance of business intelligence systems. The studied sample was the electronics industry in Taiwan. Since about one third of the reviewed companies profiled itself as being in the semi-conductor industry, the research can more or less be transposed to the situation at ASML.

The questionnaire is divided into sections concerning the different parts of the model. Specific questions are added to the questionnaire however to capture the specific situation at ASML.

1.2.5 Data collection from data sources and software uptime

A last possible method for data collection about the use and adoption of BI is to query the connection logs of the data sources and combine this with a review of software uptime (if possible).

Querying the connection logs will give an insight in how frequent and how intense the end users use the provided sources. If they copy a subset to use locally on their machine, there will be a onetime connection to a data source, if however people work directly with the, by IT managed and governed, data sources, there will be a lot of traffic visible in the connection logs (Eckerson, 2013).

The caveat with this investigation is you don't actually need a BI tool to connect to a data source. It is possible to connect using Excel and have the appearance of working the correct way. Therefore the proposition is made to combine the log query with a software uptime review. If for instance an end-user connects to a data source on a regular basis, but at those given times there is no BI tool active on that user's computer, it's reasonably to suspect the reporting is made using a non BI-tool.

If however an end-user connects to a data source once, but on multiple times during the day or week a BI tool is active but no connection to a data source is registered, it's safe to assume the user is working on a non-governed (partial) copy of a data source.

In both cases mentioned above it's possible errors will be made in the reporting due to outdated information or incorrect data use.

1.3 Selected methods for current BI adoption measurement

It is my belief that, in order to get a good result set, a survey combined with analyses of user logs (e.g. data source access, software uptime, etc.) is needed to get a good impression of the adaptation rate of BI at ASML at this point.

1.3.1 Survey

The survey will be made according to C.K. Hou as proposed above. To enrich the pool of questions the survey of Yoon, 2014 (Yoon, 2014, pp. 3761-3762) is used as well. These questions cover the same topics, sometimes with a slightly different intake. Also specific questions to tailor the questionnaire to the situation at ASML should be added. In the list of proposed survey questions in the next section,

questions in black are the original ones Hou used in his research. The questions in blue are those used from Yoon's study and the those in red are the ones specifically tailored to capture the situation at ASML.

The main focus of the specific questions is set to the facilitating conditions ASML provides for the end users of the BI systems. These conditions, according to the decomposed Theory of Planned Behavior (dTPD) (Taylor & Todd, 1995), make it easier for an individual to execute a certain behavior. These facilitating structures can augment or constrain the usage of BI software. Classic facilitating structures are money and time (which usually constrain behavior), but also education, problem solving and information sharing can aid or restrict the adoption of BI in the end-user workflow.

Some of the added questions concern 'perceived ease of use' and 'compatibility' as well. 'Perceived ease of use' was first mentioned by Davis (Davis, 1989, pp. 320) as one of two factors (the other 'perceived usability') that determine user adoption of IT. 'Perceived ease of use' tries to capture how comfortable the end user is using a given BI tool. The questions Hou proposed in his survey only concern the usage of BI tools. To capture possible roadblocks users experience using a BI tool at ASML some questions were added concerning the limitations of non-BI tools and the perceived speed of BI-tools.
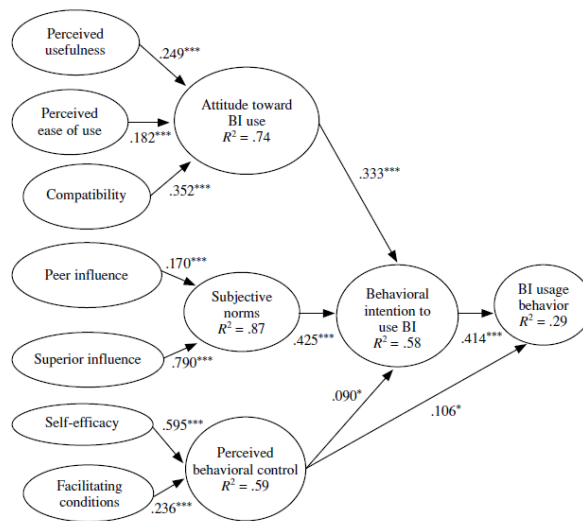

Figure 6-3: BI usage behavior model according to Hou (Hou, 2014, pp. 592)

Two questions tailored to the ASML situation were also added to the 'Compatibility' section of the survey. These questions are specifically designed to query how well the current BI infrastructure fits to the workflow of the end user, and how well the provided resources (e.g. data sources) are tailored to answer for specific business needs.

Furthermore some general questions were added as well, like e.g. 'In which department do you work' to try and find out if those factors have any result in the adoption of BI.

The last big addition of specific ASML tailored questions was to the BI usage behavior section. These questions specifically try to capture elements of the daily workflow, using (or not using) business intelligence tools. Some of the additions try to target the use of non-IT approved software for BI analyses, like for instance Excel spreadsheets or own custom built programs.

In this section also questions were added about the usage of reports and spreadsheets, in order to examine if end users use these IT provided resources in a proper way.

1.3.2   User log data

Since survey responses can be biased or untruthfully answered, survey information should be supplemented with factual data, like for instance the number of help requests from a given department to the helpdesk about BI, the access logs of the BI portal for a given department or user based access logs to data sources.

Using this data a wider view is possible on the usage of certain BI tools and of certain reports. Factual data can back up the findings of the survey. In case the data is in accordance with the survey results, an estimate can be made of the amount of time and effort it will take to implement solutions in the entire company. If however the data and survey results do not match or anomalies are detected, new information is created about possible subclusters in the total population of BI users.

This could mean that not all solutions work for the same department or the same user-type. It could be that, for instance, BI-architects are more served with online tutorials to learn new content, while BI-analysts prefer an instructor-led classroom, or vice versa.

1.4   Not selected methods

One method of the above mentioned methods could not be used for this survey. This is the cost analysis method, since it requires a baseline to start the analysis from. This means a measurement of the costs for reporting before or at the very beginning of the implementation of the BI solution. Since said baseline is not known for the investigated departments, this method can't be used for this initial survey.

Cost and time analysis however could be used for the follow up survey after the implementation of improvement measures to increase the adoption rate of BI given there is a measurement before the improvement process starts.

A cost and time analysis as a follow up survey could provide value and justification of the introduced measures to the business users.

1.5   Tools for data gathering

1.5.1   Survey question pool:

Below you can find the questions that are considered to be added in the survey. In black are the questions as they are used in the study of C.K. Hou mentioned earlier. In blue the questions mentioned in Yoon (Yoon, 2014, pp. 3761-3762) are matched to the subdivision Hou used in his study. In the subdivision title the construct Yoon used is mentioned in blue as well for reference.

In red are the questions added to the survey to capture the specific situation at ASML. The subdivisions indicate what part of the decomposed Theory of Planned Behavior, the theory that according to Hou predicts user acceptance of Bi the best, the question belongs to (Hou, 2014, pp. 592-593).

General:

- How old are you.
- In which department do you work.

Perceived Usefulness (attitude towards BI) / Extrinsic motivation:

- Using the BI system Improves my job performance.
- Using the BI system in my job increases my productivity.
- Using the BI system enhances my effectiveness in my job.
- Overall, I believe using the BI system is useful in my job.
- Using BI applications will improve my job performance.
- Using BI applications will enhance my job effectiveness.
- Using the BI application to be useful for my job.

Perceived ease of use (attitude towards BI):

- Learning to operate the BI system is easy for me in my job.
- My interaction with the BI system is clear and understandable.
- I believe it is easy to get the BI system to do what I want it to do.
- Overall, I believe the BI system is easy to use in my job.
- I use a different kind of software (e.g. Excel) for BI analyses and reporting, but notice this software has its limitations (e.g. slowing down on complex calculations, limited amount of data, etc.).
- At present, I think the performance of the provided BI tools is:.

Compatibility (attitude towards BI) / Compatability:

- Using the BI system is compatible with all aspects of my work.
- I think that using the BI system fits well with the way I like to work.
- Using the BI system fits my workstyle.
- Using BI applications will fit with the way I like to gather information at work.
- Using BI applications will be incompatible with how I like to do things in my work.

- Using BI applications will fit well with the way I like to interact with the system.
- Using the BI system I still need to manipulate (e.g. calculations, filtering, etc.) the provided data sources to be able to use them.
- The data manipulation I want to do is not possible in the BI system I use.

Attitude towards use (attitude towards BI) / Intrinsic motivation:

- Using the BI system is a good idea.
- Using the BI system would be a wise idea.
- I am satisfied with using the BI system.
- Using the BI system is a pleasant experience.
- I found using Bi applications to be enjoyable.
- The actual process of using BI applications is pleasant\
- I have fun with using BI applications.
- I use a different kind of software (e.g. Excel) for BI analyses and reporting out of habit.

Peer influence (Subjective norms) / Social influence:

- Peers who influence my behavior think that I should use the BI system.
- Peers who are Important to me think that I should use the BI system.
- My co-workers encourage my efforts to use BI applications on the job.
- My co-workers help me to further develop the skills to use BI applications.
- I use a different kind of software (e.g. Excel) instead of a BI system because a colleague taught me so.

Superior influence (Subjective norms) / Social influence:

- My superior who influences my behavior thinks that I should use the BI system.
- My superior to whom I report thinks that I should use the BI system.
- My manager views using BI applications as an important aspect of his/her job.
- My manager is supportive of efforts to apply newly acquired skills and knowledge about BI applications.

Subjective norms (Subjective norms) / Social influence:

- My peers confirm my knowledge and ability to make use of the BI system.
- My superiors confirm my knowledge and ability to make use of the BI system.
- My co-workers value using BI applications.
- My manager supports using BI applications.

Self-efficacy (perceived behavioral control):

- I feel comfortable using the BI system reasonably well on my own.
- I can use the BI system reasonably well on my own.
- I can use the BI system if there is no one around to help me.
- I prefer to make reports myself, so I can make then look as I want it.

Facilitating conditions (perceived behavioral control) / Situational constraints:

- The resources necessary (e.g. computer hardware and software) are available to me to use the BI system effectively.
- I have access to hardware, software and services needed to use the BI system.
- I will have the time necessary to strengthen my skills using BI applications.
- I have so much work, that it is difficult to apply newly acquired skills and knowledge about BI applications.
- My employer provides sufficient resources needed to try and use BI applications.
- My present job requires me to develop my skills and abilities of using BI applications.
- My company's policies and work rules allow me to participate in training.
- My company values employees learning and development activities.
- My company emphasizes the need for learning to their employees.
- Before creating reports I still need to manipulate (e.g. calculations, filtering, etc.) the provided data sources to be able to use them.
- The data manipulation I want to do is not possible in the BI system I use.
- My department uses a BI system.
- I received enough training about the available BI systems in my department and their usage.
- I know where I can find help if I have a Business Intelligence related question.
- What is your favorite way of finding help (IT SharePoint, IT direct chat, Colleague or peer, Big Data & Analytics homepage, Google).
- In my department, there is a colleague who I can ask for help concerning BI systems.
- What method would you prefer for BI-tool training (Instructor led classroom, Online video, Online self-paced instructions, webinar, asking a peer or colleague for help).
- Prior to working with BI systems at ASML, I had some experience with BI systems.
- I know how to find the data sources provided by IT.
- I know how to connect to the data sources provided by IT.
- It would be easy if I could reuse reports made by colleagues.

Perceived behavioral control (perceived behavioral control) / Complexity:

- I have the knowledge and ability to make use of the BI system.
- Using the BI system is entirely within my control.
- There is a clear and understandable process regarding how to use BI applications.
- Using BI applications will require a lot of effort.
- Using BI applications will be difficult for me.

Behavioral intention to use BI (sum attitude, norms and behavioral control) / Adoption intention:

- I intend to use the BI system more in the future.
- I want to use the BI system in the future.
- It is likely that I will use the BI system in the future.

- Assuming I have continued access to BI applications, I intend to adopt it in my work.
- Assuming I have continued access to Bi applications, I am willing to change my work activities to use the tool in my work.

BI usage behavior (results from intention) / Relative Advantage:

- How much time you spend each week using the BI system
- Using the BI application will enhance my efficiency in gathering information.
- Using BI applications will make it easier to gather information.
- Using BI applications will increase my effectiveness to gather information.
- Using BI applications will increase the quality of information that I gather.
- Please specify how much you use a BI-tool (e.g. Spotfire, SAP BW + BO, OBIEE), as a percentage based on an average workday At present, how much time each week do I spend on reporting.
- At present, I use a different software (e.g. Excel) for visualizing data.
- At present, I use a different software (e.g. Excel) for analyzing data.
- I use reports provided to me by the IT department.
- I sometimes download reports made by IT and continue to work with them locally on my computer.
- In my department the same reports are reused on a periodic scale (e.g. weekly, monthly, yearly).
- What BI tool do you use the most?
- Could you briefly specify what you use the BI tool for?
- Could you briefly specify what BI functionalities you use in Excel?
- Could you specify the usage ratio of Excel versus Bi-tools for reporting on an average workday?

1.5.2   Selected questions for the ASML survey

Of the many questions listed in the pool 40 were selected to be included in the actual survey. The majority of the not selected questions from the pool were redundant questions specifically tailored to verify if the responses to two or more similar questions generate a comparable result. Hou proved in his study all the questions resulted valid responses using the calculation of Cronbach's Alpha and Composite Reliabilities.

Yoon also proved the proposed questions resulted in a valid set of answers. Only one alteration was made to the model because 'Relative advantage' and 'Compatibility', although conceptually different, were viewed identically by the respondents. Therefore he decided to combine both in a single construct, instead of keeping them apart (Yoon, 2014, pp. 3762)

Keeping the same amount of similar questions to prove the accuracy of the responses is however not in the scope of this project, so the amount of redundant questions was significantly lowered. Only on those parts of the survey where the user's personal inclination and self-perceived success at using BI is tested, redundant questions were kept in order to verify the results.

Number	Question	Response type
1	In which department do you work?	Dropdown – single select
2	Please specify how much you use a business intelligence tool (e.g. Spotfire, SAP BW + BO, OBIEE, SAP Analysis Office), as a percentage based on an average workday.	Slider – percentage
3	What business intelligence tool do you use the most?	Select box list – multi select
4	Can you briefly specify what you use the business intelligence tool for? (Leave blank if N/A)	Textbox – open question
5	At present, I use different software then the tools mentioned in question 3 for visualizing or analyzing data.	Likert – 7 options with added N/A
6	Can you briefly specify what software and what BI functionalities you use it for? (Leave blank if N/A)	Textbox – open question
7	Using business intelligence in general is a good idea.	Likert – 7 options with added N/A
8	I have fun with using business intelligence applications.	Likert – 7 options with added N/A
9	Using business intelligence in my job increases my productivity.	Likert – 7 options with added N/A
10	My co-workers value using business intelligence applications.	Likert – 7 options with added N/A
11	My manager views using business intelligence applications as an important aspect of my job.	Likert – 7 options with added N/A
12	Using business intelligence applications will improve my job performance.	Likert – 7 options with added N/A
13	I believe it is easy to get a business intelligence tool to do what I want it to do.	Likert – 7 options with added N/A
14	Overall, I believe a business intelligence tool is easy to use in my job.	Likert – 7 options with added N/A
15	At present, I think the performance of the provided business intelligence tools is.	Likert – 7 options with added N/A
16	If applicable, could you specify what is going wrong with the current setup of business intelligence tools? (Leave blank if N/A)	Textbox – open question
17	I think that using business intelligence tools fit well with the way I like to work.	Likert – 7 options with added N/A
18	I feel comfortable using business intelligence tools reasonably well on my own.	Likert – 7 options with added N/A
19	I know how to find and connect to the data sources provided by IT.	Likert – 7 options with added N/A
20	Before creating reports I still need to manipulate (e.g. calculations, filtering, etc.) the provided data sources to be able to use them.	Likert – 7 options with added N/A
21	The data manipulation I want to do is not possible in the business intelligence system I use.	Likert – 7 options with added N/A
22	I use a different kind of software (e.g. Excel spreadsheets) for analyses and reporting out of habit.	Likert – 7 options with added N/A
23	I use a different kind of software (e.g. Excel spreadsheets) instead of BI because a colleague taught me so.	Likert – 7 options with added N/A

Number	Question	Response type
24	I use a different kind of software (e.g. Excel spreadsheets) for analyses and reporting, but notice this software has its limitations (e.g. slowing down on complex calculations, limited amount of data, etc.).	Likert – 7 options with added N/A
25	Could you specify the usage ratio of Excel spreadsheets versus business intelligence tools (e.g. Spotfire, BW/BO, OBIEE, SAP Analysis Office) for reporting on an average workday?	Slider – percentage
26	I use reports provided to me by the IT department.	Likert – 7 options with added N/A
27	I prefer to make reports myself, so I can make then look as I want it.	Likert – 7 options with added N/A
28	I sometimes download reports made by IT or a colleague and continue to work with it locally on my computer.	Likert – 7 options with added N/A
29	I sometimes see reports that are very similar to reports I've made (i.e. the content of the report)	Likert – 7 options with added N/A
30	I have access to hardware, software and services needed to use required business intelligence tools for my job.	Likert – 7 options with added N/A
31	I received enough training about the available business intelligence systems in my department and their usage.	Likert – 7 options with added N/A
32	There is a clear and understandable process regarding how to use business intelligence applications	Likert – 7 options with added N/A
33	I have so much work, that it is difficult to apply newly acquired skills and knowledge about business intelligence applications.	Likert – 7 options with added N/A
34	I know where I can find help if I have a business intelligence related question.	Likert – 7 options with added N/A
35	What is your favorite way of finding help (IT SharePoint, IT direct chat, Colleague or peer, Big Data & Analytics homepage, Google).	Dropdown – single select
36	What method would you prefer for BI-tool training (Instructor led classroom, Online video, Online self-paced instructions, webinar, asking a peer or colleague for help).	Dropdown – single select
37	I have the knowledge and ability to make use of the provided business intelligence tools.	Likert – 7 options with added N/A
38	Using business intelligence tools is difficult for me.	Likert – 7 options with added N/A
39	Assuming I have continued access to a business intelligence application, I am willing to change my work activities to use the tool in my work.	Likert – 7 options with added N/A
40	Considering using business intelligence tools, what would help you the most for optimizing your own workflow (e.g. more education, a different tool, changes in existing reports, etc.) or do you have any other remarks concerning the usage of business intelligence tools?	Textbox - open question

Table 6-1: Overview of the questions posed in the survey and their response types