DATA PRESENTATION FORMATS AND DECISION MAKING EFFECTIVENESS

AN EXPLORATORY STUDY

BY

TROND VEGARD JOHANNESSEN

MAY 2015

ABSTRACT

The purpose of this thesis is to enhance the understanding of how data presentation formats may affect managerial decision processes.

Compared to previous studies in this area, I have examined decision makers' need for data presentation formats in tasks that are more similar to real-life managerial tasks. I have focused not only on the *outcome* of a decision process, but also on *how* presentation formats may support decision makers' mental representation of the task and their decision process. In addition to studying the effects of either tables or graphs, I have included a combined presentation format. Furthermore, I have allowed the subjects to use decision aids, such as MS Excel.

In two experimental studies, I have explored how MBA students used graphs, tables and both presentation formats to solve two tasks varying in complexity and uncertainty. In Study 1, I used verbal protocols to gain insights into how the subjects applied the presentation formats. In Study 2, I also used eye tracking to complement the verbal-protocol data.

I found that decision makers need both presentation formats. Tables provide a basis for calculations to facilitate the integration of decision variables and result variables, which is important for effective decision making. Graphs provide an overview of the relationships among variables. Particularly XY graphs were important to enhance the decision makers' understanding of causal relationships among variables in the early phase of the decision process. Line graphs were primarily used to check whether the development of the result was as expected.

My findings have implications for designers of information systems. Designers should include both tables and graphs in computerised systems to support managerial decision processes. Furthermore, data presented on the web or in enterprise systems should be easy to export to a spreadsheet format so that the data can be used for additional analyses.

My research has also implications for managers. Managers using spreadsheets in their analyses should know how to programme formulas correctly, among others how to build dynamic models using relative cell references. My research has illustrated how bad spreadsheet skills may result in calculation errors that may seriously damage the outcomes of decision processes.

To Tuva and Mathias

ACKNOWLEDGEMENT

This thesis is the result of an exciting and challenging journey, upon which many people have contributed and given their support.

First and foremost I want to thank my main supervisor, Professor Anna Mette Fuglseth, whose guidance, support and encouragement has been invaluable, both as regards this PhD project and as a colleague. I am deeply grateful for all her contributions of ideas, knowledge and insight into the research area, and for her ability to see the overall picture while at the same time having a focus on every detail. Her thorough feedbacks has stimulated my reflection and her supervision has enhanced my development as a researcher. I would also like to thank my co-supervisor, Professor Kjell Grønhaug, for his valuable feedback on my manuscripts, in particular the constructive advices on structural and methodological issues, and for always being positive and encouraging.

Thanks to colleagues at the Department of Strategy and Management for valuable inputs to my work. A special thanks to all fellow doctoral students for socialising and academic support. Heidi Buverud, Ann-Kristin Elstad, Jan Vidar Håtuft, Siv Skard and Randi Vindegg have been true friends whom I could turn to for comments, advice and moral support. I want to thank the administrative staff; Elaine Pettersen, May-Britt Rød, Liz-Beth Lindanger, and particularly the office manager Paal Fennell, for help with administrative matters. I am particularly grateful to the head of the department, Paul Gooderham, for his encouragement during the process. I thank NHH Norwegian School of Economics for funding my PhD project.

Finally, I would like to express my deepest gratitude to my family for their continuous love and support. A special thanks to Fredrik for your patience, understanding and help throughout this process, and to Tuva and Mathias for bringing so much joy into my life.

Bergen, May 2015

Trond Vegard Johannessen

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICIES

1 INTRODUCTION

Data presentation formats and their effects on problem solving is an important topic in information systems research; see, for example, reviews by Kelton, Pennington and Tuttle (2010) and Dilla, Janvrin and Raschke (2010). Developments within information and communication technology have increased decision makers' possibilities to search for, collect, organise and analyse data. On the one hand, this development has increased decision makers' possibilities to enhance the understanding of their task environments and thus improve their decisions. According to McAfee and Brynjolfsson (2012), data-driven management decisions are more effective than decisions based on intuition. On the other hand, more data increase the load on the decision makers' cognitive capacity, which is a limited resource (Anderson, 2013). Therefore, understanding of the relationships between data presentation formats and decision processes may help designers of information systems design screen displays that effectively support managers in their decision processes.

Numerous studies have examined presentation format issues in connection with factors such as characteristics of the task and of the individual, see Chapter 2. Most of this research has focused on the relationship between presentation formats and the decision *outcomes* in rather simple tasks. However, the environments that most managers must handle today are characterised by complexity and uncertainty.

The purpose of this thesis is to enhance the understanding of how data presentation formats may affect managerial decision-making in real-life tasks, or at least tasks that are more similar to real-life tasks. In particular, I focus on enhancing the understanding of how presentation formats may influence decision makers' mental representation and processing of a task. By doing so, I am responding to a call for research by Kelton et al. (2010). According to these authors, the relationship between presentation formats and mental problem representation and processing is one of the least understood areas in the research on how presentation formats may influence problem solving (Kelton et al. 2010 p. 99).

In line with the recommendations by Kelton et al. (2010), I have applied process-tracing methods to assess possible mental representation differences related to various presentation formats, and my research question is:

Do data presentation formats influence decision-making effectiveness, and if so, how?

My research question is related to decision making and managerial decision making in particular. In accordance with Simon (1960), I look at managers as decision makers, and I see decision making as a process. . Simon (1960) describes decision making as a process consisting of three phases: 1) *intelligence*, that is, searching the environment for conditions that call for decisions; 2) *design*, that is, inventing, developing and analysing possible courses of action and 3) c*hoice*, that is, selecting a particular course of action from those available. In general, the order of the three phases is in the sequence they are presented above, meaning that the intelligence phase precedes the design phase, and the design phase precedes the choice phase. However, each phase in making a particular decision is itself regarded as a complex decision-making process. For example, a problem at any level may generate sub-problems that in turn have their own intelligence, design and choice phases, and so on (Simon, 1960).

Decision making and problem solving are closely related, and the two phenomena are often seen as two sides of the same matter. The three phases of intelligence, design and choice are thus regarded as phases of problem solving (see e.g. Kaufmann, 1991).

Decision-making effectiveness is related to the degree of goal attainment resulting from the decision-making process for a specific task. However, in evaluating this outcome, I take into consideration whether it is influenced by, among other things, lucky choices and calculation errors. Presentation format refers to the way data is displayed to the decision maker. The same set of data may be presented in different formats, for example, as graphs and tables.

Compared to previous research, I study the possible influence of data presentation formats on decision making in more complex tasks. Previous research has treated graphs and tables as if they were mutually exclusive. However, as discussed in Chapter 2, Paivio's dual-code theory (Paivio, 1971, 1986, 2007), implies that it is relevant to study how the two presentation formats may complement each other as well. Finally, as argued by Edwards (1992), managers are usually not unaided. Therefore, in my research I have given decision makers access to decision aids.

In order to answer my research question, I have conducted two experimental studies building on a research project developed by Fuglseth (1999). In study 1, 42 Master of Business Administration (MBA) students were asked to make decisions related to managing a summer restaurant. The tasks varied in complexity, and the presentation formats were graphs, tables and combined graphs and tables. The tasks were constructed so that they had optimal solutions. The subjects had access to decision aids, such as spreadsheets, calculators, pens and paper. The subjects' handling of the tasks was studied using verbal protocols and observation.

The results of Study 1 showed that the subjects were silent when they looked at graphs or performed calculations. Therefore, I had difficulty assessing the subjects' information processing. Furthermore, the verbal protocols indicated large variations in terms of the time used by the subjects to solve the tasks. Therefore, Study 2 was developed and conducted as a follow-up study to Study 1. In Study 2, I complemented the verbal protocols with eye-tracking data and included more accurate measures of how the subjects spent the time handling the task. Where possible, I have aggregated data from Study 1 and Study 2 to follow up on the preliminary results of Study 1.

The rest of the thesis is organised in five chapters: In Chapter 2, I present a review of the literature focusing on the previous research in the field of data presentation formats, and I position my thesis as regards my research contribution. In Chapter 3, I present theories related to answering my research question, and I present Study 1. In Chapter 4, I present Study 2, and In Chapter 5, I present the analysis of the aggregated data from Study 1 and Study 2. Finally, in Chapter 6, I address the limitations of the thesis and make recommendations for future research.

2 LITERATURE REVIEW

In this chapter, I review research issues related to my research problem. I start with the literature on graphs versus tables, including the theoretical framework used to understand most of the results in the research thus far. Second, I review the literature on the relationship between task complexity and presentation format and how it affects decision-making effectiveness. Finally, I summarise the presented literature and position my research.

2.1 Presentation formats

The presentation of data is a phenomenon that has captured researchers' attention for many years in several disciplines. Within educational psychology, for example, researchers have studied the effects of using text, animation and narration (sound) on learning (Mayer, 1990; Mayer & Anderson, 1991; Mayer & Moreno, 1998; Mayer & Sims, 1994). Another example is within engineering and system safety, where the presentation of emergency operating procedures has been investigated in areas such as aviation, navigation, chemical plants and nuclear power plants (see e.g. Xu et al., 2008). In the latter case, the presentation formats used were flow charts (2D and 3D) and success trees. Within marketing, the effects of different presentation formats on information acquisition strategies have been studied using information cards showing different sorting alternatives (e.g. Bettman & Kakkar, 1977).

Common amongst the above examples is that they study the effects of using different types of data presentation formats. A data presentation format can be described as the method of disseminating data to users (Kelton, et al., 2010). Graphs and tables are two such data presentation formats that have been studied extensively within the fields of information science, finance and accounting (see e.g. Desanctis & Jarvenpaa, 1989; Dilla, Janvrin, & Jeffrey, 2013; Dilla, et al., 2010; Dilla & Steinbart, 2005; Frownfelter-Lohrke, 1998; Gettinger, Koeszegi, & Schoop, 2012; Ives, 1982; Jarvenpaa, 1989; Shaft & Vessey, 2006; Speier, 2006; Vessey, 1991, 1994; Vessey, 2006).

Although tables and graphs can present the same data, the way the data is presented is still fundamentally different; a graphical presentation emphasises relationships among data and highlights trends, whereas a tabular presentation emphasises symbolic information. Larkin and

Simon (1987) describe such differences in presentation formats based on whether the presentation formats are informationally and computationally equivalent:

> Two representations are informationally equivalent if all of the information in the one is also inferable from the other, and vice versa. Each could be constructed from the information in the other. Two representations are computationally equivalent if they are informationally equivalent and, in addition, any inference that can be drawn easily and quickly from the information given explicitly in the one can also be drawn easily and quickly from the information given explicitly in the other, and vice versa.

A substantial part of the literature on data presentation formats has been concerned with providing guidelines for how to properly design graphs (see e.g. Bertin, 1981; Frownfelter-Lohrke & Fulkerson, 2001; Futkerson, Pitman, & Frownfelter-Lohrke, 1999; Ives, 1982; Jarvenpaa, 1989; Jarvenpaa & Dickson, 1988; Kosslyn, 1989). Ives (1982), for example, provides guidelines for the use of gridlines by suggesting that unnecessary or overly prominent grid lines should be avoided and that grid lines should not cross bars in a bar chart. Other guidelines include the use of colours, labels and axes, to mention some. Following established guidelines for proper graph design is important when studying the effects of graphs versus tables (Bertin, 1981). Nevertheless, a recent review article by Kelton et al. (2010) shows that issues regarding graphical construction have largely been ignored in much of the graph/table research.

Kumar and Benbasat (2004) investigate the effects of data presentation formats on decision time using 3D graphs compared to 2D graphs. They use three task situations characterised by differences in task complexity (i.e. number of data points, 9 or 25) and whether the user had to extract data points from the graphs. Their findings suggest that 3D graphs consistently outperform 2D graphs in the three task situations. Dull and Tegarden (1999) study the effects on decision accuracy with 2D graphs compared to rotatable 3D graphs. Their findings indicate that decision accuracy improves with the use of rotatable 3D graphs. Tractinsky and Meyer (1999) study user preferences for 2D graphs and 3D graphs and find that decision makers tend to prefer 2D graphs over 3D graphs when given the opportunity to choose presentation format.

Most studies investigate data presentation formats from the receiver’s standpoint, for example, based on how data presentation formats may improve decision-making effectiveness. An alternative approach is to investigate data presentation formats from the presenter’s perspective, for example, using data presentation formats as communicative and persuasive means for the presenter to convey a message (see e.g. Arunachalam, Pei, & Steinbart, 2002; Tractinsky & Meyer, 1999). In this respect, Tractinsky and Meyer (1999) find that participants in their experiment were more likely to violate the principles of graph design when the data itself reflected undesirably on the presenter.

Yet another research stream investigates the interactive aspects of the presentation formats in information systems (see e.g. Dilla, et al., 2010; Dull & Tegarde, 1999; Kumar & Benbasat, 2004; Tractinsky & Meyer, 1999). Dilla et al. (2010) describe an interactive data presentation (or interactive data visualisation as they call it) as a computer-supported visual representation of data that allows users to select the information they wish to view. Furthermore, according to the authors, an interactive data presentation format consists of three elements: 1) *interaction*, that is, the decision makers' dialogue with the information system, 2) *selection*, that is, the user can choose which data to display, for example by filtering and aggregating the data set and 3) *representation*, that is, the presentation format of the data, such as tables (numerical) and graphs.

The current study is based on a receiver's standpoint and investigates how data presentation formats may influence decision-making effectiveness. I do not aim to contribute to the development of graph design guidelines; however, established guidelines will be followed. I will investigate static presentation formats, such as printed graphs and tables and tables and graphs presented in a spreadsheet model on a computer screen. Thus, I do not take into consideration more interactive presentations of data in a spreadsheet, such as pivot tables, pivot charts and dashboards. I will use 2D graphs.

2.1 Graphs versus tables

At least 80 years of research has been carried out to investigate the effects of using graphs compared to tables. For instance, Washburne (1927) studied the effects of tables, line graphs, pictographs and bar graphs on the recall of information and found the effect of the presentation formats to be contingent upon the information being sought by the reader. He found that tables were best for identifying specific data values, and line graphs were best for identifying trends. Pictographs were best for making very simple comparisons, and bar graphs were best for more complex comparisons.

The relationship between the way data is presented and decision-making outcome has been studied in a wide range of special fields, such as information science, finance and accounting (see e.g. Desanctis, 1984; Speier, 2006). The comprehensive research stream assessing the effects of various display formats on problem-solving performance has led to the widely shared belief that there is not one optimal format but that the effectiveness of a specific presentation format depends on the type of task to be performed (Speier, 2006). Nevertheless, there are still many unanswered questions regarding data presentation formats (Kelton, et al., 2010).

2.1.1 Early studies – inconsistent results

The earliest studies on graphs versus tables assumed one of the presentation formats to be superior to the other in judgment and decision-making tasks, and therefore, most studies focused on finding *the most* effective presentation format among them. Several advantages of using graphs were suggested; however, there were no accepted conclusions regarding the appropriateness of specific chart types (e.g. pie chart, line chart, bar chart) and for which situations they should be used (Desanctis, 1984). As a consequence, a variety of chart types have been used in the table versus graph literature, making it hard to compare results. Furthermore, many researchers had a pre-determined assumption that graphs would outperform tables, leading to studies that were confounded by investigator bias, resulting in 'false' support of graphs as a presentation format (E. V. Wilson & Addo, 1994).

The belief that there was a most effective presentation format was not supported, and instead the research on presentation format produced a stream of contradictory findings. DeSanctis (1984) illustrated this well in her review of 29 studies comparing graphs to tables. She found that 'A total of 12 studies have found tables to be better than graphs. No meaningful difference between the two presentation modes was found in 10 studies. Only 7 have found graphs to outperform tables.'

Several researchers (e.g. Amer, 1991; Davis, 1989; Desanctis, 1984; Jarvenpaa, 1989; Jarvenpaa, Dickson, & DeSanctis, 1985; Montazemi & Wang, 1988; Vessey, 1991) addressed the conflicting results in the early graph/table research and provided guidelines for how to overcome similar inconsistencies in further studies. Some main conclusions can be drawn from reviewing the literature. First, the early research can be characterised as being *atheoretical*, that is, the research was not grounded in theories specifically related to the use of data presentation formats and thus lacked theories to guide the research effort. Some authors suggest that the atheoretical approach contributed to a lack of interpretability and generalisability of the results (see e.g. Amer, 1991; Vessey, 1994). A second problem is the great number of *different measuring instruments* being used in the early research, leading to problems with the reliability and validity of the results. For example, Jarvenpaa et al. (1985) pointed out that 'The use of different measures, even on the same construct variables, inevitably causes incomparable results and, therefore, leads to research labelled "conflicting"'. As a means to avoid problems with poor internal validity, Jarvenpaa et al. (1985) argued that future research should examine the cognitive processes underlying the decisions made by subjects in experimental settings. Techniques such as process tracing and protocol analysis were suggested as ways to discover the methods and reasoning that subjects use when solving the tasks (see e.g. Desanctis, 1984). The think-aloud method (e.g. Ericsson & Simon, 1980, 1993), that is, asking participants to verbalise their thoughts, and eye tracking (see e.g. Holmqvist et al., 2011; Russo, 1978), that is, making recordings of where participants direct their gaze, are examples of two such process-tracing techniques. I will elaborate on the think-aloud method and on eye tracking in Chapter 3 and 4.

A third problem often mentioned is the *inappropriate research design* used in many studies (e.g. Davis, 1989; Jarvenpaa, et al., 1985). Many of the early studies were overly simplistic and did not address important problems in the field, for instance, by using measures that were not directly related to decision-making effectiveness. Furthermore, many studies lacked experimental control. For example, some studies compared presentation formats that were not informationally equivalent, that is, there were content differences between the graphical and tabular presentation format being used. Jarvenpaa et al. (1985) argued that future studies need to control for factors other than presentation format that might affect decision-making effectiveness. Such factors might include the characteristics of the subjects (e.g. motivation, familiarity with the presentation formats being used, knowledge, etc.), learning effects and other correlated variables.

A final problem that characterised the early graph and table research is the *diversity in the experimental task* being used at the time. According to Jarvenpaa et al. (1985), 'The use of diverse and often unrelated and incomparable task situations makes the integration of findings across studies difficult because subjects' performance may be more a consequence of the task environment than of the use of graphics.' As a consequence, several authors called for research that took into account the characteristics of tasks (e.g. complexity, content, difficulty). Davis (1989) specifically addressed the importance of task complexity when referring to the fact that most business researchers studying graphs/tales have used tasks with a low level of complexity.

However, as Davis argued, the relative advantages of different presentation formats might only become apparent with more complex tasks.

The inability to find an optimal form of presenting data eventually led to a consistent belief among researchers that other aspects, such as task characteristics or decision makers' characteristics, needed to be taken into account in order to understand the effects of the different presentation formats on decision making (see e.g. C. Edwards, 1983; Ghani & Lusk, 1982; Ives, 1982; Jarvenpaa, 1989; E. V. Wilson & Addo, 1994). Nevertheless, subsequent studies in general continued to produce inconsistent results, with no clear guidelines on how to use different presentation formats and how they affect decision making (for a review, see Frownfelter-Lohrke, 1998).

2.1.2 Cognitive fit

Gradually, a cost–benefit theory (see e.g. Beach & Mitchell, 1978; Einhorn & Hogarth, 1981; Payne, 1982) was applied as a means to explain how presentation formats, such as graphs and tables, and task characteristics influence decision making (see e.g. Kleinmuntz & Schkade, 1993; Vessey, 1991, 1994) by facilitating some decision strategies and at the same time hampering others.

A basic assumption of the cost–benefit approach is that the process of solving a given problem can be carried out in many ways using a variety of information-processing strategies available to the decision maker. These strategies vary with regards costs, that is, the cognitive effort required to use a strategy, and benefits, that is, the ability of a strategy to produce an accurate response. Furthermore, because humans' capacity for information processing is a limited resource (Anderson, 2013), the theory suggests that humans change information processing strategy so that they minimise the joint cost of effort and error in making a decision. This means that their strategy selection is the result of a compromise between the desire to make a correct decision and the desire to minimise the effort (Beach & Mitchell, 1978; Payne, 1982). A number of factors may influence the error and effort required to make a decision, and changes in these factors may thus induce decision makers to change information processing strategy. Examples of such factors are task complexity, response mode (i.e. whether the task requires judgement or choice), time pressure and the characteristics of the presentation format (Payne, 1982).

In this setting, the term 'strategy' denotes a general approach to information processing involving elementary mental processes. Such strategies can either be holistic or analytic (Amer, 1991; Umanath & Vessey, 1994). Holistic strategies involve elementary perceptual processes, such as making associations and perceiving relationships in data. Analytic strategies involve verbal processes, such as extracting discrete data values and computations. Perceptual processes are assumed to require less effort than verbal processes, while verbal processes are assumed to give responses that are more accurate.

In the early nineties, Vessey (1991) introduced the theory of cognitive fit as a way of explaining the inconsistent results in the numerous studies examining the effects of graphs and tables on decision making. The cognitive fit theory is a special case of the more general cost–benefit theory described above.

The key element of the cognitive fit theory is to separate the problem-solving task from the problem representation (i.e. presentation format); see Figure 2.1. By doing so, it is possible to identify the distinguishing features of graphs and tables and the types of tasks for which they are useful.

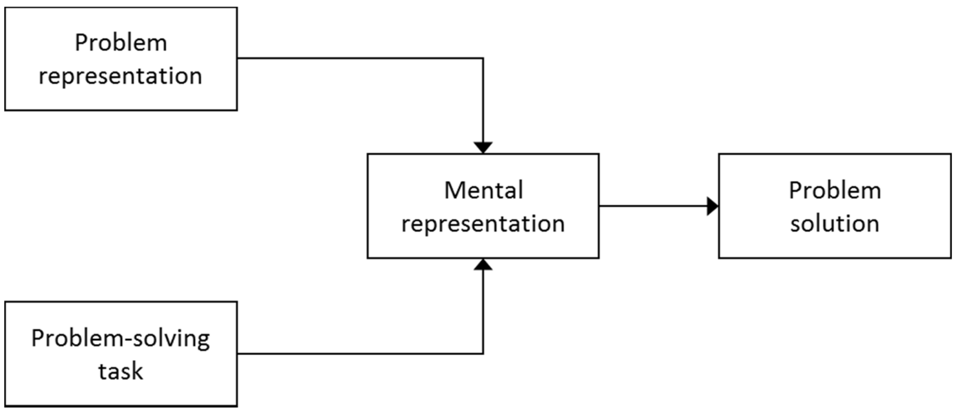


Figure 2.1. Model for cognitive fit (Vessey, 1991).

The cognitive fit theory views problem solving as an outcome of the relationship between the problem representation and the problem-solving task. The mental representation is the way the problem is represented in the human brain, and it is formulated using attributes of both the problem-solving task and the problem representation (Vessey, 1991).

The theory directly relates task type to data presentation format (Frownfelter-Lohrke, 1998), which might explain several of the results in earlier studies. Furthermore, the theory has provided much of the foundation for later studies examining the effects of data presentation formats on decision making (Speier, 2006) in simple tasks.

According to the cognitive fit theory, for the most effective and efficient problem solving to occur, data presentation formats should match the task to be accomplished (Vessey, 1994, 1991). According to Vessey and Galetta (1991), there are two basic types of tasks, spatial and analytic, based on the type of information that facilitates their solution. An example of a spatial task is as follows (Vessey & Galletta, 1991): 'In which month is the difference between deposits and withdrawals greatest?' Solving this task requires a comparison of trends, and it is, according to the authors, best accomplished using perceptual processes. An example of a symbolic task is (Vessey & Galletta, 1991): 'Provide the amount of withdrawals in April'. This task requires a specific amount as a response and is best accomplished using analytic processes.

Examples of data presentation formats are graphs and tables. Graphs are spatial presentation formats, that is, they emphasise and provide a visualisation of the relationships among the data and allow the user to view the data as an integrated whole. However, graphs do not present discreet data values directly. Tables are numeric presentation formats in the sense that they emphasise the presentation of discrete data values, but they do not present data relationships directly. According to the cognitive fit theory, graphs are the appropriate presentation format for spatial tasks, whereas tables support symbolic tasks. The argument is that when the data presentation format and the task type match (i.e. cognitive fit occurs), the decision makers can formulate a mental representation and use the same type of mental information processes to both act on the data presentation and solve the task. When the data presentation format does not match the task, similar processes cannot be used to both act on the data and solve the problem. The mental representation will be based on either the data presentation format or the task. If the mental representation is based on the data presentation format, the decision maker must transform it in order to derive a solution to the problem. If the mental representation is based on the task, the decision maker will have to transform the data derived from the presentation format into a mental representation suitable for task solution. In both cases, more cognitive effort is required for the additional mental steps. Thus, cognitive fit between the (external) presentation format and the task type is supposed to lead to an effective (accurate) and efficient (fast) problem solution (Vessey, 1994) by guiding the decision maker in their choice of decision strategy (Vessey, 1991).

The above description of cognitive fit applies elements from cost–benefit theory to explain how matching data presentation format with the task type (i.e. cognitive fit occurs) *minimises* both cognitive effort and error in the decision-making strategies being used (Vessey & Galletta, 1991). Vessey (1994) extended the original cognitive fit theory to include more complex tasks by applying cost–benefit theory to suggest that the information-processing strategy may occur as a result of a *trade-off* between error and cognitive effort. Complex spatial tasks will normally be solved using perceptual processes because this strategy will result in the least amount of effort. With the requirement for accuracy, however, decision makers may be induced to switch from perceptual to analytical processes, which are facilitated by tables. Complex symbolic tasks place significant strain on humans' cognitive resources. As the complexity of a symbolic task increases, decision makers may prefer – or may have to – use perceptual rather than analytical processes due to limited cognitive capacity. A symbolic task so complex that a decision maker may no longer use analytical processes to solve it is referred to as a limiting task. In such tasks, therefore, the appropriate data presentation format might not be a table but a graph that supports perceptual processes (Vessey, 1994).

Additional extensions to the theory of cognitive fit have been offered throughout the years. Shaft and Vessey (2006), for instance, extended the cognitive fit model to include characteristics of the decision maker by taking the decision makers' prior knowledge and understanding (e.g. experience) of the problem domain into consideration. Vessey (2006) labels such existing problem domain-specific knowledge as the 'internal problem representation', whereas the data presentation format (e.g. graphs and tables) is labelled 'external problem representation'. During problem solving and development of the mental representation, information in the internal problem representation is retrieved from memory and integrated with information from the external representations. According to Shaft and Vessey (2006), therefore, the decision makers' mental problem representation is based on the task, the external problem representation and the internal problem representation, as well as the interaction between the external and internal representations. For effective and efficient decision making to occur, there should be a fit among all these elements.

Kelton et al. (2010) present an extension of the cognitive fit model by Shaft and Vessey (2006) by including the effect of learning. They suggest that a decision task may be a repetitive (i.e. iterative) task, allowing for the decision maker to learn from feedback during the problem-solving process; see Figure 2.2.

Figure 2.2 illustrates how feedback on the problem-solving performance on a single iteration can affect the decision maker's mental representation for task solution and the internal representation of the problem domain in subsequent iterations of the same task. The bi-directional arrow between the internal representation of the problem domain and the external problem representation (i.e. data presentation format) indicates that the two constructs have a joint influence on the mental representation.

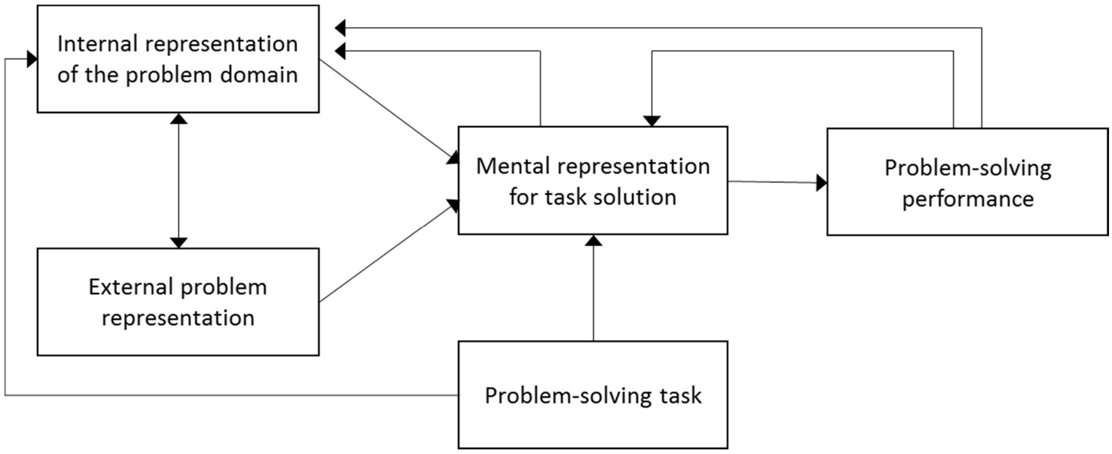


Figure 2.2. Adapted model of cognitive fit by Kelton et al. (2010).

Vessey (2006) reviewed the literature on cognitive fit and found some empirical support for the theory (e.g. Speier, 2006; Tuttle & Kershaw, 1998; Vessey, 1994; Vessey & Galletta, 1991; E. V. Wilson & Addo, 1994). The theory of cognitive fit has also been used extensively in later studies (see e.g. Dilla, et al., 2013; Dilla, et al., 2010; Gettinger, et al., 2012; Kelton, et al., 2010) to understand the effects of presentation formats on decision-making effectiveness.

The cognitive fit theory has been largely successful in explaining outcomes in fairly simple tasks involving data acquisition and well-defined evaluations (for an overview, see Speier, 2006; Tuttle & Kershaw, 1998; Umanath & Vessey, 1994; Vessey, 1994, 2006). Evaluating the results of three published graph-versus-table studies using complex tasks with performance constraints, Vessey (1994) also found empirical support for strategy shifts, that is, using perceptual rather than analytic processes in complex symbolic tasks. Speier (2006) and Speier, Vessey, and Valacich (2003) also claimed to have gained support for the cognitive fit theory when using complex tasks. In her study, Speier (2006) used tasks from the operations management domain. For example, the participants were asked to assess machine workload schedules for a set of machines over a time period, or they were asked to solve a facility location task based on the minimisation of total costs. The tasks were classified in two groups based on their complexity: simple tasks and complex tasks. The tasks differed with regard to the number of information cues (simple: 1-12, complex: 30) and the number of processing steps required (simple: 0–6, complex: 18). In addition, the complex task also required the subjects to rank their potential solutions, meaning that up to six alternatives had to be compared. However, the complex task was well-structured with clear expectations of what rules should be followed and did not involve uncertainty. Thus, even the complex task constitute an example of simple choice tasks, with just a minimum of problem clarification.

2.2 Task

As discussed above, task complexity is believed to play an important role in determining the effectiveness of presentation formats in decision making. Therefore, I will elaborate on the concept of task and relevant task characteristics, such as complexity and uncertainty.

2.2.1 Introduction

The task concept has been explained and defined in many ways throughout the literature. Fuglseth and Grønhaug (1995) define a task as a piece of work that has to be done, often within a certain time limit. Handling a task involves solving problems where variations among the problems to be solved are dependent on the task domain. A task can consist of levels of progressively smaller subtasks (Byström & Järvelin, 1995).

Wood (1986) defines tasks as consisting of three essential elements: required task acts, information cues and products. Products are operationalised as the measurable results of solving the task. Required acts are equivalent to the necessary steps required to solve the task. Information cues are measured as the pieces of information necessary to make the required judgments during task performance. Bonner (1994) provides a similar definition of tasks based on general principles for information processing. According to her, a task consists of inputs, processing and outputs. These elements correspond to Woods' information cues, required acts and products, respectively. Regardless of how the elements are conceptualised, it is apparent that there is some kind of interrelationships among them.

In the following, I find it purposeful to use Fuglseth and Grønhaug's (1995) definition of a task as 'a piece of work that has to be done'. Note that I do not include the proposed restriction of a time limit, even though I accept that some tasks are constrained by time. However, I regard time limit as a characteristic of the task rather than a part of the task definition. Furthermore, I will assume Bonner's (1994) classification of task elements as inputs, processes and outputs.

2.2.2 Task complexity

In the previous section, the terms 'simple', 'complex' and 'more complex' have been applied somewhat intuitively, which is also done in the literature (Speier, 2006; Vessey, 1994). In order to categorise tasks, I need a typology. In my research, I have used Campbell's (1988) typology of complex tasks, which is based on the work by Schroder et al. (1967). The advantage of building on Schroder et al. (1967) is that their constructs provide a common language both for analysing objective task complexity and for translating these attributes into cognitive processes (Campbell, 1988, p. 43). In line with Schroder et al. (1967), Campbell distinguishes between objective and subjective or experienced task complexity. Objective task complexity is a function of the task per se, and subjective task complexity is related to the individual's perception and handling of the task. As argued by Campbell (1988), subjective and objective task complexity are related. Subjective task complexity can explain how objective task complexity is handled. The relationship between objective and subjective task complexity can be moderated by, for example, familiarity with the task domain, the availability of decision aids and the data presentation formats (Campbell, 1988).

In accordance with Schroder et al. (1967), Campbell (1988) applies three properties of an objective complex task: (1) the number of dimensions requiring attention, (2) the number of alternatives associated with each dimension and (3) the relationships among the dimensions and alternatives, including the degree of uncertainty. Elaborating then on the relationships, Campbell (1988) distinguishes among four main types of complex tasks: choice tasks[1], judgement tasks, problem tasks and fuzzy tasks. Choice tasks involve selecting the best alternative from a set of possibilities. Judgement tasks require the subjects to evaluate diverse sources of information and then make a judgement or prediction about some future event. Problem tasks are characterised by a multiplicity of paths to a well-specified outcome, that is, they require the subject to search for and find the best way to achieve the outcome. Fuzzy tasks are characterised by the presence of both multiple desired outcomes and multiple ways of attaining each of the desired outcomes. Tasks representative of this category are often found in business contexts. Within each of these four main categories, there are subcategories related to the interdependencies and uncertainty of the linkages among the dimensions and alternatives; for a detailed discussion, see Campbell (1988).

[1] Termed decision tasks by Campbell (1988).

2.3 Positioning of my research

Based on the literature review above, I will use the remaining part of this chapter to explain how my thesis will fill a gap in the existing literature. First, I will argue for the use of more complex tasks that better relate to real-life managerial decision tasks. Second, based on the dual-code theory (Paivio, 1971, 1986, 2007), I argue for a study of the combined use of graphs and tables. Furthermore, I explain why access to decision aids should be taken into consideration, and finally I argue for the use of process-tracing methods in addition to measuring decision-making outcomes, such as economic results.

2.3.1 Task complexity

The cognitive fit theory has mainly been tested in choice tasks and judgement tasks, with little conflicting interdependence and/or uncertainty among the dimensions and alternatives. Even though some researchers (see, e.g. Speier, 2006; Speier et al., 2003) claim to have gained support for the theory of cognitive fit also when using complex tasks, the 'high-complexity' tasks used in these studies are questionable. The task presented by Speier (2006) as a complex–symbolic task involves five dimensions and six alternatives associated with each dimension (i.e. 30 information cues) and 18 rather simple calculations/comparisons. In Campbell's (1988) typology, the task is characterised as a choice task with some interdependence among the alternatives.

Most studies investigating the relationships between data presentation formats and decision quality use tasks that can be characterised as either spatial or symbolic, and they assume a decision-processing strategy that is either holistic (using mainly perceptual processes) or analytic. Real-life managerial decision tasks are, however, often 'fuzzy', as described above. They can be achieved using a variety of spatial and symbolic subtasks, and they usually require both perceptual and analytic processes. How decision makers choose to structure such tasks into subtasks may have significant implications for the accuracy of the outcome. In order to enhance the understanding of how data presentation formats may support decision makers, the cognitive fit theory should be tested in tasks that are more similar to real-life managerial decision tasks.

2.3.2 Presentation formats

In the studies performed so far, data presentation formats are usually presented as if they were mutually exclusive. Exceptions are the studies by DeSanctis and Jarvenpaa (1989), Frownfelter-Lohrke (1998) and Lucas (1981).

The theory of cognitive fit (Vessey, 1991, 1994) builds on the dual-code theory of cognition (Paivio, 2007, 1986, 1971). This theory states that human beings have developed different types of mental representations and operations that are assigned to different information-procession functions. There is one system specialised for the representation and processing of information concerning non-verbal objects and events, and there is one system specialised for dealing with language. Paivio refers to the two systems as the non-verbal or imagery system and the verbal system. According to Paivio, the two systems are independent in the sense that either system can be active without the other. At the same time, the two systems can be interconnected so that activity in one system can initiate activity in the other. I will elaborate more on the dual-code theory in the next chapter.

Vessey is only partly using the dual-code theory to understand the effects of presentation format in decision-making processes. She argues as if the verbal and non-verbal systems are independent of each other and therefore treats graphs and tables as if they are mutually exclusive. On the other hand, the description by Paivio (1971, 1986, 2007) of the two systems being interconnected supports the idea of examining the effects of a combined display of graphs and tables. In relatively simple tasks with limited strain on working memory, I expect that the decision maker can mentally visualise the relationship between variables from the tabular display and does not need the graphic display. In complex tasks placing a high cognitive load on the subject, the graphic display may give an overview but not enough detail to reach a high decision quality, while the tabular display may not give a sufficient overview to handle the details appropriately. I therefore expect that graphs will increase humans' general understanding of the relationships among variables in such tasks and that additional tables will increase the understanding of details. Thus, as opposed to other researchers who assume that the best presentation format is either a graph or a table, I will study the effect of the combined use of graphs and tables in addition to examining each display type separately.

2.3.3 Decision aids

Most research in this area is based on the notion of the unaided decision maker. However, in real-life managerial decision making, decision makers are usually not unaided. Typically, decision makers have access to a number of decision aids and additional sources of information, such as electronic databases and spreadsheet models, written information sources and other people. Therefore, I agree with Edwards (1992) that researchers should take this aspect into account in their research design and study whether access to decision aids is of significance for the effectiveness in decision-making processes. Therefore, an important part of my study will be to evaluate whether decision makers are able to utilise their available decision aids in order to increase decision-making effectiveness.

2.3.4 Mental representations

As stated in the introduction to this thesis, most studies focus on the outcome of the decision-making process. There have been few efforts to understand how the various data presentation formats influence decision makers' mental representations and processes. A focus on outcome only might tell us something about *when* a given presentation format is preferred over another but not *why*. A focus on outcome does not take into consideration things such as luck, misinterpretations of the decision problem, motivation and choice of decision strategies. Therefore, in order to control for such conditions, it is necessary to investigate the processes and strategies underlying the decision-making process. For similar reasons, some researchers have called for studies that include process-tracing techniques in addition to outcome measures (see, e.g. DeSanctis, 1984; Jarvenpaa et al., 1985; Vessey, 1994; Shaft & Vessey, 2006). With my research, I will enhance the understanding of the relationships between data presentation formats and problem-solving performance by studying how subjects represent and process tasks that are similar to real-life decision-making tasks involving the use of decision aids.

3 STUDY 1

In this chapter, I will present my first study investigating the relationship between data presentation formats and decision-making effectiveness. As indicated in the introduction, my research builds on a research project initially developed by Fuglseth (1999). Fuglseth collected data to test the stimuli developed for the project. Fuglseth and Grønhaug (2000) showed in their preliminary analysis that there were significant differences between the two experimental tasks (high and low complexity). Thus, the experimental design worked as intended.

In this study, I extend the work by Fuglseth and Grønhaug (2000) by adding more data (subjects) to the data set and by expanding the analysis of the data.

The remainder of this chapter is structured as follows. In the next section, I will present the theories used for this study. I elaborate on the dual-code theory by Paivio (1971, 1986, 2007) and the cognitive complexity theory by Schroeder et al. (1967). I also discuss the think-aloud method (e.g. Ericsson & Simon, 1980, 1993), which is used as the process-tracing method in this study. In the following section, I present my research model and research design. I then describe my data collection procedure and present my analysis. Finally, the implications of the study are discussed, together with recommendations for further research.

3.1 Theories

3.1.1 Dual-code theory

According to the dual-code theory (Paivio, 1971, 1986, 1991, 2007), human beings have developed two different types of codes of mental representation that are assigned to different information-processing functions, a *verbal* system that is specialised for language and a *non-verbal* system that deals with non-verbal objects and events (Paivio, 2007). In this setting, *mental representations* refer to internal formats of information used in memory, whereas *coding* refers to the ways the external world is captured in those internal forms. The activation of representations within and between the systems is referred to as *processing* (Sadoski & Paivio, 2001).

Figure 3.1 summaries the structural assumptions of the dual-code theory and illustrates the idea of separate but interconnected systems. The internal organisation of the two systems, including connections within and between the systems, is thus illustrated. Finally, the figure shows that the two systems are connected to sensory input and response output systems.

Figure 3.1 shows a general model of the dual-code theory.

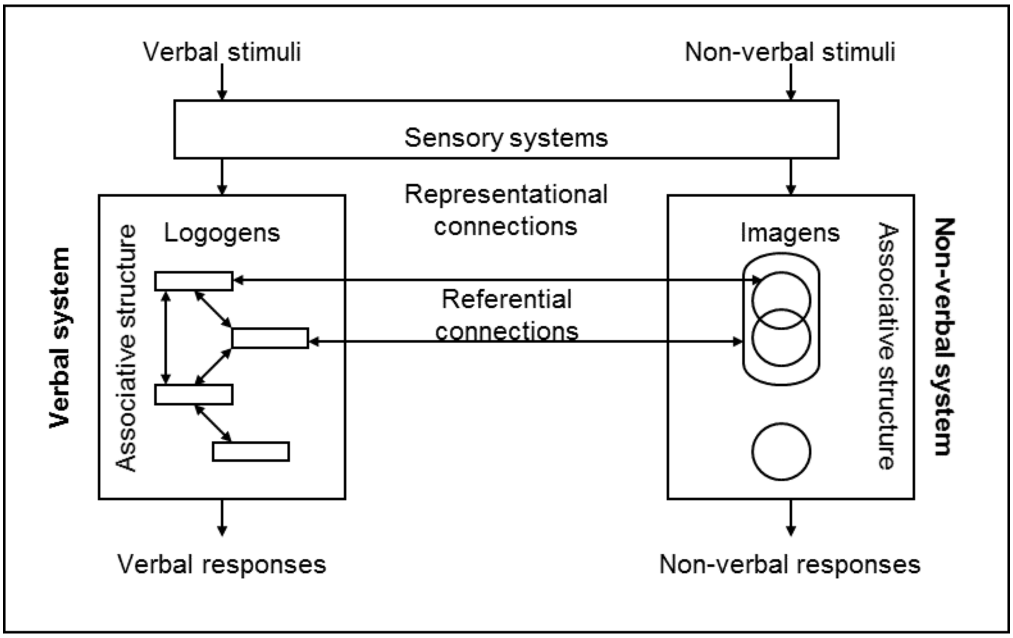


*Figure 3.1.*A schematic description of the structure of the verbal and non-verbal system (Paivio, 1986 p. 67).

In the following, I will present a summary of the dual-code theory as shown in Figure 3.1. I will elaborate on the internal organisation of the two systems, how they are activated and the processing of information within and between the two systems. The summary below builds on Paivio (1986, 1991, 2007).

The idea of two sub-systems implies a difference between the two systems regarding their structure and functionality. The dual-code theory assumes that the representational units, or building blocks, have some physical form in neural structures and pathways and that the structure of these units differs in the two systems. In the verbal system, the building blocks are referred to as logogens. In the non-verbal system, the building blocks are referred to as imagens. *Logogens* contain information that underlies our use of words. They reflect the internal organisation and variable size of language units as perceived and produced. The logogens are organised in sequential hierarchical structures in which logogens of smaller language units can be composed into larger language units or larger units can be decomposed into smaller units. For example, the word *time* consists of four separate characters, but it can also be recognised as a whole. Furthermore, words may be included in familiar phrases, such as *once upon a time*, which may have a phrasal unity that can be recognised holistically. Logogens are alternatively called verbal representations, verbal encodings, mental language and inner speech. *Imagens* contain information that generates mental images, that is, they correspond to natural objects, holistic parts of objects and natural grouping of objects. Imagens are typically part of a synchronously[2] organised hierarchical structure, that is, a nested collection of overlapping imagens, which in turn can be part of an even larger structure. For example, a pair of eyes continue with a face, and a face continues with a body and a body moves in some surroundings. Thus, the different imagens can be seen simultaneously in time. Like visual perception, however, visual imagens have a limited span, and different parts of a synchronously available representation may have to be imaged successively or 'scanned'. Imagens can be continuous, as in the example of the eyes, the face and the body, but they can also be discontinuous and discrete. For example, we can immediately shift from a scene at work, such as our office, to a scene at home, such as the kitchen. Imagens representing familiar scenes, such as our home, may be analysed from different perspectives. For example, if asked to tell the number of windows in their home, people often report taking a mental tour through the rooms in their home (Paivio, 2007). Imagens are alternatively called non-verbal representations, non-verbal encodings, mental images or imagery.

The dual-code theory defines processing as the activation of the mental representations, that is, imagens and logogens and the connection between them. Three levels of processing are defined in the model in Figure 3.1 – representational processing, associative processing and referential processing.

Representational processing refers to the relatively direct activation of logogens by linguistic stimuli and of imagens by non-verbal stimuli. The activation of verbal and non-verbal representations is a joint function of variables in the stimulus situation and relevant individual difference variables. For example, differences in background or preferred modes of thinking may influence whether an individual 'thinks' mainly in a verbal mode or a non-verbal mode. Similarly, differences in the context (the stimulus situation) may influence, for example, whether a circle is a geometrical shape, the letter 'o' or the numerical value of zero. Empirical observations indicate that the non-verbal system is more likely to be evoked and used with pictures of objects as stimuli than with words as stimuli and with concrete words rather than with abstract words. For example, the word 'dog' is more imaginable than the word

[2] Synchronously means that all the information is available at one time.

'intelligence'. The verbal system is activated when words serve as stimuli, especially words that are high in the person's acquired capacity to evoke verbal associations. Activation of the verbal system would also occur when a task demands verbal processing or when instructions are given to carry out a task verbally.

Associative processing refers to the activation of representations within either system, that is, connections between logogens and other logogens or between imagens and other imagens, accounting for the spread of association among words or among images. For example, seeing the visual word 'chair' might activate associated visual logogens such as 'armchair', 'recliner', 'bench', 'stool', etc.

Referential processing refers to activity between the two sub-systems, such as logogens to imagens or imagens to logogens. Such processing occurs whenever we name an object or form an image to its name. The activation is indirect because the object or name must first be identified, entailing direct activation of an imagen or logogen, which then activates a representation in the other system. For example, seeing the visual word 'chair' would activate its respective logogens and in turn activate an image of the object, a chair.

According to the dual-code theory, the verbal and non-verbal systems are initially separate, and they can therefore function independently of each other, meaning there can be activity in one system but not in the other. The two systems can also work in parallel, meaning that there are separate activities in both systems at the same time. For instance, one can hum a tune using the non-verbal system while at the same time performing calculations using the verbal system. Finally, the two systems can work together in a connected way so that activity in one system can initiate activity in the other.

According to the dual-code theory, all cognition – including perception, memory, meaning and knowledge – must be accounted for by the operations of the representations within and between the verbal and non-verbal codes (Sadoski & Paivio, 2001). The organisation of representations is sequential in the verbal system and non-sequential in the non-verbal system, resulting in characteristically different constraints on processing. In other words, the verbal system is specialised for sequential processing whereas the non-verbal system is specialised for synchronous or parallel processing of multiple representational units. In the remaining part of this thesis, I will use the term *analytic processes* when referring to thinking that is carried out mainly in the verbal system and the term *spatial processes* when referring to thinking that is carried out mainly in the non-verbal system.

Related to problem solving, Paivio (2007) claims that most real-life and laboratory tasks involve problem solving that requires complex mixtures of both analytic and spatial processes. Still, some tasks may be considered to be mainly analytic or spatial. As an example, Paivio (2007) points to crosswords and jigsaw puzzles. A crossword puzzle is mostly analytic (i.e. verbal), based on verbal associative knowledge activated by printed verbal cues and structural frames, that is, the number of letters a solution word must have. A jigsaw puzzle, on the other hand, is mostly spatial (i.e. non-verbal) and is based on perceptual–motor activity and imagery; pieces are found and compared with slots and tried out for fit, either by actually moving the pieces or by moving and comparing them mentally.

Kaufmann (1988) characterises the two information processing systems in the following way:

> A linguistic-propositional (verbal) representational format is strong in the sense that great precision may be achieved in the form of explicit descriptions. It is easily and quickly manipulated and contains the full range of computational operations. In contrast, imagery (non-verbal) is more ambiguous and less easily manipulated, and only comprises simple cognitive operations of a perceptual kind, like anticipations and comparisons. This may be useful and even necessary in complex task environments, where computational operations in the sense of rule-governed inferences are difficult or impossible to perform.

Therefore, in ill-structured tasks with high novelty, complexity or ambiguity, human beings seem to switch from a linguistic–propositional representation to an imagery-based representation (Kaufmann, 1984, 1991).

Implications for the current thesis

In most of the studies performed so far, displays of graphs and tables are presented as if they were mutually exclusive. However, the description by Paivio (2007, 1991, 1986, 1971) supports the idea of examining the effects of combined displays of graphs and tables. In relatively simple tasks with limited strain on working memory, I expect that the decision maker can mentally visualise the relationship between variables from the tabular display and does not need the graphic display. In complex tasks, the graphic display may give an overview but not enough details to reach a high-quality decision. The tabular display may not give a sufficient enough overview to handle the details appropriately. Therefore, I expect that graphical displays will increase decision makers' general understanding of the relationships between variables in complex tasks and that additional tables will increase the understanding of details.

3.1.2 Cognitive complexity theory

The cognitive complexity theory by Schroder et al. (1967) explains human beings' information processing based on their cognitive structures – their construct system. The authors distinguish between two information-processing variables, content variables and structure variables. The focus of the cognitive complexity theory is on the structure variable. This means that the theory explains the complexity of human information processing based on *how* individuals structure information in their construct systems, as opposed to *what* information is actually processed. Thus, any position on any topic can be held at any level of complexity (Suedfeld, 2010).

Two interdependent properties of information-processing structures are suggested, namely dimensions and integration. *Dimensions* represent the constructs or the content, in the construct system. Few or many dimensional units of information can be used to evaluate a construct. For example, one can evaluate the construct of light by the degree of brightness, saturation and hue. Thus, these three properties are dimensions related to the construct of light. Schroder et al. (1967) use the term *differentiation* to describe a person's ability to perceive several dimensions in a stimulus rather than a single dimension alone. The term *integration* refers to a person's ability to identify multiple relationships among the differentiated dimensions. Thus, integration cannot occur in the absence of differentiation. Information processing, then, can be characterised by the number of dimensional attributes processed and the degrees of freedom involved in the way these dimensions are combined. According to Schroder et al. (1967), cognitive complexity is particularly related to integration, but the greater the number of dimensions, the more likely is the development of complex integration.

Cognitive complexity is not regarded as a personality trait or a general cognitive style. Rather, it may vary across different areas, meaning that a person may function with a high level of complexity within certain areas and a low level of complexity in others. The level of information processing within an area may develop over time, depending on both the person and the situational variables. Thus, the level of information processing is not held to be static over time (Schroder, et al., 1967).

Schroder et al. (1967) describe four levels of information processing based on the degree of integration and the degree of differentiation. The four levels are ordered in terms of their complexity, from low to high complexity.

1) A low integration index is characterised by a hierarchical relationship among the dimensions, that is, the dimensions for a range of stimuli are organised in a fixed way with only a single hierarchy of rules; see Figure 3.2. The information processing of a person with low cognitive complexity is characterised by a categorical black and white way of thinking, (i.e. there is little potential for developing scaled dimensions but only yes/no categories, which they either fit or do not), minimising of conflicts, anchoring of behaviour in external conditions and generalisation of external objects and events into one's fixed setup.

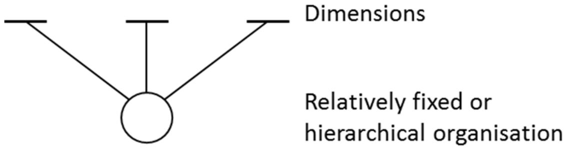


Figure 3.2. Low integration index

2) A moderately low integration index is characterised by the ability to generate and compare alternative organisations of dimensions. For example, if there are three dimensions, such a structure will provide at least two possible rules or perspectives for combining these dimensions; see Figure 3.3. The behaviour of a person with moderately low cognitive complexity is characterised by the ability to think in nuances and the ability to use simple rules of causation, that is, conditioning. However, a form of rigidity is still present in that when a particular set of integrating rules has been accepted, alternative combinations are almost completely ignored, and therefore the rules are not effectively combined.

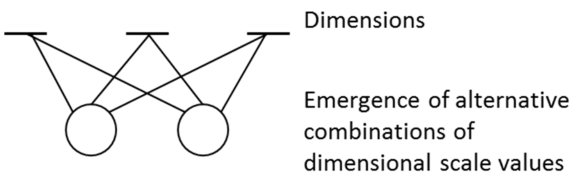


Figure 3.3. Moderately low integration index.

3) A moderately high integration index is characterised by the ability to generate more complex relationships between the dimensions, allowing various combinations of dimensions to be compared and evaluated against each other; see Figure 3.4. The uncertainty increases, not in the sense that the world is more chaotic but rather in the sense that alternatives exist. A person at this information processing level seeks more information before ‘closing’ on a particular decision, and even when a decision is made the person is still open to a number of alternative interpretations. The behaviour of a person with a moderately high integration index is characterised by being less deterministic, meaning he or she is able to increase the number of interpretations of the stimulus by combining and using two or more alternative systems of interpretation. The person is also able to understand how behaviour based on one standpoint may influence the situation based on another standpoint. Thus, the rules are minimally fixed and no longer anchored in the past.

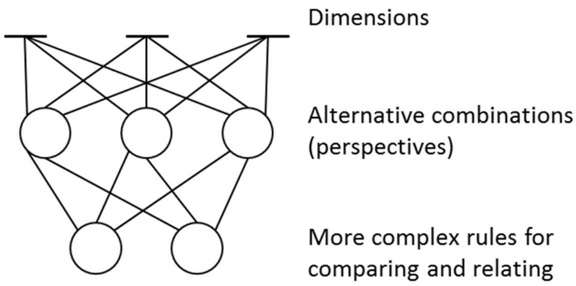


Figure 3.4. Moderately high integration index

4) A high integration index is characterised by the ability to organise several sub-construct systems in alternative ways, as illustrated in Figure 3.5. The main difference between the moderately high level of information processing and the high level is mainly the difference between an empirical and a theoretical view. At the high level of information processing, a person is able to generate or apply general laws that systematise a large and differentiated body of information. A person with a high integration index is able to generate the rules of the theory for various structures.

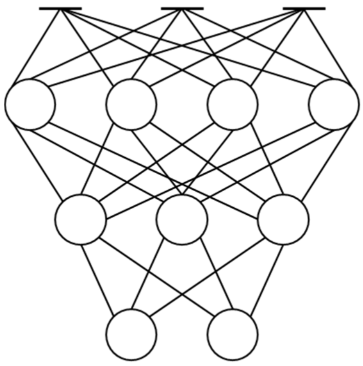


Figure 3.5. High integration index.

In line with Fuglseth (1989), I will use the term 'well-developed' to label construct systems that are characterised as having either a moderately high or a high integration index. A person with a well-developed construct system is assumed to be able to handle a complex situation in constantly changing environments more effectively than a person with a simpler construct system (Schroder, et al., 1967).

Cognitive complexity has been extensively applied by researchers to understand human information processing behaviour in a wide range of areas, including decision making (Choi, 2010).

In line with Paivio (1986, 1991, 2007), Vessey (1991, 1994, 2006) characterises information processing by the use of analytic and spatial processes only. She does not take into consideration *how* these processes are used. By applying a measure of level of information processing, I do not only characterise processes as analytic or spatial but also how they are used. For example, analytic processes can be simple, such as making a quick assessment of the situation before making a decision, or the processes can be more specific, such as including more dimensions and integrating rules before decisions are made. Thus, by assessing the subjects' level of information processing, I hope to get a better understanding of whether 'the full range of computational operations' are utilised. Related to spatial processes, I asses to what extent decision makers perform anticipations and comparisons, for example, whether they compare few or many aspects of the situation before making decisions.

3.1.3 Think aloud

Several researchers in the stream of graph/table research have called for studies to investigate the cognitive processes (e.g. decision strategies) underlying the obtained results. Even though cognitive processes and subjective thoughts cannot be observed and measured directly, several methods for inferring such phenomena exist. The think-aloud method (see e.g., Ericsson & Simon, 1980, 1993) is an example of such a method. The method is widely used (Fox, Ericsson, & Best, 2011) and is generally accepted as a central and indispensable method for studying thinking (Ericsson & Simon, 1980, 1993; Payne, 1994; T. D. Wilson, 1994).

The think-aloud method implies asking participants to 'think out loud' or verbalise their thoughts while solving a specified task. The participants are instructed to directly express what they are thinking without any explanation or interpretation, and they are allowed to say whatever comes to their mind without interruption from the observer (e.g. Ericsson & Simon, 1980). The verbalisation is recorded (usually a voice recording that is transcribed later), resulting in a verbal protocol that can be analysed. Russo et al. (1989) label the task of verbalisation as a *secondary task*, whereas the main task that the participants are asked to solve while verbalising is labelled the *primary task*. In the following, I will use the same terms when elaborating on the think-aloud method.

A distinction is made between concurrent and retrospective verbalisation (Ericsson & Simon, 1980, 1993). Concurrent verbalisation describes situations where information is verbalised at the time the subject is attending to it, and therefore the verbalised content is believed to be the information currently available in short-term memory. For such reasons, concurrent verbalisation has been frequently used to mirror cognitive processes and to understand problem-solving strategies (Hoppmann, 2009). Ericsson and Simon (1980, 1993) argue that concurrent verbalisation, if instructed properly, will not change the structure and course of the cognitive processes in the primary task. However, they do not rule out the possibility that the secondary task of verbalisation may slightly decrease the speed of task performance of the primary task. In fact, they suggest that the extent to which the verbalisation task will decrease the speed of the primary task performance is dependent on the type of information being verbalised; verbalisation involving the direct articulation of information stored in a *verbal* (i.e. analytic) code is assumed not to slow down the cognitive processes involved in solving the primary task. However, verbalisation of information represented in a complex *non-verbal* (i.e. spatial) code requires the subject to recode the information to a more easily articulated format before verbalisation can occur. The extra burden placed on the subject by the 'translation process' may slow down the performance of the primary task and may also cause the verbalisation to be incomplete or even cease completely. According to Ericsson and Simon (1980, 1993), subjects may also stop verbalising or provide less complete verbalisation when working under a heavy cognitive load. Therefore, it might often be necessary to repeat the request of asking the subject to think aloud while solving the primary task (e.g., Ericsson & Simon, 1998). This also supports the presence of an observer during the verbalisation process (Hoppmann, 2009). Still, no matter how much subjects are encouraged to think aloud, some types of tasks are not suited for the think-aloud method as the subjects report content from short-term memory. This applies to tasks involving cognitive processes that leave little or no traces in short-term memory, such as automated tasks (e.g. Ericsson & Simon, 1980). An automated task is a task where the cognitive processes necessary to solve the task have become highly practiced and eventually automated. This means that the intermediate mental steps for solving the task are carried out without being interpreted and without their inputs and outputs using the short-term memory (Ericsson & Simon, 1980). Thus, automated tasks do not tie up our attention, which is a limited resource.

In order to avoid problems like the ones mentioned above, it might be tempting to ask subjects to verbalise their cognitive processes *after* having completed the primary task. Such verbalisation is described by Ericsson and Simon (1980, 1993) as retrospective verbalisation. Ericsson and Simon (1980, 1993) suggest that the content verbalised in retrospective verbalisation is retrieved mostly from long-term memory, especially when such reporting is done a long time in retrospect, and that the content is therefore less complete. Ericsson and Simon (1980, 1993) warn against retrospectively collected protocols as being subject to both fabrication and forgetting (Russo, et al., 1989). For this reason, concurrent verbalisation is usually preferred over retrospective verbalisation (see, e.g. Ericsson & Simon, 1980, 1993; Payne, 1994).

Problems such as fabrication and forgetting, as well as other sources of errors such as reactivity, have been addressed by several authors who question the ability of the think-aloud method (both concurrent and retrospective verbalisation) to produce valid data (i.e. valid verbal reports) (see for instance Fox, et al., 2011; Nielsen, Clemmensen, & Yssing, 2002; Nisbett & Wilson, 1977; Payne, 1994; Russo, et al., 1989; T. D. Wilson, 1994). The concern that verbalisation is reactive, meaning that it changes the course of the primary task processes, is discussed in detail in Russo et al. (1989). They find that response time (i.e. the time taken to complete the primary task) is lengthened in tasks where participants are asked to think aloud concurrently. They find no sign of reactivity in a Raven's task involving a recoding from pictorial to oral code. However, in a gambles task requiring an accurate response, they do find evidence of some reactivity due to the concurrent think-aloud request (Russo et al., 1989, p. 762), and they call for research replicating/testing this result. Nonetheless, as pointed out by others (e.g. Fox et al., 2011), such reactivity problems are for the most part caused by improper instructions given prior to the task of thinking aloud. For instance, Ericsson and Simon (1998) stress that participants should be explicitly instructed to focus on the primary task while thinking aloud and merely to verbalise their thoughts rather than describe or explain them to anyone else. Fox et al. (2011) conduct a meta-analysis of 94 studies comparing performance while giving concurrent verbalisations to a matching condition without verbalisation and reach the same conclusion. Their findings show that there are no reliable differences between the think-aloud condition and the silent condition as regards performance of the primary task.

To sum up, the think-aloud method has shown to be a useful tool for studying human thinking and problem solving (e.g. Ericsson & Simon, 1998; Wilson, 1994; Payne, 1994), especially in tasks where the solution to the task cannot be found in a single step but requires several intermediate steps of reasoning or trying out (Payne, 1994; Hoppmann, 2009). Even those criticising the think-aloud method concede that the method may provide a more accurate insight into the causes of a person's behaviour than by merely observing the behaviour (e.g., Nisbett & Wilson, 1977; Russo, et al., 1989). As mentioned above, guidelines on how to collect the verbal protocol so as to minimise the potential biases exist (see, e.g. Payne, 1994; Ericsson & Simon, 1993).

3.2 Research design

Based on the discussion in section 2.4, I will use the following research model as the point of departure for the studies described in the present chapter and in the subsequent two chapters of this thesis.

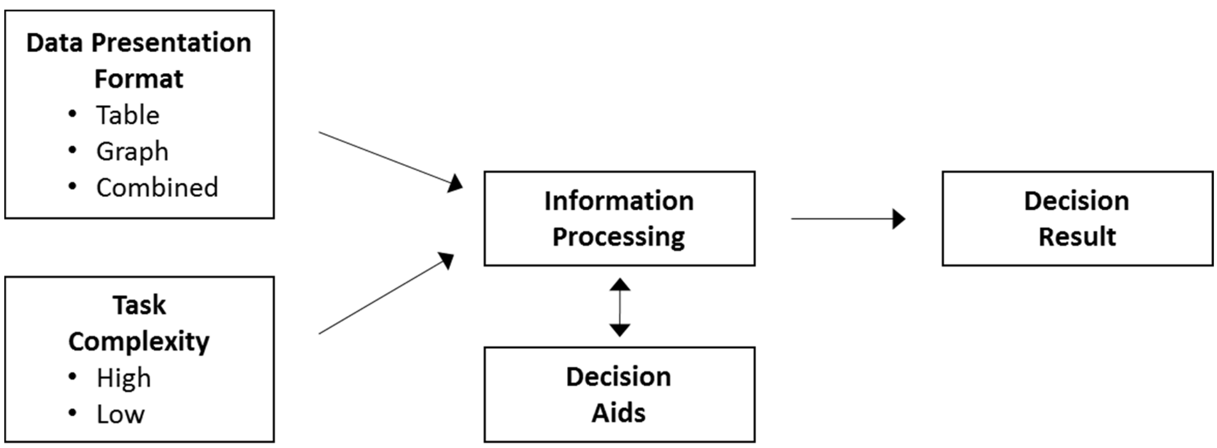


Figure 3.6. Research model.

The research model in Figure 3.6 builds on the model used by Fuglseth and Grønhaug (2000). As can be seen in the figure, the independent variables are data presentation format and task complexity. The dependent variable is decision result with information processing, including the use of decision aids, as mediating variables. The subjects could not change the format presented to them. However, as indicated by the relationship between information processing and decision aids, they could copy/extract data from the presentation format to Excel and thus extend the presentation formats to support their information processing. For example, they could make a graph from the table format, set up a table based on data values in the graph format and perform calculations on the data. They could also use a calculator, pen and paper.

As illustrated in the research model, I will study the effects of using three different presentation formats (graphs, tables or a combination of graphs and tables) on the decision- making performance. I will use two similar tasks of different complexity (low and high) to examine how differences in task complexity influence decision-making performance. The purpose of this thesis is, however, not to identify the 'best' presentation format as regards decision outcome but rather to gain more insight into *how* the data presentation formats are used in a managerial decision process. Therefore, in addition to studying the effects on decision-making performance, I will investigate whether, and possibly how, the various data presentation formats influence decision makers' mental representation and processes. This is a phenomenon that is poorly understood and incompletely documented in the literature (see, e.g. Kelton et al., 2010). I have therefore chosen an exploratory research design for my study.

In order to explore the relationships among the concepts addressed in the research model, I have conducted a laboratory experiment with six treatments, as shown in Table 3.1.

Table 3.1. Experimental treatments

		Data presentation format		
		Graph	*Table*	*Table and graph*
Task complexity	*Low*	I	II	III
	High	IV	V	VI

The findings from Fuglseth and Grønhaug's (2000) pilot study indicate that decision makers need both graphs and tables when solving problem tasks. Tables are considered important to provide details and a basis for further calculations, whereas graphs are believed to give an overview of relationships among variables. Their findings indicate this tendency quite clearly for the simple decision task. However, there was no clear pattern in their data for the complex task.

I will supplement their data set and will test the assumption that the subjects need both presentation formats. In order to do so, I build on the experimental design, including the spreadsheet models and task descriptions developed by Fuglseth (1999).

Previous studies on the effect of data presentation format on decision making are for the most part based on the assumption that decision makers are unaided. However, this is rarely the case. Therefore, in line with Fuglseth and Grønhaug (2000), I will take decision aids into account and study how access to such aids influences decision-making effectiveness.

3.2.1 Experimental setting

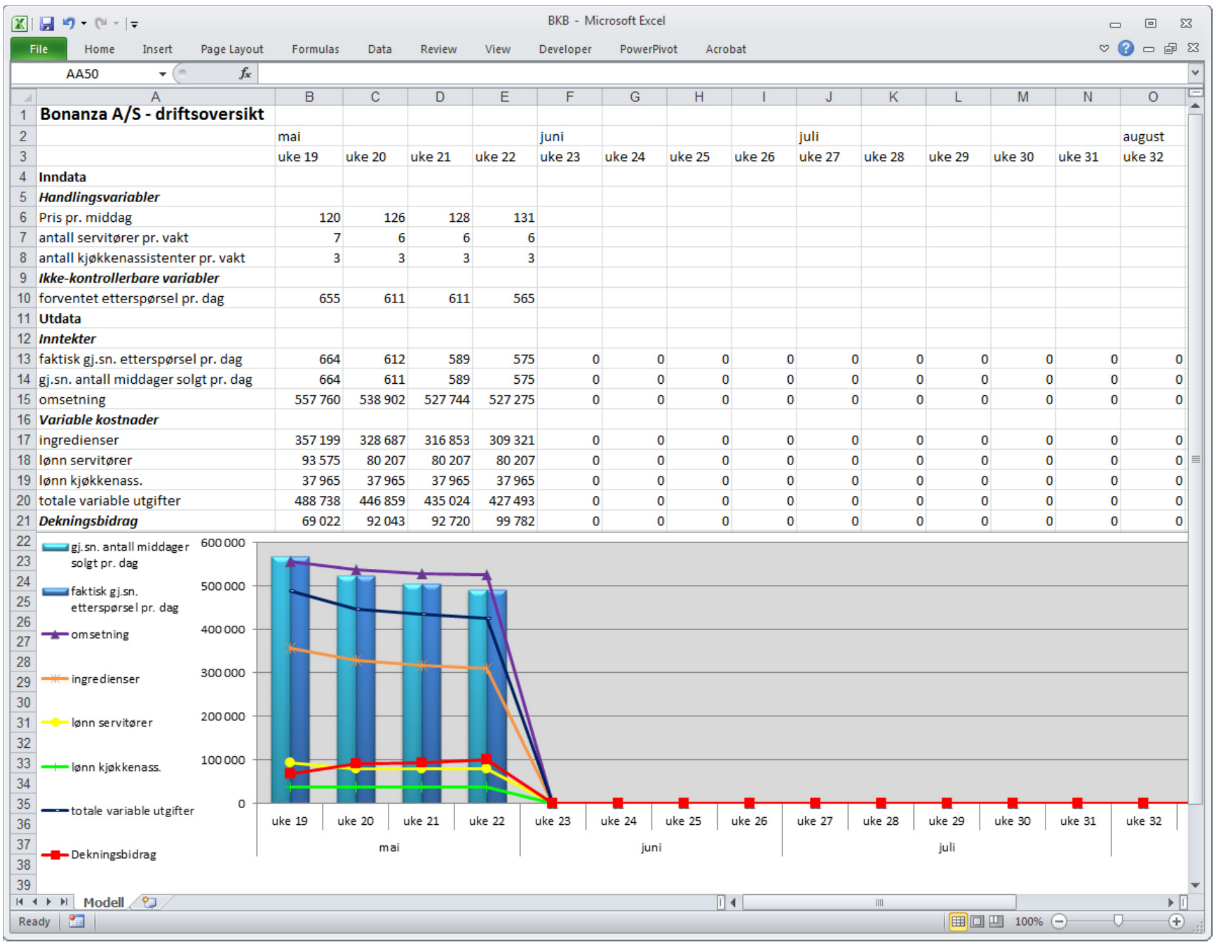
The experimental setting was the management of a summer restaurant. The subjects were given the task of running the restaurant for a period of four months (June–September) and therefore had to make 17 decisions, one each week. The subjects were told that the objective of the task was to maximise the total contribution for the four-month period; thus, the experimental task had an accuracy requirement. There was no time limit for solving the task.

A demand function was constructed for the relationship between the price of a meal and the demand for meals. The demand function was designed to create some amount of uncertainty so that the demand would not necessarily be the same each week, even though the price was held constant. The subjects were given data about the relationship between price and demand so that they were able to anticipate the demand at different prices. These data were given as prices with average demands and sales for a historical period of three years.

The sale of meals in the restaurant generated the income. Costs for ingredients and staff had to be deduced in order to calculate contribution. The number of meals sold (and with that, contribution and variable costs) was limited by actual demand and by capacity limitations in terms of waiters and assistants. If the subjects did not hire enough waiters/assistants, it would not be possible to cover the whole demand for a given week.

The subjects entered the values of the decision variables into a computerised system, which calculated and displayed the values of the result variables. The user interface of the system is a spreadsheet (see Figure 3.7 and Figure 3.8 a/b), software that the subjects were expected to be familiar with. The system displayed input data and output data for four 'historic' periods, as shown in the figures. These 'historic' data were supposed to hint at the optimal solution, for example by providing the data necessary in order to calculate the waiters' capacity (the number of meals each waiter could serve per day).

Figure 3.7 shows the user interface for one of the models used in the experimental setting.



Bonanza A/S - driftsoversikt	mai				juni				juli					august
	uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32
Inndata														
Handlingsvariabler														
Pris pr. middag	120	126	128	131										
antall servitører pr. vakt	7	6	6	6										
antall kjøkkenassistenter pr. vakt	3	3	3	3										
Ikke-kontrollerbare variabler														
forventet etterspørsel pr. dag	655	611	611	565										
Utdata														
Inntekter														
faktisk gj.sn. etterspørsel pr. dag	664	612	589	575	0	0	0	0	0	0	0	0	0	0
gj.sn. antall middager solgt pr. dag	664	611	589	575	0	0	0	0	0	0	0	0	0	0
omsetning	557 760	538 902	527 744	527 275	0	0	0	0	0	0	0	0	0	0
Variable kostnader														
ingredienser	357 199	328 687	316 853	309 321	0	0	0	0	0	0	0	0	0	0
lønn servitører	93 575	80 207	80 207	80 207	0	0	0	0	0	0	0	0	0	0
lønn kjøkkenass.	37 965	37 965	37 965	37 965	0	0	0	0	0	0	0	0	0	0
totale variable utgifter	488 738	446 859	435 024	427 493	0	0	0	0	0	0	0	0	0	0
Dekningsbidrag	69 022	92 043	92 720	99 782	0	0	0	0	0	0	0	0	0	0

Figure 3.7. User interface for the models used in the experiment.

Task complexity

The experiment was designed with two tasks of different complexity (see Appendix A). The differences were related to the number of variables and the degree of uncertainty in the relationships between the decision variables and the result variables. In the task low in complexity, the subjects were asked to make weekly decisions regarding the price of a meal and the number of waiters needed. Random variation in the demand function was set to 1% in order to introduce uncertainty. In the high-complexity task, the subjects also had to make decisions regarding the number of kitchen assistants. The random variation was set to 3% to increase the uncertainty, but still making it possible for the subjects to find the number of waiters and kitchen assistants needed in the optimal solution.

Related to Campbell’s (1988) typology of task complexity, both tasks can be characterised as problem tasks, that is, they have a well-specified outcome (maximise contribution), but there are a multiplicity of paths with interdependences among the variables involved in finding the optimal combination of meals and number of waiters/kitchen assistants.

Data presentation

The spreadsheet model was designed in three versions (see Appendix B), showing the results of the decision variables as graphs, tables or as a combination of graphs and tables. Common to all the three versions of the model used is that there is a clear distinction between input data (decision variables) and output data (result variables). The input data were presented equally in all three versions of the spreadsheet model.

Output data present the results of the decisions that the subjects enter into the model. The values of the output data are the same (given the same input data), but the presentation of the data varies in the three versions; see, for example, Figure 3.8 a) and Figure 3.8 b).

	A	B	C	D	E	F
1	**Bonanza A/S - driftsoversikt**					
2		mai				juni
3		uke 19	uke 20	uke 21	uke 22	uke 23
4	**Inndata**					
5	***Handlingsvariabler***					
6	pris pr. middag (kr)	120	126	128	131	
7	antall servitører pr. vakt	7	6	6	6	
8	antall kjøkkenassistenter pr. vakt	3	3	3	3	
9						
10	***Ikke-kontrollerbare variabler***					
11	forventet etterspørsel (måltid/dag)	655	611	611	565	
12						
13	**Utdata**					
14	***Inntekter***					
15	faktisk gj.sn. etterspørsel (måltid/dag	664	612	589	575	0
16	gj.sn. antall måltider solgt (måltid/dag	664	611	589	575	0
17	omsetning (kr/uke)	557 760	538 902	527 744	527 275	0
18	***Variable kostnader***					
19	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	0
20	servitører (kr/uke)	93 575	80 207	80 207	80 207	0
21	kjøkkenassistenter (kr/uke)	37 965	37 965	37 965	37 965	0
22	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	0
23	***Dekningsbidrag (kr/uke)***	69 022	92 043	92 720	99 782	0
24						
25						
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						

Figure 3.8a. Tabular presentation of output data.

	A	B	C	D	E	F
1	**Bonanza A/S - driftsoversikt**					
2		mai				juni
3		uke 19	uke 20	uke 21	uke 22	uke 23
4	**Inndata**					
5	***Handlingsvariabler***					
6	pris pr. middag (kr)	120	126	128	131	0
7	antall servitører pr. vakt	7	6	6	6	0
8	antall kjøkkenassistenter pr. vakt	3	3	3	3	0
9						
10	***Ikke-kontrollerbare variabler***					
11	forventet etterspørsel (måltid/dag)	655	611	611	565	0

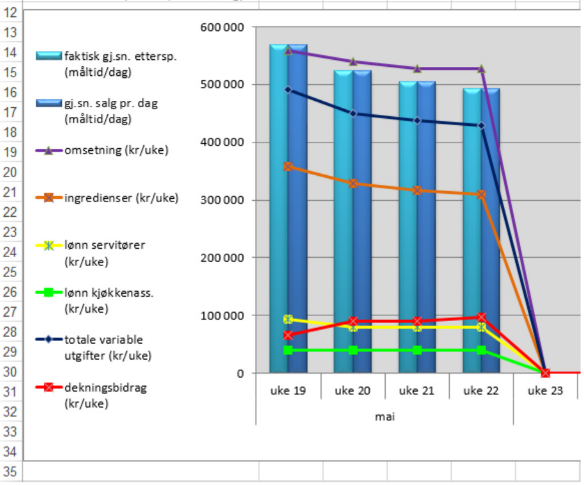


Figure 3.8b. Graphical presentation of output data.

The historical data on the relationship between price and demand was also given in different presentation formats – as a table in the version of the experiment with table displays, as an XY graph in the graphical version of the experiment and as a combination of the two in the combined version of the experiment. The historic prices ranged from NOK 120 to NOK 165 (see Appendix C).

3.2.2 Subjects

The subjects in this study were 42 Master's students from the Norwegian School of Economics. All the subjects were in their final year when the experiment was conducted. Therefore, the subjects should have the relevant background for handling the task they were given.

Invitations for attending the experiment were sent out to students in two specific Master's courses. The students were told that the experiment would give them experience in applying relevant economic theory to a practical case. They were also told that they would receive monetary compensation for attending. Students interested in participating in the experiment signed a participant list. Individual sessions were then arranged for each student. As described above, the subjects in this study were self-selected and not randomly assigned.

The subjects were expected to have sufficient knowledge of the spreadsheet program MS Excel, which was used as the user interface in the task to be solved during the experiment. All the subjects should have attended an introductory course in data processing in which the use of a spreadsheet in an economic setting is essential.

3.2.3 Measurement

The decision result was measured as total contribution for the period the restaurant was run. However, the complex and the less complex tasks did not have the same optimal solution, and therefore the decision results were not comparable across task types. Therefore, an index was created in order to make the decision results comparable. The index value is calculated by dividing the total contribution by maximum contribution.

The level of information processing was measured based on the four levels of information processing presented in the theory of cognitive complexity (Schroder et al., 1967; see section 3.1.2). By applying such a measure, I hope to obtain a better understanding of the differences in decision-making quality.

A seven-point scale was developed based on the description of the four levels of information processing in the theory (Schroder, et al., 1967, pp. 14–23) and by adapting a general manual for scoring structural properties from verbal responses (pp. 186–189). Levels 1, 3, 5 and 7 are the main levels, with rather clear rules for scoring, whereas levels 2, 4 and 6 are used when the subjects' responses indicate a development in information processing during problem solving, for example from level 3 to 5 but where level 5 is not clearly attained.

3.2.4 Data collection

Individual appointments were scheduled for each student participating in the experiment. During the experimental sessions, I was present in the room so that the subjects could ask questions if they needed help or if they were uncertain of anything. At the same time, I observed the subjects and took notes regarding their actions. The subjects were sitting at a desk in front of a 19-inch computer screen while working on the experimental task.

I started each experiment by introducing the subjects to their new 'job' as managers of the summer restaurant, Bonanza AS. I explained the task they were supposed to handle. The computerised system was also explained, that is, the decision variables and the result variables and how to run the system. In addition, I emphasised that the subjects were allowed to use decision aids and that there were no restrictions regarding how much time they were allowed to take when solving the task. In this oral introduction, I emphasised that the students had to use the system that was presented to them but that they were also allowed to carry out additional calculations in the spreadsheet model, open new spreadsheets for calculations or make their own graphs. In addition, they could use pencil, paper and a calculator. Appendix D shows the instructions used in the experiment.

After the brief introduction, I gave the subjects two detailed documents to study before they started working on the task. A task description contained all the information necessary to run the restaurant, such as what types of meals they could offer their customers, the costs involved in managing the restaurant, a description of how the demand varied throughout the day and access to labour. A document containing historical sales data showed a range of prices along with average demand and average sales for the last three years. The spreadsheet model was displayed with input data and output data for four 'historic' periods, showing how the former manager had run the restaurant.

The subjects were instructed to think aloud while solving the task. In this instruction, I followed the guidelines proposed by Ericsson and Simon (1980, 1993). For instance, the subjects were explicitly asked to focus on the primary task they had been assigned to solve but at the same time to verbalise the thoughts and feelings that naturally came to mind. To avoid potential biases from retrospective rationalisation or reconstruction effects, the subjects were never asked to describe or explain tasks performed at earlier stages of their decision-making process.

The method of data collection consisted of tape-recordings of the subjects' 'thinking' aloud while they were interpreting the data displays and making decisions. The results of using the spreadsheet system, including additional spreadsheets for calculations or graphs, were saved. The results from using paper and calculator were also saved. The tape recording and the use of the information system and decision aids were coordinated by the registration of the week number and comments about the use of decision aids on the tapes.

As I have also used data from Fuglseth and Grønhaug (2000) in my analysis, I coordinated my data collection procedures with them before I conducted the experiments. This included making sure that all the experiments followed the same procedure (i.e. using the same oral introduction, explanation of the system, handing out task descriptions and historical data, etc.) and that the different elements followed the same order. Furthermore, discussions were held about the role of the observer, for example how 'helpful' the observer should be during the experiments.

After the subjects had completed the task, I asked them to state their electives. The reason for doing so is that there may be differences in educational background due to differences in electives. The think-aloud recordings were eventually transcribed (see Appendix E) and analysed. The analysis is presented in the next section.

3.3 Data analysis

3.3.1 Presentation and explanation of the data matrix

The data set consists of data for 42 subjects. However, two of the subjects, S1-23 (index = 0.5131) and S1-39 (index = 0.5918), performed considerably worse than the rest. An inspection of the verbal protocols for these two subjects revealed that they had maximised sales instead of contribution. This misunderstanding is not related to the presentation format, and I therefore decided to remove them from the data set. Appendix F.1 shows the data set including all 42 subjects.

Table 3.2 hows my initial data analysis without subjects S1-23 and S1-39.

Table 3.2. Results of initial data analysis

No.	Task type	Pres. Format	Electives	Computation	Graph	Analytic	Spatial	Level of info.proc.	Index	Group average 1	Group average 2
S1_01	L	b	d	3		1		5	0,9993		
S1_02	L	b		2		1		4	0,9698		
S1_03	L	b		3		1		5	0,9998		
S1_04	L	b					1	3	0,9965	0,9914	
S1_05	L	g	d	3		1		4	0,9300		
S1_06	L	g	d	3		1		7	0,9954		
S1_07	L	g	d	3	2	1		4	0,9587		
S1_08	L	g					1	3	0,9864	0,9676	
S1_09	L	t	d	2		1		4	0,9788		
S1_10	L	t		2	2	1		3	0,9775		
S1_11	L	t		1	2	1		3	0,9801		
S1_12	L	t		3	2	1		5	0,9967	0,9832	0,9807
S1_13	H	b	d	1		1		4	0,9105		
S1_14	H	b	d	2		1		5	0,9026		
S1_15	H	b	d	1		1		4	0,9743		
S1_16	H	b	d	3		1		7	0,9878		
S1_17	H	b		1		1	1	4	0,7640		
S1_18	H	b					1	3	0,8083		
S1_19	H	b					1	5	0,9240		
S1_20	H	b					1	3	0,8284		
S1_21	H	b					1	4	0,8724	0,8858	
S1_22	H	g	d				1	3	0,9028		
S1_24	H	g	d	2	2	1	1	3	0,7828		
S1_25	H	g	d	2	2	1	1	3	0,8128		
S1_26	H	g	d	3		1		7	0,9358		
S1_27	H	g	d				1	3	0,8530		
S1_28	H	g					1	3	0,9331		
S1_29	H	g					1	2	0,7838		
S1_30	H	g					1	3	0,8061		
S1_31	H	g					1	3	0,8896		
S1_32	H	g					1	2	0,8274	0,8527	
S1_33	H	t	d	3	2	1	1	5	0,9843		
S1_34	H	t	d	3	2	1	1	7	0,9663		
S1_35	H	t	d	1		1		3	0,7747		
S1_36	H	t	d	3		1		5	0,9844		
S1_37	H	t	d	2		1		6	0,9184		
S1_38	H	t	d	2		1		4	0,9163		
S1_40	H	t		1		1		3	0,8394		
S1_41	H	t					1	1	0,6917		
S1_42	H	t		1		1		5	0,8782	0,8838	0,8733
	H = high	b = both	d = data	1 = simple	1 = line			1 = low			
	L = low	t = table		2 = compr.	2 = XY			7 = high			
		g = graph		3 = marg.	3 = XY, compr.						

In the following, I will explain the table, including how I have coded some of the columns based on my analysis of the verbal protocols.

- *No.* – Identifies the subject.
- *Task type* – Denotes the complexity of the decision problem given to the subject; H=high complexity and L=low complexity.

- *Pres. Format* – Denotes the data presentation format used for each subject; b = both table and graph, g = graph only and t = table only. Appendices B and C contains examples of the different presentation formats.

- *Electives* – Denotes whether the subject has an elective in business data processing; d = elective in business data processing. Having such an elective is used as a proxy for having adequate skills for handling/utilising a spreadsheet system.

- *Computation* – Denotes whether the subject performed calculations, either in a spreadsheet, on paper or by means of a calculator. The scope of the calculation is classified on a scale from 1–3.

    - The value 1 indicates that the subject carried out rather simple calculations, such as summing two numbers.

    - The value 2 indicates that the subject made comprehensive computations. An example is that the subject might have generated a table of different prices and then calculated the contribution of these prices.

    - The value 3 indicates that the subject performed advanced calculations. An example is that the subject made marginal analyses regarding the number of kitchen assistants and the number of waiters necessary for different demand alternatives. Furthermore, the subject may have performed marginal analyses in order to learn how the contribution is affected by a one-unit increase in the price.

- *Graph* – Denotes whether the subject prepared their own graphs. The degree of details in the graph is graded on a scale from 1–3, as follows:

    - The value 1 is used for rather simple line graphs without details.

    - The value 2 is used when the subject made an XY graph, where values for X and Y are plotted and a line is drawn between the plots.

    - The value 3 is used for comprehensive and detailed XY graphs. An example is that the subject developed scales for the X and Y axes and used the graph to extrapolate data values.

- *Analytic* – Denotes my evaluation of whether the subject used analytic processes in their problem solving. The variable is binomial, and the following criteria were used for analytic processing, based on sections 2.1.2 and 3.1.1:
    - Focus on discrete data values.
    - Extraction of discrete data values from the presentation formats.
    - Calculations documented in the spreadsheet or on paper.
    - Articulation of calculations – performed either mentally or by using the calculator.
    - Detailed and advanced calculations, such as marginal analyses.
    - Development of additional tables.
- *Spatial* – Denotes my evaluation of whether the subject used spatial processes in their problem solving. The variable is binomial, and the following criteria were used for spatial processing, based on sections 2.1.2 and 3.1.1:
    - Description of trends, for example, 'The demand curve is declining. ….'
    - Comparison of trends, such as, 'The sales are increasing, but the contribution is decreasing. …'
    - Ordinal comparisons, such as 'The sales in week 24 are higher than in week 25, but the contribution is lower.'
    - Ordinal anticipations, such as 'If I increase the price to 160, I believe I can reduce the number of waiters and kitchen assistants. …'
    - The subject was silent for long periods of time.
    - The subject expressed a need for graphs or created additional graphs.
- *Level of info. proc.* – The level of information processing was measured using the seven-point scale based on the four levels of information processing (Schroder et al., 1967) presented in section 3.1.2 and section 3.2.3. The subjects' verbal responses were scored as follows:

- A value of 1 is used when the subject did not use any critical judgments but exclusively made use of a 'trial-and-error' strategy. This information processing level can be characterised as a 'black and white' way of thinking (e.g. 'if it is not this, it has to be that', even though no explanations of why this is so are given), exhibiting a certainty that the chosen alternative is the best. When experiencing unexpected/bad results, the subjects uncritically tried new values for the input data without any further effort to understand the problem.

- Level 2 is used when the subject indicated there might be some causal relationships in the data, even though he/she did not pursue this thought any further.

- Level 3 is used when the subject introduced expectations of causal relationships. At this level, however, the subject considered only one causal relationship at a time, characterised by 'either/or conditions (e.g. 'If I increase the price, I expect the result to be better than in the previous week').

- Level 4 is used when the subjects indicated an understanding of causal relationships between more than two variables at a time.

- Level 5 is used when the response indicated a comprehensive understanding and evaluation of causal relationships between the variables. For example, they could tell that demand will be reduced by increasing the price and also what effects the reduced demand will have with regard to determining the number of kitchen assistants and waiters on duty, for example, 'If I increase the price by NOK 9 and at the same time reduce the number of waiters by one and keep the number of kitchen assistants unchanged, I expect the contribution to increase by NOK 14 000.'

- Level 6 is used for subjects who certainly earned level 5 but also tried to deduce functional dependencies between variables as numeric quantities.

- Level 7 is used to indicate the subject deduced functional dependencies between variables. These functions were then used to calculate the 'correct' answer to the decision problem. Compared to level 5, the subject could not only say that demand will be reduced by increasing the price by one unit but also tell how much the demand will be reduced.

Two coders coded the data on the level of information processing in this study. The principal coder coded the data that she initially collected, and developed the guidelines for how to code the subjects. When I added more subjects to the data set, I coded these subjects by following the guidelines developed. In addition, the principal coder and I recoded each other's data on the level of information processing in order to increase intercoder reliability. We mostly had the same scores for the subjects. In cases where we had different scores, we discussed the protocols and made a joint decision about how they should be coded according to the guidelines. The verbal protocols were analysed and coded independently of the subjects' performance on the contribution index.

The columns 'Analytic' and 'Spatial' were only coded once by either the principal coder or myself; however, random checks were conducted.

- *Index* – Denotes the decision results shown as an index. An optimal solution exists for the experimental tasks. However, the optimal solutions for the high-complexity task and the low-complexity task are not the same. In order to make the decision results comparable across the two tasks, an index value was calculated as the subjects' decision result divided by the optimal solution for the task.

- *Group average 1* – Denotes the average value of the contribution index for each treatment group.

- *Group average 2* – Denotes the average value of the contribution index for the low- and high-complexity tasks.

Table 3.2 shows a clear difference in the decision results between the low-complexity task and the high-complexity task. The average value of the contribution index for the low-complexity task is 0.981, and the average value of the index for the high-complexity task is 0.873 ($p < 0.0001$), indicating that the experimental tasks are perceived as having different degrees of subjective complexity.

An inspection of Table 3.2 indicates that subjects presented with both graphs and tables performed better than subjects in the other categories, both in the low-complexity task and in the high-complexity task. Table 3.3 summarises the average values of the contribution index for the different presentation formats in the low-complexity and high-complexity tasks.

Table 3.3. Average values of the contribution index for different presentation formats in the low-complexity and high-complexity tasks.

		Presentation format		
		Graph	*Table*	*Table and Graph*
Task complexity	*Low*	**0.9676** N=4, st. dev.=0.030	**0.9832** N=4, st. dev.=0.090	**0.9914** N=4, st. dev.=0.014
	High	**0.8527** N=10, st. dev.=0.059	**0.8838** N=9, st. dev.=0.100	**0.8858** N=9, st. dev.=0.075

These findings indicate that decision makers need both graphical and tabular presentation formats (see section 3.1.1). In particular, the graph format seems to be less effective in supporting the subjects' decision making. This tendency is present both in the low-complexity task and in the high-complexity task.

3.3.2 Verbal protocol analysis

In this section, I will present my analysis of the verbal protocols, which is the basis for my coding of the columns labelled Computation, Graph, Analytic, Spatial and Level of info. proc. in Table 3.2.

Analytic and spatial processes

In the low-complexity task, nine of the 11 subjects mainly used analytic processes. All these subjects also performed additional calculations using decision aids. Two of the subjects presented only with graphs (S1-06 and S1-07) extracted numeric data values from the graphs and organised these data into tables to be able to perform the calculations. Three of four subjects presented only with tables created additional XY graphs that helped them visualise the relationships among variables, that is, between price, demand and sales. Two of the 11 subjects seem to have mainly used spatial processes (S1-04 and S1-08).

In the high-complexity task, seven of the 10 subjects presented only with graphs mainly used spatial processes. None of these seven performed calculations. Two of the subjects (S1-24 and S1-25) used both analytic and spatial processes, whereas one subject (S1-26) mainly used analytic processes. The three subjects using analytic processes (S1-24, S1-25 and S1-26) performed additional calculations. S1-24 and S1-25 performed calculations using a calculator, pen and paper, whereas S1-26 used the spreadsheet. S1-26 developed a table in order to perform rather advanced calculations comprising marginality considerations. S1-26 was the highest performing subject of the subjects receiving only a graph.

Eight of the nine subjects presented only with a table used analytic processes. These eight subjects performed calculations in the spreadsheet or using a calculator, pen and paper. Three of these subjects (S1-33, S1-34 and S1-36) performed calculations that included elasticity and marginality considerations. S1-33 and S1-36 almost correctly derived the relationships among the variables before they started making decisions. These three subjects were among the highest performing subjects solving the high-complexity task.

Of the nine subjects receiving the combined format, four subjects (S1-18, S1-19, S1-20 and S1-21) mainly used spatial processes, four (S1-13, S1-14, S1-15 and S1-16) mainly used analytic processes and one (S1-17) used both spatial and analytic processes. All the subjects using analytic processes also performed additional calculations using a calculator, pen and paper or the spreadsheet.

The results in Table 3.2 show that the use of mainly spatial processes can be rather effective in the high-complexity task; see, for instance, subjects S1-19, S1-22 and S1-28. Common to these three subjects is that they performed some rough mental calculations before they started making decisions; for example, they estimated a starting price and the capacities of the waiters and kitchen assistants. They then used a systematic trial-and-error strategy, looking primarily at the development of the contribution until they found the solution. As an example, S1-19 made his first decision (week 23) with a price of NOK 140, five waiters and three kitchen assistants and then stated: 'I see that the contribution has increased, but the actual average demand is low… I need to do some calculations.' He was silent for a while, and then he continued: 'That's good!', and he entered a price of NOK 150, four waiters and two kitchen assistants into the model. He continued: 'I see that the contribution has increased and the demand has dropped.' S2-19 continued to increase the price and adjust the number of waiters and kitchen assistants for the next 12 weeks until he ended on a price of NOK 158, four waiters and two kitchen assistants.

A closer look at the contribution index in Table 3.2 shows that the highest performing subjects (S1-15, S1-16, S1-26, S1-33, S1-34 and S1-36) mainly used analytic processes. These subjects considered several alternatives before they made their decisions. Their considerations required calculations to compare and integrate data in the model. For example, most of these subjects were able to calculate the exact capacity constraints for the waiters and the kitchen assistants and relate them to the contribution. They also used the historical data showing the relationship between price and demand in order to calculate their optimal solution. A table presentation format facilitates such calculations. However, S1-33 and S1-34 also made XY graphs to get an overview of the relationships among variables. This indicates that effective decision makers needed both tables and graphs, but we note that the need for graphs might be particularly related to XY graphs.

When comparing the low-complexity task and the high-complexity task, we see a greater proportion of subjects mainly using spatial processes in the high-complexity task (42.9%) than in the low-complexity task (16.7%). This finding is in accordance with Vessey (1994), who claims that in tasks with a high degree of complexity, decision makers may prefer or may even have to use spatial rather than analytic processes due to limited cognitive capacity. This tendency is more evident for the subjects without an elective in business data processing. In the high-complexity task, 76.9% of the subjects without the elective mainly used spatial processes, but among the subjects with the elective, only 13.3% mainly used spatial processes. Thus, it seems to be a tendency that having an elective in business data processing facilitates the use of analytic processes. I will elaborate more on this finding later in the analysis.

Level of information processing

The theory of cognitive complexity focuses on the type of relationships among variables that humans use in their information processing (Schroder et al., 1967, pp. 3-5). Thus, the level of information processing can be used to assess both analytic and spatial processes. For example, it is possible for a subject using mainly spatial processes to be scored at level five on the level of information processing index as long as the subject integrates all aspects of the model when solving the task.

The level of information processing is assessed based on thorough analyses of the verbal protocols.

My data set consists of relatively few subjects, but the average values of the level of information processing index are almost the same for the low-complexity and high-complexity tasks. The average value of the level of information processing index for the low-complexity task is 4.17, whereas the high-complexity task has an average of 3.93. This is in accordance with my expectations; both tasks are designed so that information processing of level five or higher is necessary if the subjects are going to be able to calculate a solution close to the optimal solutions for the two tasks. Level five on the level of information processing scale indicates that the subjects have a comprehensive understanding of the relationships among variables in the model and also that they are able to evaluate and integrate these relationships at the same time. A subject with a score of four on the level of information processing index, however, is not able to fully integrate the relationships among variables.

Evaluating the subjects' level of information processing improves my understanding of the high and low performers as regards results in a way that the distinction between analytic and spatial processes does not. The level of information processing involves an assessment of how the subjects actually solved the task, that is, *how* they use the analytic and spatial processes and not only *what* type of processes they used in handling the tasks.

A comparison of the columns showing the level of information processing and the contribution index in Table 3.2 gives the impression that subjects who get a high score on the level of information processing index are also the ones achieving the best results (see, e.g. subjects S1-03, S1-06, S1-12, S1-16, S1-26 and S1-34). To test this impression, I generated a non-parametric correlation matrix; see Table 3.4.

The correlations matrix shows a significant correlation between the contribution index (result) and the level of information processing ($r = 0.564$, $p < 0.001$). This is in accordance with the theory of cognitive complexity that individuals with a well-developed construct system are supposed to perform more effectively (Schroder, et al., 1967, p. 10).

Table 3.4. Non-parametric correlations

	Result (index)	*Level of info. Proc.*	*Additional comp.*	*Elective*
Result (index)	1.000			
Level of info. Proc.	$.564^{***}$	1.000		
Additional comp.	$.555^{***}$	$.735^{***}$	1.000	
Elective	.160	$.445^{**}$	$.572^{***}$	1.000

Note. Spearman correlation. n=40.

** Correlation is significant at the 0.01 level (2-tailed).

*** Correlation is significant at the 0.001 level (2-tailed)

An inspection of Table 3.2 shows, however, that some individuals that I assigned a high score on the level of information processing index did not perform well on the contribution index. Subjects S1-05, S1-34 and S1-36, for instance, were given high scores on the level of information processing index but attained lower economic results than others in their groups with lower scores on the level of information processing index. Analyses of the verbal protocols for these subjects show that the rather low value on the contribution index was mainly due to calculation errors. For example, S1-34 almost correctly derived the general relationships among the variables before he started making decisions. However, he made his calculations using a calculator, pen and paper and unfortunately lost track of the development of the calculations when he decided the number of kitchen assistants in his optimal solution. Therefore, he ended up employing more assistants than necessary and thus losing money.

Table 3.2 also shows individuals receiving low scores on the level of information processing index who performed well on the contribution index. For example, subject S1-28 was the second highest performing subject in his group as regards contribution index but received only a three on the level of information processing index, that is, conditioning. An inspection of the verbal protocol for S1-28 reveals that he did not make a lot of effort to understand the relationships among variables but indicates that his good performance was mainly due to luck in his 'trial-and-error' strategy. For example, in week 27 (his fifth period), he stated: 'Let's try the same amount of people on duty, but I try to increase the price … let's say 155… Wow! The contribution is constantly increasing.'

Table 3.4 also shows significant correlations between level of information processing and additional calculations ($r = 0.735$, $p < 0.001$). According to the scoring manual presented in section 3.3.1, those who have been given high scores on the level of information processing index (i.e. level five and above) are the ones who were able to integrate the relationships among variables. An inspection of Table 3.2 shows that those who were given high scores on the level of information processing are the ones who performed marginal calculations and therefore were given high scores on additional calculations. This indicates that those who want to integrate variables have a need for marginal calculations, especially in tasks like the one being used in my experiment that demand accuracy.

The correlation matrix in Table 3.4 shows that having made additional calculations is strongly related to having an elective in business data processing ($r = 0.572$, $p < 0.001$). This is in accordance with my expectations that subjects having data processing skills are in general more capable of using the functionality of a spreadsheet program. Such skills include building dynamic models for calculations, utilising built-in functions and creating graphs based on data in the spreadsheet. Using a spreadsheet program may thus facilitate marginal calculations.

For the present purpose, having an elective in business data processing is used as a proxy for whether subjects have the necessary data processing skills for handling a spreadsheet program. Of the 40 subjects in the experiment, 20 had the elective.

According to Table 3.2 18 of the 20 subjects with an elective course in business data processing performed calculations while only eight out of 20 subjects without such an elective performed calculations.

A closer look at Table 3.2 indicates that the tendency that subjects with an elective perform more advanced calculations seems to be particularly evident for the high-complexity task. This observation is supported when running the correlation analysis for the low-complexity and high-complexity tasks separately. For the low-complexity task, the correlation analysis shows no significant relationship between additional computation and having an elective in business data processing. However, the relationship is significant when the analysis is run for the high-complexity task instead ($r=0.719$, $p<0.001$). The reason may be that individuals need the support of decision aids, especially a spreadsheet program, particularly in the high-complexity task where there are more values to consider.

When aggregating values for the level of information processing index for all the subjects in Table 3.2, we see some tendencies related to presentation format. Table 3.5 shows the aggregated data.

Table 3.5. Level of information processing

		Presentation format		
		Graph	*Table*	*Table and Graph*
Task complexity	*Low*	**4.50** N=4, st. dev.=0.957	**3.75** N=4, st. dev.=1.732	**4.25** N=4, st. dev.=0.957
Task complexity	*High*	**3.20** N=10, st. dev.=1.398	**4.32** N=9, st. dev.=1.803	**4.33** N=9, st. dev.=1.225
	Total	**3.57** N=14, st. dev.=1.555	**4.15** N=13, st. dev.=1.573	**4.31** N=13, st. dev.=1.109

When analysing the subjects independent of task complexity, there is a tendency (non-significant) that subjects with the combined format on average are scored higher on the level of information processing index than are subjects with either graphs or tables. Subjects presented with only the graph format seem to achieve the lowest average score on the level of information processing index. The same tendency is present when analysing the high-complexity task. However, in the low-complexity task, the subjects presented only with a graph seem to achieve the highest average score on the level of information processing index. This is in accordance with the above analyses of the data in Table 3.2; three of the four subjects in this treatment group performed advanced additional calculations. Two of these three subjects even copied data from the graph to additional data tables that they then used for their calculations. The above analyses reveal a significant positive correlation between level of information processing and additional calculations.

3.3.3 Adjusted results

Table 3.2 as well as the analysis of the verbal protocols, reveals that some subjects (S1-06, S1-07, S1-10, S1-11, S1-12, S1-26, S1-33 and S1-34) made adjustments to the presentation format they originally received. Three subjects presented only with graphs (S1-06, S1-07 and S1-26) prepared their own data tables in a spreadsheet based on the graphical presentation format they were given. These three subjects performed very well, and common to them is that they carried out marginal considerations/calculations of some sort. In order to do so, they needed to extract/copy data from the graphs to a data table.

Further, from the graph column in Table 3.2 it can be observed that five of the 13 subjects presented only with tables made graphs themselves (S1-10, S1-11, S1-12, S1-33 and S1-34) in order to visualise the data they received in table format. Common for these subjects is that they created XY graphs that helped them visualise the relationships between variables (e.g. price, demand and sales). This indicates that the subjects needed both graphs and tables in order to solve the decision task properly. However, none of the subjects created line graphs in order to visualise trends in the variables over time.

Common to all the subjects adjusting their original presentation format is that they actually developed additional presentation formats to support their decision-making process and thus used both tables and graphs when solving the decision task. However, the subjects' additional 'self-made' presentation format did not include as many details as the spreadsheet model showing both graphs and tables and the additional historical data they received on paper.

By taking into account the development of additional presentation formats, we can regroup the results according to what presentation formats the subjects used in their decision making. The criteria for moving a subject from the 'Graph' group to the 'Table and Graph' group were that the subject copied data from the graphs and systematically organised the data into tables with labelled columns (and then used the table for further calculations). Subjects who occasionally copied single data elements were not moved to the combined group. The criteria for moving subjects from the 'Table' group to the 'Table and Graph' group were that the subjects created a graph and then used the graph in their further decision-making process.

Table 3.6 gives an overview of the average values on the contribution index after such a regrouping, and the table now shows a more powerful effect regarding the use of a presentation format consisting of *both* tables and graphs:

Table 3.6. Cross table: presentation format and task complexity.

		Presentation format		
		Graph	*Table*	*Table and Graph*
Task complexity	*Low*	**0.9582** N=2, st. dev.=0.040	**0.9788** N=1, st. dev.=N/A	**0.9860** N=9, st. dev.=0.015
	High	**0.8435** N=9, st. dev.=0.054	**0.8576** N=7, st. dev.=0.099	**0.9049** N=12, st. dev.=0.073

A comparison of Table 3.3 and Table 3.6 shows that eight subjects that received either a table or a graph format felt a need for the combined format. The comparison also reveals that, especially in the high-complexity task, the generation of the additional format improved the outcome. In this task, the difference in contribution index value for the 'Graph' and the 'Table and Graph' cells is significant at the 0.05 level with a large effect size (Cohen's d = 0.983) (Cohen, 1988). This result indicates that decision makers need both presentation formats in problem tasks.

The regrouping of the data calls for an updated analysis of the relationship between level of information processing, task complexity and presentation format. Table 3.7 shows the average level of information processing for the six treatment groups, as well as totals for the three presentation formats.

Compared to Table 3.5, the data now shows a more consistent pattern in both tasks. There seems to be a tendency that subjects presented with a graph attain a lower level of information processing on average than do the subjects with the table format or with the combined format. The difference between the graph group and the combined group is significant for the high-complexity task (p=0.001) and for totals (p=0.001). In the low-complexity task, the results are not significant.

Table 3.7. Level of information processing after having adjusted for the presentation formats actually being used

		Presentation format		
		Graph	*Table*	*Table and Graph*
Task complexity	*Low*	**3.50** N=2, st. dev.=0.707	**4.00** N=1, st. dev.=N/A	**4,33** N=9, st. dev.=1.323
	High	**2.78** N=9, st. dev.=0.441	**3.86** N=7, st. dev.=1.676	**4.83** N=9, st. dev.=1.467
	Total	**2.91** N=11, st. dev.=0.539	**3.875** N=8, st. dev.=1.553	**4.62** N=9, st. dev.=1.396

3.4 Discussion and conclusion

3.4.1 Summary of findings

The purpose of this study was to enhance the understanding of the relationship between data presentation formats and decision-making effectiveness.

My results indicate that subjects presented with the combined presentation format on average generate the highest contribution. This tendency was enhanced when I corrected for the presentation formats actually being used by the subjects after they had complemented their original presentation formats with additional graphs and tables. Thus, the results indicate that decision makers need both presentation formats. This is in accordance with Paivio (1986, 1991, 2007), that the analytic system and the spatial system may support each other in human problem solving.

Another tendency in my results is that, in general, subjects with the highest level of information processing generated the best results. This corresponds well with Schroder et al. (1967), that people with well-developed construct systems are expected to handle complex situations more effectively than those with a simpler construct system. My findings indicate that these subjects had a need for carrying out accurate and advanced calculations. Tables seem to facilitate the use of analytic processes by providing details and a basis for performing advanced calculations. Therefore, tables seem to be important in the type of tasks I studied.

Graphs give an overview of relationships among variables and seem to be a useful supplement to the tables. However, graphs alone seem to be the least effective presentation format for the kind of tasks being used in this study. Graphs (seem to) facilitate spatial rather than analytic processes. However, as indicated above, subjects mainly using analytic processes generated the best results.

My results indicate a shift from analytic to spatial processes as the complexity level of the task increases. This is in line with the cognitive fit theory (Vessey, 1994). However, access to a spreadsheet program (i.e. a decision aid) seems to reduce the need for such a shift, provided that the decision maker possesses the skills required to master the spreadsheet program. The results in Table 3.2 show that nine of the 15 subjects having taken the elective in business data processing mainly used analytic processes, but only one of the 13 without such an elective mainly used analytic processes. An explanation for this observation may be that the subjects with the elective are better able to generate tables and thus can use analytic processes rather than spatial. Thus, with the support of a decision aid, subjects were able to achieve good results even as the task complexity increased.

3.4.2 Limitations of the study

One of the weaknesses of this study is that the number of subjects that participated in the experiment was too low. Thus, a follow-up study should be conducted in which the number of subjects is increased in order to test whether the tendencies in the current study are supported.

During the analyses of the verbal protocols, I observed several phenomena that need to be addressed in a follow-up study. First, I observed a large variation in terms of the extent to which the subjects verbalised their thoughts during the experiments. Several subjects, especially those who received a graphical presentation format, stopped verbalising while solving the experimental task. Even though I repeatedly reminded them to think aloud, they still did not comply with this request. In particular, it seemed to be hard for the subjects to think aloud while studying the graphs in the spreadsheet model and the paper graphs. This is in accordance with Ericsson and Simon (1980, 1993), that information represented in a complex non-verbal (i.e. spatial) code may be difficult to articulate because a translation to a verbal code is required. Thus, Ericsson and Simon postulate that such a condition may cause the verbalisation to be incomplete or even to cease completely.

When the subjects' verbalisations were incomplete or in periods when they became silent, it was difficult to infer the mental processes they used and to assess their level of information processing from the verbal protocols. Eye-tracking technology could provide additional data about where subjects directed their attention on the screen when they stopped verbalising. Thus, a follow-up study using eye-tracking technology should be conducted to better understand the processes underlying the decision results when the subjects were silent.

Another observation from analysing the verbal protocols was that there was great variation in terms of how much time each subject needed in order to solve the experimental task. Some subjects solved the task in less than 15 minutes, whereas others took more than one and a half hours. I also observed clear differences in the time the subjects spent before they made their first decision. Some subjects started making decisions and entered values into the spreadsheet model immediately after having read the task description, without any further effort to clarify and define the problem. Others spent most of their time trying to understand the task and infer relationships among the variables so that they could find an optimal solution before they entered any values at all into the decision model. The experimental task was designed so that decisions had to be made every week for a period of 17 weeks. However, when the students indicated that they thought they had found optimal values for that week's decision, variables were entered into the model and then copied throughout the remaining weeks in the model. The subjects who spent most of their time working on the problem before they started making decisions often found their final solution early in the game and thus required very little time after having made their first decision until completing the game.

Thus, another limitation of the study is that I did not include accurate measurements of the subjects' time consumption. Such measurements should include the total time used by each subject and how this time was allocated in different phases, such as the amount of time spent until the first decision was made. By including such measurements of time consumption, I would have been able to look for possible trade-offs between accuracy and effort (cf. Vessey, 1994). Therefore, a follow-up study should provide better control of time consumption and time allocation. Furthermore, such a study should also check how many weeks the subjects needed to find their final solution.

When I analysed the verbal protocols, I noticed that a few of the subjects were not particularly motivated to try to solve the task. For these subjects, it seemed to be more important to finish the experiment so that they could receive the reward of NOK 200 for attending. Most others, however, seemed to live up to the role as a manager being motivated to maximise his/her total earnings more than being motivated by the promised reward. Many of these highly motivated students were very concerned with finding the optimal solution before they started entering values into the model so that they did not risk losing money on non-optimal decisions early in the game.

The above analysis revealed a limitation related to the subjects and their motivation for attending and performing well on the experimental task. A measure of motivation should have been developed and applied in the experiment.

In the current study, having an elective in business data processing was used as a proxy for the subjects' spreadsheet skills. In addition, I evaluated the subjects' spreadsheet skills subjectively during my observations. However, such an evaluation presupposes that the subjects have an observable use of the spreadsheet during the experiment. In a follow-up study, the subjects should be asked about their experiences using a spreadsheet.

Based on the above, there is a need for a follow-up study where I take the following into consideration: I will increase the number of subjects in order to test whether the tendencies in the current study are strengthened. Eye-tracking technology and web-cam recordings of the participants will be used to get a better understanding of the subjects' behaviour during silent periods, and I will develop and implement more accurate measures of time consumption during the experiments. Furthermore, I will incorporate a questionnaire to measure the subjects' motivation for performing well and describe their experiences and relevant courses relating to handling spreadsheets.

4 STUDY 2

In this chapter, I present my second study investigating the relationship between data presentation format and decision-making effectiveness. The study is a follow-up study based on Study 1. I use the same research design and experimental setting as in the previous study. However, I take into account and adjust for some of the limitations of the first study.

In this study, I include eye tracking as a process-tracing method to monitor the subjects' eye movements. By doing so, I hope to better understand the decision-making processes, especially those of the silent subjects. I also hope that eye-movement data will help me to better understand how subjects utilise the combined presentation format consisting of both graphs and tables. I include more accurate measures of how much time the subjects need to solve the experimental task and how they allocate this time into two different phases, such as problem definition and problem solving. I also include a measure of the subjects' motivation for working on the tasks to see if such motivation influences the results. Finally, by adding more data to the existing data set from Study 1, I want to see if the tendencies from Study 1 are strengthened or not.

This chapter is organised as follows. In the next section, I introduce eye tracking as a data collection method. Next, I describe extensions to my research design and data collection procedure from Study 1. In the section after that, I present the data set for Study 2, as well as my analysis of the eye-movement data for extremes, that is, high and low performers for the six treatment groups in the experiment. Finally, I discuss the implications of the study and make recommendations for further research.

4.1 Introduction to eye-tracking and eye-movement analyses

4.1.1 Perception through the human eye

As discussed above, including data about what the subjects look at during the experiments may help us to better understand the cognitive processes underlying decision making. Eye tracking is the process of determining where individuals are looking. In the next sections, I will elaborate on eye tracking as a process-tracing method. However, before doing so I will first discuss some basic characteristics of the human visual system.

Human eyes cover a visual field of about 180 degrees horizontally (90 degrees to the left and 90 degrees to the right) and 90 degrees vertically. The term visual field denotes the part of the surroundings that open eyes perceive when both the eyes and the head are kept still. However, due to the way the human eye works, only a small area of the visual field appears clear at a time. This area covers about two degrees in the centre of the visual field and is commonly referred to as the fovea or foveal vision; see Figure 4.1.

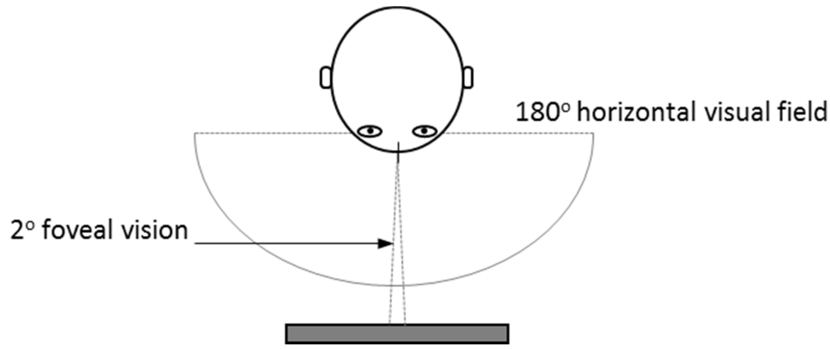


Figure 4.1. The human visual field

The resolution of the visual system changes gradually and systematically from this central fovea to the periphery, becoming increasingly blurry (Nuthmann, 2014). Therefore, in order to see things clearly, we need to move our eyes so that the object of focus is centred in the high-resolution foveal field of view. The inhomogeneity of the visual system is an important perception filtering mechanism; if everything were in focus all at once, our brains would be overloaded with information (Bojko, 2013).

As described above, eye movements are necessary for people to accurately perceive the outside world. Saccadic eye movements, consisting of saccades and fixations, are the most common and best understood type of eye movement (see, e.g. Holmqvist et al., 2011; Duchowski, 2007).

The repositioning of the fovea to a new location in the visual environment is referred to as a *saccade* (Duchowski, 2007). Both the velocity and the duration of a saccade are dependent on how much the eye moves during the saccade (Rayner, 1998). According to Duchowski (2007), a saccade typically lasts for 10 to 100 milliseconds. Each saccade requires approximately 30 milliseconds for a 5 degree visual angle and 2 milliseconds longer for each additional degree (Salvucci and Anderson, 2001). We do not obtain any new data during a saccade because the eyes are moving so quickly across the stable visual stimulus that only a blur would be perceived (Rayner, 1998; Holmqvist, 2011).

The stabilisation of the eye over a stationary object of interest is referred to as a *fixation* (Duchowski, 2007). However, fixations are not completely still. After a period of trying to keep still, the eye makes so-called microsaccades, that is, smaller and slower versions of saccades. The fixation is considered to end when the next saccades starts, while a micro-saccade or two do not put an end to a fixation (Holmqvist, et al., 2011).

An important argument for studying human eye movements is the assumption that our eye movements indicate something about what we are thinking about or paying attention to. A great deal of research has investigated this relationship between attention and eye movements (see, e.g. Rayner, 1998), and it is commonly accepted that where we place our gaze (i.e. direct our visual attention) is typically associated with what we pay attention to and think about, especially when concentrating and looking at something with a goal in mind (e.g. Anderson and Salvucci, 2001; Rayner, 1998; Hyrskikari 2006). This is called the eye–mind hypothesis (see e.g. Anderson, Bothell, & Douglass, 2004).

4.1.2 Eye tracking

Data about eye movements have long been used as a means to investigate and analyse (and better understand) human cognition, for example, the cognitive processes underlying human decision making and problem solving (e.g., Glaholt & Reingold, 2011; Payne, Braunstein, & Carroll, 1978; Rayner, 1998, 2009). Such eye movement data can be recorded using different kinds of eye-tracking techniques (discussed in more detailed below). As with verbal protocols, the collection of eye-movement data is a straightforward method of obtaining process data. However, eye tracking (especially in more recent years) has been associated with several challenges compared to other process-tracing methods, such as verbal protocols (see e.g., Payne, et al., 1978; Reisen, Hoffrage, & Mast, 2008). The eye-tracking equipment necessary to collect eye-movement data has traditionally been, and still is, very expensive. Particularly in the early years, the equipment was cumbersome and uncomfortable for the subjects. To obtain sufficient accuracy of resolution in recording eye fixations, the head of the subject might still need to be immobilised using a chin rest (a bit bar was frequently used earlier on) or something similar. Last but not least, one of today's most common technologies for eye-movement tracing is not suitable for subjects with certain characteristics, which I will describe in more detail later. Therefore, when running an eye-movement experiment, some subjects will provide unreliable eye-movement data. This limitation will be discussed in later sections.

Nevertheless, despite the fact that there are drawbacks with eye tracking as a process-tracing method, there is no doubt that eye-movement data are useful when studying human attentive behaviour. For instance, Russo (1978) compared several process-tracing methods and came to the conclusion that eye tracking has advantages not offered by other methods, and he further argued for the combined use of eye tracking and verbal protocols. Furthermore, the improvement of eye-tracking hardware has been tremendous within the last two decades, allowing for the unobtrusive measuring of humans' eye movements that typically does not have significant effects on observed behaviour (see e.g. Stüttgen, Boatwright, & Monroe, 2012). Eye-movement studies with both tracking and non-tracking conditions for the same task have shown no significant differences in performance between conditions (Salvucci & Anderson, 2001).

Eye-tracking techniques

Several techniques for recording eye movements exist. For instance, electro-oculography (EOG) tracks eye movements by measuring the skin's electric potential differences using electrodes placed around the eyes (Duchowski, 2007). Another eye-tracking method is scleral contact lens/search coil, which involves measuring eye movements by attaching a mechanical or optical reference object mounted on a contact lens that is then worn directly on the eye. The scleral contact lens/search coil method is one of the most precise ways of measuring eye movements but is also the most intrusive method as wearing the lens causes discomfort for the subject (Duchowski, 2007). Both the EOG method and the scleral contact lens/search coil method measure eye position relative to head position, which means that these methods are generally not suitable for measuring point of regard (i.e. gazepoint). To measure the point of regard, either the head must be fixed so that the eye's position relative to the head and point of regard coincide or the position of the head must also be measured (i.e. using a head tracker).

Today, a growing body of research involving eye movement measures is being conducted using the combined pupil and corneal reflex method (Duchowski, 2007; Holmqvist, et al., 2011) in video-based eye-tracking systems. The principle behind this method is as follows. When the eye is hit by a light source, the eye's cornea, which is transparent and covers the eye, reflects the light as a corneal reflex (also known as the Purkinje reflex or Purkinje images; (Duchowski, 2007)). Furthermore, the relative position of the corneal reflex and the pupil changes as the eye moves; see Figure 4.2.

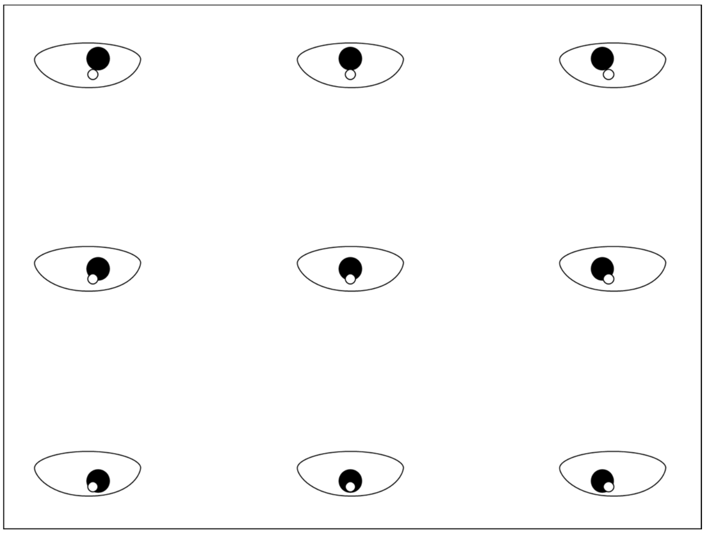


Figure 4.2. Relative position of the corneal reflexes and pupils as seen by the eye tracker (based on Duchowski, 2007, p. 58).

The eye tracker consists of an infrared (IR) camera and an IR illumination source placed in front of the subject. The IR light is reflected in the subject's cornea. IR light is used to avoid interference from other illumination sources. The IR camera records a video of the subject's eyes and then uses advanced image-processing techniques to identify the centre points of both the pupil and the corneal reflex based on the video. The relative location of the pupil and the corneal reflex are determined and then used to calculate the gaze position. A calibration process is usually required to 'link' the calculated gaze position to locations in the actual stimuli. Due to individual differences in the size and shape of the human eye and variations regarding the use of eyewear such as glasses and contact lenses, the calibration process needs to be conducted for each participant. The calibration process provides the eye tracker with some examples of calibration points in the stimulus and their corresponding pupil and corneal reflexions (Holmqvist, et al., 2011).

Several types of video-based eye trackers exist, but for the most part they fall into two main categories depending on the location of the eye-tracking device during the tracking process, which is either remote or wearable. Remote eye trackers are placed in a fixed position in front of the subject and are thus 'contact-free'. Wearable eye trackers are worn on the subject's head (e.g. on a helmet or built into glasses). A remote eye tracker is regarded as less obtrusive for subjects than wearable eye trackers. However, the downside of a remote system is that the subjects need to be positioned in front of the eye tracker, and thus only limited movements of the head are acceptable. According to Stüttgen, Boatwright, and Monroe (2012), recent improvements in software and hardware now allow subjects to move freely in an area of at least

25x25x25 inches (i.e. approximately 60x60x60 cm). This area is often referred to as the head box (e.g. Bojko, 2013, p. 60).

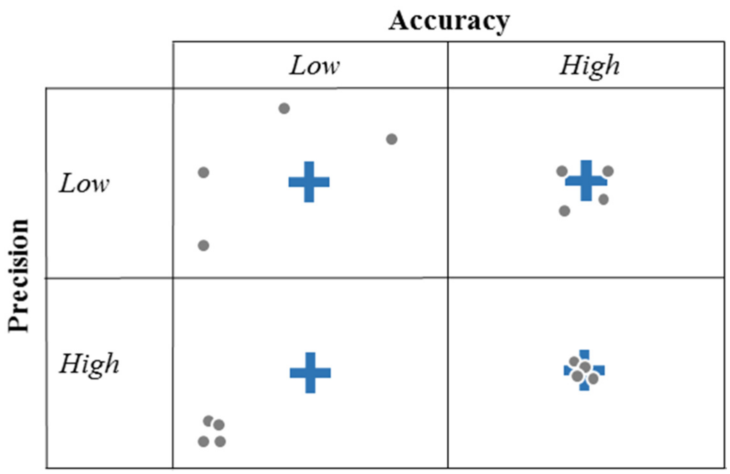
In the current study, I use a remote eye tracker. In the following, I focus on the use of remote eye trackers.

Eye-movement data

Eye trackers records subjects' gaze positions as time-stamped XY coordinates. In addition to gaze position, they can also measure the characteristics of the eye itself, such as the size of the pupil (i.e. pupil diameter). The sampling rate of eye trackers, measured in hertz (Hz), tells the number of times per second that they make recordings. This means that a 50Hz eye tracker, like the one being used in this study, records 50 data points every second for each tracked eye. Therefore, even short eye-tracking sessions of 10 minutes generate very large sets of eye-movement data.

The quality of the data produced by an eye-tracker is dependent on properties related to both the eye-tracker (e.g. it's sampling rate and the resolution of the eye camera) and the subjects (e.g. the use of eyewear or mascara). Accuracy and precision are two properties of an eye tracker that influence the data quality. Bojko (2013) uses the following table (Table 4.1) to illustrate the distinction between eye-tracking accuracy and precision:

Table 4.1. The relationship between eye-tracking accuracy and precision.



The accuracy of the eye tracker is the average difference between what the eye tracker recorded as the gaze position and what the gaze position actually was. Precision is a measure of how well the eye tracker is able to reliably reproduce a measurement, even when the position of the eye changes between the two measures.

Holmqvist et al. (2011) describe several subject-specific characteristics that may influence the data quality. The use of mascara (i.e. dark makeup worn on the eyelashes) is a problem because the eye-tracker software that identifies the dark pupil (and relates it to the corneal reflex) is confused by the other large dark areas in the immediate vicinity of the pupil caused by the mascara. Therefore, subjects should always be advised to remove mascara before attending an eye-tracking session. Droopy eyelids and downward eyelashes may also influence the data quality because the eyelids or the eyelashes cover the pupil in the lower gaze directions (see Figure 4.2). This may lead to very large downward offsets in the data or even complete data loss (i.e. glitches) for the lower part of the visual field. Glasses are problematic in many ways, for instance because the light from the corneal reflection can be reflected back to the eye and give a second corneal reflection. Similarly, the use of contact lenses may lead to multiple corneal reflections due to small air bubbles that may occur between the lens and the eye. Wet eyes (e.g. due to tears or allergic reactions) may also cause multiple corneal reflections. Multiple corneal reflections are problematic because the relative location of the pupil and the corneal reflex cannot be determined and thus the gaze position cannot be calculated.

4.1.3 Analysing eye-movement data

As already mentioned, recordings from an eye tracker typically generate very large sets of eye-movement data, and in order to consistently analyse these data, the data may need to be coded into some more manageable form.

The majority of previous eye-tracking studies were carried out by showing subjects sequences of still images, with or without sound, and having the subjects sit more or less still in front of an eye-tracker as they observe the stimuli (Holmqvist et al., 2011). Furthermore, according to Holmqvist et al. (2011), almost all software used to analyse eye movements is written for this particular set-up. Eye-tracking data from the set-up described above consists of time-stamped XY coordinates that relate to a coordinate system covering the stimuli. It follows that, when using the same fixed stimuli for all subjects, any data samples (across subjects) with a given coordinate (e.g. x=100, y=139) relate to the exact same element/location in the stimuli. Thus, comparing subjects' attention to different elements in the stimuli can easily be done (e.g.

analysing how much time subjects have spent looking at a given object or determining how many times they have looked at a certain area of the stimuli).

The above description is an example of fixation-based eye-movement analysis that investigates events such as fixations and saccades, which is the case for the majority of eye-movement analysis being conducted. In order to conduct such analyses, the raw data (i.e. the time-stamped XY coordinates) needs to be transformed into fixations using a fixation identification algorithm. One such algorithm, the dispersion-based fixation detection algorithm, determines fixations based on two thresholds – minimum fixation duration and maximum dispersion. For example, if the minimum fixation duration is set to 100 milliseconds and maximum dispersion is set to 100 px, a fixation is recognised whenever raw data samples stay within 100 px for at least 100 milliseconds.

When eye-movement events have been defined, the data can be further analysed, for example by using areas of interest or by visualising the events. An area of interest (AOI) is a user-defined area in a stimulus. Determining an AOI allows a researcher to define what regions in the stimulus he/she might be interested in gathering data about (Holmqvist et al., 2011). For example, much of the previous research on user interfaces and usability has been based on dividing the interface (i.e. the stimulus) into predefined AOIs and then measuring transitions into and from these areas, as well as the time spent in each area (Goldberg & Kotval, 1999). Transitions between different AOIs and total AOI dwell time are just two examples of the vast amount of AOI measures that can be computed. Other examples are time to first fixation on an AOI, the percentage of subjects who fixated on an AOI, the number of fixations on an AOI, etc. For an overview of (some) AOI measures, see Holmqvist et al. (2011). Typically, AOI measures aggregate eye-movement data for the AOI being investigated. When using static stimuli (e.g. a static picture), the AOI measures can easily be aggregated across subjects. Therefore, AOIs are useful in conducting quantitative analyses.

The visualisation of fixation-based eye movements may prove useful both when conducting qualitative analyses of the data and when communicating quantitative and qualitative findings (Holmqvist, 2011; Bojko, 2013). Bojko (2013) classifies eye-movement visualisations by three dimensions – *the amount of data shown* (i.e., individual or aggregated data), *the format of the visualisation* (i.e. static or dynamic) and the *type of information shown*. Some visualisations convey only spatial information (the location of gaze activity on the stimulus), while others convey both spatial and temporal information (not only *where* but also *when* gaze activity occurred). Bojko uses the following table to compare different eye-movement visualisations.

Table 4.2. Comparison of eye-movement visualisations.

Visualisation type	*Amount of Data Shown*	*Format of Visualisation*	*Type of Information Shown*
Gaze plot/Scanpath	Individual	Static	Spatial and temporal
Gaze Video	Individual	Dynamic	Spatial and temporal
Bee Swarm	Individual and aggregate	Dynamic	Spatial and temporal
Heat map	Aggregate	Static	Spatial only
Focus Map/Gaze Opacity Map	Aggregate	Static	Spatial only
Dynamic Heat map	Aggregate	Dynamic	Spatial only

In the current study, gaze plots/scanpaths are used to illustrate eye-movement data. A type of gaze video was used when conducting the eye-movement analyses. Therefore, I limit my description of eye-movement visualisations to these two visualisation types. Please see, for example, Bojko (2013) and Holmqvist et al. (2011) for descriptions of the other visualisation types in Table 4.2.

A gaze plot or scanpath is an image showing an individual's fixations represented as dots and their saccades represented as lines. The size of a dot is proportional to the duration of the fixation, with larger dots representing longer fixations (Bojko, 2013). Figure 4.3 shows an example of a scanpath from the current study.

As has been done in Figure 4.3, sequence numbering can be printed on the dots to show the temporal order of the fixations in the scanpath. However, including sequence numbering tends to reduce the readability of the scanpath. Therefore, I have mostly not included sequence numbering in the scanpaths included in the dissertation.

The number and size of fixations in a scanpath is dependent on the settings used for the fixation identification algorithm. Thus, when comparing gaze patterns for a stimulus across several subjects, the parameters for the fixation identification algorithm should be equal for all subjects.

In the following, I use the terms scanpath, gaze plot and scanpath diagram interchangeably. A scanpath shows the background as a static picture, and is thus useful when studying static stimuli, such as pictures and text.

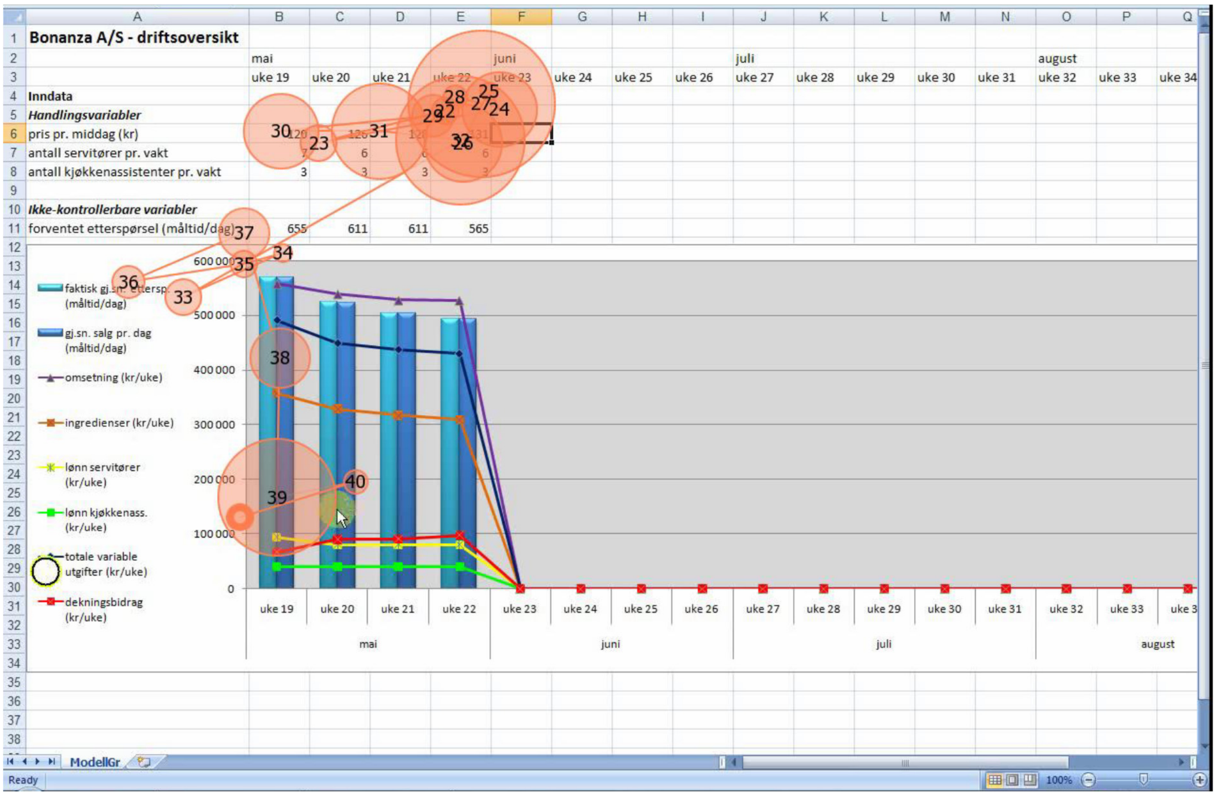


Figure 4.3. An example of a scanpath.

If the stimuli are dynamic in nature, other visualisations, such as a gaze video, might be more suitable. For example, when subjects interact with a spreadsheet user interface like the one being used in Study 1 and Study 2, the stimuli constantly changes during the session. Content is added/removed in the spreadsheet, the mouse cursor is moving, menu elements and dialog boxes appear/disappear and the user may scroll up/down/back/forth in the spreadsheet and switch between different spreadsheet/programs. Thus, by recording a screen video of the subject's interaction with the spreadsheet, a dynamic stimulus is 'created' during the eye-tracking session. A gaze video is a dynamic visualisation of an individual's point of gaze represented by a moving dot or a cross and overlaid on top of the stimulus, such as the above screen video (Bojko, 2013). However, when watching a gaze video, one must realise that the eye moves very fast and that it is hard to analyse data by watching the video played at full speed. For that reason, a visualisation type that combines gaze videos with scanpaths may be used. This implies projecting 'moving scanpaths' onto the gaze video. Each fixation and saccade is shown for a limited amount of time, for example five seconds, before they fade away as new fixations and saccades appear. Real-time gaze observations, like the one described above, may be especially useful when subjects have trouble verbalising their thoughts or when tasks are supposed to be performed in silence (Bojko, 2013).

4.1.4 Eye-tracking protocols and verbal protocols

Both verbal protocols and eye-movement protocols have been used to study cognitive processing. However, according to Lohse and Johnson (1996), verbal protocols provide a more direct measure of information acquisition and cognitive processes than do eye-movement protocols. It is widely accepted that attention and eye movements are closely connected (see, e.g. Glaholt & Reingold, 2011; Rayner, 1998). Data about eye movements can reveal information acquisition behaviour but not necessarily that this information has been processed in a certain manner, as opposed to verbal protocols (Payne, Braunstein, & Carroll, (1978). Therefore, the researcher must infer underlying cognitive processes and decision strategies from information acquisition data, a process that is not necessarily straightforward. For example, eye tracking does not account for the elements that people never looked at with their foveal vision but noticed in the part of the visual field just outside the foveal vision. Furthermore, eye-tracking data may show that a person looked at certain elements in a stimulus but not *why* they looked at these elements or whether they *understood* what they were looking at. Thus, there are several questions that eye-tracking data alone may not answer. However, eye-tracking data may be used to augment other process-tracing methods, such as verbal protocols. By filling in the gaps between observable events and think-aloud statements, eye tracking may improve the interpretation of why people say or do certain things. In the same way, verbal protocol data and other process-tracing data may be necessary to interpret and qualify eye-tracking findings, as illustrated above.

4.2 Research design

As the current study is a follow-up study based on Study 1, I will use the same research design as in the first study. Thus, the research model, the experimental treatment groups and the experimental setting are identical to the ones presented in section 3.2. However, based on the above discussion, I will add more control variables to the current study. These include the time used to solve the task and how this time was allocated, motivation, age, gender and the number of weeks (of the total 17 weeks in the experiment) each decision maker needed in order to solve the task.

The subjects in this study were 37 Master’s students from the Norwegian School of Economics and thus should possess the same qualifications as the subjects participating in Study 1.

4.2.1 Measurement

The decision result, that is, total contribution, was measured in the same way as in Study 1; see section 3.2.3. More control variables have been included, such as age, gender, time and need for cognition (NFC).

As discussed in section 3.4.2, I noticed that there seemed to be large variations in terms of how much time each subject needed to solve the experimental task. In addition, I noticed clear differences in terms of how the subjects allocated their time into two phases when solving the task. The first phase is related to Simon's (1960) intelligence phase, that is, to what extent the subjects tried to understand the task and infer relationships among the variables so that they could find an optimal solution before they entered values into the decision model. The second phase is related to Simon's design and choice phases.

In order to distinguish between these two phases, I have defined Problem definition as the phase from the subjects had finished reading the task description until they had entered data for the first decision (i.e. week 23) into the spreadsheet model. Problem solving is defined as the phase from the subjects had entered their first decision into the spreadsheet until they had entered the last decision (i.e. week 39). Problem attention includes both the problem definition and the problem-solving phases.

Level of information processing is measured in the same way as in Study 1, that is, based on the four levels of information processing presented in the theory of cognitive complexity (Schroder, et al., 1967); see section 3.1.2. However, level of information processing is measured separately for the problem-definition phase and the problem-solving phase.

The subjects' motivation for working on the experimental task was measured using NFC (Cacioppo & Petty, 1982; Cacioppo, Petty, & Chuan Feng, 1984). The NFC is a relatively stable personality trait that reflects an individual's tendency to engage in and enjoy effortful cognitive activity (Cacioppo, Petty, Feinstein, & Jarvis, 1996). NFC is usually perceived not as the ability to think but as an intrinsic motivation to think and it therefore correlates well with various measures of intrinsic motivation (Furnham & Thorne, 2013). NFC is measured by asking subjects to describe the extent to which they agree with a set of statements using a nine-point scale (-4=strongly disagree, 0=neither agree nor disagree, +4=strongly agree). The original list of statements consists of 34 items (Cacioppo & Petty, 1982), but a revised and shortened list of only 18 items is proposed in Cacioppo, Petty, and Kao (1984). The shortened list has proved to measure NFC just as well as the original list (Cacioppo, et al., 1996). Therefore, in the current study, I have used the revised list of 18 statements. The statements have been translated to Norwegian; see Appendix G. Nine of the statements are reverse-scored. Based on the subjects' scoring on the 18 statements, the NFC index was calculated. The index is a bipolar index, ranging from low to high (Cacioppo, et al., 1996).

4.3 Data collection

4.3.1 Procedure

As in Study 1, individual appointments were scheduled for each student participating in the experiment. Because Study 2 included recordings of the students' eye movements, the students were told not to wear mascara during the experimental setting.

I was present in the room during the whole experiment, and the subjects were able to ask questions if they needed help or were uncertain of anything. During Study 1, I noticed that the subjects verbalised more freely if I did not sit right next to them but rather sat a few meters away. The feeling of being observed was obviously less intrusive. Thus, in this second study the experimental room was furnished so that I (the observer) sat behind a desk at one end of the room while the subjects sat at a desk at the other end of the room; see Figure 4.4.

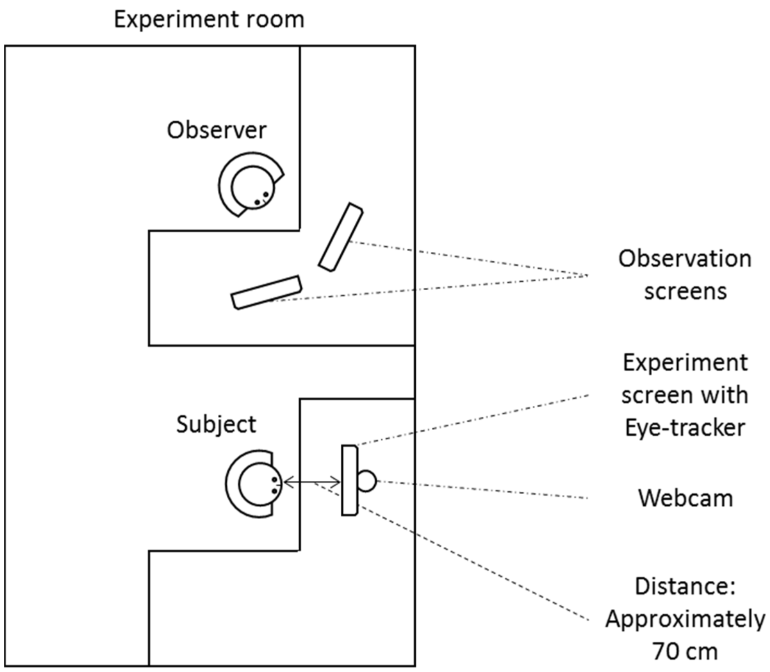


Figure 4.4. Setup for the eye-tracking experiments.

As shown in the figure, the two desks were arranged at 90 degree angles. The setup shown in Figure 4.4 will be explained in detail below.

An observation form (see Appendix H) was developed for use in the experiments. On the form, I took notes regarding the subjects' time consumption, both for the problem-definition phase and the problem-solving phase. Notes were also taken regarding the subjects' general approach to the task using the following categories: random trial and error, trial and error based on experience, optimisation of the capacity constraints and other approaches. The observation form included my assessments of the subjects' general economic knowledge for handling the task, their ability to master spreadsheet software, to what extent the subjects indicated a need for additional presentation formats and to what extent they made additional graphs or additional tables, classifications of any additional calculations being made (simple, comprehensive, marginal) and their general progression during the experiment. Finally, notes were taken regarding the subjects' use of contact lenses, glasses and mascara during the experiment.

I followed the procedure for data collection described in Study 1. Thus, the experimental task was introduced and explained as in Study 1. After I introduced and explained the task as in Study 1, a nine-point calibration process was conducted to ensure that the eye tracker produced accurate eye-movement recordings. The calibration process was conducted after the subjects had read the task description but before they started to study the spreadsheet model. Because none of the subjects left their chair during the experiment, the calibration process was not repeated for any of the subjects.

The criterion for a successful calibration was set to matches that were less than 0.5 cm (0.5^{o} visual angle) from the actual target location.

The subjects were told that the eye tracker would monitor their eyes during the experiment. Humans cannot control the size of their pupils while looking at something but being aware that their eye movements are monitored may influence on their gaze pattern. Therefore, I told the subjects before the experiment that the eye tracker would record the size of their pupils and that I was curious about whether the size of their pupils was affected by the presentation format being used. After the subjects had finished the experiment, however, I told them that I had also recorded their eye movements and explained the reason for doing so. I also asked for permission to keep and analyse the eye-movement data. All the subjects agreed.

When the subjects completed the experiment, they were asked to fill out a questionnaire (see Appendix G). In the questionnaire, the subjects were asked to specify control variables, such as age, gender, eyewear (glasses, contact lenses, none) typically worn and worn during the experiment, electives they had taken, grades for the introductory course in data processing and questions regarding spreadsheet skills and skill level in working with graphs. The questionnaire also included an 18-item questionnaire designed to measure the subjects' NFC.

4.3.2 Methods of data collection

The methods of data collection were tape-recordings of the subjects' thinking aloud during the experiments, recordings of the subjects' eye movements, recordings of the computer screen and video recordings of the subjects themselves.

Think-aloud recordings

The think-aloud procedure described in Study 1 was followed. In addition, after having completed the experiment, subjects were asked for a short cued retrospective report about their overall decision strategy. The subjects were instructed to look back at the decisions they made in the spreadsheet model and at the same time explain how they had reasoned during the process. According to Hoppman (2009), the reasons for decisions being made is often only explained in such retrospective follow-up questions. I do not rule out the possibility that such retrospective verbalisation may suffer from fabrication or forgetting, but I consider the retrospective reports to be a useful supplement to the concurrent verbal reports. A portable voice recorder was used to record the subjects' thinking aloud.

Eye movement recordings

Moreover, an eye tracker was applied to record the subjects' eye movements whenever they looked at the computer screen. The eye movements were registered as time-stamped XY coordinates relative to the image on the computer screen.

I used an iViewX^{TM} RED eye tracker manufactured by SensoMotoric Instruments (SMI; Teltow, Germany) in the experiment. The eye tracker recorded subjects' eye positions at a frequency of 50 Hz (i.e. a sampling rate of 50 samples/second), a tracking resolution of 0.03^{o} and gaze position accuracy of 0.4^{o}. Only the left eye was recorded and the gaze position was determined using the pupil and corneal reflection method (discussed in section 4.1.2). The eye tracker was attached to a 19-inch thin-film-transistor liquid-crystal display (TFT-LCD) screen used for stimulus presentation, and the screen resolution was 1280x1024. The distance between the participant's eyes and the screen was about 70 cm. The recommended operating distance for the iViewX^{TM} RED eye tracker is 60–80 cm. At this distance, the head-tracking range (i.e. the range in which the head can be moved freely while at the same time achieve accurate recordings) is 40 cm (sideways movements) x20 cm (up/down). The illumination of the screen was kept constant, and neither room lighting nor direct sunlight interfered with the recording capabilities of the eye tracker.

During the experiments, I monitored the ability of the eye tracker to record eye movements on an observer screen. If the eye tracker lost sight of a subject's eyes, for example due to incorrect positioning (distance/angle) of the subject relative to the eye tracker, I asked for a correction in order to obtain the eye-movement recording. It was only rarely necessary to ask for a such correction.

Recordings of the computer screen (stimulus)

The SMI Experiment Center 2 software was used to record a gaze video of the subjects. The video shows the subjects' gaze positions as a small moving dot projected onto the screen image during the experiment. Thus, every change being made in the spreadsheet was also recorded.

Video and audio recordings of the subjects

A webcam was used to record the subjects while solving the experimental task. My experience interpreting the verbal reports from Study 1 shows that it was sometimes hard to establish what the subjects had been doing during different parts of the experiment. For example, did they look at the spreadsheet model on the computer screen, read the task description, make calculations on the calculator or write notes using pen and paper? Even though I had observed the subjects in Study 1, the observation notes were sometimes not good enough to infer subjects' behaviour. Webcam recordings are included in this study so that establishing subjects' behaviour is easier.

A webcam with a microphone was used to record a video of the subjects and their thinking aloud during the experiment. By including sound recordings in the videos, I was able to synchronise the webcam videos with the recordings of subjects' think-aloud statements.

4.4 Presentation and explanation of the data matrix

The data set for Study 2 consists of 37 subjects. However, as was also the case in Study 1, some subjects performed considerably worse than the rest of their group members (subjects S2-12 (index=0.6483), S2-24 (index=0.5893) and S2-26 (index=0.8384)). An inspection of the verbal protocols for these three subjects shows that they had maximised sales instead of contribution. These misunderstandings are not related to the presentation format, and I have therefore removed them from the data set. Appendix F.2 shows the data set including all 37 subjects.

Table 4.3 shows my initial data analysis without subjects S2-12, S2-24 and S2-26.

Table 4.3. Results

No.	Task type	Pres. Format	Electives	Computation	Graph	Analytic	Spatial	Level of info. proc. p.d.	Level of info. proc. p.s.	Index	Group average 1	Group average 2	P.D. min.	Group average P.D. min.	P.S. min.	Group average P.S. min.	P.A. min.	Group average P.A. min.	NFC
S2_04	L	b	d	2	2	1	1	6	6	0,9696			49,00		4,25		53,25		0,8889
S2_32	L	b	d	1		1		3	5	0,9828			5,83		19,85		25,68		0,0000
S2_34	L	b		2		1		6	6	0,9982			26,78		3,92		30,70		0,1667
S2_35	L	b	d	3	3	1	1	6	6	0,9932			12,17		3,47		15,63		-0,2778
S2_42	L	b		2		1		4	4	0,9455	0,9779		10,37	20,83	50,03	16,30	60,40	37,13	-0,0556
S2_06	L	g					1	2	3	0,9381			1,62		58,48		60,10		-0,2778
S2_31	L	g		1	2	1		4	5	0,9945			33,35		12,15		45,50		-0,1667
S2_36	L	g	d	1		1		5	6	0,9755			24,98		21,63		46,62		-0,1667
S2_37	L	g		3		1		5	5	0,9980			17,68		3,32		21,00		-0,1111
S2_38	L	g		1		1		4	4	0,9922			35,57		7,95		43,52		-0,3333
S2_40	L	g	d	2		1		4	4	0,9419	0,9734		10,10	20,55	8,02	18,59	18,12	39,14	0,0000
S2_30	L	t		1		1		3	3	0,9640			24,18		16,17		40,35		0,1111
S2_33	L	t		2		1		3	4	0,9933			5,50		3,08		8,58		0,2778
S2_39	L	t		3		1		4	4	0,9982			15,08		2,52		17,60		0,1667
S2_41	L	t	d	1		1		3	4	0,9912	0,9867	0,9784	6,20	12,74	17,62	9,85	23,82	22,59	0,2222
S2_02	H	b	d	2		1	1	3	6	0,9017			4,17		46,25		50,42		0,8889
S2_05	H	b	d	3		1	1	5	5	0,9815			10,92		4,00		14,92		-0,3889
S2_07	H	b		1		1	1	1	3	0,8865			2,97		14,18		17,15		-0,1111
S2_10	H	b					1	2	4	0,8635			3,05		14,70		17,75		0,1667
S2_18	H	b					1	1	3	0,6828			0,67		18,33		19,00		0,2222
S2_19	H	b		2		1		5	5	0,8720			20,50		13,00		33,50		-0,1667
S2_23	H	b					1	1	3	0,8216	0,8585		2,35	6,37	19,50	18,57	21,85	24,94	0,5000
S2_03	H	g		1		1	1	3	4	0,8445			26,00		63,35		89,35		-0,0556
S2_13	H	g	d	2		1		4	5	0,8695			10,20		52,02		62,22		0,5556
S2_15	H	g					1	2	3	0,6721			4,93		3,18		8,12		-0,2222
S2_16	H	g		2		1		4	5	0,8894			6,35		24,00		30,35		0,5000
S2_22	H	g		2	2	1		6	5	0,8320	0,8215		6,05	10,71	58,05	40,12	64,10	50,83	0,2778
S2_01	H	t	d				1	1	3	0,8406			1,83		23,33		25,17		-0,8333
S2_08	H	t	d				1	1	2	0,8387			0,72		17,25		17,97		0,5000
S2_09	H	t		1		1		3	4	0,8909			21,47		3,75		25,22		0,0556
S2_14	H	t		2		1		3	5	0,9766			3,13		12,27		15,40		-0,3889
S2_17	H	t		1		1	1	2	3	0,8063			8,92		17,40		26,32		0,2222
S2_20	H	t		1	2	1	1	3	4	0,8081			30,95		28,38		59,33		0,2222
S2_25	H	t		2	3	1	1	4	5	0,9145	0,8679	0,8522	42,70	15,67	23,88	18,04	66,58	33,71	0,2222
	H = high L= low	b = both t = table g = graph	d = data	1 = simple 2 = compr. 3 = marg.	1 = line 2 = XY 3 = XY, compr.	1= true	1= true	1=low 7=high	1=low 7=high										

In the following, I will explain the table, including how I coded the data in the columns based on my analysis of the verbal protocols. Some of the columns are identical to the ones in Table 3.2 in Study 1 and are therefore not explained again. These columns are: No, Task type, Pres. Format, Electives, Computation, Graph, Analytic, Spatial Index, Group average 1 and Group average 2. For an explanation of these columns, see section 3.3.1.

- ***Level of info. proc. p.d*** – Level of information processing, problem definition – Denotes the score on the level of information processing index during the problem-definition phase. The measure is scored as in Study 1 (see section 3.3.1). An ordinal scale was used for scoring the variable, applying 1 as the lowest and 7 as the highest value.
- ***Level of info. proc. p.s*** – Level of information processing, problem solving – Denotes the score on the level of information processing index for the problem-solving phase. An ordinal scale was used, applying 1 as the lowest value and 7 as the highest value.
- ***P.D. Min.*** **–** Problem-definition minutes – denotes a measure of the time used in the problem-definition phase.
- ***Group average P.D. min*** – Group average problem-definition minutes – denotes the average problem-definition time for each treatment group.
- ***P.S. Min.*** – Problem-solving minutes – denotes a measure of the time used in the problem-solving phase.
- ***Group average P.S. min.*** – Group average problem-solving minutes – denotes the average problem solving time for each treatment group.
- ***P.A. min.*** – Problem-attention minutes – denotes the amount of time the subjects spent working on the task, that is, problem-definition minutes + problem-solving minutes.
- ***Group average P.A. min.*** – Group average problem-attention minutes – denotes the average problem-attention time for each treatment group.
- ***NFC*** – Denotes a measure of the subjects' need for cognition (i.e. motivation for solving the kind of task being used in the experiment). See section 4.2.1.

As in Study 1, the results in Table 4.3 show a clear difference in the decision result (contribution index) between the low-complexity task and the high-complexity task. The average value of the contribution index for the low-complexity task is 0.9786, and the average value of the index for the high-complexity task is 0.8522 ($p < 0,000$).

Table 4.4 summarises the average values of the contribution index for the six treatment groups.

Table 4.4. Average values of the contribution index for different presentation formats in high- and low-complexity tasks.

Task complexity	Presentation format: *Graph*	*Table*	*Table and Graph*
Low	**0.9737** N=6, st. dev.= 0.027	**0.9867** N=4, st. dev.= 0.015	**0.9779** N=5, st. dev.= 0.021
High	**0.8215** N=5, st. dev.= 0.086	**0.8679** N=7, st. dev.= 0.063	**0.8585** N=7, st. dev.= 0.092

As opposed to the results from Study 1 (Table 3.3) indicating that subjects with the combined format performed better than subjects in the other categories, Table 4.4 gives the impression that the subjects with tables only performed better than the subjects with the other two formats in both the low-complexity and high-complexity tasks. This result supports one of the conclusions from Study 1, namely that raw data (i.e. presented in tables) are necessary in order to make the detailed and accurate calculations necessary to achieve a good result.

The findings from Study 1 indicate that the graph format seemed to be least effective in supporting the subjects' decision making. The same tendency is found in the current study, thus supporting the findings from Study 1.

In addition to the variables studied in Study 1, the current study also includes control variables for gender, age and motivation (measured as NFC). A one-tailed bivariate correlation analysis (Pearson correlation) was generated to check for effects of the control variables; however, no significant correlations were found. See Table 4.5.

Table 4.5. Correlation matrix of control variables.

	Index	*NFC*	*Age*	*Gender*
Index	1			
NFC	-.131	1		
Age	.055	.031	1	
Gender	-.127	.022	.184	1

Note. n=34.

4.5 Data analysis of high and low performers

4.5.1 Introduction

In this section, I will present my analyses of the high and low performers for each presentation format, see also Andrienko, Andrienko, Burch and Weiskopf (2012). In order to evaluate how the two data collection methods, that is, verbal protocols and eye movements, may complement each other, I will select subjects for whom I have satisfactory eye-tracking data. As described in section 4.1.2, some eye-tracking data sets may be useless for analysis due to subjects' wearing of eyewear or mascara, etc.

The section is organised in the following way. I discuss subjects that are not included in the analysis due to dissatisfactory eye-movement data. For each experimental group, I will then present one high performer and one low performer. I will distinguish between the subjects' problem-definition[3] and problem-solving phases. For each phase, I will analyse the subjects' information processing, paying particular attention to how the verbal protocols and the eye-movement data complement my interpretation of the data.

The data set

Eye-movement data were recorded for the 37 subjects who participated in Study 2. Three subjects were removed as outliers (ref. section 4.4). Seven of the remaining 34 subjects showed to be useless for further eye-movement analyses. Three female subjects (S2-16, S2-30 and S2-32) wore mascara during the experiment even though they were asked not to when they were recruited. The subjects were asked to remove the mascara but did not comply with this request.

S2-31 got very poor results on the calibration test, even though she did not wear mascara or eyewear. Subject S2-34 wore glasses, and the IR light from the eye-tracker reflected in the glasses, making the eye-tracker unable to track the subject's eye movements. One subject (S2-38) had a visual impairment and squinted strongly. He did not perform well on the calibration test. Subject S2-40, who had long, dark eyelashes, also did not perform well. Even though I tried to adjust the eye-tracker camera angle and the subject's seating position, the eye lashes ended up covering much of the eye, making it impossible for the eye tracker to determine

[3] Problem definition has been defined as 'the time from when the subjects read the task description until they entered data for the first decision (i.e. week 23) into the spreadsheet model'. The time spent on calibration of the eye tracker that took place after the subjects read the task description is not included in the problem-definition time.

where the subject looked. All seven subjects discussed above were excluded for further eye-movement analysis.

I was unable to analyse eye-movement data for subject S2-17. The analysis software crashed every time I tried to analyse the data, no matter what I did to fix the problem. The data was loaded into the analysis software multiple times, in different manners and with different parameter settings, but nothing solved the problem.

4.5.2 The eye-movement dataset

The eye-movement data has been analysed using SMI BeGaze 3.4.52, an eye movement analysis software developed by the manufacturer of the SMI RED eye-tracker that was used in the experiment. BeGaze uses a dispersion-based fixation identification algorithm to combine raw data samples (i.e. the time stamped XY coordinates) into fixations. For the current study, minimum fixation duration was set to 100 milliseconds and maximum dispersion was set to 100 px (i.e. a fixation is recognised whenever raw data samples stay within 100 px for at least 100 milliseconds).

BeGaze was used to produce scanpath videos for each subject. These scanpath videos show animated projections of a subject's eye movement events (i.e. fixations and saccades) vis-à-vis the stimulus (i.e. the subject's gaze video). Each fixation and saccade was shown for 10 seconds before they were faded away. This means, for example, that a fixation starts as a tiny dot that then increases in size as the fixation continues to last in time. When the fixation ceases, it still appears as the dot in its final size for 10 more seconds before it is faded away. Thus, the term 'moving scanpaths' can be used to describe the content in the scanpath videos.

Optimally, the scanpath videos should have been synchronised and merged with the corresponding think-aloud recordings. However, technical problems with BeGaze made such a merge impossible. Instead, the scanpath videos have been synchronised with the transcribed think-aloud protocols and in some cases manually replayed along with the corresponding sound recordings.

The experimental task implies some expected problem-solving steps. For example, I expect that the subjects establish the capacities per waiter in both tasks and the capacity per kitchen assistant in the high-complexity task. This can be done by comparing the values of the variables in the spreadsheet model (i.e. by comparing actual sales and expected demand to the number of employees at work). Thus, a comparison of the subjects' scanpaths with the expected viewing behaviour when following expected problem-solving steps was used to establish the relative effectiveness of each presentation format.

4.5.3 Low-complexity task, both presentation formats

S2-35 – High performer

Table 4.6. Result and control variables for S2-35.

	S2-35	*Group average*
Contribution index value	0.9932	0.9779
Problem attention	15.63	37.13
Problem definition	12.17	20.83
Problem solving	3.47	16.30
Need for cognition	-0,2778	0.0768[4]

I do not have satisfactory eye-tracking data for the best performing subject in this group, that is. S2-34, as his wearing of glasses caused data loss. Subject S2-35, however, is the second best performing subject for the low-complexity task with a combined presentation format. He scored 0.9932 on the contribution index and took 15.63 minutes to complete the task. The group as a whole scored 0.9779 on average and took 37.13 minutes to solve the task. S2-35 spent 12.17 minutes on problem definition and 3.47 minutes on problem solving. The group average was 20.83 min. and 16.30 minutes, respectively.

Problem definition

Verbal protocol analysis

The verbal protocol shows that S2-35 had many silent periods and that I constantly reminded him to think aloud. The protocol further reveals that he created a mathematical model to predict the contribution at different prices before making his first decision, for example:

> 'I'm creating an overview of the contribution, first per day and then per week, based on different prices, demands and variable costs. I know how many guests each waiter can

[4] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

handle, and therefore I can calculate the number of waiters needed per week for the various alternatives.'

S2-35 used the historic data of the average prices and demands for the last three years to find the demand for each price alternative in his model. He received the data both as a table and as graphs, but the verbal protocol shows that he used the table format. Furthermore, S2-35 created a graph to visualise the contribution at the different price alternatives in his model. Figure 4.6 shows S2-35's calculations and graph. He then concluded:

'The optimal solution is to sell at a price of NOK 148 with three waiters. I just enter those values into the model.'

Eye-movement analysis

The eye-tracking data shows that S2-35 had horizontal eye movements, indicating that he was relating the variable labels to their corresponding values. Furthermore, in line with the verbal protocol analysis, the eye-tracking data show that S2-35 had vertical eye movements between the price, the number of waiters and the expected demand in the table format. However, he only spent about a minute studying the spreadsheet model before he started creating his own model in a new spreadsheet, where he spent most of his time.

Developing the spreadsheet model in Figure 4.6, the verbal protocol did not indicate how S2-35 calculated the waiter capacity, but the eye-tracking data show that he found the waiter capacity based on the average demand instead of the actual sales; see Figure 4.5 a) and b). Because of this error, his calculations of the contribution at the various sales prices were not correct.

S2-35 focused mostly on the numeric data in the table of the presentation format, but the scanpath video shows that he occasionally glanced at the graph (e.g. at time 13 minutes, 7 seconds).

S2-35 created a graph based on his calculation model to visualise the price with the highest contribution. He looked very quickly at the graph but spent a long time looking at the numeric data table he had developed, comparing the various contributions (profits); see Figure 4.6.

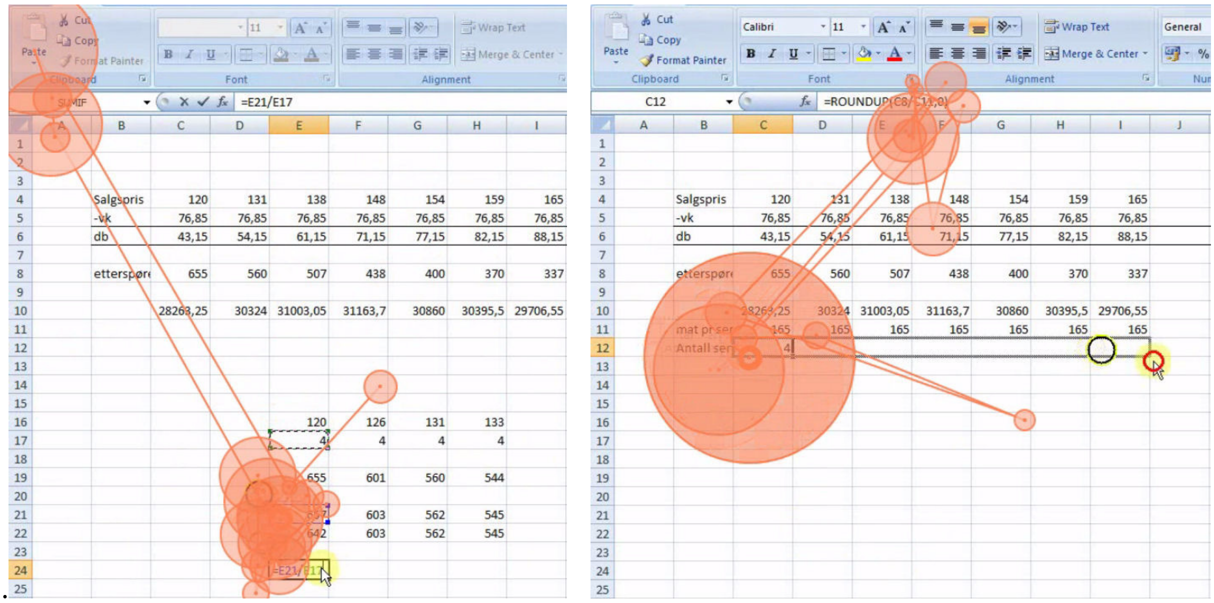


Figure 4.5. a) S2-35 used actual demand as the base for calculating waiter capacity; b) automatic calculation of the need for waiters based on demand/capacity.

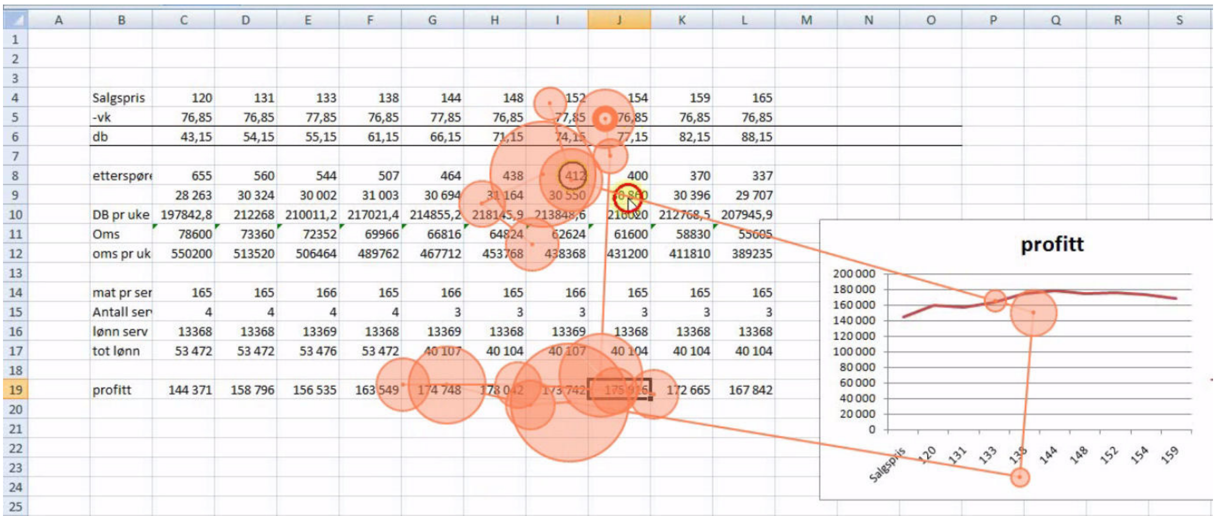


Figure 4.6. S2-35's focus on numeric data.

Later scanpaths reveal that he did not focus more on his own graph but continued to look at the numeric data in the model he had created. He selected and made the column (column H in Figure 4.6) with the highest contribution bold and focused on the number of waiters (3) and on the price (NOK 148) in this column.

Problem solving

Verbal protocol analysis

The verbal protocol reveals that S2-35 was not completely certain that he had found the optimal solution. He tried to increase the number of waiters, and he experimented twice (in two periods) with a higher price before he finally decided on the solution he had calculated in the problem-definition phase:

> 'No, I have approximately the demand that I want, and I earn good money. Yes, that's fine!'

Eye-movement analysis

In the problem-solving phase, the eye-tracking data reveal that S2-35 focused mostly on the price, the number of waiters and the average demand (e.g. at time 16 minutes and 23 seconds in the scanpath video). After having made his first five decisions, he returned to his calculations and looked at the two alternatives with a higher price (152 and 154) in his model. He did not look at prices lower than 148. Having entered his decisions into the spreadsheet model, he used the table format and sequentially compared the contribution for the periods he had managed the restaurant; see Figure 4.7.

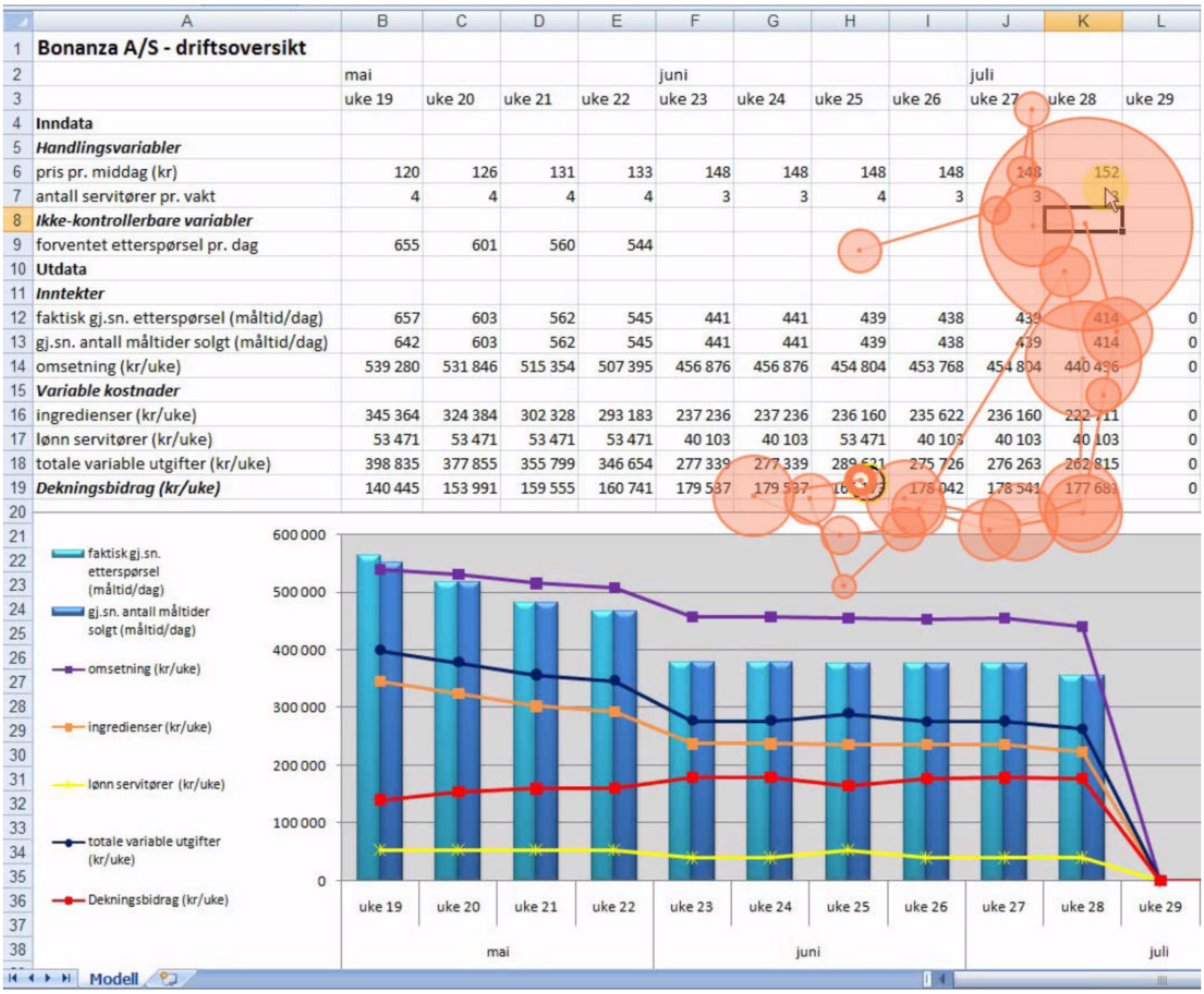


Figure 4.7. S2-35's sequential comparison of contributions.

Summary

Performance: S2-35 attained a good result because he understood the dependencies among the decision variables and the output variables before he made his decisions. He developed a relevant spreadsheet model. However, there were two reasons why he did not find the optimal solution: 1) he calculated the waiter capacity based on the wrong variable, that is, expected demand instead of actual sales and 2) his model only included the 'historic' prices he was given on paper (see Appendix C), that is, he had not extrapolated sales prices between NOK 144 and NOK 148.

Use of presentation format: The eye-tracking data show that S2-35 mainly used the table part of the combined presentation format. When he developed his spreadsheet model, it was also based on the table data of the historic presentation formats that he had received on paper. S2-35 attempted to develop an XY graph to get an overview of the relationship between sales prices and profit, but due to poor graph design, he did not seem to rely on the graph.

Complementarity of data collection methods: The verbal protocol shows that S2-35 had many silent periods in the problem-definition phase. In these periods, the eye-movement data are useful for understanding what he was doing; for example, they reveal that he used the wrong variable when calculating the waiter capacity. I would not have noticed this error without the eye-movement data. Finally, the eye-movement data reveal how S2-35 utilised the combined presentation format in the model and that he focused on the table.

S2-42 – Low performer

Subject S2-42 is the lowest performing subject in the group that received the low-complexity task and a combined presentation format. She attained a contribution index value of 0.9455, whereas the average value for her group was 0.9779; see Table 4.7. S2-42 spent 10.37 minutes trying to define the problem before she started making decisions, and she spent 50.03 minutes from her first decision until she finished the task. The average time consumption for her group was 20.83 minutes and 16.30 minutes, respectively. S2-42 got a score of -0.0556 on the NFC scale, whereas the average score for all subjects was 0.0768.

Table 4.7. Result and control variables for S2-42.

	S2-42	*Group average*
Contribution index value	0.9455	0.9779
Problem attention	60.40	37.13
Problem definition	10.37	20.83
Problem solving	50.03	16.30
Need for cognition	-0.0556	0.0768[5]

Problem definition

Verbal protocol analysis

The verbal protocol shows that S2-42 formulated the task in the following way:

> '… So the whole point is to optimise the price and demand in terms of the number of waiters too. Those are the three elements …'

Furthermore, the verbal protocol reveals that S2-42 attempted to find the capacity of the waiters during the problem-definition phase. She expressed an expectation of the capacity, i.e. NOK

[5] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

163, but she did not explain how she arrived at that figure. However, she seemed to forget her expectation or she did not rely on the estimated capacity, because she later expressed that she wanted to know the capacity of the waiters:

> 'If I had only known what the capacity was! (…) I don't get this!'

Rather than focusing on the capacity of the waiters, she directed her attention to the price of the meals, expressed through statements such as

> 'There is no use in just sitting here… I try to increase the price… maybe without reducing [the number of waiters] first.'

She then entered the price of NOK 138, but she did not explain why she chose this price.

Eye-movement analysis

The eye-tracking data reveals that S2-42 mainly focused on the price, variable costs and contribution variables. *Figure 4.8* illustrates a point in time when S2-42 focused on the contribution for the four historic periods. The historic data in the model shows a clear gap between the demand and the sales in the first week the restaurant was managed; see spreadsheet cells B12:B13 and the two leftmost blue bars in *Figure 4.8*.

However, the eye-tracking data show that S2-42 did not pay attention to these specific areas; nor did she compare sales versus actual demand in other weeks. Thus, the eye-tracking data reveal that S2-42 did not focus on the areas in the presentation format that she needed to calculate the waiter capacity correctly. Instead, she focused on the expected demand in the input data. This indicates why she arrived at a capacity of NOK 163.

The eye-tracking data show that she paid almost no attention to the graph during the experiment. The only exception was once during problem definition when she briefly looked at the demand in the historic periods.

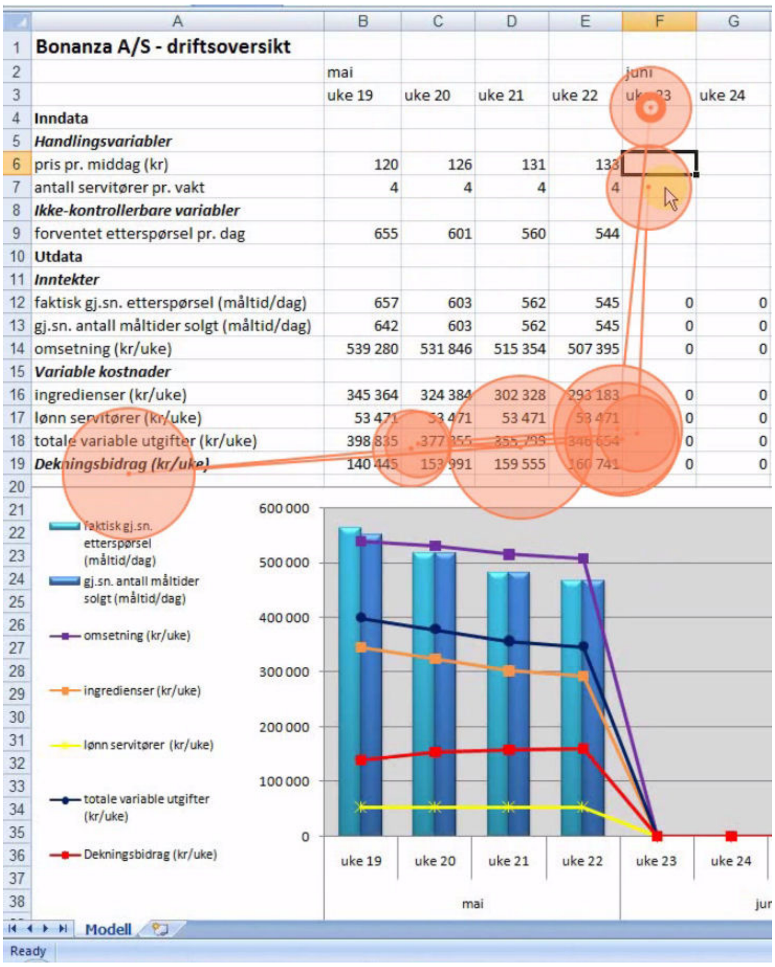


Figure 4.8. Horizontal eye-movements by S2-42.

Problem solving

Verbal protocol analysis

Also during the problem-solving phase, the protocol analysis shows that S2-42 still did not succeed in finding the waiter capacity:

> 'Is it usual to have four waiters here? I do not believe that one waiter can serve 337 guests at a time? … The question is whether two waiters can manage 348, for example. I do not know that. Well, I can try, but since the model does not give any answers to that ...'

S2-42 was silent for long periods while she developed the table shown in Figure 4.9. During the silent period, she performed calculations both on paper and in the spreadsheet. She used the historic sales/demand data to create a model to predict the optimal solution, and she used her model to experiment with three and two waiters. She decided that she wanted to continue with two waiters without any further explanation:

> 'What I am trying now is … it is no use having three waiters. I can use two instead of the former four. (…) I am uncertain of how many they can handle … I can try and see if they can handle 200 each.'

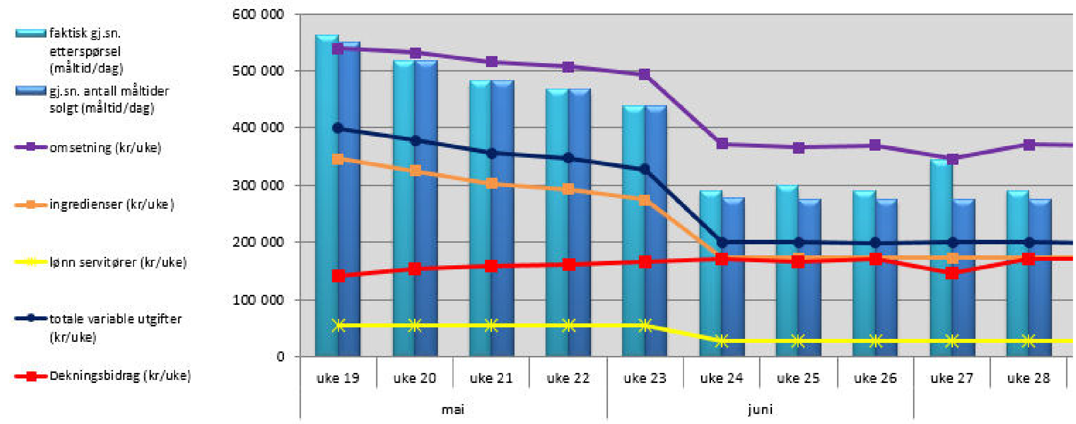
She tried various prices but ended up with a price of NOK 165, which she checked in one period, and then used for the remaining weeks; see Figure 4.9. The figure shows the decisions made by S2-42 and the calculations she made to reach her final solution. Weeks 28–39 all have a price of NOK 165 and 2 waiters; the last 11 weeks are not shown in the figure.

Eye-movement analysis

In the periods when S2-42 was silent, the screen video and eye-tracking data were useful for understanding what she was doing. The screen video reveals that S2-42 did not master the spreadsheet very well; she only partly used relative cell references in her formulas when referring to values in other cells. Furthermore, she did not use absolute cell references when referring to a fixed cell from other cells in a column and thus had do a lot of manual work when creating her model. As a result of her poor spreadsheet skills, S2-42's model was not dynamic as regards changes in its assumptions.

By showing the sequence of her calculations, the screen video also reveals a major calculation error that S2-42 seemed to overlook. When calculating the contribution for the case of three waiters, she multiplied the salary cost per waiter by five instead of three. This error explains why it was not an option for her to consider the alternative of three waiters, as shown in the analysis of the verbal protocol.

A	B	C	D	E	F	G	H	I	J	K
Bonanza A/S - driftsoversikt										
	mai				juni				juli	
	uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28
Inndata										
Handlingsvariabler										
pris pr. middag (kr)	120	126	131	133	138	165	163	165	154	165
antall servitører pr. vakt	4	4	4	4	4	2	2	2	2	2
Ikke-kontrollerbare variabler										
forventet etterspørsel pr. dag	655	601	560	544	510	337	349	337		
Utdata										
Inntekter										
faktisk gj.sn. etterspørsel (måltid/dag)	657	603	562	545	510	339	349	337	402	338
gj.sn. antall måltider solgt (måltid/dag	642	603	562	545	510	322	321	320	321	321
omsetning (kr/uke)	539 280	531 846	515 354	507 395	492 660	371 910	366 261	369 600	346 038	370 755
Variable kostnader										
ingredienser (kr/uke)	345 364	324 384	302 328	293 183	274 355	173 220	172 682	172 144	172 682	172 682
lønn servitører (kr/uke)	53 471	53 471	53 471	53 471	53 471	26 736	26 736	26 736	26 736	26 736
totale variable utgifter (kr/uke)	398 835	377 855	355 799	346 654	327 826	199 956	199 418	198 880	199 418	199 418
Dekningsbidrag (kr/uke)	140 445	153 991	159 555	160 741	164 834	171 954	166 843	170 720	146 620	171 337



	76,85	1910						
oms	ant	serv	Pris	var.kost	serv.kost	DB	3 serv	
78600	655	2		50336,75	3820	24443,25	26353,25	
75765	601	2		46186,85	3820	25758,15	27668,15	
73373	560	2		43036	3820	26517	28427	
72397	544	2		41806,4	3820	26770,6	28680,6	
69914	507	2		38962,95	3820	27131,05	29041,05	
66816	464	2		35658,4	3820	27337,6	29247,6	
64770	438	2		33660,3	3820	27289,7	29199,7	
62650	412	4		31662,2	7640	23347,8	25257,8	
61600	400	2		30740	3820	27040	28950	
59955	382	2		29356,7	3820	26778,3	28688,3	
58863	370	2		28434,5	3820	26608,5	28518,5	

Figure 4.9. S2-42's decisions and spreadsheet calculations.

Summary

Performance: S2-42 was a low performer because she did not understand the dependencies between the sales price, the demand/sales and the waiter capacity. She attempted to develop a spreadsheet model (see Figure 4.9), but the model reflected that she did not understand the dependencies among the variables. Furthermore, the model contained several constants hidden in the formulas, and S2-42 did not detect that she made a major error when she calculated the costs of three waiters. Her handling of the task was characterised by trial and error.

Use of presentation format: The eye-tracking data show that S-42 mainly used the table part of the combined presentation format both in the problem-definition phase and the problem-solving phase. When she developed her spreadsheet model, she used the table data of the historic presentation formats that she had received on paper.

Complementarity of data collection methods: The verbal protocol shows that S2-42 was silent in the problem-solving phase while she developed her spreadsheet model. In these periods, the screen video and the eye-tracking data were useful for understanding what she was doing. The eye-tracking data reveals that she focused on the expected demand instead of the sales when she estimated the waiter capacity. The screen video reveals the major calculation error related to the costs of the waiters.

Comparison of high and low performers

Performance: The main difference in the performance between S2-35 and S2-42 is the understanding of the dependencies among decision variables and the result variable. They both had analytic processes and developed a spreadsheet model to support their calculations of the optimal solution. However, S2-42 did not manage to represent the relationships among the variables correctly. They both made calculation errors, but S2-42's errors were more severe. S2-42's handling of the task was characterised by trial and error, while S2-35's approach was systematic.

Use of presentation format: Both S2-35 and S2-42 mainly used the table part of the combined presentation format. When they developed spreadsheet models, they both used the table data of the historic presentation formats that they had received on paper. S2-35 developed an XY graph to get an overview of the relationship between sales prices and profit, but he was not able to develop a graph that was useful for his decisions.

Complementarity of data collection methods: Both S2-35 and S2-42 were silent, particularly when they developed their spreadsheet models. In these periods, the screen video and the eye-tracking data were useful for understanding what they were doing, particularly as regards the revealing of their calculation errors. Furthermore, the eye-tracking data support the interpretation of S2-35's and S2-42's level of information processing.

4.5.4 Low-complexity task, graphical presentation format

S2-37 – High performer

Subject S2-37 is the highest performing subject in his group, with a contribution index value of 0.9980. Table 4.8 shows that S2-37 spent most of the problem-attention time on the problem-definition phase.

Table 4.8. Result and control variables for S2-37.

	S2-37	*Group average*
Contribution index value	0.9980	0.9737
Problem attention	21.00	39.14
Problem definition	17.68	20.55
Problem solving	3.32	18.59
Need for cognition	-0.1111	0.0768[6]

Problem definition

Verbal protocol analysis

According to the verbal protocol, S2-35 quickly found the waiter capacity from the 'historic' periods before he started to make decisions:

> 'I'll start with the week where they have maximised their capacity, that is, week 19. Four waiters have served approximately 650 guests. That gives approximately 160 guests per waiter.'

He continued:

> 'It will probably pay off to maximise the capacity of the waiters. I am going to test that. (…). I am going to find the variable unit cost per served plate (…). When maximising the waiter capacity, the unit cost should be equal since each waiter serves the same number of guests.'

The verbal protocol reveals that S2-37 then extracted numerical values for the 'historical' prices and demand from the graph he had received on paper. He had silent periods while doing so. According to my notes from observing the session, the numerical values were read from the

[6] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

graph very accurately with the help of a ruler and then entered into a table on a sheet of paper. The table had columns showing price, demand and sale. Unfortunately, I have not found the table in the file for S2-37.

S2-37 utilised the table to find the prices that maximised the capacity for two, three and four waiters. Then he calculated the contribution for the three alternatives using a calculator and concluded:

> 'Three waiters are better than four waiters, and three waiters are also better than two waiters.'

According to the verbal protocol, S2-37 expected a price of NOK 143 to result in a demand for 480 meals per day. Consequently, he entered this price and three waiters into the model for week 23.

The verbal protocol reveals that S2-37 was silent for long periods, especially when he was looking at the graphs, and that he was asked to think aloud several times during the experiment.

Eye-movement analysis

The eye-tracking data show that S2-37 carefully studied the graphs during these silent periods. The eye movements shifted horizontally between the different graphic elements representing the variables of the model and the axes. The screen video shows that he often scrolled the spreadsheet window to see the axis on the right-hand side of the graph. Figure 4.10 a), for instance, shows a scanpath when S2-37's visual attention was focused on the average demand for meals per day.

Furthermore, S2-37 repeatedly hovered over graph points with the mouse pointer so that Excel displayed the exact numeric values for these points; see Figure 4.10 b).

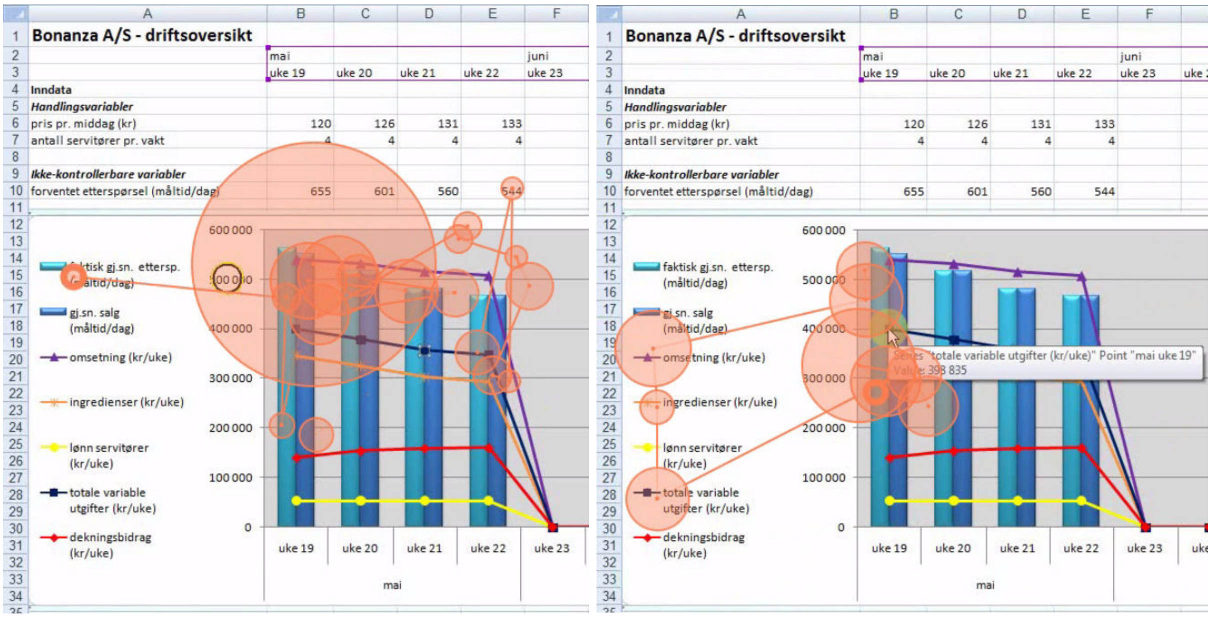


Figure 4.10. S2-37's a) horizontal shifts: focus on the average demand per day; b) S2-37's extraction of exact numeric values from the graphs.

Problem solving

Verbal protocol analysis

Inspecting the results from week 23, S2-37 realised that the demand had only been 475 meals and not the 480 that he was expecting. He therefore decided that he would reduce the price to NOK 142 in week 24. After doing so, he concluded that he had optimised the waiters' capacity and thus chose to continue with this combination for the remaining weeks in the experiment:

> 'Now the capacity is fully utilised.'

Eye-movement analysis

The eye-tracking data show that S2-37 carefully studied the results of his decisions and that he compared input and output variables. Figure 4.11 illustrates that he focused on the demand for meals and the actual sales, as well as on the price and the number of employees in week 23.

The figure also shows that he focused on the numeric values from the graph. Later scanpaths reveal that he extracted numeric values for the demand and sales per day, the sales per week and the variable costs per week. However, Figure 4.11 illustrates that S2-37 did not focus on the contribution or try to extract the data values of the contribution.

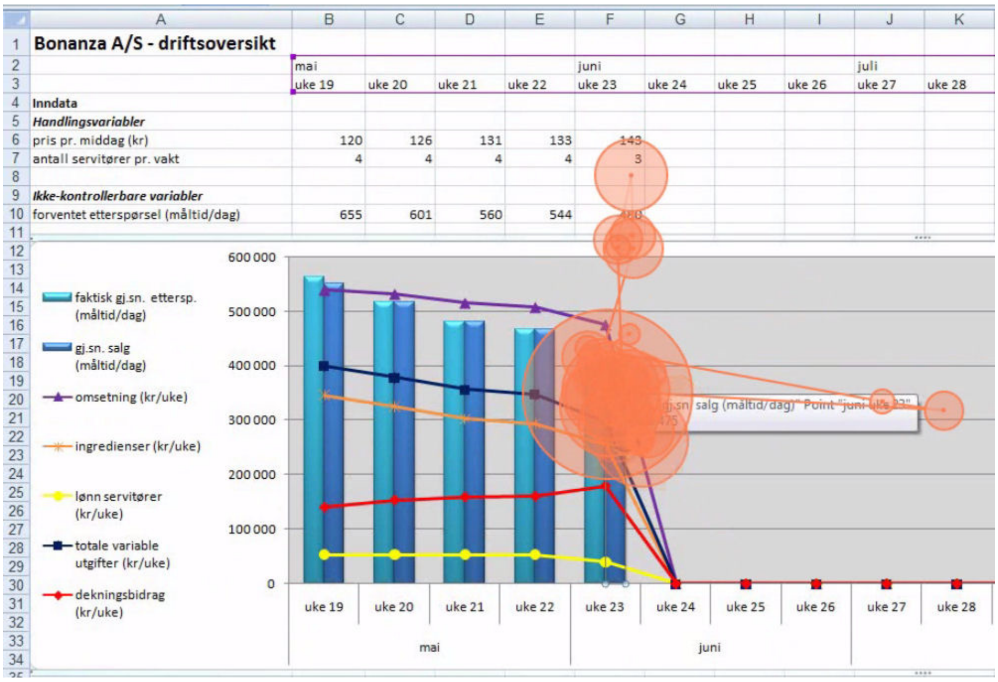


Figure 4.11. S2-37's focus on demand and sales per day and extraction of exact numeric values from the graph.

Summary

Performance: S2-37 attained a good result because he understood the dependencies among the decision variables and the output variables before he made his decisions but also because he was lucky. He did not really understand how to find the optimal solution, as he focused on fully utilising the waiter capacity. The combination of fully utilising the waiter capacity at the price of NOK 142 is close to the optimal solution because the profit curve is rather flat around the optimal price.

Use of presentation format: S2-37 was given the graphical presentation format. However, in the problem-definition phase he mainly used the graphical presentation format to create a data table representing the dependencies among the variables. He carefully read out accurate values from the graphs and developed an additional presentation of numerical data in a table. The verbal protocol shows that he used this table when making his decisions. The eye-tracking data further reveal that he used the numeric data points in the graph when he evaluated the results of his decisions. Thus, he had a need for accurate data to handle the task.

Complementarity of data collection methods: S2-37 was silent for long periods, and the eye-tracking data were useful supplements to the verbal protocol. The eye-tracking data shows that he fixated on the numeric values that Excel provided in order to have accurate data. The screen video shows that he hovered over different parts of the graph with the mouse cursor in order to have Excel reveal screen tips with accurate data. In addition, the eye-tracking data confirms the impression gained from the verbal protocol that he did not focus on the contribution when evaluating the results of his decisions.

S2-06 – Low performer

Subject S2-06 is the lowest performing subject in the group that received the low-complexity task and a graphical presentation format. He scored 0.9381 on the contribution value index, whereas the group average was 0.9737. As shown in Table 4.9, S2-06 spent only 1.62 minutes before making his first decision but then spent 58.48 minutes in the problem-solving phase.

Table 4.9. Result and control variables for S2-06.

	S2-06	*Group average*
Contribution index value	0.9381	0.9737
Problem attention	60.10	39.14
Problem definition	1.62	20.55
Problem solving	58.48	18.59
Need for cognition	-0.2778	0.0768[7]

Problem definition

Verbal protocol analysis

The verbal protocol reveals that S2-06 did not utilise the historical price/demand data that he had access to. Furthermore, the protocol shows that S2-06 almost immediately entered values into the spreadsheet model. He entered a price of NOK 133 into the model after a brief review of the ‘historical’ trends. He also reduced the number of waiters from four to three:

> ‘I see that there has been a drop in the sales but also a strong reduction in costs, hence the contribution has increased. The salary cost is the same since they are just as many at work. My immediate thought is that it is tempting to reduce the waiters by one – since contribution has increased and there has been a drop in sales.’

[7] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

Eye-movement analysis

The eye-tracking data show that S2-06 had horizontal eye movements between the variable labels and their corresponding lines/bars in the diagram. He also studied the development of the variables over time. Furthermore, the eye-tracking data reveal that he did not pay attention to the parts of the graph showing a gap between actual sales and average demand, which could be used to find the capacity of the waiters.

Problem solving

Verbal protocol analysis

The verbal protocol shows that S2-06 was asked to think aloud several times during the problem-solving phase.

According to the protocol, S2-06 continued to try different prices in the problem-solving phase, mostly with three waiters:

> ‘I think I will just try to increase the price with the same number of waiters, just to see if it might have any effect. I guess it has, but how much? (…) The contribution went up. I sell just as much as before. I’ll try a bit more in that direction. Maybe a really high price this time since it didn’t seem to affect the sales.’

The quote presented above illustrates that S2-06 used ordinal considerations like ‘it went up’, ‘it went down’ and ‘it’s about the same as before’.

S2-06 focused mainly on the price, for example:

> ‘I try to find the price that maximises the contribution. I would like to try with an even higher price to check that the contribution goes down. I try to find the price that gives the highest contribution, and the increased price almost closed the gap [between sales and actual demand].’

S2-06 experimented with both five and two waiters; see Figure 4.12. In week 28, for instance, he tried a price of NOK 120 and five waiters and then concluded that the contribution in that week was the lowest of all the weeks. Furthermore, he reasoned:

> ‘The costs increased dramatically so this didn’t do any good. Clearly, I don’t manage to sell enough here.’

The combination of a price of NOK 120 and four waiters (i.e. one less) had, however, already been used in the first of the four historic weeks (week 19) and had resulted in the lowest contribution until S2-06's decision in week 28.

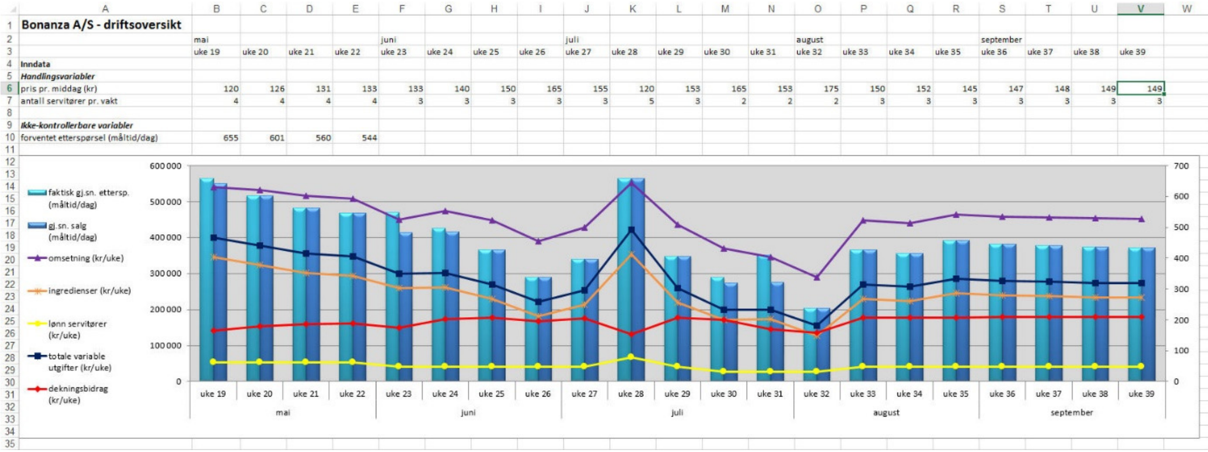


Figure 4.12. S2-06's decisions.

S2-06 did not perform any calculations on paper or in the spreadsheet, and he had only one numerical comparison when he found the capacity of the waiters in week 30:

> 'Let's see. In week 24, three people employed and they managed to handle about 400 to 450, let's say 475 guests. So if I divide 475 by 3. That is about 158 guests per waiter.'

However, he did not use this capacity in his decision-making process. He continued to lower the price from NOK 165 to NOK 153 in week 31, resulting in an even bigger gap between average demand and actual sales. S2-06 reasoned:

> 'The demand for meals has increased from week 30 to 31, but the sale remains the same. How come the contribution has dropped when both the sales and the costs are the same?'

S2-06 tried another period with two waiters and a higher price of NOK 175 before he decided to increase the number of waiters to three and tried with prices ranging from NOK 145 to NOK 152 in the remaining seven periods. In week 34, he stated:

> 'I'm going nowhere (…). Now it is trial and error.'

Eye-movement analysis

S2-06 did not focus on the sales/demand during the problem-definition phase. However, during the problem solution phase, the eye-tracking data reveal that S2-06 focused on the area showing a gap between the sales and the average demand in weeks 23 and 24; see Figure 4.13. The verbal protocol, however, indicates that he only focused on the sales and that he did not compare the sales to the actual demand. He was concerned only with the sales.

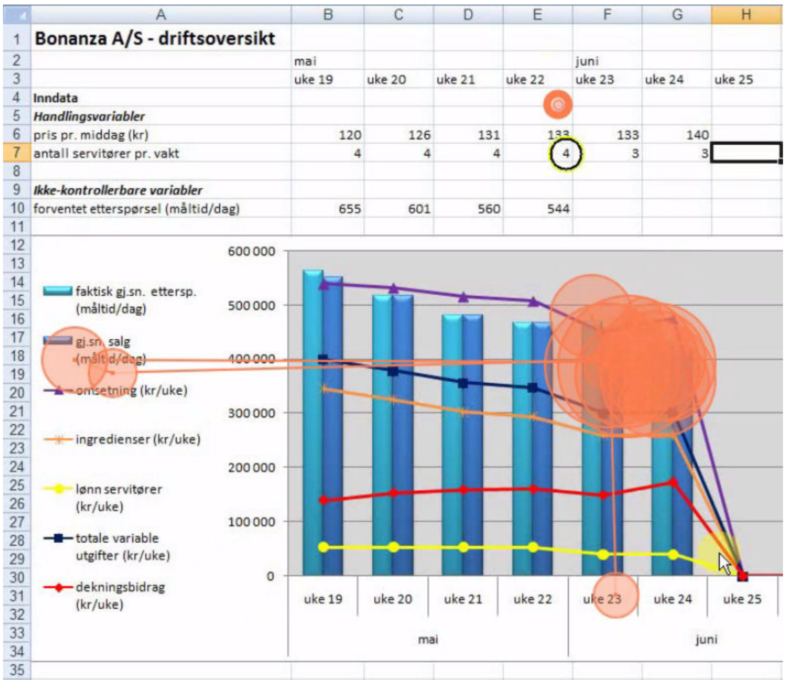


Figure 4.13. S2-06's fixating on average sales.

The eye-tracking data further reveal that S2-06 did not have many systematically vertical shifts between two or more variables. Instead, the eye movements appear rather unsystematic (e.g. at time 00:30:27 – 00:30:33 in the scanpath video).

Summary

Performance: S2-06 was a low performer because he used a trial-and-error strategy without learning from earlier trials. He understood the objective of the task, that is, to maximise the contribution. In the problem-definition phase, he attempted to understand the development of the values of the output variables, but he did not link the development explicitly to the values of the decision variables. In his handling of the task, he never developed an understanding of the relationships among the decision variables and the output variables. He performed only one calculation, that is, in week 30 when he found the waiter capacity. His argumentation was characterised by ordinal considerations.

Use of presentation format: S2-06 did not use the XY graph to learn about the relationship between prices and 'historic' demand. He only used the graph in the Excel model in his handling of the task.

Complementarity of data collection methods: The eye-tracking data reveals few vertical shifts among variables and thus complements the impression from the verbal protocol that S2-06 did not develop an understanding of the relationships among input and output variables. However, the verbal protocol also complements the eye-tracking data. The eye-tracking data shows that S2-06 focused his visual attention on the area covering the gap between the sales and the demand in weeks 23 and 24, but the verbal protocol shows that he only used the sales data in his handling of the task. Thus, the eye-tracking data, presented for example in Figure 4.13, need to be interpreted in relation to the verbal protocol data.

Comparison of high and low performer

Performance: The main difference in the performance between S2-37 and S2-06 relates to how well the two candidates understood the dependencies among the decision variables and the result variables. S2-37 spent most of his time in the problem-definition phase and handled the task mainly using analytic processes. S2-06, on the other hand, did not put much effort into defining the task before making his decisions. He handled the task using mainly spatial processes and ordinal considerations.

Use of presentation format: S2-37 complemented his original XY graph with an additional table that represented the dependencies among variables as accurate numeric values. He used that table when handling the task and performing accurate calculations and marginal considerations. He used accurate numeric data points in the line graph when evaluating the results of his decisions. S2-06 did not use the XY graph showing the relationship between price and demand in his handling of the task. He only used the line graph in Excel.

Complementarity of data collection methods: The eye-tracking data were useful complements for understanding the behaviour of both S2-37 and S2-06. S2-37 had many silent periods, and the eye-tracking data show that he focused on accuracy and numeric values. For S2-06, the eye-tracking data complements my impression from the verbal protocols that he did not understand the relationships among input and output variables, but the verbal protocol also complements my understanding of the eye-tracking data by showing that S2-06 did not use all the data present in the areas he was looking at.

4.5.5 Low-complexity task, tabular presentation format

S2-39 – High performer

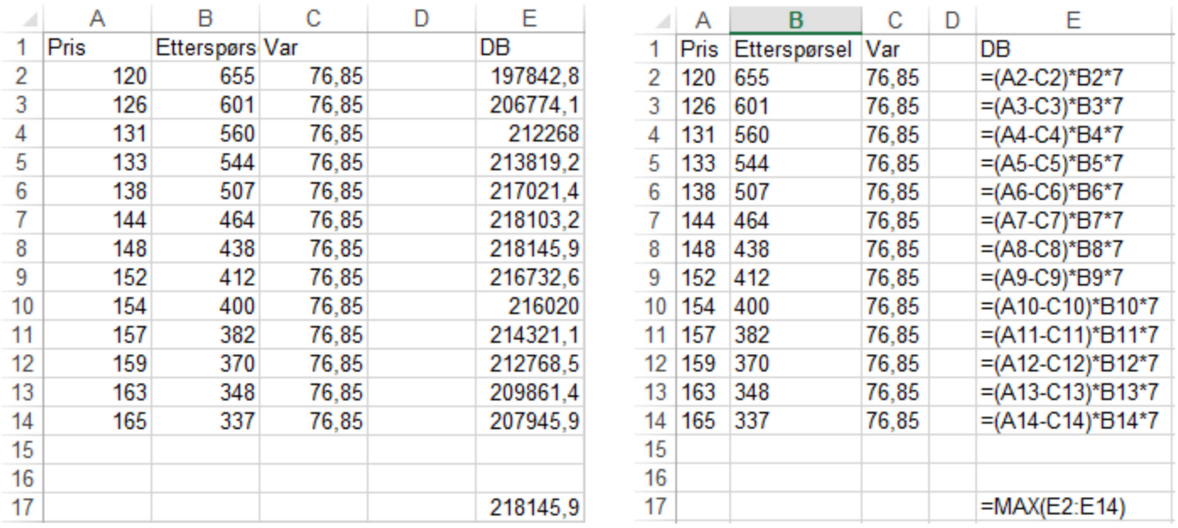
Table 4.10. Result and control variables for S2-39.

	S2-39	*Group average*
Contribution index value	0.9982	0.9867
Problem attention	17.60	22.59
Problem definition	15.08	12.74
Problem solving	2.52	9.85
Need for cognition	0.1667	0.0768[8]

Problem definition

Verbal protocol analysis

According to the verbal protocol, S2-39 spent most of time during the experiment defining the problem before he started making decisions (see Table 4.10). He studied the spreadsheet model and the 'historic' prices and demand data he had available on paper (see Appendix C). He was very careful to understand the calculation of the waiters' wages correctly. The verbal protocol reveals that he was silent for 3.12 minutes when he developed the spreadsheet model shown in Figure 4.14.



	A	B	C	D	E
1	Pris	Etterspørs	Var		DB
2	120	655	76,85		197842,8
3	126	601	76,85		206774,1
4	131	560	76,85		212268
5	133	544	76,85		213819,2
6	138	507	76,85		217021,4
7	144	464	76,85		218103,2
8	148	438	76,85		218145,9
9	152	412	76,85		216732,6
10	154	400	76,85		216020
11	157	382	76,85		214321,1
12	159	370	76,85		212768,5
13	163	348	76,85		209861,4
14	165	337	76,85		207945,9
15					
16					
17					218145,9

	A	B	C	D	E
1	Pris	Etterspørsel	Var		DB
2	120	655	76,85		=(A2-C2)*B2*7
3	126	601	76,85		=(A3-C3)*B3*7
4	131	560	76,85		=(A4-C4)*B4*7
5	133	544	76,85		=(A5-C5)*B5*7
6	138	507	76,85		=(A6-C6)*B6*7
7	144	464	76,85		=(A7-C7)*B7*7
8	148	438	76,85		=(A8-C8)*B8*7
9	152	412	76,85		=(A9-C9)*B9*7
10	154	400	76,85		=(A10-C10)*B10*7
11	157	382	76,85		=(A11-C11)*B11*7
12	159	370	76,85		=(A12-C12)*B12*7
13	163	348	76,85		=(A13-C13)*B13*7
14	165	337	76,85		=(A14-C14)*B14*7
15					
16					
17					=MAX(E2:E14)

Figure 4.14. S2-39's own spreadsheet model, normal view and formula view.

[8] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

He stated:

> ‘Yes, now I thought that if I took the price minus the variable unit costs and looked at what gives the highest contribution from the meals. And when I did that, it looked as if a price of NOK 148 would be great.’

He used the MAX function in the spreadsheet to find the row with the highest contribution; see cell E17 in Figure 4.14.

He indicated that he expected a demand of 438 meals at the price of NOK 148 and then considered the number of waiters necessary to cover the demand:

> ‘The first week of May shows that four waiters have handled a demand of 657 guests per day. That is 164 meals per waiter.’

He asked questions to make sure he had understood the calculation of the waiters’ wages correctly, and he probably then understood that his calculation of the waiter capacity was not quite correct. However, he considered that three waiters would be able to cover a demand of about 438 and entered the price of NOK 148 and three waiters into the model for week 23.

Eye-movement analysis

Corresponding to S2-39’s attempts to understand the task, the eye-tracking data shows many horizontal eye movements between the variable labels and the attached values. Furthermore, the data show that S2-39 sequentially compared the values for the sales, demand, contribution and total variable costs. For instance, Figure 4.15 shows a moment in time when S2-39 paid attention to the area of the table showing the average sale and demand values in the four historic weeks of the model.

After having made his own spreadsheet model, he returned to the spreadsheet where he was going to enter his decisions. The eye-tracking protocol now reveals more vertical eye movements comparing values across variables than he had had before he created his model. Figure 4.16 shows that S2-39 shifted his attention between the areas of the table showing the number of waiters and the average demand and sales at the same time as the verbal protocol shows that he calculated the waiter capacity. Notice that the fixations are shifted upwards in the screen video compared to where S2-39 actually focused his attention.

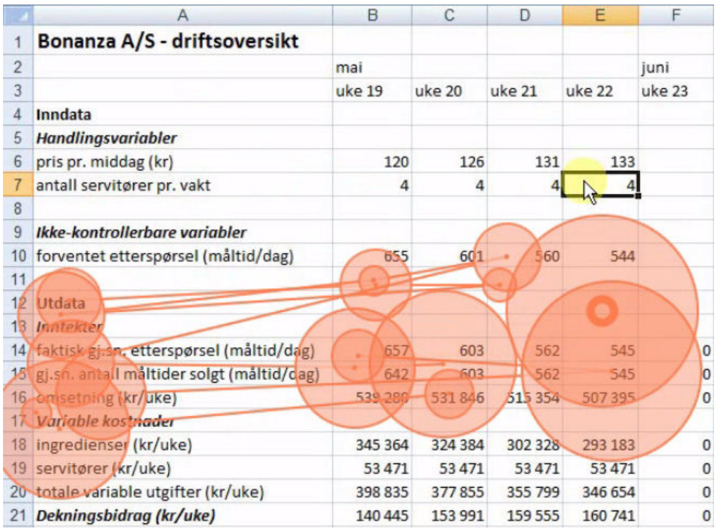


Figure 4.15. Horizontal eye-movements for S2-39.

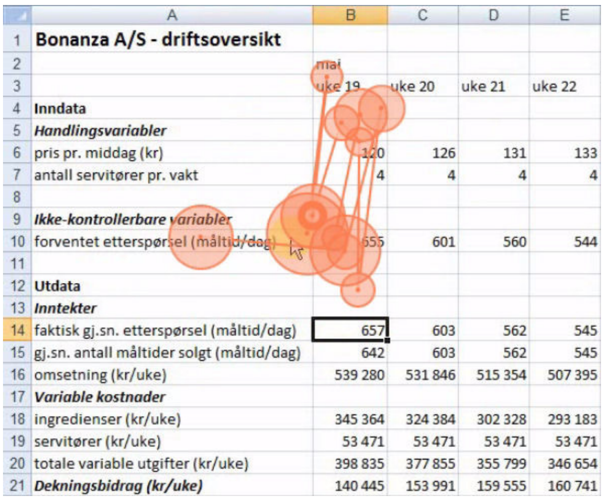


Figure 4.16. Vertical eye-movements (offset in the eye-tracking data) for S2-39.

Problem solving

Verbal protocol analysis

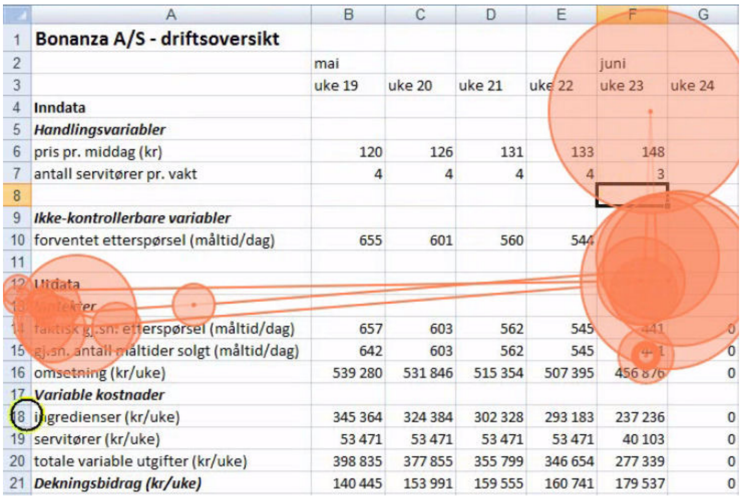
Inspecting the results of his decisions for week 23, the verbal protocol shows that S2-39 considered whether he should reduce the price and whether he should reduce the number of waiters. He concluded as follows:

> 'It seems to me that if I had reduced the number of waiters to two, I would probably have a problem. Therefore, I think this is the best solution that I can get.'

S2-39 copied his decision from week 23 throughout the remaining weeks of the model. In a brief discussion with me, he expressed that he was satisfied.

Eye-movement analysis

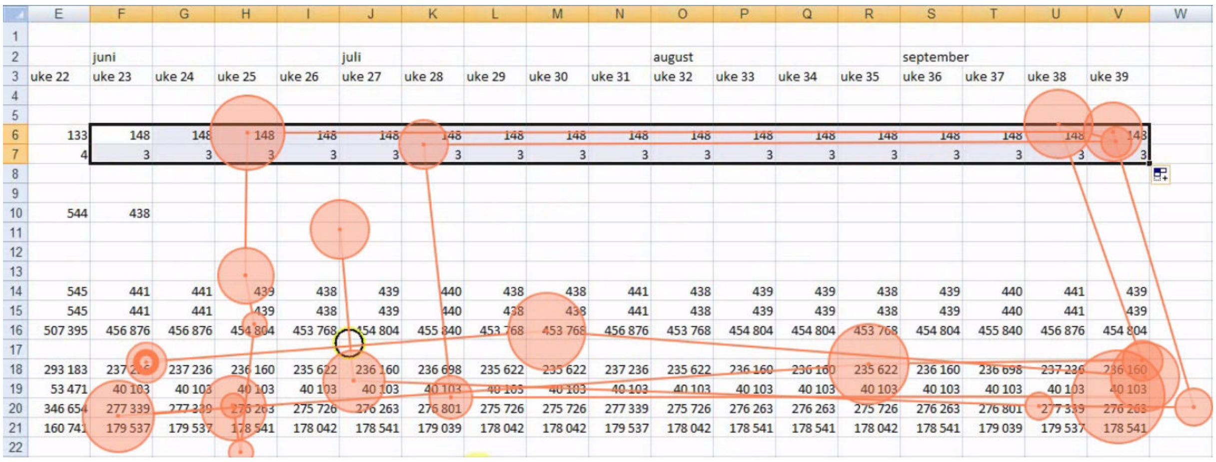
After S2-39 had made his first decision, the eye-tracking data show that he focused his visual attention on the area of the output data showing the average demand and average sales; see Figure 4.17. He also focused on the three waiters that he had employed. Later scanpaths reveal that he compared the contribution from his first week as manager to the contribution in the four historic weeks.



	A	B	C	D	E	F	G
1	**Bonanza A/S - driftsoversikt**						
2		mai				juni	
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24
4	**Inndata**						
5	***Handlingsvariabler***						
6	pris pr. middag (kr)	120	126	131	133	148	
7	antall servitører pr. vakt	4	4	4	4	3	
8							
9	***Ikke-kontrollerbare variabler***						
10	forventet etterspørsel (måltid/dag)	655	601	560	544		
11							
12	**Utdata**						
13	***Inntekter***						
14	faktisk gj.snt. etterspørsel (måltid/dag)	657	603	562	545	441	0
15	gj.sn. antall måltider solgt (måltid/dag)	642	603	562	545	441	0
16	omsetning (kr/uke)	539 280	531 846	515 354	507 395	456 876	0
17	***Variable kostnader***						
18	ingredienser (kr/uke)	345 364	324 384	302 328	293 183	237 236	0
19	servitører (kr/uke)	53 471	53 471	53 471	53 471	40 103	0
20	totale variable utgifter (kr/uke)	398 835	377 855	355 799	346 654	277 339	0
21	***Dekningsbidrag (kr/uke)***	140 445	153 991	159 555	160 741	179 537	0

Figure 4.17. Focus on average sales and demand for S2-39.

The screen video reveals that he compared contributions for the weeks he had managed the restaurant; see Figure 4.18.



	E	F	G	H	I	J	K	L	M
1									
2		juni				juli			
3	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30
4									
5									
6	133	148	148	148	148	148	148	148	148
7	4	3	3	3	3	3	3	3	3
8									
9									
10	544	438							
11									
12									
13									
14	545	441	441	439	438	439	440	438	438
15	545	441	441	439	438	439	440	438	438
16	507 395	456 876	456 876	454 804	453 768	454 804	455 840	453 768	453 768
17									
18	293 183	237 236	237 236	236 160	235 622	236 160	236 698	235 622	235 622
19	53 471	40 103	40 103	40 103	40 103	40 103	40 103	40 103	40 103
20	346 654	277 339	277 339	276 263	275 726	276 263	276 801	275 726	275 726
21	160 741	179 537	179 537	178 541	178 042	178 541	179 039	178 042	178 042
22									

	N	O	P	Q	R	S	T	U	V	W
1										
2		august				september				
3	uke 31	uke 32	uke 33	uke 34	uke 35	uke 36	uke 37	uke 38	uke 39	
4										
5										
6	148	148	148	148	148	148	148	148	148	
7	3	3	3	3	3	3	3	3	3	
8										
9										
10										
11										
12										
13										
14	441	438	439	439	438	439	440	441	439	
15	441	438	439	439	438	439	440	441	439	
16	456 876	453 768	454 804	454 804	453 768	454 804	455 840	456 876	454 804	
17										
18	237 236	235 622	236 160	236 160	235 622	236 160	236 698	237 236	236 160	
19	40 103	40 103	40 103	40 103	40 103	40 103	40 103	40 103	40 103	
20	277 339	275 726	276 263	276 263	275 726	276 263	276 801	277 339	276 263	
21	179 537	178 042	178 541	178 541	178 042	178 541	179 039	179 537	178 541	
22										

Figure 4.18. S2-39's evaluation the result.

Summary

Performance: S2-39 attained a good result because he understood the dependencies among the decision variables and the output variables before he made his decisions. He developed a relevant spreadsheet model. However, there were two reasons why he did not find the optimal solution: 1) He calculated the waiter capacity based on the wrong variable, that is, expected demand instead of actual sales. 2) His model only included the 'historic' prices he was given on paper (see Appendix C), that is, he had not extrapolated sales prices between NOK 144 and NOK 148.

Use of presentation format: S2-39 used the tables, both the one in the spreadsheet model and the one on paper showing historic prices and demands, when solving the task. He did not attempt to create a graph based on the tables.

Complementarity of the data collection methods: The verbal protocol shows that S2-39 had silent periods when he made his spreadsheet model shown in Figure 4.14. In these periods, the scanpath video was useful for showing the logical development of the model. There is consistency between the verbal protocol and the eye-tracking data. Both process-tracing methods show that S2-39 used the wrong variable when calculating the waiter capacity.

S2-41 – Low performer

Table 4.11 gives an overview of the result and control variables for subject S2-41. He participated in a treatment group where I had only four subjects. Three of the subjects (S2-39, S2-41 and S2-33) performed rather well (index values ranging from 0.9912–0.9982), whereas the last subject (S2-30) performed significantly worse (index value of 0.9640). Unfortunately, I do not have valid eye-movement data for S2-30, and in the following analysis of the low performer, I have therefore used S2-41 (second lowest performer). This explains why his value on the contribution index is actually higher than the group average.

Table 4.11. Result and control variables for S2-41.

	S2-41	*Group average*
Contribution index value	0.9912	0.9867
Problem attention	23.82	22.59
Problem definition	6.20	12.74
Problem solving	17.62	9.85
Need for cognition	0.2222	0.0768[9]

[9] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

Problem definition

Verbal protocol analysis

The verbal protocol shows that S2-41 studied the historical data, both in the historic periods in the spreadsheet model and the average price and demand data shown on paper. He reasoned:

> 'I see that the sale is higher with low prices. However, the contribution seems to have increased as the prices have been increased. I'm being judged by the contribution so that's more important than a high sale at a low price.'

The protocol shows that S2-41 wanted to increase the price and that he expected the demand to drop. Therefore, he also wanted to reduce the number of employees from four to three. However, he was worried that the employees would have too much work to do:

> 'This is a horrible decision to make considering the three poor waiters which is then left in the restaurant.'

Furthermore, S2-41 calculated the capacity of the waiters:

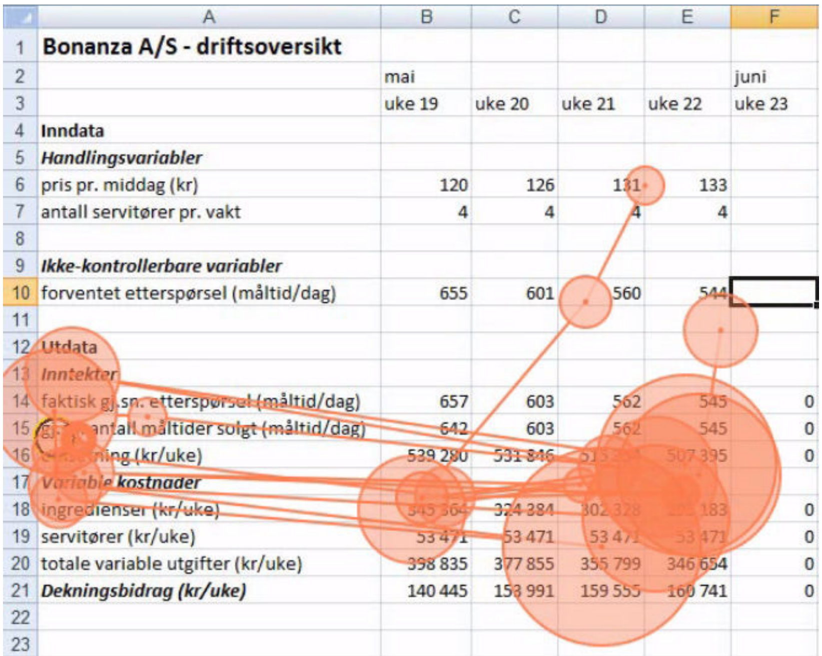
> 'Four waiters have been able to handle 655 guests during one day. Hence, three waiters should at least be able to handle 491 guests per day.'

The statement above shows, however, that S2-41 based his calculations of the waiter capacity on the expected demand per day in week 19 and not on the actual sales per day; therefore, the capacity of the waiters was not completely correct.

S2-41 used the historic price/demand data to find the price that historically would provide a demand of about 491 guests. He then entered a price of NOK 140 and three waiters into the model.

Eye-movement analysis

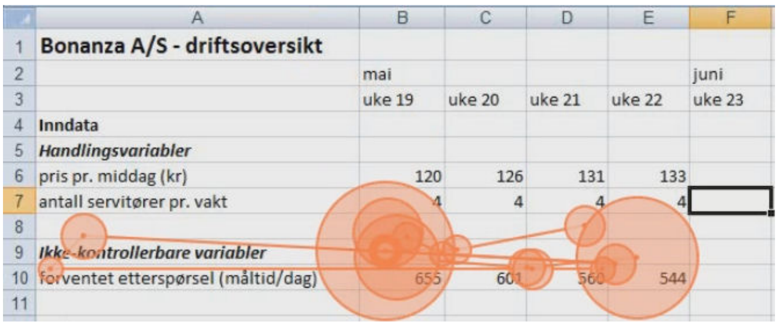
The eye-tracking data show that S2-41 had many horizontal shifts in his eye movements early in the problem-definition phase. He repeatedly shifted his attention between the variable labels and the corresponding data values; see Figure 4.19. Such shifts might indicate that he tried to learn the model.



	A	B	C	D	E	F
1	**Bonanza A/S - driftsoversikt**					
2		mai				juni
3		uke 19	uke 20	uke 21	uke 22	uke 23
4	**Inndata**					
5	***Handlingsvariabler***					
6	pris pr. middag (kr)	120	126	131	133	
7	antall servitører pr. vakt	4	4	4	4	
8						
9	***Ikke-kontrollerbare variabler***					
10	forventet etterspørsel (måltid/dag)	655	601	560	544	
11						
12	**Utdata**					
13	***Inntekter***					
14	faktisk gj.sn. etterspørsel (måltid/dag)	657	603	562	545	0
15	[illegible] antall måltider solgt (måltid/dag)	642	603	562	545	0
16	[illegible]ning (kr/uke)	539 280	531 846	[illegible]	507 395	0
17	***Variable kostnader***					
18	ingredienser (kr/uke)	345 364	324 384	302 328	[illegible] 183	0
19	servitører (kr/uke)	53 471	53 471	53 471	53 471	0
20	totale variable utgifter (kr/uke)	398 835	377 855	355 799	346 654	0
21	***Dekningsbidrag (kr/uke)***	140 445	153 991	159 555	160 741	0
22						
23						

Figure 4.19. Horizontal eye movements during the problem-definition phase for S2-41.

In addition, S2-41 sequentially focused his visual attention on weekly values for the different variables. In particular, he focused on contribution, expected demand, actual demand, actual sales, price and the number of waiters. He did not direct his visual attention much to the cost variables. Figure 4.20 shows an example where S2-41 sequentially compared the expected demand for meals over time to the historic period.



	A	B	C	D	E	F
1	**Bonanza A/S - driftsoversikt**					
2		mai				juni
3		uke 19	uke 20	uke 21	uke 22	uke 23
4	**Inndata**					
5	***Handlingsvariabler***					
6	pris pr. middag (kr)	120	126	131	133	
7	antall servitører pr. vakt	4	4	4	4	
8						
9	***Ikke-kontrollerbare variabler***					
10	forventet etterspørsel (måltid/dag)	655	601	560	544	
11						

Figure 4.20. Horizontal eye movements for S2-41, focusing on expected demand for meals.

S2-41 also had vertical eye movements, comparing values across variables. When doing so, he focused on the values for the first (19) and the last (22) week in the historic period. Notice that the fixations are a bit offset compared to the real fixation location. In the upper right corner, the fixation data is actually higher up on the screen than shown in Figure 4.21.

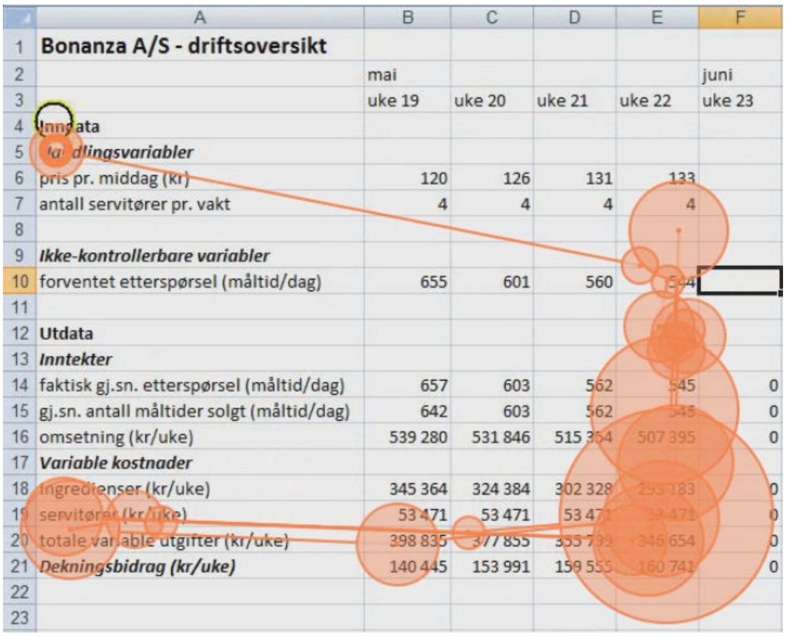


Figure 4.21. Vertical eye movements for S2-41, comparing values across variables.

Problem solving

Verbal protocol analysis

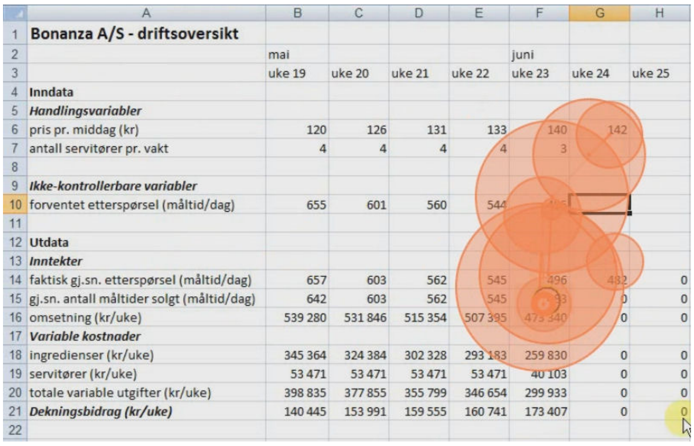
Interpreting the results of his first decision, S2-41 realised that he had not been able to handle all the guests. He therefore decided to run another week with a price of NOK 142 and still keep three waiters. He then reasoned:

> 'I managed to handle all the guests. Good! (…) I wonder what happens if I increase the price even more and reduce the number of waiters to two. I am afraid that the sales will drop too much and hence the contribution will be reduced. However, it should not stand untested. Let's try.'

Based on his calculated capacity of the waiters, S2-41 reasoned that two waiters would handle 320 guests per day. Furthermore, he used the historic price/demand data to decide the price as NOK 170. He realised that the contribution dropped and decided to go back to his former decision of three waiters. In the following six periods, he tested prices ranging from NOK 142 to NOK 144 and concluded that the price of NOK 143 provided the highest contribution. Thus, in the last eight weeks of the experiment he used a price of NOK 143 and three waiters.

Eye-movement analysis

The eye-tracking data shows that having made his first decision, S2-41 focused on the difference between demand and the sales, and he compared his results with those variables against the four historic periods. In addition, as Figure 4.22 shows, when he was entering the price for week 24, he focused a lot on the area showing a gap between the sale of and demand for meals in week 23. When he had made his decisions for week 24, he focused on the area showing the contribution and compared his results in the two first weeks that he managed the restaurant.

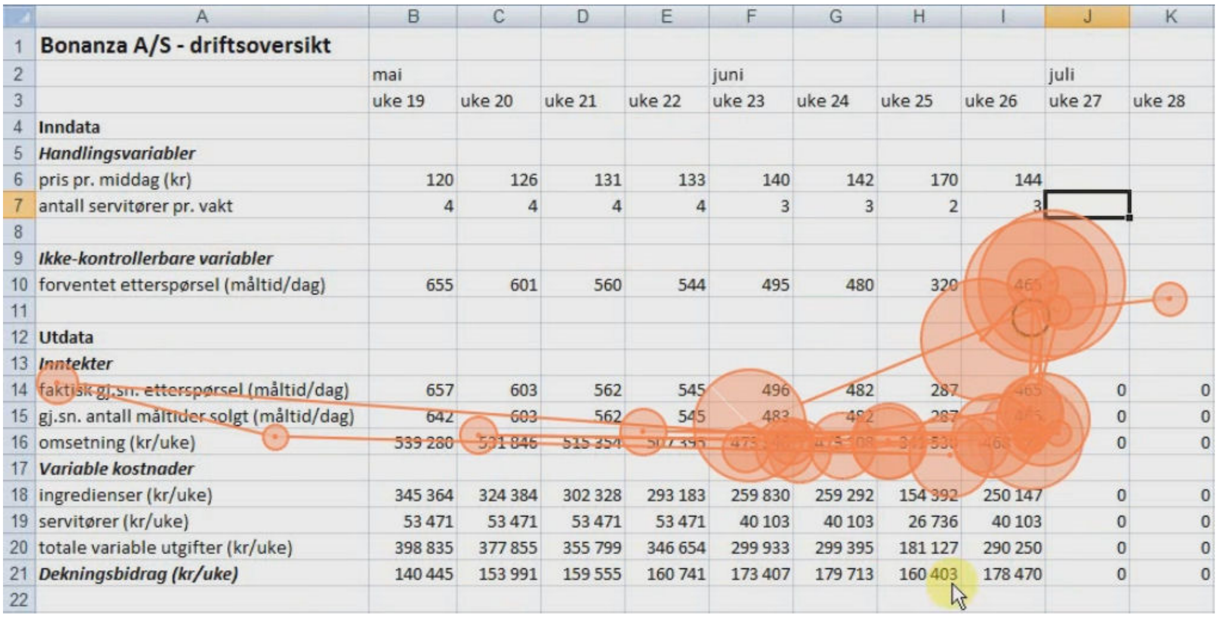


	A	B	C	D	E	F	G	H
1	**Bonanza A/S - driftsoversikt**							
2		mai				juni		
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25
4	**Inndata**							
5	***Handlingsvariabler***							
6	pris pr. middag (kr)	120	126	131	133	140	142	
7	antall servitører pr. vakt	4	4	4	4	3		
8								
9	***Ikke-kontrollerbare variabler***							
10	forventet etterspørsel (måltid/dag)	655	601	560	544	[illegible]		
11								
12	**Utdata**							
13	***Inntekter***							
14	faktisk gj.sn. etterspørsel (måltid/dag)	657	603	562	545	496	482	0
15	gj.sn. antall måltider solgt (måltid/dag)	642	603	562	545	[illegible]	0	0
16	omsetning (kr/uke)	539 280	531 846	515 354	507 395	473 340	0	0
17	***Variable kostnader***							
18	ingredienser (kr/uke)	345 364	324 384	302 328	293 183	259 830	0	0
19	servitører (kr/uke)	53 471	53 471	53 471	53 471	40 103	0	0
20	totale variable utgifter (kr/uke)	398 835	377 855	355 799	346 654	299 933	0	0
21	***Dekningsbidrag (kr/uke)***	140 445	153 991	159 555	160 741	173 407	0	0
22								

Figure 4.22. S2-41's focus on the gap between the actual sale and demand for meals in week 23.

According to the verbal protocol, S2-41 concluded that a price between NOK 142 and NOK 144 was the optimal price. Furthermore, he spent five weeks to test these price alternatives. The eye-tracking data show that while doing so, S2-41 had several horizontal shifts along the values of the contribution variable, as well as vertical shifts between the contribution variable and the area showing the difference between average demand and actual sales; see, for example, Figure 4.23.

Furthermore, the eye-tracking data show that S2-41 had vertical eye movements between price, demand, sales and contribution when making his decisions.



	A	B	C	D	E	F	G	H	I	J	K
1	**Bonanza A/S - driftsoversikt**										
2		mai				juni				juli	
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28
4	**Inndata**										
5	***Handlingsvariabler***										
6	pris pr. middag (kr)	120	126	131	133	140	142	170	144		
7	antall servitører pr. vakt	4	4	4	4	3	3	2	3		
8											
9	***Ikke-kontrollerbare variabler***										
10	forventet etterspørsel (måltid/dag)	655	601	560	544	495	480	320	465		
11											
12	**Utdata**										
13	***Inntekter***										
14	faktisk gj.sn. etterspørsel (måltid/dag)	657	603	562	545	496	482	287	465	0	0
15	gj.sn. antall måltider solgt (måltid/dag)	642	603	562	545	483	482	287	465	0	0
16	omsetning (kr/uke)	539 280	531 846	515 354	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0
17	***Variable kostnader***										
18	ingredienser (kr/uke)	345 364	324 384	302 328	293 183	259 830	259 292	154 392	250 147	0	0
19	servitører (kr/uke)	53 471	53 471	53 471	53 471	40 103	40 103	26 736	40 103	0	0
20	totale variable utgifter (kr/uke)	398 835	377 855	355 799	346 654	299 933	299 395	181 127	290 250	0	0
21	***Dekningsbidrag (kr/uke)***	140 445	153 991	159 555	160 741	173 407	179 713	160 403	178 470	0	0
22											

Figure 4.23. Comparisons of values of sales and demand over time for S2-41.

Summary

Performance: S2-41 is a rather high performer because he had a good understanding of the general development of the relationships among variables. He understood the objective of the task, to maximise the contribution, but he did not understand how to achieve that objective. S2-41 performed only simple calculations. He did not try to calculate contribution at different waiter capacities, nor did he try to extrapolate the demand for sales prices between NOK 144 and NOK 148. This explains why he did not find the optimal solution (NOK 146). S2-41's handling of the task was characterised by a systematic trial-and-error strategy where he tried to gradually improve his solution.

Use of presentation format: S2-41 used the tables, both the one in the spreadsheet model and the one on paper, when solving the task. He did not attempt to create a graph based on the tables.

Complementarity of data collection methods: The eye-tracking data supports the interpretation of the verbal protocols suggesting that S2-41 tried to understand the general development of the relationships among the variables in the spreadsheet model.

Comparison of high and low performer

Performance: A main difference in the performance of S2-39 and S2-41 is the understanding of the dependencies among the decision variables and the result variables. S2-39 spent most of his time on problem definition, whereas S2-41 put less effort into defining the problem before making his decisions. S2-39 developed a relevant spreadsheet model and performed marginal calculations, whereas S2-41 performed only a few simple calculations. Both S2-39 and S2-41 handled the task using mainly analytic processes, but S2-39 had a more systematic approach to the task than did S2-41.

Use of presentation format: Both S2-39 and S2-41 used the table when handling the task. Neither of the subjects attempted to create a graphical representation of the data in the table.

Complementarity of data collection methods: The eye-tracking data for both S2-39 and S2-41 support my interpretations of their verbal protocols.

4.5.6 High-complexity task, both presentation formats

S2-05 – High performer

Subject S2-05 is the highest performing subject in the high-complexity task with a contribution index value of 0.9815 (see Table 4.12). He received the combined format as shown in Figure 4.24.

Table 4.12. Result and control variables for S2-05.

	S2-05	*Group average*
Contribution index value	0.9815	0.8585
Problem attention	14.92	24.94
Problem definition	10.92	6.37
Problem solving	4.00	18.57
Need for cognition	-0.3889	0.0768[10]

Problem definition

Verbal protocol analysis

The verbal protocol reveals that S2-05 spent most of the time in the experiment defining the problem before he started to make decisions. He carefully studied the graph and table with historic prices and demand data that he had available on paper. These data are essential to find the optimal solution. Then S2-05 stated:

[10] Average NFC for all subjects in Study 2. The NFC measure is not dependent on presentation format.

> ‘I just need to understand these graphs (…). It seems to me that the optimal price is around NOK 150–151. That would provide me with a demand of 430 guests.’

The verbal protocol reveals that he found the capacity of the waiters from the ‘historic’ periods and that he made an estimate of the capacity for the kitchen assistants:

> ‘It seems that six waiters can handle 600 guests. That gives a ratio of 1:100, right (…). I believe that the kitchen assistants can handle 200 guests each.’

Based on the capacities he found, he reasoned that he would need five waiters and three kitchen assistants to handle a demand of 430. He entered a price of NOK 151, five waiters and three kitchen assistants into week 23 of the model.

S2-05 did not estimate the kitchen assistants’ capacity correctly (their real capacity is 250) and he therefore would only have needed two assistants in week 23. However, in the complex task, the historic data do not provide enough information for the subjects to accurately calculate the kitchen assistants’ capacity. For that, he would have had to use a ‘test week’.

Eye-movement analysis

The eye-tracking data support the interpretations of the verbal protocols. Figure 4.24 shows 10 minutes of scanpaths projected onto a still picture. The figure shows S2-05’s eye movements until he made the first decision.

Figure 4.24 seems ‘messy’, but the point is to illustrate that S2-05 had many horizontal shifts in order to understand the variables and the accompanying values. Furthermore, he had many vertical shifts, indicating that he tried to understand the capacity constraints and the causal relationships among the variables. Figure 4.24 shows that S2-05 mainly used the table format in the problem-definition phase. However, as illustrated in the quotation above, he mainly used the graphical presentation format while studying the historic price and demand data on paper.

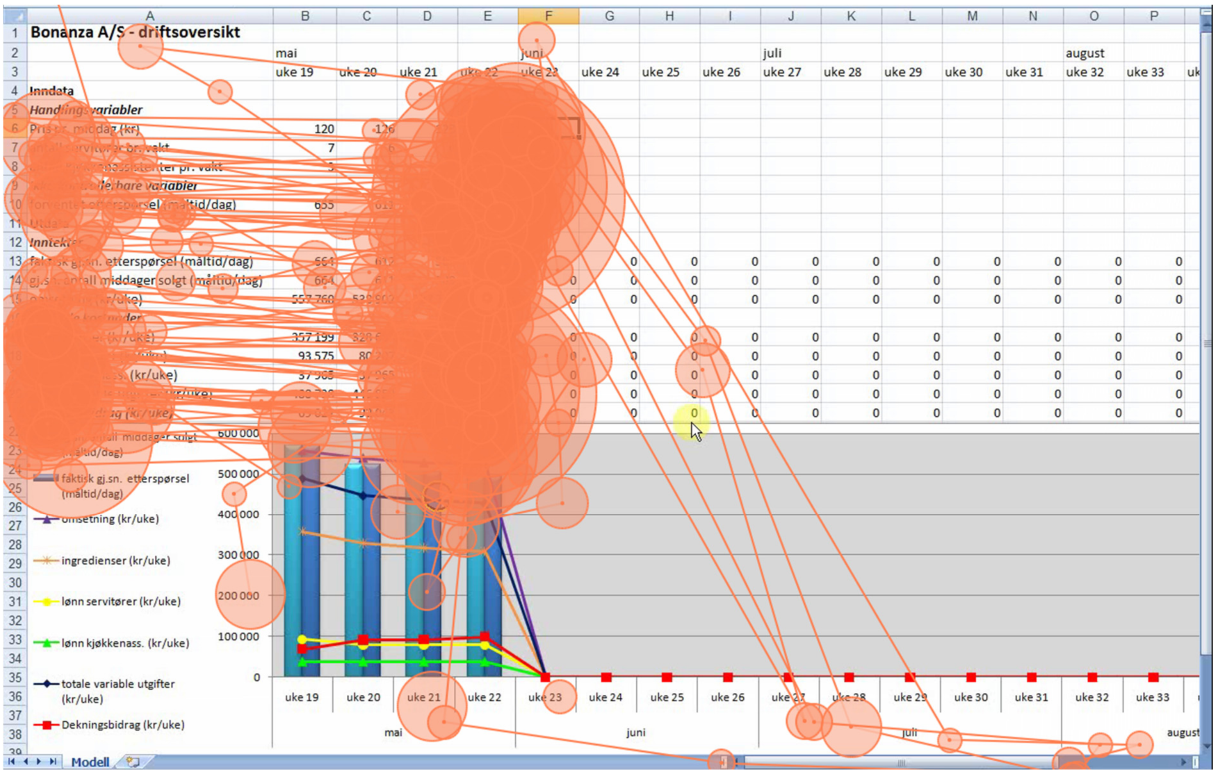


Figure 4.24. Eye movements of S2-05 during the problem-definition phase.

Problem solving

Verbal protocol analysis

While inspecting the result of week 23, the verbal protocol shows that S2-05 performed calculations on marginal costs and profits. For example, he found that reducing the number of kitchen assistants and waiters together with an increase in the price of meals would give a higher profit, that is, he understood the interdependencies among the variables:

> ‘It seems I have to hire this extra waiter [the fifth] only because of the last 20 meals. The question is whether I would benefit from firing the last waiter and rather increase the price (…). If so, I would only need two kitchen assistants.*’*

Based on the historic price–demand data, he reasoned that he would need to increase the price to NOK 155 in order to lower the demand so that four waiters could handle it. He then entered NOK 155, four waiters and two assistants into the input cells for week 24 in the model.

S2-05 actually found the optimal solution after the second trial (in week 24) but used the following three decisions to test whether he could improve the solution.

Eye-movement analysis

An inspection of the eye-tracking protocol in the problem-solving phase reveals that S2-05 used the graph to evaluate his solutions immediately after having entered decisions for a period into the model; see Figure 4.25a. He then used the table to perform a more accurate inspection of the results of the decision. For instance, Figure 4.25b shows a moment in time when he evaluated whether he had managed to cover the demand.

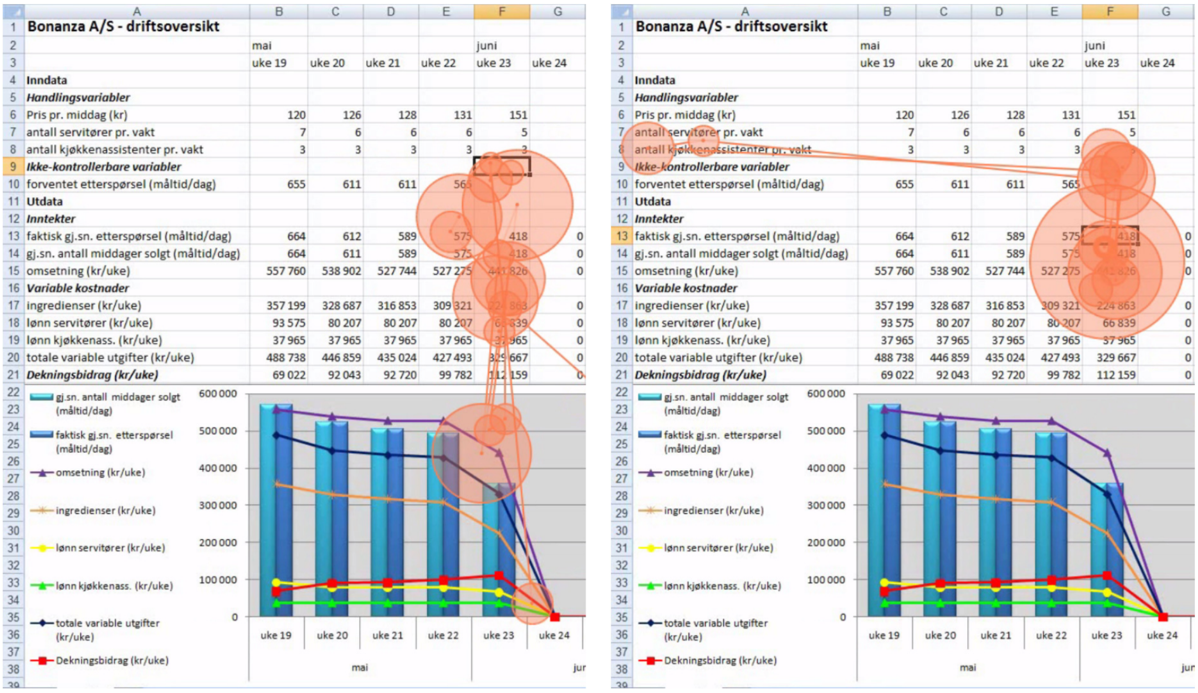


Figure 4.25. a) S2-05 used the graph to quickly get an overview of decisions and to evaluate them. b) S2-05 used the table to perform accurate comparisons, e.g, of demand and sales.

Even though S2-05 looked at both the graph and the table, most of the problem-solving time was spent looking at the table. This is not unexpected, as a graph gives a quick overview of trends whereas comparing numeric values and performing calculations require more time.

Summary

Performance: S2-05 attained a good result because he understood the dependencies among the decision variables and the output variables before he made his decisions. He found the optimal solution in the second week as a manager. The high-complexity task is constructed so that an extra week is needed to find the capacity of the kitchen assistants.

Use of presentation format: The eye-tracking data show that S2-05 mainly used the table part of the combined presentation format when he developed his understanding of the relationships among the variables; see Figure 4.24. However, the verbal protocol reveals that he mainly used the graph part of the data that he had received on paper when he studied the relationship between the sales prices and demand. In the problem-solving phase, S2-05 used the graph part of the presentation format to check whether the development in output variables was as expected, and he used the table part to check details.

Complementarity of data collection methods: The verbal protocol does not reveal how S2-05 integrated the values of the decision variables to arrive at his first decision. Thus, the eye-movement data were useful to document that S2-05 really made an effort to integrate the capacities with sales prices and demand to arrive at the optimal solution. The eye-tracking data shows many vertical eye-movements indicating that he attempted to understand the relationships among the variables during the problem-definition phase; see, for example, Figure 4.24. The data collection methods also complement each other, documenting how S2-05 used both presentation formats as described above.

S2-18 – Low performer

Subject S2-18 is the second lowest performing subject in Study 2 and the lowest performing subject in the treatment group that received the combined presentation format. As shown in Table 4.13, she spent only 0.67 minutes until she made her first decision but then 18.33 minutes in the problem-solving phase.

Table 4.13. Result and control variables for S2-18.

	S2-18	*Group average*
Contribution index value	0.6828	0.8585
Problem attention	19.00	24.94
Problem definition	0.67	6.37
Problem solving	18.33	18.57
Need for cognition	0.2222	0.0768[11]

Problem definition

Verbal protocol analysis

The verbal protocol shows that S2-18 did not spend much time trying to define the problem before she started making decisions. She asked about the number of weeks she was going to

[11] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

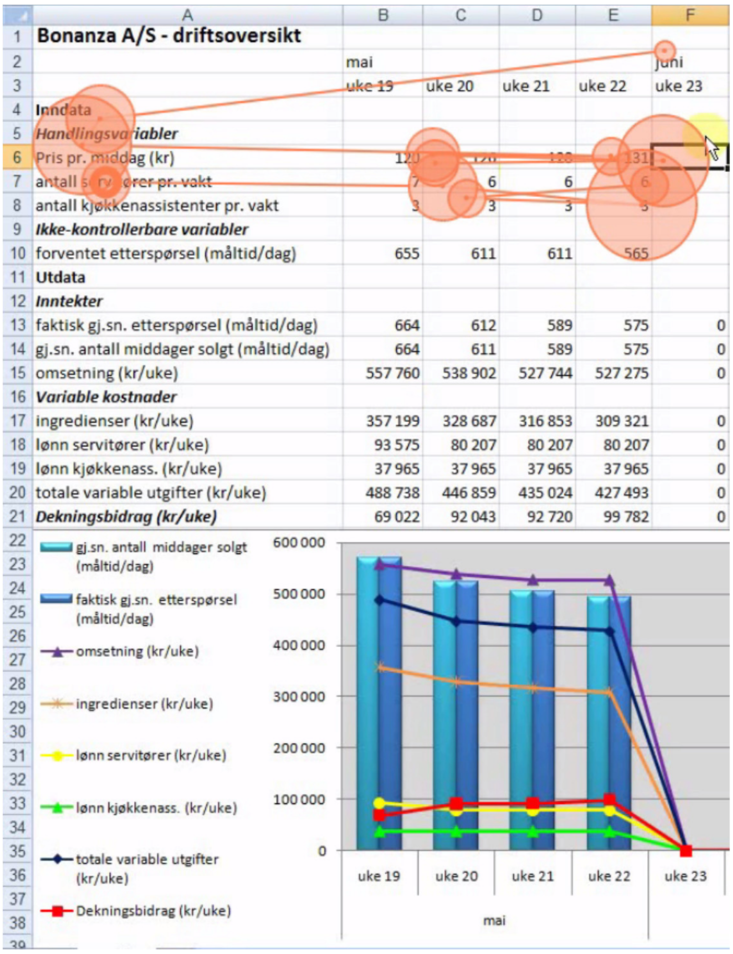
manage the restaurant and about what to do if she found a solution that she was satisfied with before she had reached her final week. She stated:

> 'Very well, I will try my way out. I am very good at using the trial and error method. Let me see.'

She then entered a price of NOK 128, seven waiters and three assistants into the first week of the model without any further explanation.

Eye-movement analysis

The eye-tracking protocol shows that S2-18 mostly focused on the input variables and the contribution variable. As shown Figure 4.26, she had horizontal eye-movements between the variable labels and their corresponding values. She also compared the values for each input variable over time. She did not pay attention to the area of the historic periods showing a gap between the demand for meals and the sale of meals.



	A	B	C	D	E	F
1	**Bonanza A/S - driftsoversikt**					
2		mai				juni
3		uke 19	uke 20	uke 21	uke 22	uke 23
4	**Inndata**					
5	***Handlingsvariabler***					
6	Pris pr. middag (kr)	120	[illegible]	[illegible]	131	
7	antall servitører pr. vakt	7	6	6	6	
8	antall kjøkkenassistenter pr. vakt	3	3	3	3	
9	***Ikke-kontrollerbare variabler***					
10	forventet etterspørsel (måltid/dag)	655	611	611	565	
11	**Utdata**					
12	***Inntekter***					
13	faktisk gj.sn. etterspørsel (måltid/dag)	664	612	589	575	0
14	gj.sn. antall middager solgt (måltid/dag)	664	611	589	575	0
15	omsetning (kr/uke)	557 760	538 902	527 744	527 275	0
16	***Variable kostnader***					
17	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	0
18	lønn servitører (kr/uke)	93 575	80 207	80 207	80 207	0
19	lønn kjøkkenass. (kr/uke)	37 965	37 965	37 965	37 965	0
20	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	0
21	***Dekningsbidrag (kr/uke)***	69 022	92 043	92 720	99 782	0

Figure 4.26. S2-18 had horizontal eye movements during problem identificiation.

Problem solving

Verbal protocol analysis

According to the verbal protocol, S2-18 concluded that her first decision had resulted in a lower contribution compared to the last week (week 22) of the historic period. She stated:

> 'I need to look at the ratio between the number of waiters and the number of kitchen assistants. In the beginning, he [the former manager] hired many employees, the prices were low and hence the capacity was higher and they earned a lot of money. However, I cannot increase the price and continue with that same capacity; the capacity will simply not be fully utilised. The demand will not be that high.'

In week 24, the verbal protocol shows that S2-18 decided to use the 'employee ratio' from week 22 and to increase the price to NOK 133. When comparing week 24 to week 22, she concluded that week 22 provided a higher contribution and thus that her price in week 24 was too high:

> 'Let's see. The contribution [in week 24] is 98 399 and the price was 133. Week 22, where the price is 131, is the optimal so far [contribution of 99 782]. I will try a price in between [i.e. 132].'

The above statement illustrates that S2-18 compared input data to the output data presented as a table. However, the verbal protocol reveals that she also used the graphs sometimes:

> 'The purple one is slowly declining ... But I'm not going to look at sales, right? (…) I'm just wondering where… Oh, there it is. The red one. Yes, I just did not see where the red one was.'

Figure 4.27 shows S2-18's decisions. According to the figure, S2-18 used a price of NOK 132 in 12 of the following 15 weeks. The verbal protocol reveals that she had many silent periods during the problem solving and therefore did not explain very well what she was thinking in the different weeks. She used a price of NOK 131 in weeks 32–34 and reasoned:

> 'I think this week [no. 32] is good, but maybe not optimal compared to former weeks. But that might not necessarily mean anything (…) Let's see, maybe there is a trend … High, high, low ... It's hard to choose … The demand has increased … But I haven't had as high as there [week 29]. I think 132 is smoother.'

S2-18 settled on a price of NOK 132 in week 35 and used that same price for the remaining weeks of the experiment. Therefore, the last four weeks are not shown in Figure 4.27.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	**Bonanza A/S - driftsoversikt**																	
2		mai				juni				juli					august			
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32	uke 33	uke 34	uke 35
4	**Inndata**																	
5	***Handlingsvariabler***																	
6	Pris pr. middag (kr)	120	126	128	131	128	133	132	132	132	132	132	132	132	131	131	131	132
7	antall servitører pr. vakt	7	6	6	6	7	6	6	5	6	6	6	6	6	6	6	6	6
8	antall kjøkkenassistenter pr. vakt	3	3	3	3	3	3	3	3	4	3	3	3	3	3	3	3	3
9	***Ikke-kontrollerbare variabler***																	
10	forventet etterspørsel (måltid/dag)	655	611	611	565													
11	**Utdata**																	
12	***Inntekter***																	
13	faktisk gj.sn. etterspørsel (måltid/dag)	664	612	589	575	584	551	566	558	562	565	566	558	553	570	575	564	554
14	gj.sn. antall middager solgt (måltid/dag)	664	611	589	575	584	551	566	505	562	565	566	558	553	570	575	564	554
15	omsetning (kr/uke)	557 760	538 902	527 744	527 275	523 264	512 981	522 984	466 620	519 288	522 060	522 984	515 592	510 972	522 690	527 275	517 188	511 896
16	***Variable kostnader***																	
17	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	314 163	296 410	304 480	271 665	302 328	303 942	304 480	300 176	297 486	306 632	309 321	303 404	298 024
18	lønn servitører (kr/uke)	93 575	80 207	80 207	80 207	93 575	80 207	80 207	66 839	80 207	80 207	80 207	80 207	80 207	80 207	80 207	80 207	80 207
19	lønn kjøkkenass. (kr/uke)	37 965	37 965	37 965	37 965	37 965	37 965	37 965	37 965	50 619	37 965	37 965	37 965	37 965	37 965	37 965	37 965	37 965
20	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	445 702	414 582	422 651	376 468	433 154	422 113	422 651	418 348	415 658	424 803	427 493	421 575	416 196
21	***Dekningsbidrag (kr/uke)***	69 022	92 043	92 720	99 782	77 562	98 399	100 333	90 152	86 134	99 947	100 333	97 244	95 314	97 887	99 782	95 613	95 700

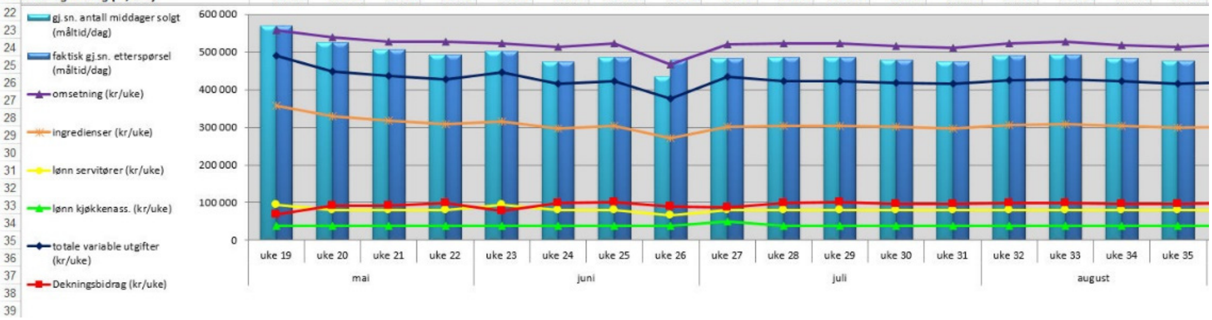


Figure 4.27. S2-18's decisions.

Figure 4.27 also shows that S2-18 reduced the number of waiters from six to five in week 26 and that she increased the number of kitchen assistants from three to four in week 27. The verbal protocol reveals no justifications for her decisions regarding the number of employees in these weeks:

> 'No, at least that [week 26] was not the optimal solution. The only thing I can try then is to do the opposite in week 27.'

She then increased the number of kitchen assistants from three to four and reasoned:

> 'No, that did not work either. I presume that the former [i.e. six waiters and three kitchen assistants in week 25] is the optimal solution.'

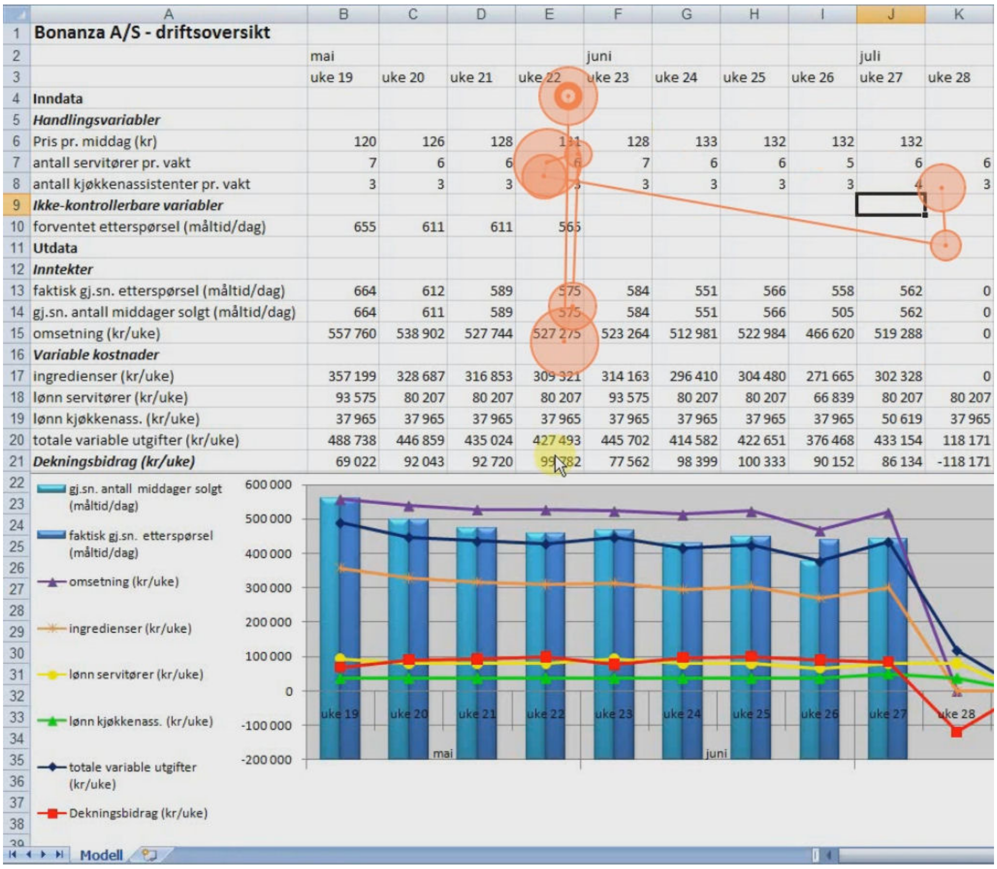
The protocol shows that she did not know the capacities of the employees while making her decisions:

> 'I don't know how much available capacity I have left.'

According to the verbal protocol, S2-18 performed only a few simple calculations using a calculator, pen and paper. For instance, she multiplied the number of employees by the salary cost per employee in week 26. She did not try to calculate the capacity of either the waiters or the kitchen assistants.

Eye-movement analysis

The eye-tracking data reveal that S2-18 focused mainly on the input variables, the variable costs and the contribution during the problem-solving phase. She looked at the demand and sales a few times, for example in week 28 when she looked back at week 22 (see Figure 4.28), but she did not use the data to find estimates of the capacity of the employees.



	A	B	C	D	E	F	G	H	I	J	K
1	**Bonanza A/S - driftsoversikt**										
2		mai				juni				juli	
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28
4	**Inndata**										
5	***Handlingsvariabler***										
6	Pris pr. middag (kr)	120	126	128	1[illegible]1	128	133	132	132	132	
7	antall servitører pr. vakt	7	6	6	6	7	6	6	5	6	6
8	antall kjøkkenassistenter pr. vakt	3	3	3	[illegible]	3	3	3	3	4	3
9	***Ikke-kontrollerbare variabler***										
10	forventet etterspørsel (måltid/dag)	655	611	611	565						
11	**Utdata**										
12	***Inntekter***										
13	faktisk gj.sn. etterspørsel (måltid/dag)	664	612	589	575	584	551	566	558	562	0
14	gj.sn. antall middager solgt (måltid/dag)	664	611	589	575	584	551	566	505	562	0
15	omsetning (kr/uke)	557 760	538 902	527 744	527 275	523 264	512 981	522 984	466 620	519 288	0
16	***Variable kostnader***										
17	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	314 163	296 410	304 480	271 665	302 328	0
18	lønn servitører (kr/uke)	93 575	80 207	80 207	80 207	93 575	80 207	80 207	66 839	80 207	80 207
19	lønn kjøkkenass. (kr/uke)	37 965	37 965	37 965	37 965	37 965	37 965	37 965	37 965	50 619	37 965
20	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	445 702	414 582	422 651	376 468	433 154	118 171
21	***Dekningsbidrag (kr/uke)***	69 022	92 043	92 720	99 782	77 562	98 399	100 333	90 152	86 134	-118 171

Figure 4.28. S2-18's focus on number of employees, demand and sales of meals.

In general, the eye-movement data show that S2-18 had few vertical shifts in her gaze pattern. However, she had many horizontal eye movements where she sequentially looked at data values for given variables in the table format; see Figure 4.29.

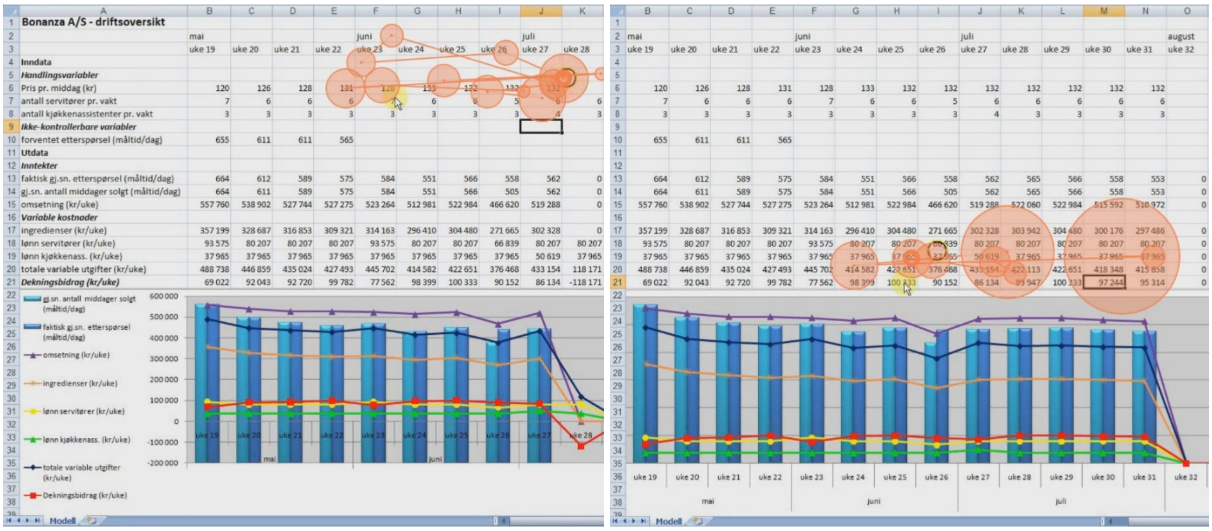
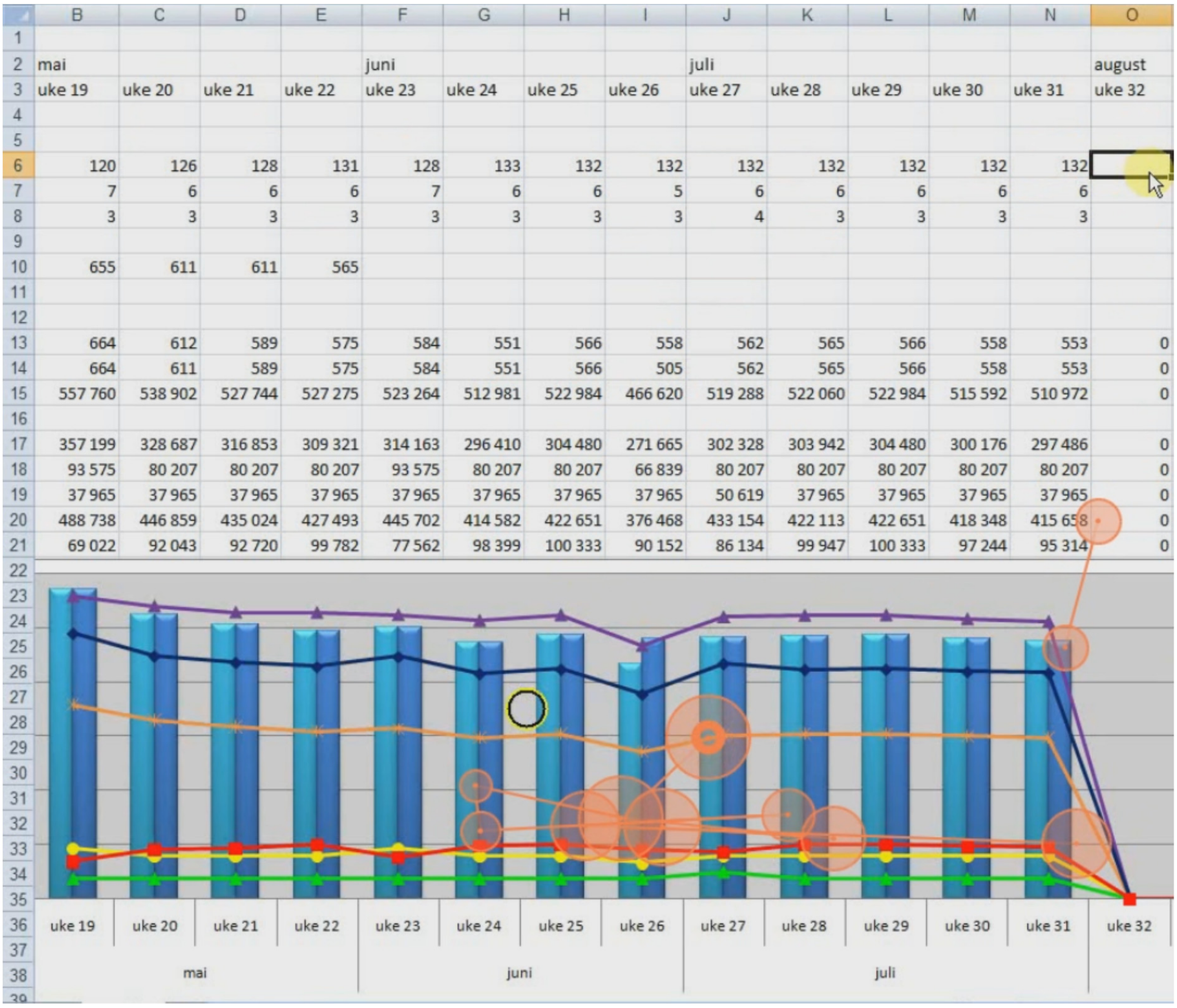


Figure 4.29. a) S2-48's horizontal eye movements on the input data. b) S2-18's horizontal eye-movements on the output data.

The eye-tracking protocol shows that S2-18 used the graphical presentation format in addition to the table format. For instance, Figure 4.30 shows the scanpath 18 seconds after the scanpath in Figure 4.29b, when she compared the result of week 31 against previous periods. Subsequent parts of the scanpath video show that she also used the table to compare numerical values for the contribution for the same periods. According to the verbal protocol, S2-18 was silent when she looked at the graph shown in Figure 4.30.



	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1														
2	mai				juni				juli					august
3	uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32
4														
5														
6	120	126	128	131	128	133	132	132	132	132	132	132	132	
7	7	6	6	6	7	6	6	5	6	6	6	6	6	
8	3	3	3	3	3	3	3	3	4	3	3	3	3	
9														
10	655	611	611	565										
11														
12														
13	664	612	589	575	584	551	566	558	562	565	566	558	553	0
14	664	611	589	575	584	551	566	505	562	565	566	558	553	0
15	557 760	538 902	527 744	527 275	523 264	512 981	522 984	466 620	519 288	522 060	522 984	515 592	510 972	0
16														
17	357 199	328 687	316 853	309 321	314 163	296 410	304 480	271 665	302 328	303 942	304 480	300 176	297 486	0
18	93 575	80 207	80 207	80 207	93 575	80 207	80 207	66 839	80 207	80 207	80 207	80 207	80 207	0
19	37 965	37 965	37 965	37 965	37 965	37 965	37 965	37 965	50 619	37 965	37 965	37 965	37 965	0
20	488 738	446 859	435 024	427 493	445 702	414 582	422 651	376 468	433 154	422 113	422 651	418 348	415 658	0
21	69 022	92 043	92 720	99 782	77 562	98 399	100 333	90 152	86 134	99 947	100 333	97 244	95 314	0

Figure 4.30. S2-18 used both the table and the graphical presentation format.

Summary

Performance: S2-18 was a low performer because she did not attempt to understand the dependencies between the variables. She did not use the historic price–demand data. She used a trial-and-error strategy where she attempted to improve the contribution gradually by comparing it with previous results. She made few calculations and never explored the possibilities of increasing the contribution considerably. Her information processes were mainly spatial.

Use of presentation format: The eye-tracking data show that S-18 used both the table part and the graph part of the combined presentation format. However, there is no systematic pattern in how she used the two presentation formats interchangeably.

Complementarity of data collection methods: The eye-tracking data support the impression from the verbal protocol that S2-18 did not attempt to understand the relationships among the input and the output variables. S2-18 had few vertical eye movements among the variables. The eye-tracking data show mainly horizontal eye movements between either input or output variables when S2-18 compared values with previous decisions; see, for example, Figure 4.29 a) and Figure 4.29 b).

Comparison of high and low performer

Performance: The main difference in the performance between S2-05 and S2-18 is related to the understanding of the dependencies among the decision variables and the result variable. S2-05 developed an understanding of the relationships among the variables before he made his first decision, and he only needed one additional week to find the optimal solution. S2-18 relied on a trial-and-error strategy. Even though she attempted to improve her result by comparison with previous values of the decision variables, her trials were characterised by 'anchoring and insufficient adjustment' (Tversky & Kahneman, 1974). S2-05 used analytic processes with marginal considerations related to the number of waiters and kitchen assistants. S2-18 used mainly spatial processes.

Use of presentation format: S2-05 and S2-18 used both the table part and the graph part of the combined presentation format. S2-05 used the XY graph presenting historic prices and demand to find a relevant starting point in the problem-definition phase. However, the eye-tracking data show that S2-05 mainly used the table part of the combined presentation format when he developed his understanding of the relationships among the variables; see Figure 4.24. In the problem-solving phase, S2-05 used the graph part of the presentation format to check whether his decisions gave the expected results. S2-18 used the graph part of the presentation format when she compared the contribution for various weeks, but she also used the table.

Complementarity of data collection methods: The verbal protocols for both S2-05 and S2-18 did not reveal the rather large differences in S2-05's and S2-18's handling of the task. The eye-movement data documented the effort that S2-05 made to understand the dependencies among the variables and how he used the presentation formats. The eye-movement data for S2-18 supports the interpretation of the verbal protocol that she did not make an effort to understand the relationships among the input and the output variables.

4.5.7 High-complexity task, graphical presentation format

S2-13 – High performer

I do not have satisfactory eye-tracking data for the best performing subject in this group, S2-16, as her wearing of mascara resulted in poor eye-movement data quality. Subject S2-13 is the second best performing subject for the high-complexity task with the graphical presentation format. He scored 0.8695 on the contribution index and spent 62.22 minutes to complete the task (see Table 4.14). The group as a whole scored 0.8215 on average and spent 50.83 to solve the task. The eye-tracking data for S2-13 are satisfactory until halfway into the problem-solving phase. After this point in time, the eye-tracking data is missing a lot of the time. However, the screen video was still helpful for understanding the behaviour of S2-13.

Table 4.14. Result and control variables for S2-13.

	S2-13	*Group average*
Contribution index value	0.8695	0.8215
Problem attention	62.22	50.83
Problem definition	10.20	10.71
Problem solving	52.02	40.12
Need for cognition	0.5556	0.0768[12]

Problem definition

Verbal protocol analysis

The verbal protocol shows that S2-13 asked some clarifying questions about the employees before he started making a calculation model in the spreadsheet based on the information in the task description. The protocol shows that he had silent periods while he worked on his model and that I reminded him to think out loud. S2-13 stated:

> 'I am trying to create an overview of the different costs, and I will also see if I can figure out what demand we can handle per day.'

S2-13 expressed a need for numeric values:

[12] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

> 'I should really like to have some numeric values in addition to this graph … It is hard to figure out what price I shall use and thus also what the demand will be.'

The protocol shows that S2-13 decided to try some weeks in the model. He entered a price of NOK 140, six waiters and three kitchen assistants into the model.

Eye-movement analysis

The eye-tracking data shows that S2-13 had horizontal eye movements relating variable labels to their corresponding values, both in the input data and in the graph showing the output data. He also had some vertical eye movements between the price and the contribution. The eye-tracking data shows that S2-13 focused on reading out accurate values from the graph by hovering over data points in the graph with the mouse cursor; see Figure 4.31.

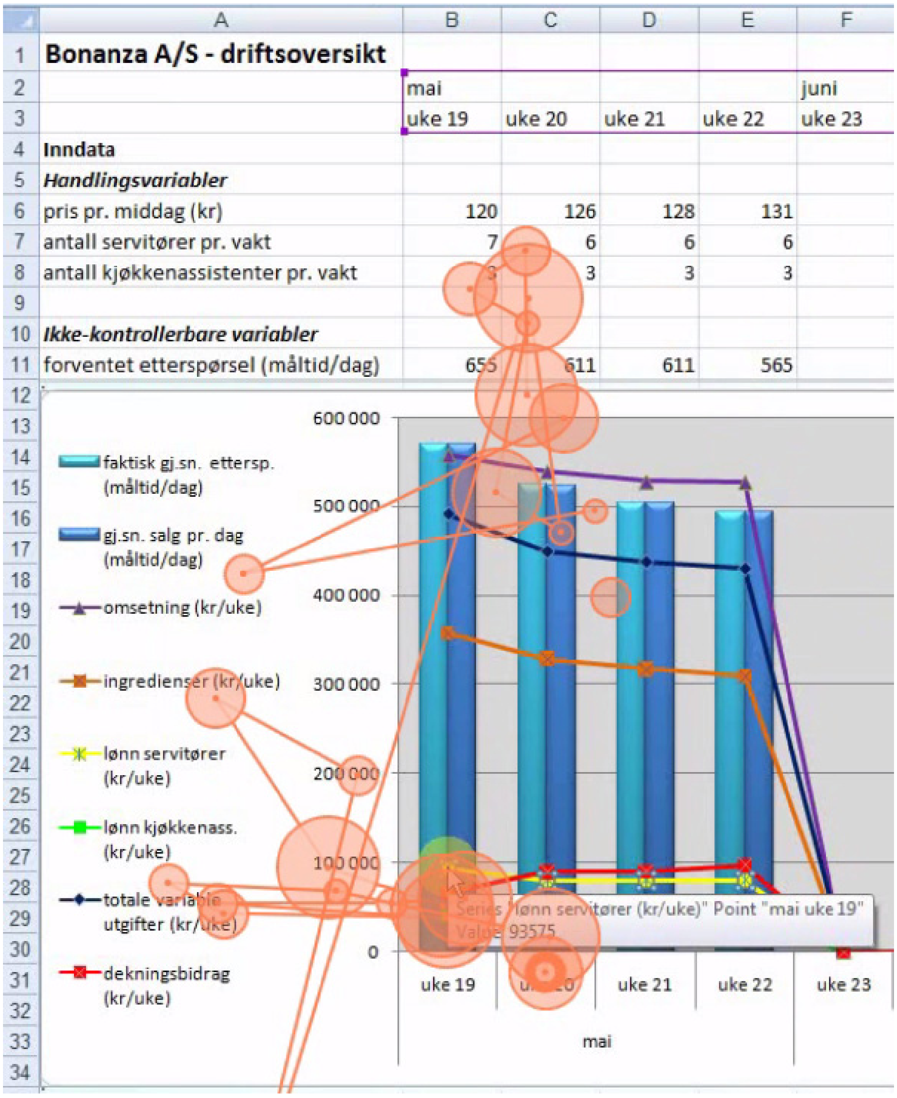


Figure 4.31. S2-13 focused on reading out accurate values from the graph.

Problem solving

Verbal protocol analysis

Based on his first decision, S2-13 noted that the sales dropped. Therefore, he decided to try another week, namely week 24, with a price of NOK 140, four waiters and two kitchen assistants. This implies that he reduced the number of waiters by two and the number of kitchen assistants by one. He concluded:

> ‘We didn’t manage to handle the demand, but the contribution increased.’

In week 25, he continued with a price of NOK 140, five waiters and three kitchen assistants. In week 26, he increased the price to NOK 145 and kept the number of employees unchanged. He then reasoned:

> ‘No, now the contribution dropped. OK. Then I’ll have to make some calculations on what I have got.’

The verbal protocol shows that S2-13 continued building his calculation model in Excel and that he had silent periods while doing so. He used the paper-based XY graph to read out combinations of price and demand from the historic data. He expressed that he needed to know how many guests each waiter and kitchen assistant could handle. He continued:

> ‘Let’s see. In week 24, I reduced the number of employees … Four waiters and two kitchen assistants. Expected demand was 500, but I only managed to handle 405. So I assume that a waiter can handle 100 guests and each kitchen assistant can handle 200 guests.’

Note that the estimated capacity for waiters is correct but the capacity for the kitchen assistants is underestimated. The ‘real’ capacity for kitchen assistants is 250.

S2-13 continued by entering different combinations of prices and demands, read from the XY graph, into his calculation model. He adjusted the number of employees according to his capacity estimates, but the verbal protocol does not reveal how these adjustments were made. Based on his trials in his model, he reasoned that he could not find a better price than the one he had found in week 26. However, he concluded that he could reduce the number of kitchen assistants by one. Thus, in week 27 he entered a price of NOK 145, five waiters and two kitchen assistants into the model. He noticed that the contribution increased. He then continued trying out other combinations in his calculation model:

> 'It seems like a price of NOK 165 and a demand of 340, which is the lowest demand in the historic data, provides an even higher contribution. Therefore, I'll try a price of NOK 165, three waiters and two kitchen assistants.'

He entered the values into week 28 in the spreadsheet model but concluded that his expectations were not met:

> 'No, that did not work. The contribution dropped.'

S2-13 continued by trying to express the demand as a function of the price, and he used the solver function in Excel to try to optimise the contribution. He realised that his calculations were not correct, but he concluded that a price of about NOK 149 would be the optimal price. For this price, the calculation model suggested five waiters and two kitchen assistants. S2-13 entered the values into week 29 in the model. He then used the next two weeks to test whether he could improve the solution. In week 30, he used a price of NOK 150 and in week 31, he used a price of NOK 148. He then continued with a price of NOK 149, five waiters and two kitchen assistants in the remaining eight weeks.

Eye-movement analysis

The eye-tracking data shows that S2-13 focused on the contribution and whether he managed to handle the demand when he made his decisions in weeks 23–26. He had vertical eye movements between price and contribution and between the number of waiters/kitchen assistants and demand/sales. He also had horizontal eye movements where he related the variable labels to their corresponding values. After having made a decision, he read out values for the contribution for the different weeks. He also read out accurate values from the graph when he calculated the capacities for the waiters and the kitchen assistants; see Figure 4.32.

The verbal protocol did not show how S2-13's calculation model adjusted the number of employees according to the demand. However, the scanpath video from the eye-tracking recordings reveals how the capacities were incorporated. S2-13 created formulas in Excel to calculate how many waiters and kitchen assistants he needed, based on the demand. He used the round-up formula to round up the number that he got from dividing the expected demand, that is, the number in cell B12 in Figure 4.33, by the capacity for waiters and the capacity for kitchen assistants; see cells B16 and B17 in Figure 4.33. He used one decimal digit in the formulas, that is, 430 divided by 100 would be rounded up to 4.3, but when he formatted the cells, he chose to show whole numbers only. Therefore, the value of 4.3 was shown as the number 4 in the spreadsheet (i.e. in cell B16). This explains why the calculated contribution did not match the contribution in the graph in the spreadsheet model.

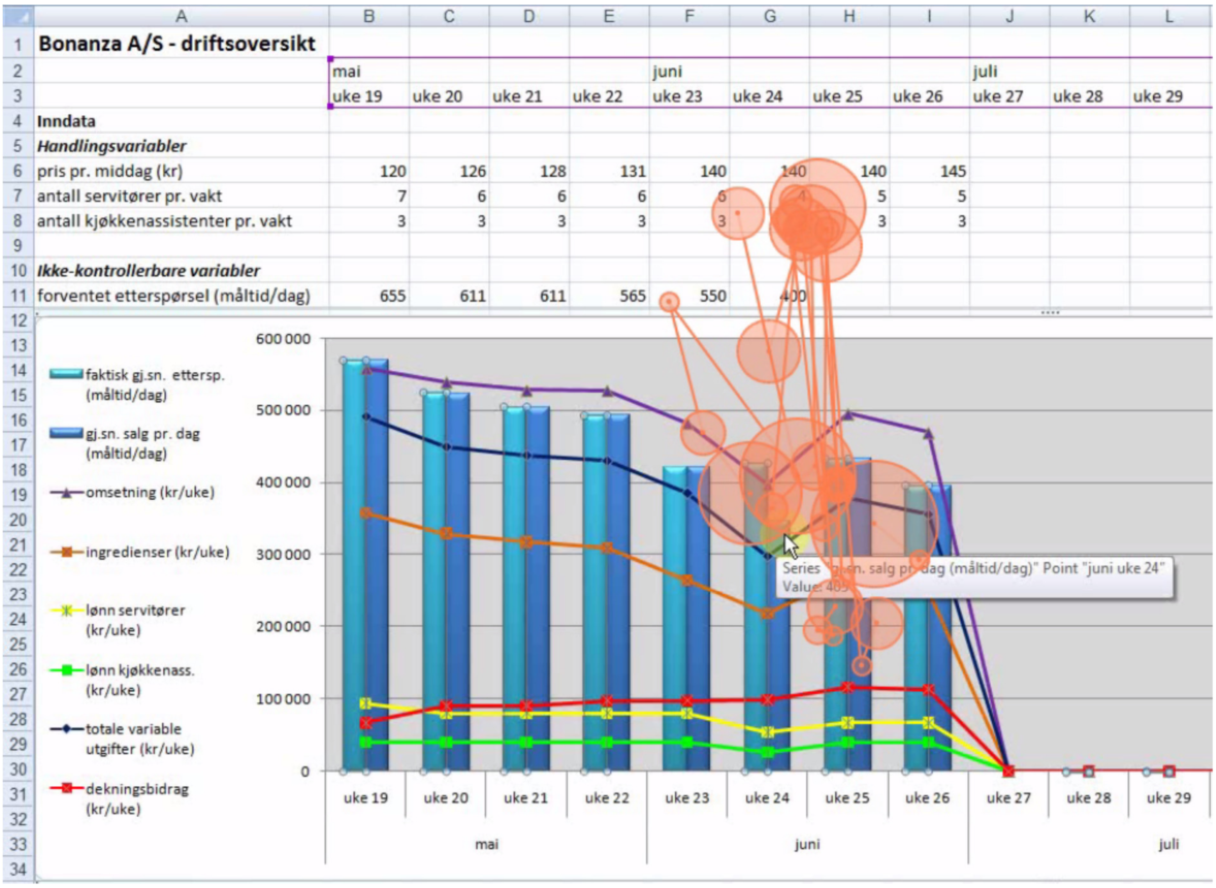


Figure 4.32. The moment in time when S2-13 calculated the capacity for the waiters and the kitchen assistants.

The eye-tracking data shows that S2-13 continued to focus on the contribution throughout the whole experimental session. He also focused on reading out and comparing numerical values from the graph.

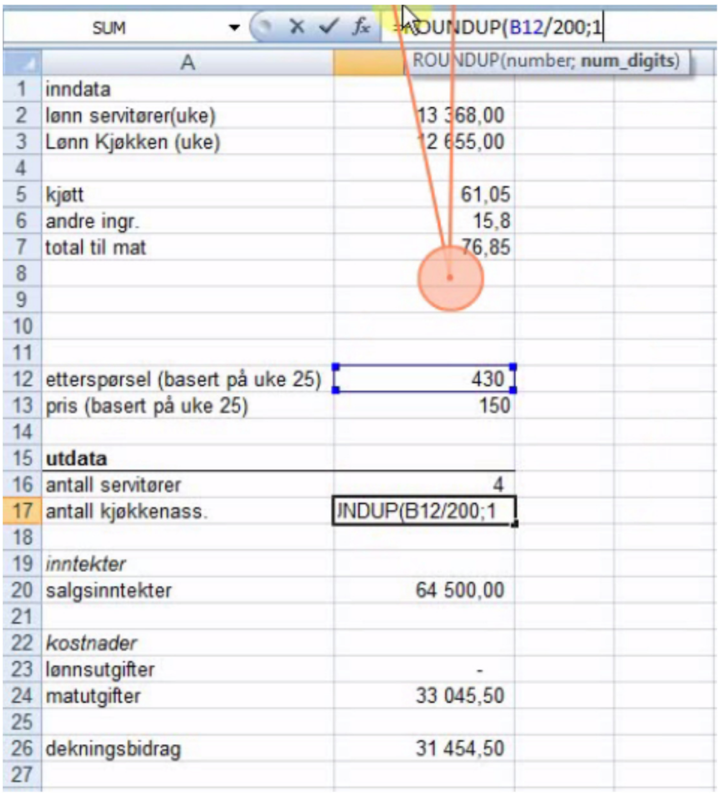


Figure 4.33. S2-13's calculations were based on hidden and incorrect assumptions.

Summary

Performance: S2-13 performed well because he had a good understanding of the dependencies among the decision variables and the output variables. He started out by using a systematic trial-and-error strategy but continued to define the problem more thoroughly after the fourth week. There are two reasons why S2-13 did not find the optimal solution: 1) He used an incorrect estimate of the kitchen assistants' capacity. 2) He calculated salary costs based on decimal numbers of employees, even though his model indicates that he used whole numbers of employees. Thus, S2-13's model was based on incorrect and hidden assumptions, which caused the calculated contribution for a given combination of price, number of waiters and number of kitchen assistants to differ from the 'real' contribution shown in the graph when using the same combination of decision variables.

Use of presentation format: S2-13 was given the graphical presentation format. However, after the fourth week in the experiment, he used the graphical presentation format to create a data table representing the dependencies among the variables. The table shows only one combination of demand and price at a time, and S2-13 had to replace these values when he evaluated new price and demand combinations. Therefore, it was harder for him to compare different price and demand solutions than if he had represented multiple combinations at a time.

Complementarity of data collection methods: S2-13 had several silent periods, and the eye-tracking data were useful supplements to the verbal protocols. For example, the eye-tracking data show that he fixated on numeric values that Excel provided in order to have accurate data. He hovered over different parts of the graph with the mouse cursor in order to have Excel reveal screen tips with accurate data. The eye-tracking data also confirm the impression from the verbal protocol that he focused on the contribution when evaluating the results of his decisions. Finally, the screen video uncovered shortcomings in the assumptions underlying the calculation model, causing the model to calculate the wrong result. I would not have noticed this error without the scanpath video.

S2-15 – Low performer

Subject S2-15 is an example of a low-performing subject. As Table 4.15 shows, he attained a contribution index value of 0.6721, which is well below the average. Furthermore, the table reveals that he spent considerably less time (8.12 minutes) working on the task compared to the other subjects in the treatment group (50.83 minutes on average).

Table 4.15. Result and control variables for S2-15.

	S2-15	*Group average*
Contribution index value	0.6721	0.8215
Problem attention	8.12	50.83
Problem definition	4.93	10.71
Problem solving	3.18	40.12
Need for cognition	-0.2222	0.0768[13]

Problem definition

Verbal protocol analysis

Before making his first decision, S2-15 studied the spreadsheet model and asked questions about how to know whether he employed enough waiters. I explained to him that the blue columns in the graph show demand and sales per day. He stated:

> 'OK, yes, I see that the demand per day is higher than the sales per day. They have not been able to sell to all. Fine, then that is what I will have to look at then.'

[13] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

He then entered a price of NOK 130, six waiters and three kitchen assistants into the model with no further explanation. The verbal protocol reveals that he had long silent periods, both during problem definition and problem solving, even though I encouraged him to verbalise his thoughts.

Eye-movement analysis

Figure 4.34 shows a scanpath diagram for 10 seconds of S2-15's eye fixations at the beginning of the session, that is, during the problem-definition phase.

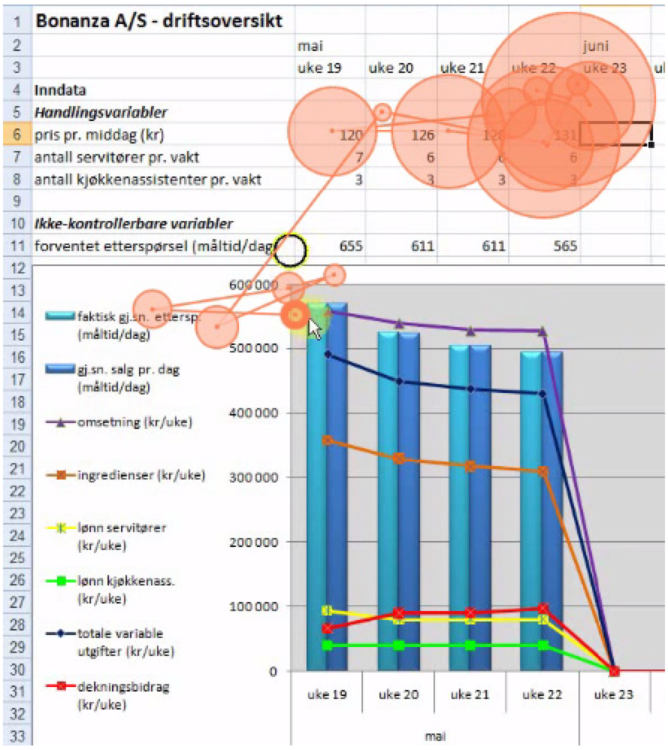


Figure 4.34. S2-15 had mostly horizontal eye movements during problem definition.

The figure illustrates that S2-15 mainly had horizontal eye movements in the problem-definition phase and that he paid particular attention to the prices of the meals (large circles).

Problem solving

Verbal protocol analysis

According to the verbal protocol, S2-15 was silent after having made his decision in week 23. He did not comment on the result but instead continued with the next three weeks. He used three kitchen assistants in the weeks he managed the restaurant. In week 24, he proceeded with the same price (NOK 130) as in week 23, but he reduced the number of waiters to five. Then, in weeks 25 and 26, he returned to six waiters and increased the price to NOK 131 and NOK 132, respectively. I reminded him to think aloud, and he then stated:

> 'I've tried to reduce the price and continued with the same number of employees [in week 23]. By doing so, I handled the demand, but the contribution dropped. Then I tried another week [24] with the same price but with one less waiter, but then I did not manage to handle the demand. I increased the price and the capacity back [referring to week 23], and I got a satisfying result; I handled the demand, and the contribution increased. Finally, I tried the same number of employees, but I increased the price to see if it had any effect, and as you can see, the contribution dropped.'

In week 27, S2-15 reduced the price to NOK 131 and concluded:

> 'This is a reasonable solution that I think I will continue to use. However, there is not much economic theory behind my decision making … there is a lot of trial and error. But I think it's a good solution.'

He then copied his decision for week 27 for the remaining 12 weeks.

The above transcript reveals that S2-15 was aware that a capacity constraint existed, at least for the waiters. Furthermore, the transcript shows that he focused on the contribution and whether he managed to handle the demand. My observational notes and the verbal protocol show, however, that he did not perform any calculations either in the problem-solving phase or during the problem-definition phase.

Eye-movement analysis

During the problem-solving phase, the eye-tracking data show that S2-15 shifted his visual attention between the input data and the area of the graph showing the daily sales and demand for meals (see Figure 4.35). The eye-tracking protocol also reveals that S2-15 did not relate the graph elements to the axes of the graph. When he made his decisions for a week, he only took one quick look at the line showing the contribution for that specific week. He did not try to 'read' a numeric figure for the contribution. In fact, the screen video reveals that S2-15 was very fast in his decision making. He only spent about 20 seconds from the time he made his decision in week 23 until he finished his decisions in week 26.

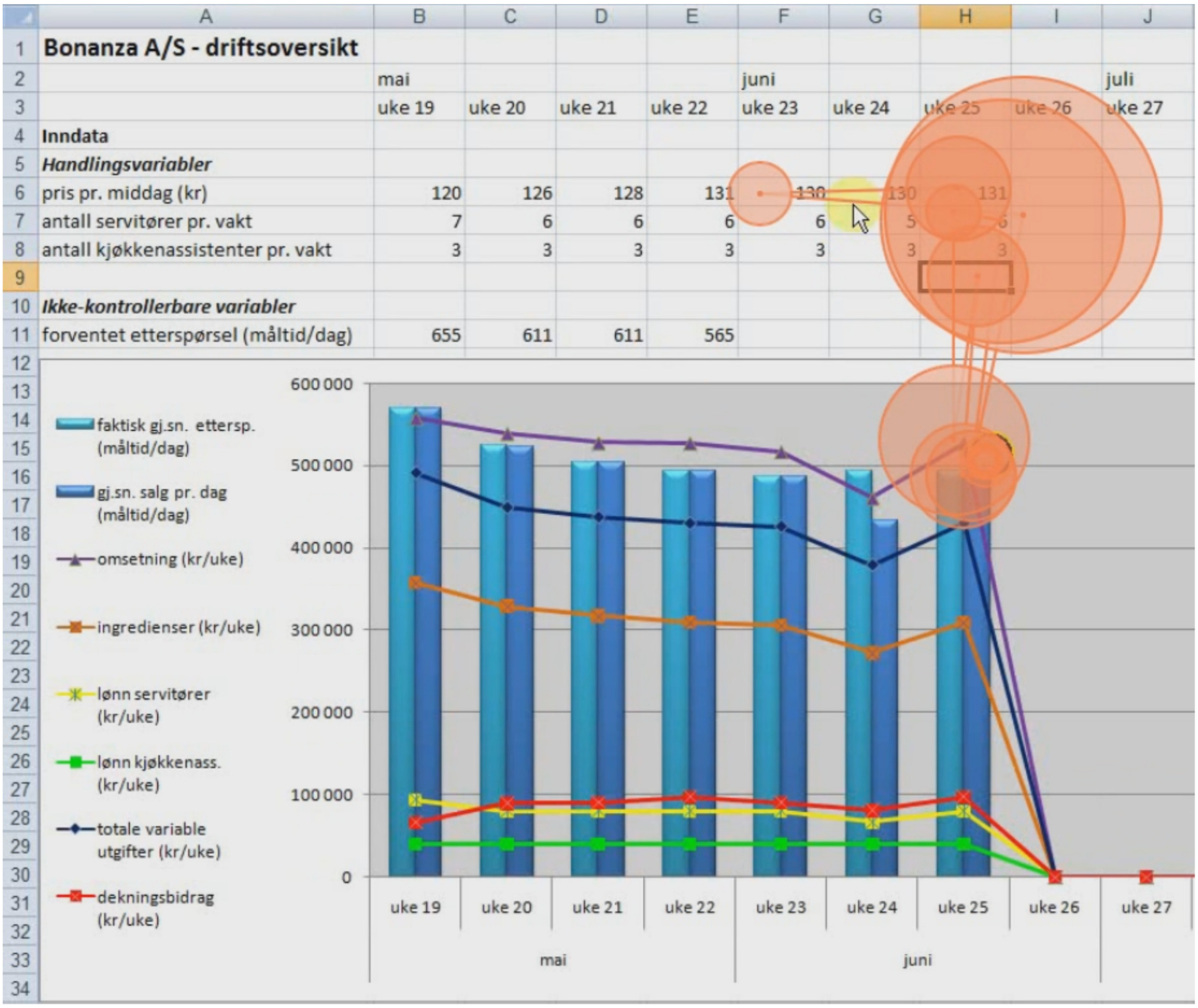


Figure 4.35. Vertical eye movements of S2-15.

Summary

Performance: S2-15 was a low performer because he did not attempt to understand the dependencies between the variables in the model. He did not utilise the XY graph showing the price–demand relationship data, and he used a very limited price range. S2-15 handled the task using trial and error and mainly spatial processes. He expressed that he wanted to utilise the waiters' capacity, but he never tried to calculate this capacity.

Use of presentation format: S2-15 used the graph in the spreadsheet model when handling the task. He did not use the paper graphs showing the relationship between price and demand. He did not attempt to read out values from the graph, and he did not make any additional tables.

Complementarity of data collection methods: S2-15 had silent periods where the eye tracking complements the verbal protocol. For example, the eye-movement data supports the interpretation of the verbal protocol that he did not try to understand the relationships among variables. He had very few vertical shifts between input and output data.

Comparison of high and low performer

Performance: The main difference in the performance between S2-13 and S2-15 is related to their understanding of the dependencies among decision variables and the result variables. S2-13 started out using a systematic trial-and-error strategy, but after the fourth period he developed an understanding of the relationships among the variables before he made his next decisions. He developed a relevant calculation model even though he did not manage to represent the relationships between the employee's capacities and the demand correctly. S2-15 did not attempt to understand the dependencies among the variables in the model; instead, he tried to improve his result by a comparison with previous periods. As seen for other subjecs, such as S2-18, his trials were characterised by 'ancoring and insufficient adjustment' (Tversky & Kahneman, 1974). S2-13 mainly used analytic processes, whereas S2-15 hadled the task mainly using spatial processes.

Use of presentation format: Both S2-13 and S2-15 were given the graphical presentation format. However, the way they used the data in the presentation format was fundamentally different. S2-13 focused on reading out accurate numeric values from the graph so that he could perform accurate numeric comparisons and calculations, and he developed a tabular representation of the relationships among the variables. He read out accurate values from both the XY graph showing historic prices-demand data, and the data in the spreadsheet model when making his decisions. S2-15, on the other hand, did not use the XY graph, and he only performed ordinal comparisons based on the graph in the spreadsheet model.

Complementarity of the data collection methods: Both S2-13 and S2-15 were quite silent, and in these periods the eye-tracking data were useful for understanding what they were doing. The eye-tracking data for S2-13 reveals his focus on accuracy, and the scanpath video shows how he erroneously calculated the salary costs for his employees. For S2-15, the lack of vertical shifts between decision variables and results variables in the eye-movement data supports my interpretation of the verbal protocols that he did not try to understand the relationships among variables.

4.5.8 High-complexity task, tabular presentation format

S2-14 – High performer

S2-14 is the highest performing subject in the group with the high-complexity task and the tabular presentation format. He was the fastest subject to complete the task within his treatment group, and, as Table 4.16 shows, he spent only 15.40 minutes working on the task, whereas the average time for the group was 33.71 minutes.

Table 4.16. Result and control variables for S2-14.

	S2-14	*Group average*
Contribution index value	0. 9766	0.8679
Problem attention	15.40	33.71
Problem definition	3.13 (8,75)[14]	15.67
Problem solving	12.27 (6,65)[15]	18.04
Need for cognition	-0.3889	0.0768[16]

Problem definition

Verbal protocol analysis

The verbal-protocol analysis reveals that S2-14 did not put much time into defining the task before he started making decisions and that he was silent most of the time until he made his first one. Finally, he stated:

> 'I think I will try a tactic where I try my way. I see in the weeks leading up to now that he has gained a higher contribution by increasing the price.'

He made his first decision after only 3.13 minutes and entered a price of NOK 133, six waiters and three assistants.

Eye-movement analysis

The eye-tracking data shows that S2-14 had both horizontal and vertical eye movements before he made any decision. The data shows horizontal eye movements between variable labels and their corresponding values, as well as vertical eye movements between the different data values.

[14] The first decision was a 'test decision'. Thus, the problem-definition phase continued until the decision for the second week was made.
[15] Based on the above footnote, the actual problem-solving phase only lasted 6.65 min.
[16] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

However, eye-movement data also reveal that S2-14 did not focus on the gap between demand and sales in week 20, which would be necessary to calculate the capacity of the waiters before making his first decision.

Problem solving:

Verbal protocol analysis

S2-14's strategy for the first week (week 23), however, did not work well, and the verbal protocol reveals that he therefore started to define the task properly in the problem-solving phase. He had several silent periods during the problem-solving phase. He calculated the capacity of the waiters and the kitchen assistants before making his next decision:

> 'By dividing actual demand by the number of waiters they have used, I see that a waiter handles approximately 100 meals. One kitchen assistant can handle about 200 meals.'

Notice that the above statement does not show the figures that S2-14 used to calculate the capacities.

He then concluded that he needed to reduce the number of waiters and the number of assistants if he were to increase the price:

> 'So if I try to reduce the number of waiters to five and the number of assistants to two, then the kitchen assistants should be able to handle about 450 meals. I think that the five waiters would handle at least the same. Based on the table, to make a demand of about 450, I think maybe I can go up to a price of approximately NOK 145.'

The quote shows that S2-14 used historic price–demand data to set a price according to his expected capacity of two assistants. He noticed that the contribution increased and decided to try to increase the price even more to NOK 154 in week 25. He reduced the number of waiters to four and kept the number of assistants at two. S2-14's combination in week 25 is actually the optimal solution for the high-complexity task, as can be seen in Figure 4.39.

The verbal protocol shows that S2-14 made marginal considerations and mental calculations regarding the possibility of increasing the contribution by reducing the number of employees and at the same time increasing the price; for example:

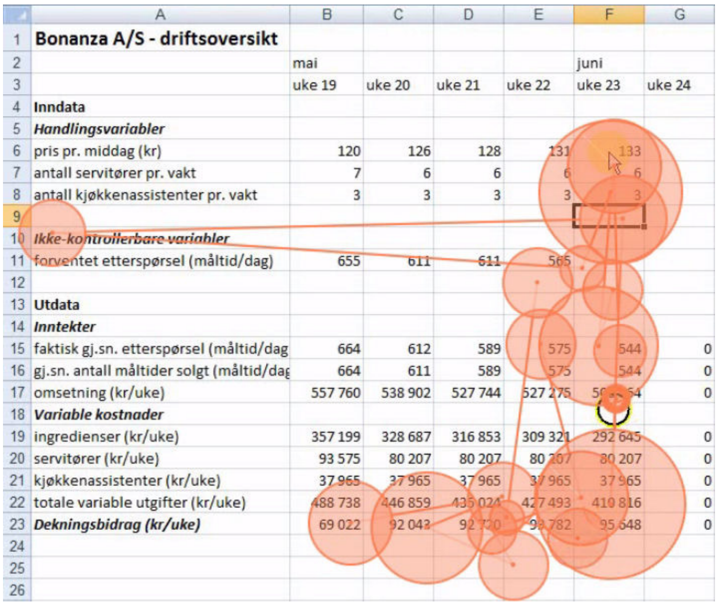
> 'If I should reduce by one more assistant, I would have to increase price far beyond the range in this (historic) table in order to have a low enough demand.'

S2-14 used the two following periods to test whether he could improve the solution even more by setting the price to NOK 155 in week 26 and to NOK 153 in week 27. He realised that he had found the optimal solution in week 25 and therefore continued with a price of NOK 154, four waiters and two kitchen assistants for the remaining weeks of the experiment.

Eye-movement analysis

The eye-tracking data shows that S2-14 sequentially compared the contribution from his first decision against the contribution in the four historic periods; see Figure 4.36.

The eye-tracking data shows that S2-14 focused on the gap between the sales and the demand for meals in the moments just before the verbal protocols stated that he had calculated capacities for waiters. Figure 4.37 a) shows the vertical eye movements between the number of employees and the sales of/demand for meals. Similar eye movements were observed at the moments when S2-14 estimated the capacity of the kitchen assistants.



	A	B	C	D	E	F	G
1	**Bonanza A/S - driftsoversikt**						
2		mai				juni	
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24
4	**Inndata**						
5	***Handlingsvariabler***						
6	pris pr. middag (kr)	120	126	128	131	133	
7	antall servitører pr. vakt	7	6	6	6	6	
8	antall kjøkkenassistenter pr. vakt	3	3	3	3	3	
9							
10	***Ikke-kontrollerbare variabler***						
11	forventet etterspørsel (måltid/dag)	655	611	611	565		
12							
13	**Utdata**						
14	***Inntekter***						
15	faktisk gj.sn. etterspørsel (måltid/dag	664	612	589	575	544	0
16	gj.sn. antall måltider solgt (måltid/dag	664	611	589	575	544	0
17	omsetning (kr/uke)	557 760	538 902	527 744	527 275	[illegible]	0
18	***Variable kostnader***						
19	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	292 645	0
20	servitører (kr/uke)	93 575	80 207	80 207	80 207	80 207	0
21	kjøkkenassistenter (kr/uke)	37 965	37 965	37 965	37 965	37 965	0
22	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	410 816	0
23	***Dekningsbidrag (kr/uke)***	69 022	92 043	[illegible]	[illegible]	95 648	0
24							
25							
26							

Figure 4.36 - Horizontal eye-movements for S2-14: comparison of contribution.

Figure 4.37 b) shows vertical eye movements between the number of kitchen assistants and sales/demand data and horizontal eye movements between the numeric values S2-14 was focusing on and their corresponding labels.

The verbal protocol shows that S2-14 wanted to fully utilise the capacity of his employees when he aimed at the optimal solution. The eye-tracking data reveals that he had many vertical eye movements between the number of employees and the sales/demand data, similar to the ones shown in Figure 4.37.

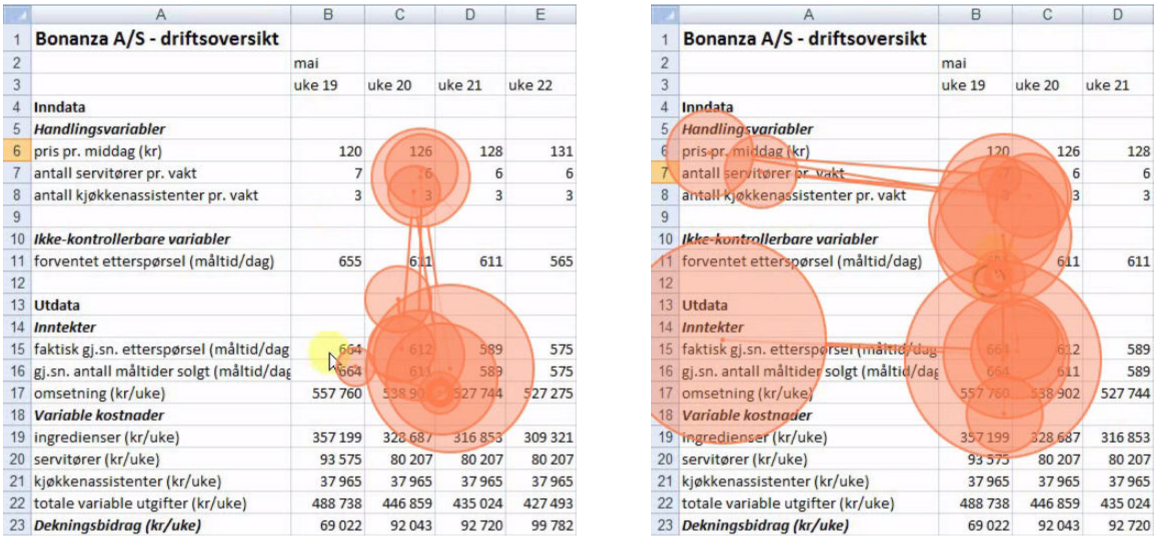


Figure 4.37. a) Eye movements for S2-14 when finding the waiters' capacity. b) Eye movements for S2-14 when finding the capacity for the kitchen assistants.

Furthermore, the eye-tracking data clearly shows that S2-14 systematically focused his visual attention on previous periods in the model as he made his decisions. He particularly focused on input data, the contribution and whether he managed to cover the demand; see Figure 4.38 and Figure 4.39.

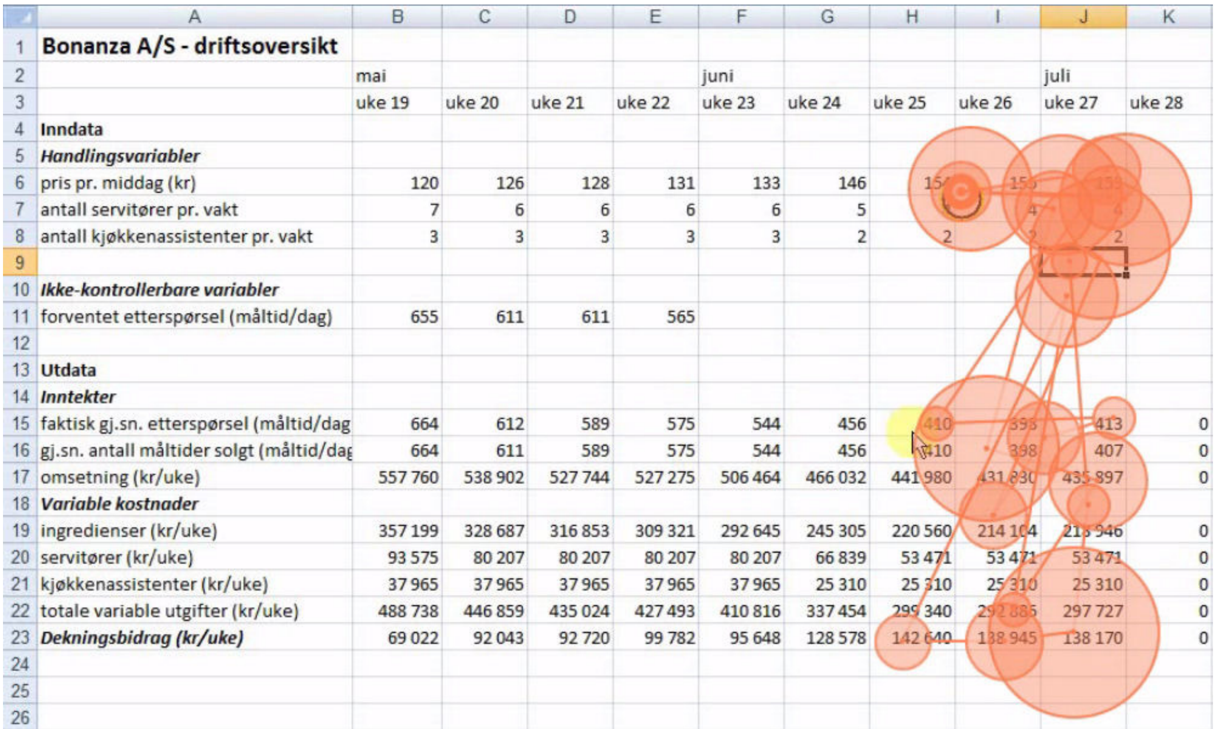
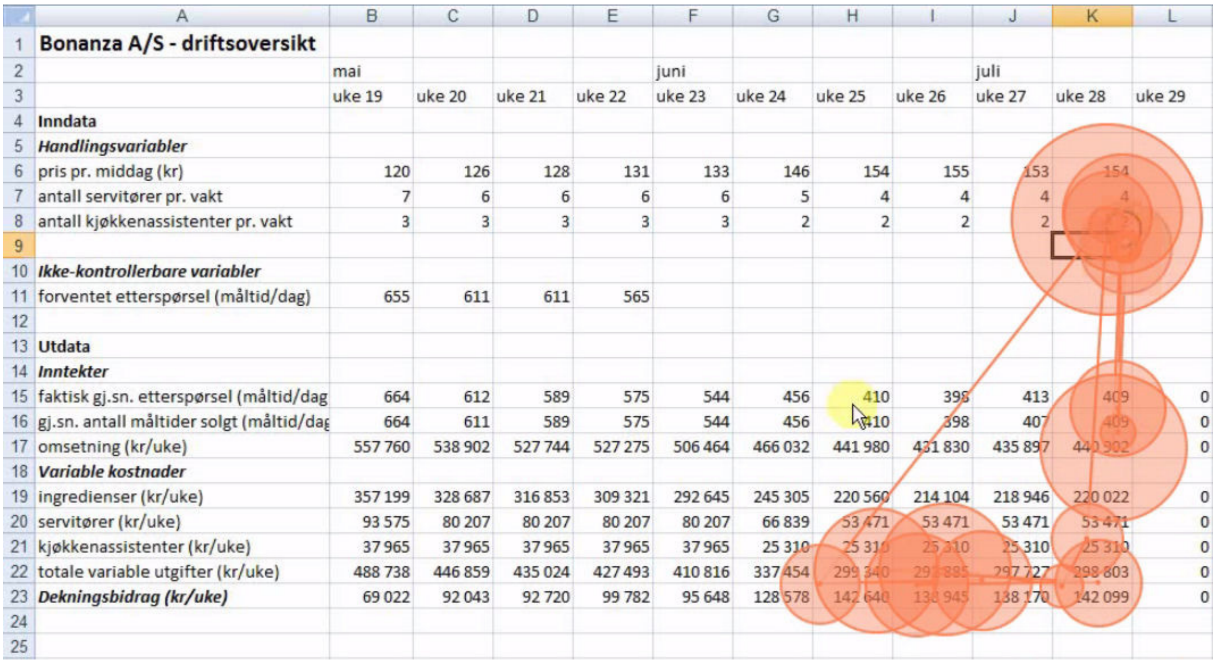


Figure 4.38. Attention directed to input data, sales/demand and contribution by S2-14.



	A	B	C	D	E	F	G	H	I	J	K	L
1	**Bonanza A/S - driftsoversikt**											
2		mai				juni				juli		
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29
4	**Inndata**											
5	***Handlingsvariabler***											
6	pris pr. middag (kr)	120	126	128	131	133	146	154	155	153	154	
7	antall servitører pr. vakt	7	6	6	6	6	5	4	4	4	4	
8	antall kjøkkenassistenter pr. vakt	3	3	3	3	3	2	2	2	2	2	
9												
10	***Ikke-kontrollerbare variabler***											
11	forventet etterspørsel (måltid/dag)	655	611	611	565							
12												
13	**Utdata**											
14	***Inntekter***											
15	faktisk gj.sn. etterspørsel (måltid/dag	664	612	589	575	544	456	410	398	413	409	0
16	gj.sn. antall måltider solgt (måltid/dag	664	611	589	575	544	456	410	398	407	409	0
17	omsetning (kr/uke)	557 760	538 902	527 744	527 275	506 464	466 032	441 980	431 830	435 897	440 902	0
18	***Variable kostnader***											
19	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	292 645	245 305	220 560	214 104	218 946	220 022	0
20	servitører (kr/uke)	93 575	80 207	80 207	80 207	80 207	66 839	53 471	53 471	53 471	53 471	0
21	kjøkkenassistenter (kr/uke)	37 965	37 965	37 965	37 965	37 965	25 310	25 310	25 310	25 310	25 310	0
22	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	410 816	337 454	299 340	[illegible]	297 727	298 803	0
23	***Dekningsbidrag (kr/uke)***	69 022	92 043	92 720	99 782	95 648	128 578	142 640	138 945	138 170	142 099	0
24												
25												

Figure 4.39. Attention directed to the contribution for the last five periods by S2-14.

Summary

Performance: S2-14 started out using a trial-and-error strategy in the first week, but he then made an effort to understand the dependencies among the decision variables and the output variables before he made his next decision. Therefore, he achieved a good result. He calculated the capacity of the waiters, estimated the capacity of the kitchen assistants and utilised the historical price and demand data to optimise the contribution. S2-14 found the optimal solution in his third week as manager. He did not extrapolate sales prices and demands in the range close to his optimal price, but he used two periods to test whether prices above/below his optimal price would provide even better solutions than he had already found. S2-14's handling of the task was characterised by a systematic and structured approach.

Use of presentation format: S2-14 was given the table format. He used both the historical data showing the relationship between price and demand and the table in the spreadsheet model. He did not attempt to develop a graphical representation of the data, for example, to visualise trends over time, but the eye-tracking data show that he sequentially compared variable values in the table over time; see Figure 4.36 and Figure 4.39. This might indicate that he mentally represented how the variable values developed over time.

Complementarity of the data collection methods: The eye-tracking data were useful supplements to the analysis of the verbal protocol, especially in the periods where S2-14 was silent, showing what he was focusing his visual attention on. For example, the eye-movement data improved my understanding of how S2-14 calculated the capacities of the employees and further confirms my impression from the verbal protocol that he had a good understanding of the relationships among input and result variables in the model.

S2-20 – Low performer

I was not able to analyse the eye-tracking data for the worst performing subject, S2-17, in this group. The eye-movement analysis software crashed whenever I tried to load S2-17's gaze data into the software; see section 4.5.1. Therefore, I analysed the second lowest performing subject, S2-20, instead.

Table 4.17. Result and control variables for S2-20.

	S2-20	*Group average*
Contribution index value	0.8081	0.8602
Problem attention	59.33	28.23
Problem definition	30.95	11.17
Problem solving	28.38	17.06
Need for cognition	0.2222	0.0768[17]

As shown in Table 4.17, S2-20 spent considerably more time working on the task compared to the average for his group. Figure 4.40 shows S2-20's decisions through week 32. The decisions for weeks 33–39 are the same as in week 32 and thus have been omitted from the figure.

Problem definition

Verbal protocol analysis

The analysis of the verbal protocol shows that S2-20 spent more than half of his session trying to clarify and define the problem before he made his first decision. During the problem definition, he studied historical price–demand data that he had access to on paper and asked clarifying questions such as whether he could trust that the historic data were still valid and whether his decisions would apply for a whole week or a day at a time. Furthermore, he performed some calculations in the spreadsheet. He calculated the percentage distribution of

[17] Average NFC for all subjects in Study 2. The NFC measure is not related to the presentation format.

the salary costs for waiters, kitchen assistants, and the ingredients' percentage shares of the total variable costs. However, he did not explain why he performed these calculations.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	**Bonanza A/S - driftsoversikt**														
2		mai				juni				juli					august
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32
4	**Inndata**														
5	***Handlingsvariabler***														
6	pris pr. middag (kr)	120	126	128	131	126	126	131	138	180	170	139	139	145	139
7	antall servitører pr. vakt	7	6	6	6	6	5	6	6	2	2	5	5	5	5
8	antall kjøkkenassistenter pr. vakt	3	3	3	3	2	3	3	3	1	1	3	2	2	2
9															
10	***Ikke-kontrollerbare variabler***														
11	forventet etterspørsel (måltid/dag)	655	611	611	565	610	610								
12															
13	**Utdata**														
14	***Inntekter***														
15	faktisk gj.sn. etterspørsel (måltid/dag	664	612	589	575	601	609	575	512	190	293	512	504	458	509
16	gj.sn. antall måltider solgt (måltid/dag	664	611	589	575	500	506	575	512	190	204	512	504	458	509
17	omsetning (kr/dag)	79 680	76 986	75 392	75 325	63 000	63 756	75 325	70 656	34 200	34 680	71 168	70 056	66 410	70 751
18	omsetning (kr/uke)	557 760	538 902	527 744	527 275	441 000	446 292	527 275	494 592	239 400	242 760	498 176	490 392	464 870	495 257
19	***Variable kostnader***														
20	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	268 975	272 203	309 321	275 430	102 211	109 742	275 430	271 127	246 381	273 817
21	servitører (kr/uke)	93 575	80 207	80 207	80 207	80 207	66 839	80 207	80 207	26 736	26 736	66 839	66 839	66 839	66 839
22	kjøkkenassistenter (kr/uke)	37 965	37 965	37 965	37 965	25 310	37 965	37 965	37 965	12 655	12 655	37 965	25 310	25 310	25 310
23	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	374 492	377 006	427 493	393 602	141 601	149 132	380 234	363 276	338 530	365 965
24	***Dekningsbidrag (kr/uke)***	69 022	92 043	92 720	99 782	66 508	69 286	99 782	100 990	97 799	93 628	117 942	127 116	126 340	129 292
25															
26															
27	Prosenter, variable kostnader														
28	Ingredienser	0,730859	0,735551	0,728356	0,723571	0,71824	0,722011	0,723571	0,699769	0,721821	0,735869	0,724371	0,746339	0,727797	0,748204
29	Servitører	0,191462	0,17949	0,184373	0,187622	0,214175	0,177289	0,187622	0,203777	0,18881	0,179275	0,175784	0,18399	0,197439	0,182638
30	Kjøkkenassistenter	0,077679	0,084959	0,08727	0,088808	0,067584	0,1007	0,088808	0,096454	0,08937	0,084857	0,099845	0,069671	0,074764	0,069159
31															

Figure 4.40. S2-20's decision making until the final solution was reached.

Acording to the verbal protocol, S2-20 formulated the objective of the task as follows:

> 'I need to set the price according to what fits optimally to the waiters … how to fully utilise them and the kitchen assistants. That's the task.'

The statement above shows that S2-20 aimed at fully utilising the capacity of his employees. However, such a strategy does not necessarily maximise the contribution, and S2-20 thus did not really understand how to find the optimal solution. Nevertheless, the protocol also shows that S2-20 focused on maximising the contribution:

> 'OK, so when he [the manager] is going to reward us, he will look for the highest contribution. So that's what I have to focus on.'

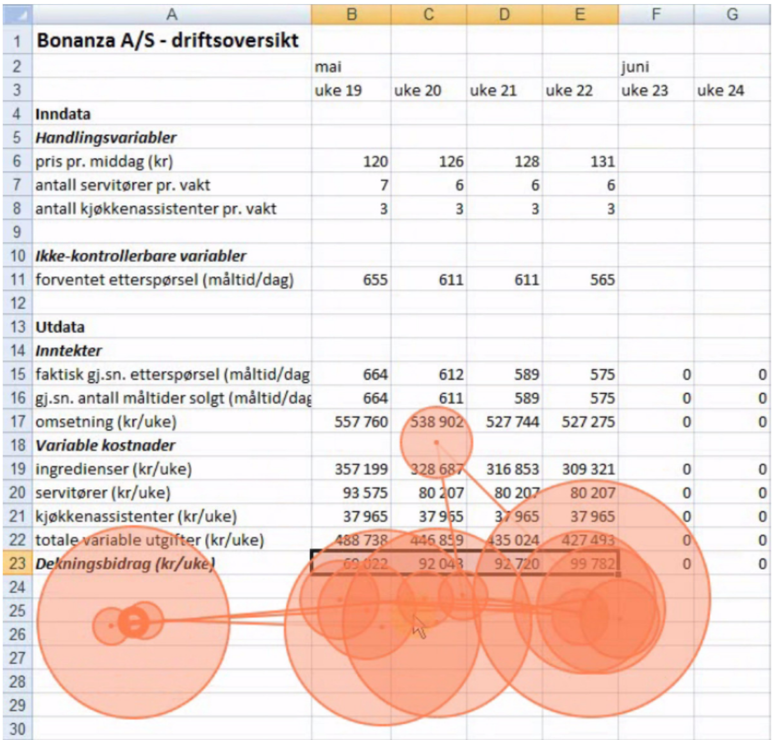
The verbal protocol reveals that S2-20 calculated the capacity of the waiters based on the historic periods in the spreadhseet before making any decisions:

> 'Let's see; 611 divided by 6 waiters equals 101. That is, the waiters handle about 100 each.'

S2-20 did not perform calculations regarding the capacity of ther kitchen assistants. Instead, he decided to try to use the same decision as in the second histoic week, except that he would reduce the number of kitchen assistants to two. He entered a price of NOK 126, six waiters and two kitchen assistants into the model. He entered an expectation of 610 meals per day.

Eye-movement analysis

The eye-tracking data shows that S2-20 had many horizontal eye-movements where he compared values for variables over time. For example, Figure 4.41 shows that S2-20 compared contribution in the four historic weeks during the problem-definition phase. Note that, at this point in time, the scanpaths are skewed downwards compared to the 'real' positions where S2-20 looked.



	A	B	C	D	E	F	G
1	**Bonanza A/S - driftsoversikt**						
2		mai				juni	
3		uke 19	uke 20	uke 21	uke 22	uke 23	uke 24
4	**Inndata**						
5	***Handlingsvariabler***						
6	pris pr. middag (kr)	120	126	128	131		
7	antall servitører pr. vakt	7	6	6	6		
8	antall kjøkkenassistenter pr. vakt	3	3	3	3		
9							
10	***Ikke-kontrollerbare variabler***						
11	forventet etterspørsel (måltid/dag)	655	611	611	565		
12							
13	**Utdata**						
14	***Inntekter***						
15	faktisk gj.sn. etterspørsel (måltid/dag	664	612	589	575	0	0
16	gj.sn. antall måltider solgt (måltid/dag	664	611	589	575	0	0
17	omsetning (kr/uke)	557 760	538 902	527 744	527 275	0	0
18	***Variable kostnader***						
19	ingredienser (kr/uke)	357 199	328 687	316 853	309 321	0	0
20	servitører (kr/uke)	93 575	80 207	80 207	80 207	0	0
21	kjøkkenassistenter (kr/uke)	37 965	37 965	37 965	37 965	0	0
22	totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	0	0
23	***Dekningsbidrag (kr/uke)***	69 022	92 043	92 720	99 782	0	0

Figure 4.41. S2-20 comparing values for contribution; skewed scanpaths.

The scanpaths also reveal that S2-20 associated values with their corresponding variable labels.

S2-20 had some vertical eye movements during the problem-definition phase, where he compared values across variables. For examplee, Figure 4.42 a) shows that he compared prices against contribution in the historic period. Furthemore, the eye-tracking protocol show that S2-20 focused a lot on the area in the spreadsheet model showing sales of and demand for meals.

Figure 4.42 b) shows the moment in time when he calculated the capacity for the waiters. The figures shows that he related the number of waiters to the actual demand for and sales of meals.

Again it must be noticed that the fixations are skewed downwards compared to what S2-20 actually looked at in both Figure 4.42 a) and Figure 4.42 b).

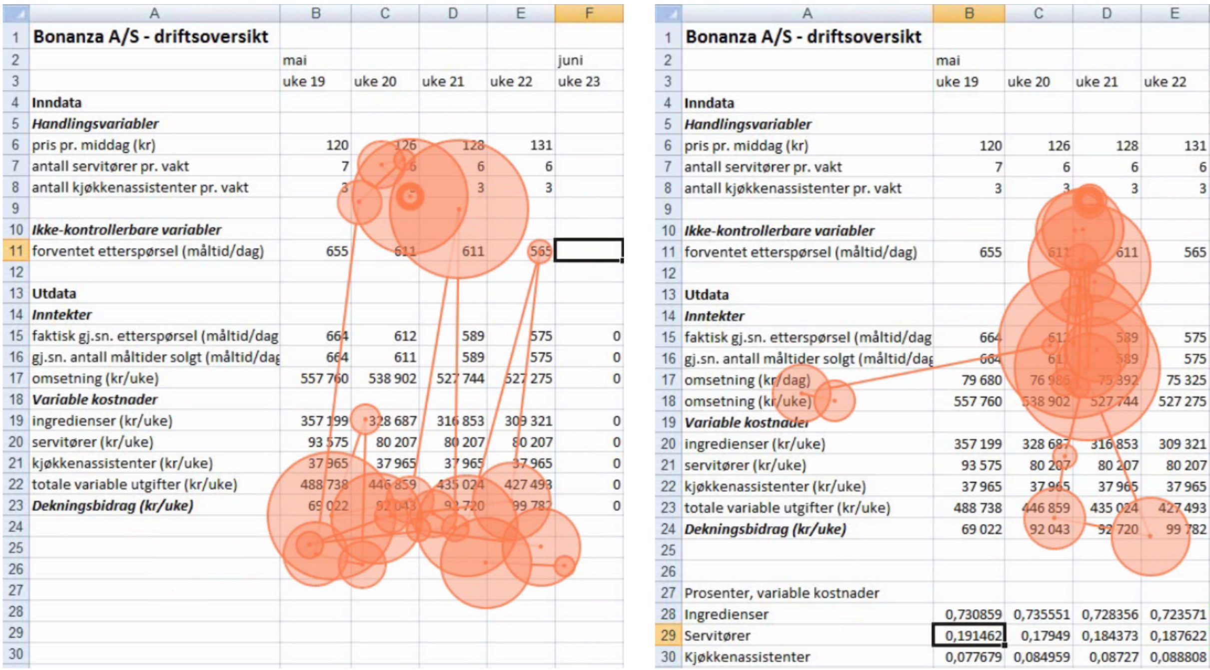


Figure 4.42. Vertical eye movements of S2-20: a) when comparing contribution to input variables and b) when calculating waiter capacity.

Problem solving

Verbal protocol analysis

The verbal protocol shows that S2-20 observed a drop in the contribution and reasoned:

> 'I have not been able to handle the demand. That's why it went so bad! (...) I've always thought that it was the waiters who could not serve all the guests, but it [the inability to handle all the guests] might just as well be because of the kitchen assistants.*'*

S2-20 had access to a sheet of paper showing a table with prices ranging from NOK 120 to NOK 165 and the average actual demand and sales (price x demand) for the last three years (see Appendix C). Instead of using the table data from the piece of paper, the verbal protocol shows that S2-20 created a graph to represent the data; see Figure 4.43.

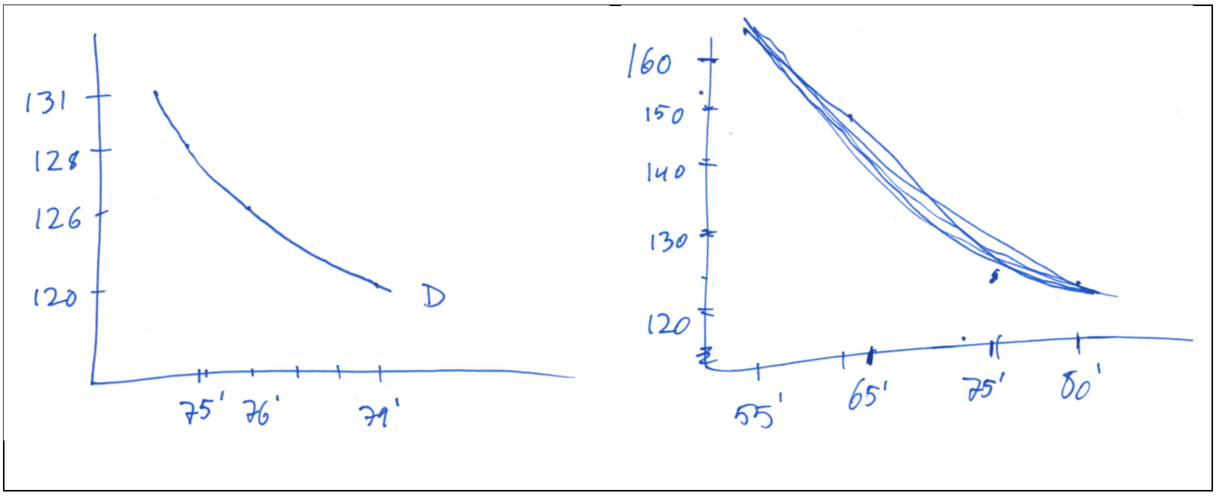


Figure 4.43. S2-20 had a need for a graphical representation of the table data.

When creating the graph, however, S2-20 made a mistake and based it on the sales values instead of the demand data. Nevertheless, he labeled the line in the graph as D (demand) and used the graph to predict demand for his price decisions.

S2-20 decided to try another week (week 24) with a price of NOK 126, three kitchen assistants and only five waiters. He did not explain why he reduced the number of waiters from six to five; however, the verbal protocol shows that he got confused when the contribution did not increase more than it did and that he was unable to handle all the demand:

> 'I've done something wrong here ... switched price and ... no ... I get really confused now, really stressed out! (…). It should be some simple calculation procedure here that I don't see. I don't understand why I don't get it.'

He tried another week with a price of NOK 131, six waiters and three kitchen assistants. He noticed that the contribution increased and decided to increase the price even more to NOK 138 in week 26. In the two next weeks, he first tried to increase the prices even more to NOK 180, and then reduced it to NOK 170, with two waiters and one kitchen assistant. Because the historic price–demand data did not show prices higher than NOK 165, S2-20 tried to predict the demand based on the data he already had:

> 'Let's see. From 120 to 160. How much is the increase? Yes, 40% increase in the price in regard to … Price+40%, what is the difference? 697, yes … The demand went down by 50%. Let's see if that holds for the next interval as well. I am supposed to reduce the demand by 50% from 500 to 250. And there I had 140 multiplied by 1.4 ...'

Notice that the above statement is not complete. S2-20 only partly verbalised his thoughts during these calculations.

S2-20 concluded that neither a price of NOK 180 nor NOK 170 improved his contribution, and in week 29 he looked back on his previous weeks and decided to use a decision close to the one made in week 26. He entered a price of NOK 139, five waiters and two kitchen assistants. The verbal protocol reveals that at this point in time, S2-20 had given up trying to calculate an optimal solution and that he had decided to just try his way out. In week 30, he kept the price at NOK 139 and the number of waiters at five but adjusted the number of kitchen assistants from three to two. He noticed that this adjustment gave him the highest contribution of all weeks and pretty much decided that he wanted to continue with this solution. However, he tried one week with a price of NOK 145, but because the previous price of NOK 139 reuslted in a higher contribution, he continued with that solution for the remaining weeks.

Eye-movement analysis

An analysis of the scanpaths shows that after S2-20 made his first decision, he focused a lot on the area showing actual sales and average demand for meals in his first week as manager. The verbal protocol reveals that at the same point in time he commented on the gap between demand and sales; thus, the two protocol methods complement each other. S2-20 also compared his sales and demand data to eqvivalent data in the previous periods. However, the eye-movement data does not show that he related the gap between sales and demand to the number of employees.

Summary

Performance: S2-20 is a low performer because he did not understand the relationship between input and result variables in the model, even though he made an effort to do so during the problem-definition phase. He did not really understand how to find the optimal solution, as he focused on fully utilising the capacities of the waiters and the kitchen assistants. Furthermore, when he attempted to represent the relationship between price and demand, he made an error that confused him and caused him to solve the task using trial and error. S2-20 handled the task using both analytic and spatial processes.

Use of presentation format: S2-20 was given the tabular presentation format. He tried to represent the historic price–demand data on a graph but when doing so he unintentionally based his graph on sales values instead of demand values. He was not aware of his error, but he eventually understood that the graph did not provide him with the data he needed. Therefore, he discarded the graph as a basis for further decision making.

Complementarity of the data collection methods: The eye-tracking data supports my interpretation of the verbal protocols that S2-20 did not really understand the relationships among input and decision variables in the spreadsheet model. He seems to have had an unstructured gaze pattern, especially in the early phase of the experiment, and the eye-movement data show that he did not systematically investigate the relationships between variables in the model. In his two first weeks as a manager, S2-20 ended up not being able to handle the demand for meals. In the first week, this was due to a lack of both waiters and kitchen assistants. In the second week, the inability to serve all the customers was due to a lack of waiters. S2-20 did not seem to detect this.

Comparison of high and low performer

Performance: The main difference in the performance between S2-14 and S2-20 relates to the understanding of the relationships among the input variables and the result variables. S2-14 used one 'test week' before he developed an understanding of the dependencies among the variables. Then he only needed one additional week to find the optimal solution. S2-20 did not understand how to find the optimal solution, and he failed in representing the relationship between price and demand. S2-14 used analytic processes with marginal considerations related to the number of employees. S2-20 used both analytic and spatial processes, and his handling of the task was characterised by trial and error.

Use of presentation format: S2-14 and S2-20 were given the tabular presentation format. S2-20 attempted to create a graphical representation of the relationship between price and demand; however, this representation was not correct as he unintentionally used the wrong input variable for the demand. S2-20 did not realise his error and got confused. He discarded the graph for further decision making, and thus, both S2-14 and S2-20 based their decision making on the tabular displays.

Complementarity of the data collection methods: The eye-tracking data for S2-14 and S2-20 supports my interpretation of their verbal protocols. For S2-14, the eye-movement data were useful for understanding what he was doing during his silent periods and how he performed his calculations. For S2-20, the eye-movement data confirms my impression that he did not understand the relationships among the variables in the model. The eye-tracking data supports the interpretation of S2-14's and S2-20's level of information processing.

4.6 Discussion

This study is a follow-up study based on Study 1. The purpose of the study was to enhance the understanding of whether and possibly how data presentation formats influence the decision-making process by including eye tracking as process-tracing method in addition to think aloud. As mentioned in section 3.4.2, it was difficult to use the verbal protocols to assess the subjects' level of information processing when they were silent or verbalised incompletely. Also, accurate measures of how the subjects allocated their time into problem definition and problem solving were included to better understand possible trade-offs between accuracy and effort (see Vessey, 1994).

In the following, I will discuss the findings from section 4.5. I will compare similarities and differences between the high and low performers for each presentation format, and I will evaluate whether the differences may be related to the presentation formats. Next, I will discuss how the eye-tracking data and the scanpath videos complement my analyses of the verbal protocols. I will also describe the difficulties I had in my attempts to coordinate the two data collection methods in my analyses. Finally, I will discuss limitations of the eye-tracking part of the study and propose implications for further research. Other limitations of my research will be addressed in Chapter 6.

4.6.1 Summary of findings

Graphical presentation format

Related to the graphical presentation format, the highest performers (S2-37 and S2-13) seemed to emphasise the development of an understanding of the relationships among the variables in the model. The starting point for this development was a careful study of the paper-based XY graph showing the relationship between price and demand. These subjects seemed to have a need for accuracy, and they made the effort of reading out accurate values from the graph and constructed a data table so that they could perform accurate calculations. Both the highest performing subjects mainly used analytic processes to solve the task.

The lowest performers (S2-06 and S2-15) did not make use of the XY graph showing the relationship between price and demand. There are no indications that they used the decision aids to perform calculations in terms of spreadsheets and documents. Furthermore, the verbal protocols do not indicate that they performed mental calculations. They did not seem to fully understand the relationships among variables, but the verbal protocols indicate that they detected the general development of the relationships among variables in the first four periods in the model. For example, when the prices increased in the first four periods, the contribution also increased. They knew that they were supposed to maximise the contribution, but they did not know how to attain this goal. Compared to the highest performers, the lowest performers used the graph directly in their decision-making. They did not attempt to extract numeric values from the graph. The lowest performers mainly used spatial processes when solving the task.

Above, I have contrasted the highest and lowest performers. However, when analysing the verbal protocols for the other subjects, I see the same patterns. The verbal protocols reveal that several of the subjects who received only a graphical presentation format, expressed a need for numeric values. These subjects are S2-03, S2-13, S2-16 and S2-22.

The above findings of the high and low performers with the graphical presentation format are in accordance with the cognitive fit theory by Vessey (1991, 1994). The experimental tasks used in the experiment have both analytic and spatial elements. However, the tasks may be characterised as being mainly analytic. Thus, the graphical presentation format (spatial) did not perfectly match the task (analytic) so that when creating their mental representation of the problem, the subjects could not use the same mental processes to act both on the task and on the presentation format (cf. Vessey, 1994).

It seems as though the high performers have based their mental problem representation mainly on the task (analytic) and therefore needed to transform the graphical presentation format into a format that matched the task, that is a tabular format. By doing so, they could use analytic processes to act both on the analytic elements of the task and on the presentation format, and they were able to solve the task accurately. Thus, cognitive fit occurred (Vessey, 1991, 1994), that is, the cognitive efforts used in carrying out the decision-making strategies were minimised.

The low performers seem to have relied mainly on the graphical data presentation format when creating their mental representation of the problem. Thus, there was a mismatch between the task and the data presentation format being used. The task required accuracy, but the lowest performers did not seem to be able to use analytic processes to solve it. As mentioned above, some subjects who wanted to use analytic processes were limited to use spatial processes due to the presentation format (S2-03, S2-13, S2-16 and S2-22 as shown above).

Tabular presentation format

The highest performers with the tabular presentation format (S2-39 and S2-14) used the tables showing historical data, both in the model and on paper, in order to develop an understanding of the relationships among the variables in the model. They performed accurate calculations including marginal considerations. S2-39 performed his calculations in the spreadsheet whereas S2-14 used a calculator, pen and paper when performing his calculations. They both mainly used analytic processes in a systematic approach to solving the task. Neither S2-39 nor S2-14 attempted to create a graphical presentation of the table data. However, S2-14 had many horizontal eye-movements where he sequentially compared variable values in the table over time. These movements may indicate that he mentally represented how the variable values developed over time.

The lowest performers (S2-20 and S2-41) did not seem to understand the relationships among the variables, but they understood the general development of the relationships among the variables in the model. They both performed some simple calculations, such as finding the waiter capacity, S2-20 using the spreadsheet and S2-41 using a calculator. S2-20 wanted to create an XY graph, but unfortunately used the wrong input variable and thus failed to represent the relationship between price and demand graphically. Subject S2-41 mainly used analytic processes, whereas S2-20 used both analytic and spatial processes. They both handled the task using a systematic trial-and-error strategy based on experience, in an attempt to constantly improve the result. They attained relatively high scores on the level of information processing because they attempted to integrate the decision variables and the result variables in the model.

Related to Vessey (1991, 1994), the above description of the subjects presented with a table format supports my findings from Study 1 that for tasks requiring accuracy, a tabular presentation format is more suitable than graphs. In my analyses, I did not detect the strategic shift from analytic to spatial processes (cf. Vessey, 1994) even though my experimental tasks were more complex than the tasks used by Vessey (1994). I believe that the access to decision aids may have delayed such a shift.

Combined presentation format

The highest performers who had the combined presentation format (S2-05 and S2-35) seem to have solved the task mainly using the table. The line graphs were used differently than I had expected. Seemingly, the subjects did not use the graphs in the model to get an overview of the relationships among the variables before they started making decisions. Instead, the graphs were used to check whether the results of their decisions were according to their expectations (see, e.g. S2-05). Both S2-05 and S2-35 also applied spatial processes in their decision-making. A surprising finding is that of the two types of graphs, that is, line graphs and XY graphs, these subjects seem to have a particular need for the XY graph showing the relationship between price and demand.

Of the lowest performers with the combined presentation format (S2-42 and S2-18), S2-42 used primarily the table. She mainly used analytic processes and performed calculations in the spreadsheet. However, there were errors in her calculations. S2-18 used both the table and the line graphs in the model, but she did not make use of the historic price-demand data available on paper. She mainly used spatial processes. Compared to the highest performers, the lowest performers handled the task primarily using a trial-and-error strategy. S2-42's approach was systematic, that is, she constantly tried to improve the result based on the outcome of her previous decisions. S2-18's approach, on the other hand, seemed to be more random, that is, she did not keep track of her results from previous weeks, and thus did not manage to improve the result as the experiment progressed.

Related to Paivio (1971, 1986, 2007), my findings show that the subjects with the combined format used both the table and the graph part of the presentation format, and that they used both analytic and spatial processes when handling the task. The table data seem to support the use of analytic processes and performance of accurate calculations, whereas particularly the XY graphs seem to be important for using spatial processes to understand the (causal) relationship between variables. For example, the historic price-demand data only show some 'historic' values (e.g. 133, 138, 144) for the price variable, and the XY graph seems to have been used to extrapolate the demand values for the prices missing in the data.

Differences related to the presentation formats

Comparing the highest and the lowest performers, the above analyses show that the main difference in performance within each presentation-format group is related to the subjects' understanding of the dependencies among decision variables and the result variables. The highest performers seem to have made an effort to understand the task before they made their first decision. The lowest performers seem to have relied more on trial-and-error strategy in an attempt to gradually improve their result. These differences among the highest and lowest performers are consistent with the amount of time used by the highest and lowest performers on problem definition and problem solving. The highest performers solved the task quicker than the lowest performers, probably because they had a better understanding of how the task should be solved before they started making decisions. The highest performers spent 51% of their time on the problem-definition phase, that is, before they started making decisions, and less time on the actual problem-solving phase. The lowest performers spent most of their time in the problem-solving phase and less time (24%) on the problem-definition phase. This supports my impressions from Study 1.

The above analyses indicate, in line with the findings from Study 1, that subjects who had the graphical format only, seem to have been limited to mainly using spatial processes. However, the analyses also indicate that subjects with the necessary spreadsheet skills were able to compensate for this limitation by complementing the graphical presentation format with a tabular format that supported the use of analytic processes.

The above analyses support my findings from Study 1 that a tabular presentation format facilitates the use of analytic processes and accurate calculations. Numeric values seem necessary to develop accurate understanding of the relationships. Graphs may give an overview of the relationships among variables, as discussed above for the relationship between price and demand.

4.6.2 Complementarity of the process tracing methods

A limitation of the verbal protocols is that they did not provide any clues about the subjects' behaviour in the periods when they became silent, for example, when they created models in the spreadsheet, performed calculations or looked at graphs. In addition, the verbal protocols were inadequate when the subjects' verbalisations were incomplete, for example, when they did not say anything about what they were looking at. In these periods, the eye-tracking data were useful for filling in the gaps between observed events and the think-aloud statements in the verbal protocols of the subjects. In general, the eye-movement data reveal the subjects' visual attention, both when they were silent and when they verbalised, providing a more complete picture of how they handled the task. More specifically, the eye-tracking data have been beneficial for a) better understanding calculations, b) better understanding how the presentation formats were used, c) support for my interpretation of the verbal protocols

The eye-movement data were helpful to better understand the calculations carried out by the subjects. They reveal errors both when calculations were based on the wrong variables and when the subjects made general calculation errors they were not aware of. For example, both S2-39 and S2-35 calculated the waiter capacity based on expected demand instead of on actual demand, an error I would not have noticed without the eye-tracking data. S2-42, on the other hand, had a general formula error in her calculation of the waiter costs; she increased the waiter costs by adding an extra waiter instead of subtracting one. I noticed this error by watching the screen video of her session.

Eye-movement data also reveal how the subjects *used* the presentation formats. For example, the line graphs in the combined presentation format were used differently than I had expected. They seemed not to be used to get an overview of the problem during the problem-definition phase. Instead, they were used to check whether the results of their decisions were according to their expectations during the problem-solving phase. On the other hand, the eye-movement data have drawn my attention to the fact that XY graphs seem to be important presentation formats for the type of tasks used in my experiments. Furthermore, eye-movement data helped me better understand how high performers used the graph format. For example, S2-35 and S2-13 fixated on the screen tips showing numeric values from the graph and used these data to create a table that could be used for accurate calculations. Thus, the eye-tracking data support my interpretation of the verbal protocols.

An implication of including eye tracking in Study 2 is that I am better able to assess the level of information processing, especially when the subjects are silent or verbalise incompletely. I therefore also have a better understanding of the relationship between the level of information processing and the result (i.e. contribution). For example, I have a better understanding of the situations where calculation errors, rather than a low level of information processing, cause poor results, such as described for S2-42.

4.6.3 Limitations of eye tracking as process tracing method

I encountered some difficulties when using eye tracking as a process-tracing method in the current study. The problems can be related to the experiment, the subjects and the equipment being used.

In terms of the experiment, the spreadsheet models were not initially designed for an eye-tracking experiment. In my experiments, the subjects were allowed to scroll up/down and sideways. If I had designed the experiments with static screen displays, it would have been easier to aggregate eye-movement data for subjects having the same presentation format and detect differences and similarities among the subjects. Such analyses might have enhanced the understanding of, for example, when and how line graphs were used by subjects with the combined format.

Another limitation related to the experimental design is that I used paper-based presentations of data in addition to the data presented on the computer screen. The eye tracker recorded eye movements when the subjects looked at the computer screen. However, I do not have eye-movement recordings when the subjects looked at the historic price-demand data. Instead of using paper-based graphs and tables to present these data, I should have presented the data on the computer screen so that the eye tracker would also track the subjects' eye movements as they were looking at these data.

In terms of the subjects, I experienced problems with the data quality. As mentioned in section 4.1.2, the quality of the eye-tracking data might be influenced by the characteristics of the subjects. I found that some of my eye-tracking data could not be analysed due to factors such as long eyelashes, drooping eyelids, squinting or the use of glasses, contact lenses or mascara. As a result, I was only able to analyse eye-movement data for 26 of the 34 subjects in the data set for Study 2, see section 4.5.1. For future eye-tracking studies, I recommend a more thorough control of the subjects prior to their participation. In addition, I would recommend to have as a requirement that mascara must be removed for the subjects participating in the experiments.

I have also experienced problems with glitches in the eye-tracking data, that is, the eye tracker did not track the subjects' eyes during some periods in the experiments. Such glitches were most often due to the subjects repositioning themselves in relation to the computer screen/eye tracker. I was able to detect such eye-tracking glitches when I was monitoring the eye-tracking process on the observer screen. However, I often had to stay close to the subjects when they asked questions or otherwise required my attention. Thus, correction for such glitches was often not made. For future eye-tracking studies, I recommend having an additional observer whose only task is to monitor the eye-tracking process and another observer whose task is to observe the subject and to answer questions.

Finally, the equipment used in the experiment caused some difficulties. The problem relates to the synchronisation of the think-aloud recordings with the eye-tracking/screen-video recordings so that simultaneous analysis of the two types of recordings could be coordinated. At the time the eye-tracking sessions were conducted, the current version of the eye-tracking software (BeGaze) did not offer sound recording functionality. Thus, the recordings of the voice (think aloud) were made separately. My intention was to merge the sound recordings with the corresponding screen videos (including scanpaths). However, due to errors in the software that recorded the screen videos, periods where there were no eye movements and no activity on the screen, (e.g. mouse movements or keystrokes) were randomly omitted from the videos. As a result, the length of the screen video recordings and the sound recordings did not match; therefore, the two recordings could not be merged. Instead, I had to manually synchronise two recordings using statements about the weekly decisions (including the sound of keystrokes) in the sound recordings and relate them to the periods in the screen video showing that values were entered into the worksheet.

A possible limitation relates to the speed of the eye-tracker, that is, the sampling rate (Hz), see sections 4.1.2 and 4.3.2. As mentioned above, the eye-movement data show that subjects mainly used the line graphs to check whether the results of their decisions were according to their expectations. However, the fixations on the line graphs were limited, in both numbers and durations. It is thus possible that the eye-movement data do not provide valid insight into how the subjects acquired data from the line graphs. The eye tracker used in the current study had a sampling rate of 50 Hz and is considered a slow eye tracker (Holmqvist, et al., 2011, p. 30). Therefore, for future studies, I recommend using an eye tracker recording at a higher sampling rate in order to detect and measure more precisely how the line graphs are used.

Another limitation that might explain why there were so few/short fixations on the line graphs in the combined presentation format, relates to how the eye tracker works and the anatomy of the human eye. As described in sections 4.1.1 and 4.1.4, the eye-tracker records only what is being focused on with the foveal vision, that is, what is being fixated on, but humans are also capable of noticing elements just outside the foveal vision (see e.g. Bojko, 2013). Maybe the subjects looking at the line graphs were able to grasp a holistic picture of the lines representing the development of variable values without such use of the graph being shown as fixations on the graph. Future research should therefore include questions to the subjects about whether and, if so, how they used the line graphs during the experiment.

5 STUDY 1 + 2

This thesis attempts to enhance the understanding of whether, and possibly how, presentation formats affect managerial decision processes, that is, decision makers' mental representation of the problem, and how they solved it. I have collected data in two experimental studies, using verbal protocols in the first study and verbal protocols and eye-tracking in the second study.

In this chapter, I will present the analyses and findings of the aggregated data from Study 1 and Study 2. First, I present the data set from Study 1 and Study 2. Then I discuss whether the tendencies found in Study 1 are strengthened when aggregating the data from both studies. Furthermore, I discuss whether the analyses of data from Study 2 have improved the understanding of the findings in the aggregated data set. Finally, I relate my findings to previous research.

5.1 Presentation of the aggregated data matrix and outcomes

The aggregated data set consists of data for 40 subjects from Study 1 (see Table 3.2) and 34 subjects from Study 2 (see Table 4.3). The aggregated data set for Study 1 + 2 can be found in Appendix F.3. This data set consists of the same columns as in Table 3.2, described in section 3.3. In Study 1, the level of information processing was assessed once for each subject. In Study 2, I evaluated the level of information processing both in the problem-definition phase and in the problem-solving phase. The column labelled *Level of info. proc.* in the aggregated data set contains data about the level of information processing for the problem-solving phase in Study 2.

Table 5.1 shows the combined results from Study 1 and Study 2, measured as the average contribution–index values. The table shows the outcomes based on the data presentation formats that the subjects received at the start of the experiment.

Despite the lack of significance, the table indicates that the subjects that received the graphical presentation format generated a lower contribution than subjects presented with a table or a combination of a table and a graph. Furthermore, in both task types, the subjects presented with the table format actually performed best. There are few differences between the results from Study 1 and the results shown in Table 5.1.

Table 5.1. Average results based on original presentation format.

		Presentation format		
		Graph	*Table*	*Table and Graph*
Task complexity	*Low*	**0.9711** N=10, st. dev.=0.027	**0.9850** N=8, st. dev.=0.012	**0.9839** N=9, st. dev.=0.019
	High	**0.8423** N=15, st. dev.=0.068	**0.8768** N=16, st. dev.=0.083	**0.8739** N=16, st. dev.=0.081

5.2 Protocol analyses

In this section, I will present my analyses of the verbal protocols following the same structure as in section 3.3.2. I will emphasise whether the results of the aggregated data support the tendencies found in Study 1. In addition, I will focus on how the analyses of high and low performers in Study 2 have contributed to an enhanced understanding of the results.

5.2.1 Analytic and spatial processes

Study 1 indicates that the use of analytic processes seems to be important for attaining a good result in the tasks used in the experiments, that is, tasks requiring accuracy (see section 3.3.2). The same pattern is found in the aggregated data set. The average contribution index for the subjects using mainly analytic processes is 0.9465, whereas the average value for the subjects using mainly spatial processes is 0.8457 ($p < 0.001$). This difference is particularly evident in the high-complexity task group, where the average contribution index for subjects using mainly analytic processes is 0.9013 and 0.8244 for the subjects using mainly spatial processes ($p = 0.002$). A possible explanation of this finding is that the high complexity task involves more variables that need to be integrated, which might be difficult to do accurately with spatial processes.

As in Study 1, the aggregated results show that the low-complexity task was handled mainly using analytic processes. Of the 27 subjects handling the low-complexity task, 22 (81%) mainly used analytic processes, three (11%) mainly used spatial processes and two (7%) used both analytic and spatial processes. Among the 47 subjects handling the high-complexity task, 17 (36%) mainly used analytic processes, 18 (38%) mainly used spatial processes and 12 (26%) used both analytic and spatial processes. An example of a subject using both analytic and spatial processes is S2-05, who received the combined presentation format. He used analytic processes when he performed calculations on marginal costs and profits based on the table format, and he used spatial processes when he checked on the line graphs whether the development of his decisions were as expected.

The analyses of data from Study 2 have enhanced the understanding of how the data presentation formats may influence the subjects' use of analytic and spatial processes, see section 4.6.2. For example, the eye-tracking data show that subjects with the combined presentation format mainly used the table part of the format. Furthermore, the analyses in Study 2 show that some of the subjects who had a graph only, wanted to use analytic processes and perform additional calculations (S2-03, S2-13, S2-16, S2-37 and S2-40). By offering decision aids in my study, some of these subjects (e.g. S2-16 and S2-40) managed to read out numeric values from the graph so that they could perform rather simple calculations. Others, (e.g. S2-13 and S2-37) read out very accurate numeric values from the graph and created a table that they used to perform rather advanced calculations. However, the analyses of the verbal protocols also reveal that some of the subjects with the graph format only, expressed a need to use analytic processes, but that they were limited to using spatial processes due to the lack of numeric values in the graph format. Examples of such expressions were those produced by S2-03: 'Why isn't there any numbers here?', 'I need numbers, otherwise I cannot calculate the contribution' and 'Shouldn't numbers appear here now?'

As illustrated above, the graph format either required additional effort to extract accurate values from the graph so that analytic processes could be used, or the graph format limited the subjects' use of analytic processes, forcing them to use spatial processes instead. Thus, a graphical presentation format only may reinforce the need to use spatial processes as the level of task complexity increases. The availability of a table format seems to facilitate the use of analytic processes.

In line with the finding from Study 1, the aggregated data show that fewer subjects were able to use analytic processes as the level of task complexity increased. This result may indicate a shift from analytic to spatial processes. Such a shift seems to be particularly evident for the graph group. In that group, 80% (8 out of 10) of the subjects mainly used analytic processes when solving the low-complexity task, whereas only 27% (4 out of 15) mainly used analytic processes when solving the high-complexity task.

This result from my studies, that is. the shift from analytic to spatial processes, is consistent with the cognitive fit theory (Vessey, 1991, 1994). Thus, my research supports this theory for tasks that are more complex than the tasks used in previous studies (see, e.g. Vessey, 1994; Speier et al., 2003; Speier, 2006).

5.2.2 Level of information processing

In the previous section, I have focused on the distinction between analytic and spatial processes. In this section, I discuss *how* the subjects have applied these processes according to the theory of cognitive complexity (Schroder et al., 1967).

Study 1 reveals several tendencies related to the level of information processing, see Table 3.4. Table 5.2 presents a corresponding correlation matrix for the aggregated data set. A comparison of Table 3.4 and Table 5.2 shows that the same tendencies are also present in the aggregated data set.

Table 5.2. Non-parametric correlations.

	Result (Index)	*Level of info. Proc.*	*Additional comp.*	*Elective*
Result (Index)	1.000			
Level of info. proc.	$.535^{***}$	1.000		
Additional comp.	$.544^{***}$	$.717^{***}$	1.000	
Elective	.151	$.329^{**}$	$.407^{***}$	1.000

Note. Spearman correlation (2-tailed), n=74.
** Correlation is significant at the 0.01 level
*** Correlation is significant at the 0.001 level

In line with Schroeder et al. (1967) and with the results from Study 1, the correlation matrix for the aggregated data set shows a significant correlation between the contribution index (result) and the level of information processing ($r=0.535$, $p < 0.001$). The subjects mainly using analytic processes achieved the highest scores, both with regard to results, as described above, and as regards the level of information processing index. This tendency from Study 1 is strengthened in the aggregated data. The average value of the level of information processing for subjects mainly using analytic processes is 4.62, whereas the equivalent value for the subjects mainly using spatial processes is 2.95 ($p < 0.001$). This difference is present both for the low- and high-complexity tasks, but is more evident in the high-complexity task. Thus, subjects that were able to integrate the variables of the experimental task, attained a higher contribution (profit). Subjects that used analytic processes were able to find the values of these variables that gave this high profit, particularly in the high-complexity task involving more variables to integrate. Such integration is difficult to achieve accurately with spatial processes.

The analyses in Study 2 have enhanced the understanding of the relationship between the economic result and the level of information processing. The eye tracker provided additional data about what the subjects were looking at, both when they verbalised and when they were silent. These data helped explain why subjects who were given high scores on the level of information processing index did not attain a high contribution. For example, the eye-tracking data showed that S2-35 and S2-39 used expected demand instead of actual sale when calculating the waiter capacity. The eye movements also revealed that S2-42 made calculation errors in her attempt to create a model that predicted contribution at different prices. On the other hand, the eye-tracking data also helped me understand why subjects with a low score on the level of information processing index achieved a good result. For example, S2-37 misunderstood how to find the optimum contribution, but his lucky combination of price and full utilisation of the waiter capacity gave a rather high contribution.

All the highest performing subjects (S1-03, S2-37, S2-39, S1-16, S1-26 and S1-36) in the six treatment groups mainly used analytic processes. All these subjects performed additional calculations at level 3, that is, marginal calculations of some sort, see section 3.3.1. In order to do this, S2-37 and S1-26, who had the graph format only, created an additional table format by reading out accurate values from the graphs. Among the lowest performing subjects, the low- and high-complexity task groups differ with regard to the use of analytic and spatial processes. In the low-complexity task group, all the lowest performers (S2-42, S1-05 and S2-30) mainly used analytic processes, and they performed additional calculations. However, due to calculation errors, none of these subjects were able to integrate the variables in the model correctly. Thus, they ended up handling the task using a systematic trial-and-error strategy, that is, they gradually improved their solutions based on their past experiences. In the high-complexity task, all the lowest performers (S2-18, S2-15 and S1-41) mainly used spatial processes, and they handled the task using an unsystematic trial-and-error strategy. By unsystematic trial-and error strategy, I mean that they did not seem to learn from their experiences as the experiments progressed.

The distinction between a problem-definition phase and a problem-solving phase in Study 2 has revealed the importance of understanding the task before making decisions. All the highest performing subjects made an extra effort to understand the relationships among the variables in the model during the problem-definition phase. They mainly used analytic processes during this phase (e.g. S2-35, S2-37, S2-39, S2-13, S2-14 and S2-05). Some subjects created additional graphs and tables during the experiments. These additional presentation formats were made mostly during the problem-definition phase. The results from Study 2 show that the subjects who mainly used analytic processes spent an average of 17.85 minutes (53.3%) of their time in the problem-definition phase, while the subjects who mainly used spatial processes spent an average of 2.17 minutes (8.9%) of their time on problem-definition activities ($p = 0.002$).

The eye-tracking data from Study 2 also supports my interpretation of the verbal protocols. Low performers seem to have mainly used a trial-and-error strategy (e.g. S2-18, S2-15, S2-06 and S2-20). The eye-tracking data show horizontal eye movements relating variable values over time and relating the variable values to their appropriate labels. They show fewer vertical eye movements relating the variables in the model to each other, see for example Figure 4.29. The eye movements of the high performers indicate a more systematic attempt to understand the relationships among variables. In addition to the horizontal eye movements described for low performers, the high performers' eye movements are characterised by more vertical movements among the model variables, indicating that they tried to understand the relationships among these variables, see for example Figure 4.38.

As in Study 1, Table 5.2 shows a significant positive correlation between the level of information processing and additional calculations ($r = 0.717$, $p < 0.001$).

As discussed above, the importance of analytic processing has been strengthened in Study 2. The eye-tracking data have enhanced the understanding of how the subjects used the graph format, especially the subjects who only had a graph. The results show that the high-performing subjects in the graph group attained a high contribution mainly because they used analytic processes when handling the task. They read out accurate variable values from the graph and performed marginal calculations so that they could integrate the values in their solution. The lowest performers, on the other hand, mainly used spatial processes and thus had problems integrating the variables. They made few or simple calculations, such as dividing the sales by the number of employees. They handled the task mainly using a trial- and-error strategy, some (e.g. S2-41) more systematically than others (e.g. S2-18 and S2-15).

As regards the relationship between having performed additional calculations and having an elective in business data processing, Table 5.2 shows the same pattern as in Study 1, that is, a significant, positive correlation ($r = 0.407$, $p < 0.001$).

Study 2 provides additional insights into the importance of having proper spreadsheet skills when solving tasks like the ones being studied in this thesis. For example, S2-42 had poor spreadsheet skills. She wanted to create a spreadsheet model for estimating the contribution at different prices. However, she did not manage to enter formulas with correct (relative) cell addresses. Consequently, her model produced other results than the experimental task model, causing her to handle the task using trial-and-error instead of relying on her calculations. Several subjects (e.g. S2-20 and S2-25) attempted to make additional XY graphs, but due to errors and poor graph design the graphs were not used. Some subjects did not have the proper spreadsheet skills to read out accurate numeric values from the graphs as illustrated by S2-03 above. Therefore, these subjects were not able to use analytic processes and perform additional calculations when handling the task, even though they wanted to.

On the other hand, the analyses show that the subjects with proper spreadsheet skills were able to modify the original presentation format to include a presentation of the data in a format that suited their needs. This was particularly evident in the graph group, where the analyses show that proper spreadsheet skills were necessary to read out data accurately from the graph.

5.3 Adjusted results

Table 5.3 presents the aggregated results when I have adjusted the results to the presentation formats the subjects actually used while handling the tasks. The criteria for moving a subject from the 'graph' group or from the 'table' group to the 'table and graph' group are described in Study 1, section 3.3.3.

A comparison of Table 5.1 and Table 5.3 shows that 11 subjects that received either a table or a graph format used decision aids to complement the presentation format they received in the experimental treatment. Six subjects that received the table format (three in each task category) generated an XY graph during the session. Five subjects (three in the low-complexity task group and two in the high-complexity task group) copied data from the graphs and generated data tables for their analyses.

Table 5.3. Average results based on presentation format as adapted by subjects.

Presentation format

		Graph	*Table*	*Table and Graph*
Task complexity	*Low*	**0.9655** N=7, st. dev.=0.028	**0.9851** N=5, st. dev.=0.014	**0.9841** N=15, st. dev.=0.017
	High	**0.8330** N=13, st. dev.=0.067	**0.8588** N=13, st. dev.=0.081	**0.8882** N=21, st. dev.=0.077

The comparison also shows that for the high-complexity task in particular, the generation of the additional format improved the outcomes. In this task, the difference in the contribution index value for the 'graph' and the 'table and graph' cells is significant at the 0.05 level with a rather large effect size (Cohen's d= 0.78) (Cohen, 1988). This result is in line with the findings in Study 1.

As discussed above, the analyses of the verbal protocols show that additional subjects, such as S2-03 and S2-20, expressed a need for the combined format, although they were not able to develop the presentation format that they missed.

The results indicate that decision makers need both presentation formats in the tasks used in my experiments. This finding is in accordance with my expectations based on Paivio (1971, 1986, 2007), see section 3.1.1. The combined presentation format supports both the verbal and the non-verbal mental representation systems and thus facilitates information processing in the two sub-systems.

Study 2 provides additional insights into how the combined presentation format was actually used. The subjects mainly used the table part of the combined presentation format, and the line graphs were used differently than I had expected. The subjects did not use the line graphs to get an overview of the general development of the relationships among the variables in the model before they started making decisions. Instead, they used the line graphs to check whether the results of their decisions were as expected. To the extent that they used the graphs during the problem-definition phase, they used the paper-based XY graphs showing the relationship between price and demand.

Subjects with a table format that created additional graphs all created XY graphs. None of these subjects created line graphs. Thus, the combined presentation format, as presented to the subjects, may not have been exactly what the subjects needed. For a combined format, table data in addition to XY graphs seem to be the important parts of the presentation format.

The adjustments in Table 5.3 compared to Table 5.1 were to move subjects from the graph and table groups to the combined format group. Analyses of eye-tracking data from Study 2 indicate that two subjects in the combined group in Table 5.3 should perhaps have been moved to the table group. S2-10, who received the combined presentation format, seems to have focused attention on the table format in his handling of the task. S2-37 received the graph format. She seems to have used the graph primarily to read out accurate numerical values in order to develop a table, which she then used in her handling of the task. However, I do not have data to conclude that the additional format was not applied. Though, if I move these two subjects, the difference between the contribution for the combined group and the graph group increases for the high complexity task.

Thus, the conclusion from Study1 that subjects need both presentation formats is strengthened in the aggregated set, especially for the high complexity task.

6 SUMMARY AND CONCLUSIONS

This thesis attempts to improve the understanding of whether and possibly how presentation formats support managerial decision processes.

Previous research has mostly studied effects of different presentation formats on the *outcome* in simple choice tasks (see section 2.4.1). I have responded to a call for research to enhance the understanding of how presentation formats support decision-makers' mental representation of a problem, and how they solve it (Kelton et al. 2010). Furthermore, I have responded to calls for research on effects of presentation formats in more complex tasks (e.g. Davis, 1989; Vessey, 1994; Kelton et al. 2010).

I have examined decision-makers' need for data presentation formats in tasks that are rather similar to real-life managerial tasks, that is, the relationships among decision variables and outcome variables are characterised by interdependencies and uncertainties. As opposed to most previous research that have focused on either tables or graphs, I have also studied the use of a combined presentation format. Another characteristic that distinguishes my studies from previous studies is that the subjects were allowed to use decision aids, such as spreadsheets and a calculator.

I have chosen an exploratory design because I do research in an area that is badly understood. Furthermore, the purpose of my research is not to find out which presentation format is 'best' in some sense, but rather to enhance the understanding of *how* decision makers apply presentation formats in their decision processes.

In two experimental studies, I have examined how MBA students used graphs and tables, either in isolation or combined, to solve two tasks varying in complexity and uncertainty. The experimental tasks have well-defined optimal solutions so that I can compare the outcome of the subjects' decision processes. The problem tasks designed for the experimental studies had both spatial and analytic subtasks. The tasks concerned time–series data, that is, they included spatial elements of detecting and comparing trends in the output data. They included analytic elements of calculating capacities and dependencies. In Study 1, I used verbal protocols to gain insights into how the subjects applied the presentation formats. In Study 2, I also used eye tracking to complement the verbal-protocol data.

6.1 Summary of findings

My results show that analytic processes are important for attaining a good result in the experimental tasks, particularly in the high-complexity task. Subjects presented with the table format generated on average the highest contribution. However, when I correct for the presentation formats actually *used* by the subjects, the combination of tables and graphs give the highest contribution in the high-complexity task. Tables seem to facilitate analytic processes by presenting discrete variable values. The graph format, on the other hand, provides a visualisation of the relationships among the variables and allows the user to view the data as an integrated whole. Thus, the combined format has qualities to support both the analytic and spatial elements of the experimental tasks.

Subjects presented with graphs only, generated the lowest contribution. As regards the low performers, the graphs seemed to limit these subjects' use of analytic processes, forcing them to use spatial processes. However, the highest-performing subjects presented with the graph format made the extra effort to 'transform' the graph into a table format which they then used in their problem solving. Thus, in the same way as the high performers with a table presentation format, they could make a mental representation and use the same kind of mental processes, that is analytic, to act both on the task and the presentation format.

These findings are in accordance with the cognitive fit theory (Vessey, 1991), stating that a fit between the presentation format and the task is supposed to lead to accurate decisions.

The low-complexity task was handled mainly using analytic processes, whereas fewer subjects were able to use analytic processes as the task complexity increased. This result may indicate a shift from analytic to spatial processes. Such a shift seems to be particularly evident for the graph group. The shift from analytic to spatial processes due to an increase in task complexity, is consistent with the cognitive-fit theory (Vessey, 1991, 1994). Thus, my research extends the application of the cognitive fit theory to tasks that are more complex than the tasks used in previous studies.

By including evaluation of the subjects' level of information processing (Schroder et al. 1967), my research has enhanced the understanding of *how* analytic and spatial processes were used by the subjects. The results show that subjects that were able to integrate the variables of the experimental tasks, and thus were given a high score on the level of information processing index, also attained a higher contribution (profit). This was particularly the case for the subjects that used analytic processes.

The highest performers presented with the table format used accurate numeric data to perform advanced marginal calculations, see section 3.1.1. Thus, they managed to integrate the variables accurately, that is, they were able to find the values of the decision variables that gave high profits. As described above, the highest performers presented with the graph created additional tables which they then used as described for the high performing subjects.

High performers presented with the graph format, who did not make an additional table, attempted to integrate variables, but their integration was not so accurate as the high performers presented with the table format. Low performers, irrespective of the presentation format, seemed to handle the task using trial-and-error.

The eye tracking data in Study 2 enhanced the understanding of the relationship between the level of information processing and the economic result. The eye tracker provided additional data about what the subjects were looking at, both when they verbalised and when they were silent. These data helped explain why subjects who were given high scores on the level of information processing index did not attain a high contribution and vice-versa.

My research indicates that decision makers need both presentation formats. The high-performing subjects mainly used the table when handling the task as described above, but they also used the graphs. However, the graphs were used differently than expected (see section 4.6.1). I expected the subjects to use the line graphs to get an overview of the development of the relationships among the variables before they made decisions. However, the tendency was to use the line graphs mainly to perform quick checks of whether the development of the result variable was as expected after they had made a decision. All the subjects that made additional graphs, made XY graphs. It seems that the subjects needed the XY graphs to visualise causal relationships among variables, such as sales prices and demand.

In sum, my findings have enhanced the understanding of how XY graphs and line graphs are used, and how these graphs complement the table format. Building on Paivio's dual-code theory (Paivio, 1971, 1986, 2007), my research documents *how* combined presentation formats may support managers in their decision-making processes by activating both the verbal and the non-verbal mental representation systems.

The distinction between a problem-definition phase and a problem-solving phase in Study 2 has revealed the importance of understanding the task before making decisions. All the highest performing subjects made an extra effort to understand the relationships among the decision variables and the result variable during the problem-definition phase. Almost all of them developed additional presentation formats in this phase. XY graphs seemed to be important to understand causal relationships between variables, whereas tables were used to integrate the variables accurately and thus attain a high profit. With a detailed understanding of the relationships among decision variables and the result variable, the highest-performing subjects were almost able to find the optimal solution before they started making decisions. The lowest performers spent most of their time in the problem-solving phase. Their handling of the task was mainly characterised by more or less systematic trial-and-error.

My results illustrate how important it is for decision makers to be able to utilise decision aids in order to increase decision-making effectiveness in the tasks used in my studies. In particular, the subjects' spreadsheet skills seemed to be important to complement the presentation format they received with the format they needed. The subjects in my study were MBA students in their final year. I expected them to know how to handle a spreadsheet. This expectation was only partly confirmed. Some students were not able to use the spreadsheet to support their analyses. Some students with the table format explicitly expressed the need for a graph but did not manage to generate it. Several students with the graph format expressed a need for numeric values, but most of them did not manage to read out accurate values to create a table. Therefore, these subjects were not able to perform the additional calculations they wanted to. On the other hand, students with proper spreadsheet skills were able to modify the original presentation format to include a presentation of the data in a format that suited their needs. This was particularly evident in the graph group, where the analyses show that proper spreadsheet skills were necessary to 'transform' the graph into a table. Thus, my results indicate that spreadsheet skills are important to solve the task accurately.

Finally, my findings illustrate the complementarity of combining verbal protocols, eye-tracking and observational data, thus providing a more complete picture of how the subjects used the presentation formats when handling the tasks. In particular, the eye-tracking data provided additional understanding of the subjects' calculation errors and errors caused by using the wrong variables. Also, as illustrated above, eye tracking improved the understanding of how the various presentation formats were used. Furthermore, eye tracking complemented and supported interpretations of the verbal protocols, for example by enhancing the understanding of situations where calculation errors, rather than a low level of information processing, caused poor results. Similarly, the verbal protocol data were also important to support the interpretations of the eye-tracking data. For example, the eye-movement data for one of the subjects (S2-06) show that he focused his visual attention on an area in the model covering a gap between sales and demand, but the verbal protocol shows that he only used the sales data in his handling of the task.

6.2 Implications for system designers and managers

My research has implications for designers of information systems. Since the introduction of the graphical user interface, there has been increased use of graphical formats to present data. My research shows that the table presentation format is essential. Tables provide a basis for calculations to support the integration of decision variables and result variables, and thus support effective decisions. In computerised systems designed to support managers in real-life decision situations, designers should include both tables and graphs. Particularly XY graphs should be included as they seem to be important for decision makers to enhance their understanding of causal relationships among variables. Spreadsheet software such as MS Excel allows decision makers to create additional presentation formats of their data and is by far the most popular tool for decision support used by managers (see e.g. Watson, 2009). Therefore, data presented on the web or in enterprise systems should be easy to export to a spreadsheet format so that decision makers can use the data in additional analyses. However, as illustrated in my studies, even with education in the use of MS Excel not all decision makers are able to generate the graphs they need from table data. Therefore, designers should include guidance into their managerial decision support systems, for example by adding templates or wizards for creation of XY graphs.

My research has also implications for managers. Managers using the spreadsheet to make additional analyses should know how to programme formulas correctly, among others how to build dynamic models using relative cell references. Study 2 has illustrated how poor spreadsheet skills may result in calculation errors that may seriously damage the outcomes of decision processes.

6.3 Limitations and suggestions for further research

My experiments are exploratory. However, our insights are still limited. Therefore, my studies should be followed up by exploratory experiments with an increased number of subjects in order to enhance our insights and to increase the possibilities to achieve more significant results.

My experimental tasks were not originally designed for conducting eye-tracking research. Use of static stimuli might have simplified and allowed for more analyses of the eye-movement data. Instead of presenting some data in paper-based graphs and tables, all data resources should have been presented on the computer screen so that eye movements on these resources could also have been recorded. These weaknesses are described in more detail in section 4.6.3.

My research has drawn the attention to the fact that research on presentation formats should not only distinguish between broad format categories, such as graphs and tables. More research is needed to enhance the understanding of, for example, how different types of graphs support managers' understanding of their decision environments, and how they support them in the different phases of a decision process. Study 2 indicates that decision makers need XY graphs to improve their understanding of causal relationships between variables. However, there is a possibility that the line graphs were used more extensively than my eye-tracking data indicate. As mentioned in Chapter 4, studies should be designed to test whether eye trackers register when subjects perceive a development in line-graph variable values holistically, or whether such visual perception is outside the foveal vision and not captured by eye trackers.

Furthermore, the tasks in my studies are problem tasks that require basic economic insights. They are closer to real-life decision making than other studies within this area. Still, they only reflect aspects of the complexity and uncertainty that managers must handle in real-life decision situations characterised by global competition and cost pressures. There is a need to study how managers actually use the variety of data presentation formats they are introduced to on the web, on dashboards and other decision aids.

REFERENCES

Amer, T. (1991). An experimental investigation of multi-cue financial information display and decision making. *Journal of Information Systems, 5*(2), 18-34.

Anderson, J. R. (2013). *The architecture of cognition*. Mahwah, New Jersey: Psychology Press.

Anderson, J. R., Bothell, D., & Douglass, S. (2004). Research article eye movements do not reflect retrieval processes limits of the eye-mind hypothesis. *Psychological Science (Wiley-Blackwell), 15*(4), 225-231.

Andrienko, G., Andrienko, N., Burch, M., & Weiskopf, D. (2012). Visual analytics methodology for eye movement studies. *Ieee Transactions on Visualization and Computer Graphics, 18*(12), 2889-2898.

Arunachalam, V., Pei, B. K. W., & Steinbart, P. J. (2002). Impression management with graphs: Effects on choices. *Journal of Information Systems, 16*(2), 183-202.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review, 3*(3), 439-449.

Bertin, J. (1981). *Graphics and graphic information-processing*. New York: Walter de Gruyter.

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research, 3*(4), 233-240.

Bojko, A. (2013). *Eye tracking the user experience - A practical guide to research*. Brooklyn, New York: Rosenfeld Media, LLC.

Bonner, S. E. (1994). A model of the effects of audit tax complexity. *Accounting Organizations and Society, 19*(3), 213-234.

Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management, 31*(2), 191-213.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality & Social Psychology, 42*(1), 116-131.

Cacioppo, J. T., Petty, R. E., & Chuan Feng, K. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197-253.

Campbell, D. J. (1988). Task complexity - A review and analysis. *Academy of Management Review, 13*(1), 40-52.

Choi, K. (2010). Development of a model of group cognitive complexity: A combined view of a group compositional and a group-as-a-whole perspectives. *Seoul Journal of Business, 16*(1), 119-147.

Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). New York: Psychology Press.

Davis, L. R. (1989). Report format and the decision maker's task: An experimental investigation. *Accounting, Organizations & Society, 14*(5/6), 495-508.

Desanctis, G. (1984). Computer-graphics as decision aids - Directions for research. *Decision Sciences, 15*(4), 463-487.

Desanctis, G., & Jarvenpaa, S. L. (1989). Graphical presentation of accounting data for financial forecasting: An experimental investigation. *Accounting, Organizations & Society, 14*(5/6), 509-525.

Dilla, W. N., Janvrin, D. J., & Jeffrey, C. (2013). The impact of graphical displays of pro forma earnings information on professional and nonprofessional investors' earnings judgments. *Behavioral Research in Accounting, 25*(1), 37-60.

Dilla, W. N., Janvrin, D. J., & Raschke, R. (2010). Interactive data visualization: New directions for accounting information systems research. *Journal of Information Systems, 24*(2), 1-37.

Dilla, W. N., & Steinbart, P. J. (2005). Using Information display characteristics to provide decision guidance in a choice task under conditions of strict uncertainty. *Journal of Information Systems, 19*(2), 29-55.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. London: Springer-Verlag.

Dull, R. B., & Tegarde, D. P. (1999). A comparison of three visual representations of complex multidimensional accounting information. *Journal of Information Systems, 13*(2), 117-131.

Edwards, C. (1983). The communication of computer generated management information - A review of the empirical literature. *Accounting & Business Research (Wolters Kluwer UK), 13*(52), 263-272.

Edwards, W. (1992). Discussion: Of human skills. *Organizational Behavior and Human Decision Processes, 53*(2), 267-277.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: processes of judgment and choice. *Journal of Accounting Research, 19*(1), 1-31.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*(3), 37.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis - Rev'd edition: Verbal reports as data*. Cambridge, MA: Bradford Books/ MIT Press.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture and Activity, 5*(3), 178-186.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137*(2), 316-344.

Frownfelter-Lohrke, C. (1998). The effects of differing information presentations of general purpose financial statements on users' decisions. *Journal of Information Systems, 12*(2), 99-107.

Frownfelter-Lohrke, C., & Fulkerson, C. L. (2001). The incidence and quality of graphics in annual reports: An international comparison. *Journal of Business Communication, 38*(3), 337-358.

Fuglseth, A. M. (1989). *Beslutningsstøtte : metode for diagnose av lederes informasjon- og situasjonsoppfatninger*. Bergen: Norges Handelshøyskole.

Fuglseth, A. M. (1999). Research project: The effectiveness of different types of data presentation. Universitetet i Bergen, Department of Information Science.

Fuglseth, A. M., & Gronhaug, K. (1995). Task characteristics and expertise. In G. Kaufmann, T. Helstrup & K. H. Teigen (Eds.), *Problem solving and cognitive processes: a festschrift in honour of Kjell Raaheim* (pp. 513-529). Bergen: Fagbokforlaget.

Fuglseth, A. M., & Grønhaug, K. (2000). *The effectiveness of data presentation forms.* Paper presented at the IAREP / SABE 2000 - Fairness & Cooperation Baden, Vienna/Austria.

Furnham, A., & Thorne, J. D. (2013). Need for cognition: Its dimensionality and personality and intelligence correlates. *Journal of individual Differences, 34*(4), 230-240.

Futkerson, C. L., Pitman, M. K., & Frownfelter-Lohrke, C. (1999). Preparing financial graphics. *CPA Journal, 69*(6), 28.

Gettinger, J., Koeszegi, S. T., & Schoop, M. (2012). Shall we dance? — The effect of information presentations on negotiation processes and outcomes. *Decision Support Systems, 53*(1), 161-174.

Ghani, J., & Lusk, E. J. (1982). Human information-processing research - ITs MIS-design consequences. *Human Systems Management, 3*(1), 32-40.

Glaholt, M. G., & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, & Economics, 4*(2), 125-146.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics, 24*(6), 631-645.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. v. d. W. (2011). *Eye tracking: A comprehensive guide to methods and measures*. New York: Oxford University Press.

Hoppmann, T. K. (2009). Examining the 'point of frustration'. The think-aloud method applied to online search tasks. *Quality & Quantity, 43*(2), 211-224.

Ives, B. (1982). Graphical user interfaces for business information systems. *MIS Quarterly, 6*(5), 15-47.

Jarvenpaa, S. L. (1989). The effect of task demands and graphical format on information-processing strategies. *Management Science, 35*(3), 285-303.

Jarvenpaa, S. L., & Dickson, G. W. (1988). Graphics and managerial decision making: Research based guidelines. *Communications of the ACM, 31*(6), 764-774.

Jarvenpaa, S. L., Dickson, G. W., & DeSanctis, G. (1985). Methodological issues in experimental IS research: Experiences and recommendations. *MIS Quarterly, 9*(2), 141-156.

Kaufmann, G. (1984). Mental imagery in problem solving. In A. Sheikh (Ed.), *International review of Mental Imagery* (Vol. 1, pp. 23-55). New York: Brandon House.

Kaufmann, G. (1988). Problem solving and creativity. In K. Grønhaug & G. Kaufmann (Eds.), *Innovation: A Cross-Disciplinary Perspective* (pp. 87-137). Oslo: Norwegian University Press.

Kaufmann, G. (1991). Problem solving and creativity. In J. Henry (Ed.), *Creative management* (pp. 103-134). London: Sage Publications.

Kelton, A. S., Pennington, R. R., & Tuttle, B. M. (2010). The effects of information presentation format on judgment and decision making: A review of the information systems research. *Journal of Information Systems, 24*(2), 79-105.

Kleinmuntz, D. N., & Schkade, D. A. (1993). Information displays and decision processes. *Psychological Science (Wiley-Blackwell), 4*(4), 221-227.

Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology, 3*(3), 185-226.

Kumar, N., & Benbasat, I. (2004). The effect of relationship encoding, task type, and complexity on information representation: An empirical evaluation of 2D and 3D line graphs. *MIS Quarterly, 28*(2), 255-281.

Larkin, J. H., & Simon, H. A. (1987). Why a aiagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Lucas Jr, H. C. (1981). An experimental investigation of the use of computer-based graphics in decision making. *Management Science, 27*(7), 757-768.

Mayer, R. E. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715-726.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*(4), 484-490.

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312-320.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*(3), 389-401.

Montazemi, A. R., & Wang, S. (1988). The effects of modes of information presentation on decision-making: A review and meta-analysis. *Journal of Management Information Systems, 5*(3), 101-127.

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). *Getting access to what goes on in people's heads? - Reflections on the think-aloud technique*. Paper presented at the NordiCHI.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know - Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston [reprinted by Lawrence Erlbaum in 1979].

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255-287.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. New York London: Psychology Press.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin, 92*(2), 382-402.

Payne, J. W. (1994). Thinking aloud: Insights into information processing. *Psychological Science (Wiley-Blackwell), 5*(5), 241-248.

Payne, J. W., Braunstein, M. L., & Carroll, J. S. (1978). Exploring predecisional behavior: An alternative approach to decision research. *Organizational Behavior and Human Performance, 22*(1), 17-44.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506.

Reisen, N., Hoffrage, U., & Mast, F. W. (2008). Identifying decision strategies in a consumer choice situation. *Judgment and Decision Making, 3*(8), 641-658.

Russo, J. E. (1978). Eye fixations can save the world: A critical evaluation and comparison between eye fixations and other information processing methodologies. *Advances in Consumer Research, 5*(1), 561-570.

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory & Cognition, 17*(6), 759-769.

Sadoski, M., & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Salvucci, D. D., & Anderson, J. R. (2001). Automated eye-movement protocol analysis. *Human-Computer Interaction, 16*(1), 39-86.

Schroder, H. M., Driver, M. J., & Streufert, S. (1967). *Human information processing*. New York: Holt, Rienhart and Winston Inc.

Shaft, T. M., & Vessey, I. (2006). The role of cognitive fit in the relationship between software comprehension and modification. *MIS Quarterly, 30*(1), 29-55.

Simon, H. A. (1960). *The new science of management decision*. New York: Harper.

Speier, C. (2006). The influence of information presentation formats on complex task decision-making performance. *International Journal of Human-Computer Studies, 64*(11), 1115-1131.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences, 34*(4), 771-797.

Stüttgen, P., Boatwright, P., & Monroe, R. T. (2012). A satisficing choice model. *Marketing Science, 31*(6), 878-899.

Suedfeld, P. (2010). *The complexity construct in political pshychology*. Toronto, Canada.

Tractinsky, N., & Meyer, J. (1999). Chartjunk or goldgraph? Effects of representation objectives and content desirability on information presentation *MIS Quarterly, 23*(3), 397-420.

Tuttle, B. M., & Kershaw, R. (1998). Information presentation and judgment strategy from a cognitive fit perspective. *Journal of Information Systems, 12*(1), 1-17.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science (New York, N.Y.), 185*(4157), 1124-1131.

Umanath, N. S., & Vessey, I. (1994). Multiattribute data presentation and human judgment: A cognitive fit perspective. *Decision Sciences, 25*(5/6), 795-824.

Vessey, I. (1991). Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences, 22*(2), 219-240.

Vessey, I. (1994). The effect of information presentation on decision making: A cost-benefit analysis. *Information & Management, 27*(2), 103-119.

Vessey, I. (2006). The theory of cognitive fit - One aspect of a general theory of problem solving. In P. Zhang & D. Galletta (Eds.), *Human-Computer Interaction and Management Information Systems: Foundations* (pp. 141-183). New York: M.E. Sharpe, Inc.

Vessey, I., & Galletta, D. (1991). Cognitive fit: An empirical study of information acquisition. *Information Systems Research, 2*(1), 63-84.

Washburne, J. N. (1927). An experimental study of various graphic, tabular and textual methods of presenting quantitative material. *The Journal of Educational Psychology, 18*(6), 15.

Watson, H. J. (2009). Tutorial: Business intelligence - Past, present, and future. *Communications of the Association for Information Systems, 25*, 487-510.

Wilson, E. V., & Addo, T. B. A. (1994). An investigation of the relative presentation efficiency of computer-displayed graphs. *Information & Management, 26*(2), 105-115.

Wilson, T. D. (1994). The proper protocol: Validity and completeness of verbal reports. *Psychological Science (Wiley-Blackwell), 5*(5), 249-252.

Wood, R. E. (1986). Task complexity - Definition of the construct. *Organizational Behavior and Human Decision Processes, 37*(1), 60-82.

Xu, S., Song, F., Li, Z., Zhao, Q., Luo, W., He, X., et al. (2008). An ergonomics study of computerized emergency operating procedures: Presentation style, task complexity, and training level. *Reliability Engineering & System Safety, 93*(10), 1500-1511.

APPENDICIES

A.1 Task description for the low complexity task

Bonanza A/S

Du har fått sommerjobb i biffrestauranten Bonanza som assistent for restaurantsjefen. På grunn av sykdom har restaurantsjefen overtatt ledelsen av et hotell i København på kort varsel, og du er bedt om å overta den delen av restaurantsjefens jobb som angår økonomiske beslutninger.

Bonanza er en sommerrestaurant knyttet til en fornøyelsespark. Gjestene i restauranten er besøkende i parken. Restauranten er i drift hvert år fra mai til og med september.

I Bonanza serveres biff av høy kvalitet. Gjestene har flere muligheter for å variere menyen, for eksempel velge mellom bakt potet, pommes frites, fløtegratinerte poteter, ris, og de kan velge mellom forskjellige typer grønnsaker, brød, sauser og dressinger. Dessuten kan de avtale hvordan de vil ha biffen tilberedt. For å redusere administrasjonen av restauranten brukes en enhetspris.

Det er flere serveringssteder knyttet til parken, men de er ikke i direkte konkurranse med hverandre. Andre serveringssteder tilbyr hamburgere, pølser, kebab, og. Serveringsstedene er uavhengige av hverandre når det gjelder selve driften. De samarbeider likevel på den måten at barn som ikke vil ha biff, kan gå til et annet utsalgssted og kjøpe for eksempel en hamburger og spise den i restauranten.

I tilknytning til alle utsalgsstedene, også Bonanza, er det salg av fatøl, brus og mineralvann. Prisen på disse drikkevarene er den samme i hele parken, og regnskapet fra salg av drikkevarer holdes atskilt fra middagsserveringen.

Serveringen i restauranten foregår på den måten at gjestene tar kontakt med en av personalet. Hvis det er kapasitet i restauranten, anvises et bord. Hvis det ikke er ledig kapasitet, blir gjestene informert om når det forventes å bli ledig kapasitet, og de får da muligheten for å reservere plass. Restauranten er stor, og det er mulighet for å stenge av deler av den. Kapasiteten i restauranten er derfor ikke bestemt av antall bord, men av antall servitører til å betjene gjestene. Alle henvendelsene blir registrert og brukes som et estimat for etterspørselen.

Den faste staben består av restaurantsjefen og to hovmestre. På kjøkkenet har kjøkkensjefen ansvaret for kvaliteten av driften. Han har fem erfarne kokker med seg og seks kjøkkenassistenter. Restauranten har avtale med husmødre i nærheten om å ta vakter som servitør. Slike vakter avtales for en uke om gangen. Totale utgifter til ett vaktskifte pr. uke, dvs. en servitør på hver vakt hver dag, er kr 13.368.

Bonanza åpner hver dag kl. 10 om formiddagen og er i drift til parken stenger kl. 23. Det tas ikke inn nye gjester etter kl. 22. Det er en topp i antall besøkende mellom kl. 16.30 og kl. 17.30. De ansatte arbeider på skift. Første skift møter kl. 10 og slutter kl. 17.30. Annet skift møter kl. 16 og slutter kl. 23.30, slik at begge skiftene er på vakt når den daglige etterspørselen er størst. Erfaringsmessig fordeler etterspørselen seg ellers jevnt utover dagen. På grunn av at gjestene overveiende er turister, er det jevn fordeling av etterspørselen over hele uken.

Kjøkkensjefen har avtale med faste leverandører om levering av prima oksekjøtt, ferske grønnsaker, osv. På grunnlag av disse avtalene har han satt opp følgende kostnadskalkyle for en biffmiddag i Bonanza:

kjøtt	61.05
andre ingredienser	15.80
totale variable enhetskostn.	76.85

tabell 2 - kostnadskalkyle

Restautantsjefen har drevet restauranten i to år. Han har tatt kurs i bedriftsøkonomi og i bruk av personlig datamaskin. I de to årene har han tatt i bruk regneark til registrering av etterspørsel, inntekter og utgifter. Han har dessuten laget en enkel modell som skal støtte ham ved beslutninger om hvilken pris han skal ta for en middag, og hvor mange servitører og kjøkkenassistenter han skal kalle inn hver uke. Når du overtar, får du tilgang til både data og modellen.

A.2 Task description for the high complexity task

Bonanza A/S

Du har fått sommerjobb i biffrestauranten Bonanza som assistent for restaurantsjefen. På grunn av sykdom har restaurantsjefen overtatt ledelsen av et hotell i København på kort varsel, og du er bedt om å overta den delen av restaurantsjefens jobb som angår økonomiske beslutninger.

Bonanza er en sommerrestaurant knyttet til en fornøyelsespark. Gjestene i restauranten er besøkende i parken. Restauranten er i drift hvert år fra mai til og med september.

I Bonanza serveres biff av høy kvalitet. Gjestene har flere muligheter for å variere menyen, for eksempel velge mellom bakt potet, pommes frites, fløtegratinerte poteter, ris, og de kan velge mellom forskjellige typer grønnsaker, brød, sauser og dressinger. Dessuten kan de avtale hvordan de vil ha biffen tilberedt. For å redusere administrasjonen av restauranten brukes en enhetspris.

Det er flere serveringssteder knyttet til parken, men de er ikke i direkte konkurranse med hverandre. Andre serveringssteder tilbyr hamburgere, pølser, kebab, og. Serveringsstedene er uavhengige av hverandre når det gjelder selve driften. De samarbeider likevel på den måten at barn som ikke vil ha biff, kan gå til et annet utsalgssted og kjøpe for eksempel en hamburger og spise den i restauranten.

I tilknytning til alle utsalgsstedene, også Bonanza, er det salg av fatøl, brus og mineralvann. Prisen på disse drikkevarene er den samme i hele parken, og regnskapet fra salg av drikkevarer holdes atskilt fra middagsserveringen.

Serveringen i restauranten foregår på den måten at gjestene tar kontakt med en av personalet. Hvis det er kapasitet i restauranten, anvises et bord. Hvis det ikke er ledig kapasitet, blir gjestene informert om når det forventes å bli ledig kapasitet, og de får da muligheten for å reservere plass. Restauranten er stor, og det er mulighet for å stenge av deler av den. Kapasiteten i restauranten er derfor ikke bestemt av antall bord, men av antall servitører på vakt, og av kapasiteten på kjøkkenet til å betjene gjestene. Alle henvendelsene blir registrert og brukes som et estimat for etterspørselen.

Den faste staben består av restaurantsjefen og to hovmestre. På kjøkkenet har kjøkkensjefen ansvaret for kvaliteten av driften. Han har fem erfarne kokker med seg.

Restauranten har avtale med husmødre i nærheten om å ta vakter som servitør eller kjøkkenassistent. Slike vakter avtales for en uke om gangen. Totale utgifter til ett vaktskifte pr. uke, dvs. en person på hver vakt hver dag, er kr 13.368 for servitører og kr 12.655 for kjøkkenassistenter.

Bonanza åpner hver dag kl. 10 om formiddagen og er i drift til parken stenger kl. 23. Det tas ikke inn nye gjester etter kl. 22. Det er to pressperioder i løpet av dagen. Første gang mellom kl. 12 og kl. 14. Deretter fra kl. 16 til kl. 20, med en topp i antall besøkende mellom kl. 16.30 og kl. 17.30. De ansatte arbeider på skift. Første skift møter kl. 10 og slutter kl. 17.30. Annet skift møter kl. 16 og slutter kl. 23.30, slik at begge skiftene er på vakt når den daglige etterspørselen er størst. Erfaringsmessig fordeler etterspørselen seg utover dagen som illustrert i tabell 1:

Tidsrom	%-vis fordeling
kl. 10 - 12	5 %
kl. 12 - 14	20 %
kl. 14 - 16	10 %
kl. 16 - 18	30 %
kl. 18 - 20	20 %
kl. 20 - 22	10 %
kl. 22 - 23	5 %
	100 %

tabell 1 – prosentvis fordeling av etterspørselen i løpet av dagen

På grunn av at gjestene overveiende er turister, er det jevn fordeling av etterspørselen over hele uken.

Kjøkkensjefen har avtale med faste leverandører om levering av prima oksekjøtt, ferske grønnsaker, osv. På grunnlag av disse avtalene har han satt opp følgende kostnadskalkyle for en biffmiddag i Bonanza:

kjøtt	61.05
andre ingredienser	15.80
totale variable enhetskostn.	76.85

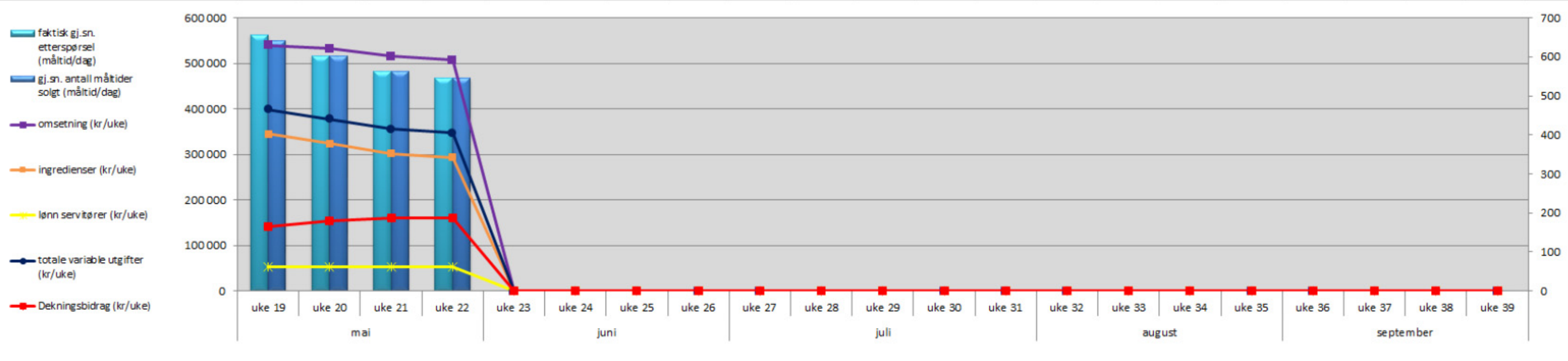
tabell 2 - kostnadskalkyle

Restautantsjefen har drevet restauranten i to år. Han har tatt kurs i bedriftsøkonomi og i bruk av personlig datamaskin. I de to årene har han tatt i bruk regneark til registrering av etterspørsel, inntekter og utgifter. Han har dessuten laget en enkel modell som skal støtte ham ved beslutninger om hvilken pris han skal ta for en middag, og hvor mange servitører og kjøkkenassistenter han skal kalle inn hver uke. Når du overtar, får du tilgang til både data og modellen.

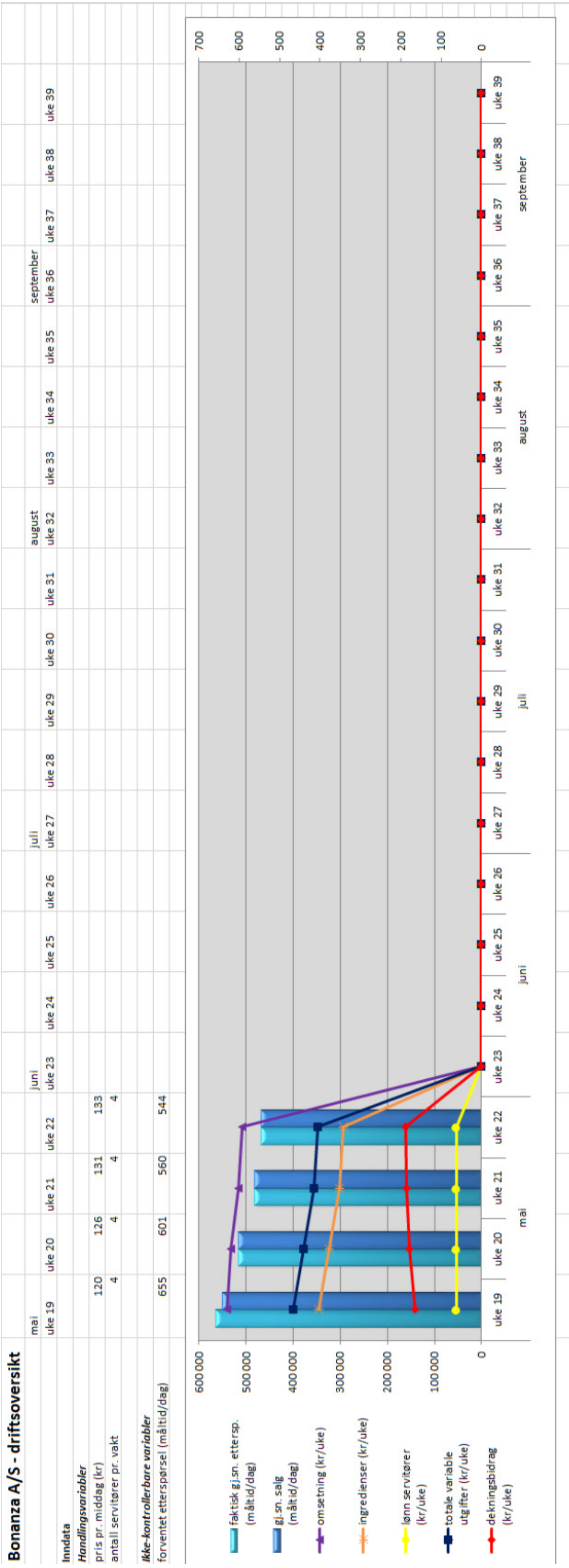
B.1 User interface for the low complexity task, combined presentation format

Bonanza A/S - driftsoversikt

	mai				juni				juli					august				september			
	uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32	uke 33	uke 34	uke 35	uke 36	uke 37	uke 38	uke 39
Inndata																					
Handlingsvariabler																					
pris pr. middag (kr)	120	126	131	133																	
antall servitører pr. vakt	4	4	4	4																	
Ikke-kontrollerbare variabler																					
forventet etterspørsel pr. dag	655	601	560	544																	
Utdata																					
Inntekter																					
faktisk gj.sn. etterspørsel (måltid/dag)	657	603	562	545	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
gj.sn. antall måltider solgt (måltid/dag)	642	603	562	545	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
omsetning (kr/uke)	539 280	531 846	515 354	507 395	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Variable kostnader																					
ingredienser (kr/uke)	345 364	324 384	302 328	293 183	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
lønn servitører (kr/uke)	53 471	53 471	53 471	53 471	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
totale variable utgifter (kr/uke)	398 835	377 855	355 799	346 654	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dekningsbidrag (kr/uke)	140 445	153 991	159 555	160 741	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



B.2 User interface for the low complexity task, graphical presentation format



Bonanza A/S - driftsoversikt

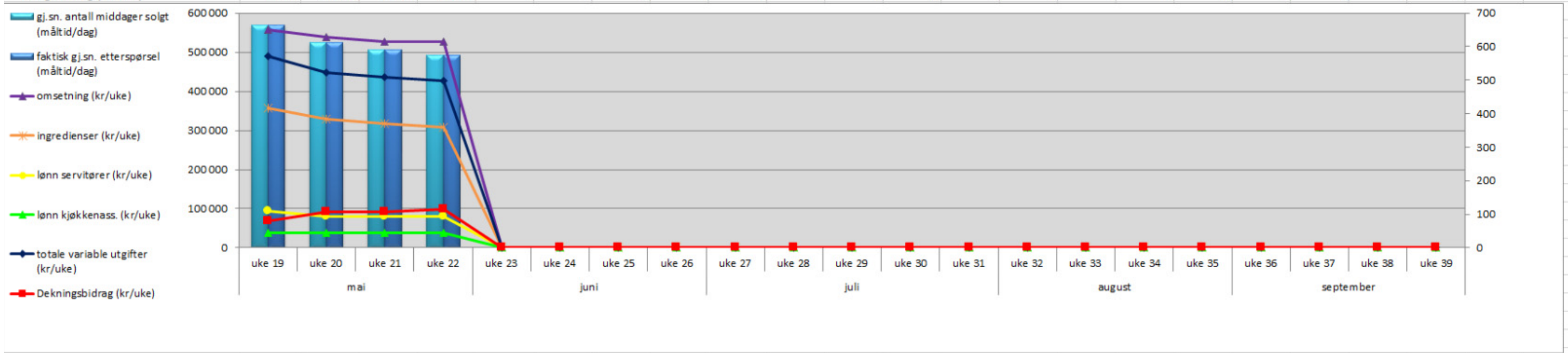
	mai				juni				juli					august				september			
	uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32	uke 33	uke 34	uke 35	uke 36	uke 37	uke 38	uke 39
Inndata																					
Handlingsvariabler																					
pris pr. middag (kr)	120	126	131	133																	
antall servitører pr. vakt	4	4	4	4																	
Ikke-kontrollerbare variabler																					
forventet etterspørsel (måltid/dag)	655	601	560	544																	

B.3 User interface for the low complexity task, tabular presentation format

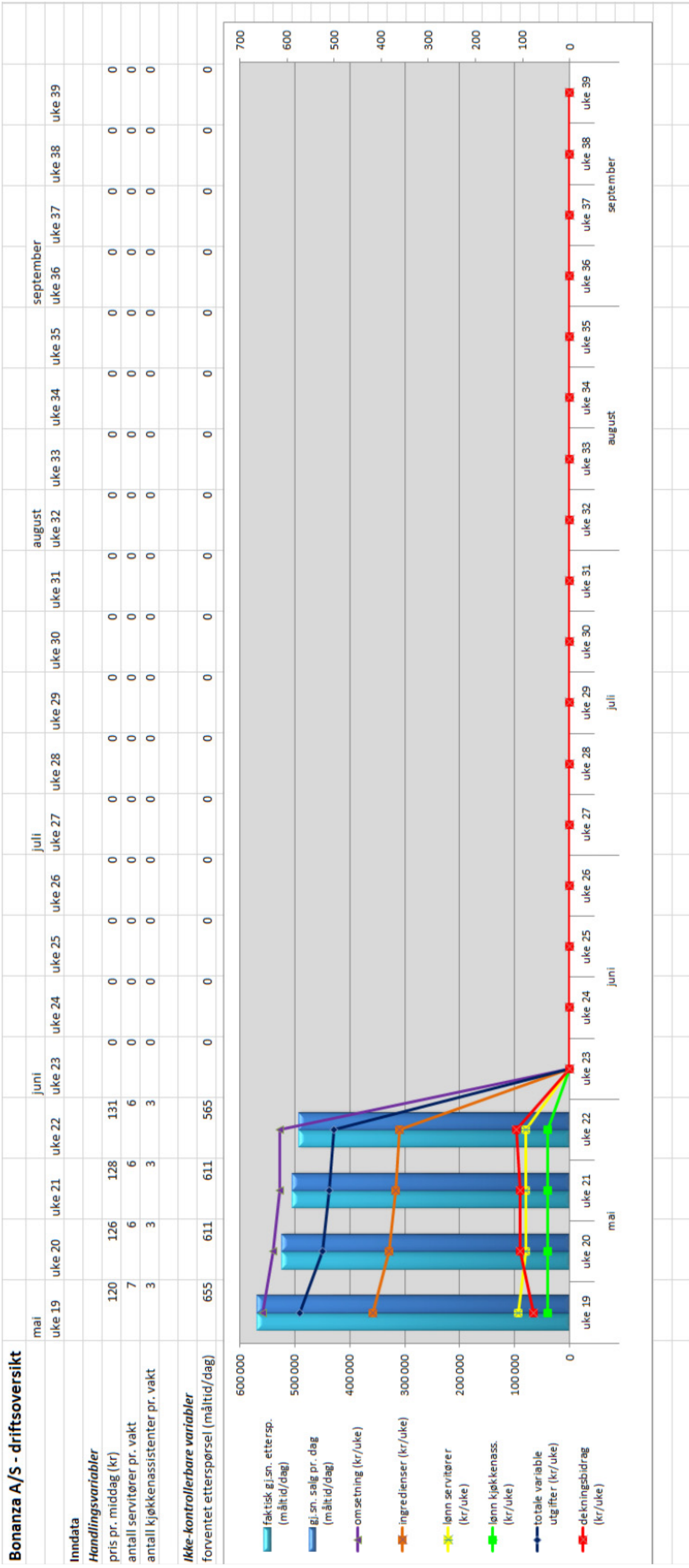
Bonanza A/S - driftsoversikt																					
	mai				juni				juli					august				september			
	uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32	uke 33	uke 34	uke 35	uke 36	uke 37	uke 38	uke 39
Inndata																					
Handlingsvariabler																					
pris pr. middag (kr)	120	126	131	133																	
antall servitører pr. vakt	4	4	4	4																	
Ikke-kontrollerbare variabler																					
forventet etterspørsel (måltid/dag)	655	601	560	544																	
Utdata																					
Inntekter																					
faktisk gj.sn. etterspørsel (måltid/dag)	657	603	562	545	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
gj.sn. antall måltider solgt (måltid/dag)	642	603	562	545	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
omsetning (kr/uke)	539 280	531 846	515 354	507 395	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Variable kostnader																					
ingredienser (kr/uke)	345 364	324 384	302 328	293 183	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
servitører (kr/uke)	53 471	53 471	53 471	53 471	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
totale variable utgifter (kr/uke)	398 835	377 855	355 799	346 654	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dekningsbidrag (kr/uke)	140 445	153 991	159 555	160 741	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

B.4 User interface for the high complexity task, combined presentation format

Inndata																						
Handlingsvariabler																						
Pris pr. middag (kr)	120	126	128	131																		
antall servitører pr. vakt	7	6	6	6																		
antall kjøkkenassistenter pr. vakt	3	3	3	3																		
Ikke-kontrollerbare variabler																						
forventet etterspørsel (måltid/dag)	655	611	611	565																		
Utdata																						
Inntekter																						
faktisk gj.sn. etterspørsel (måltid/dag)	664	612	589	575	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
gj.sn. antall middager solgt (måltid/dag)	664	611	589	575	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
omsetning (kr/uke)	557 760	538 902	527 744	527 275	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Variable kostnader																						
ingredienser (kr/uke)	357 199	328 687	316 853	309 321	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
lønn servitører (kr/uke)	93 575	80 207	80 207	80 207	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
lønn kjøkkenass. (kr/uke)	37 965	37 965	37 965	37 965	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Dekningsbidrag (kr/uke)	69 022	92 043	92 720	99 782	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



B.5 User interface for the high complexity task, graphical presentation format



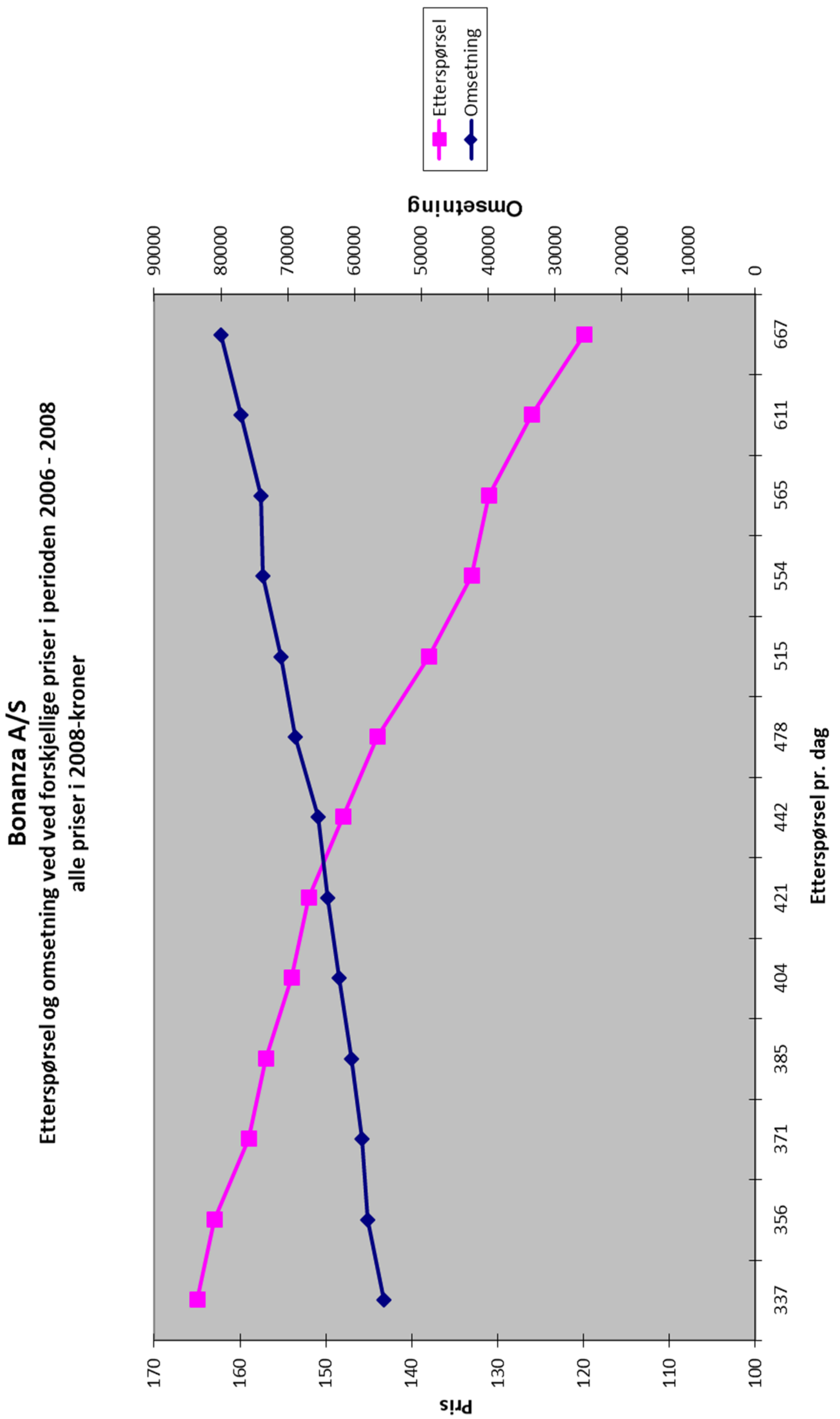
B.6 User interface for the high complexity task, tabular presentation format

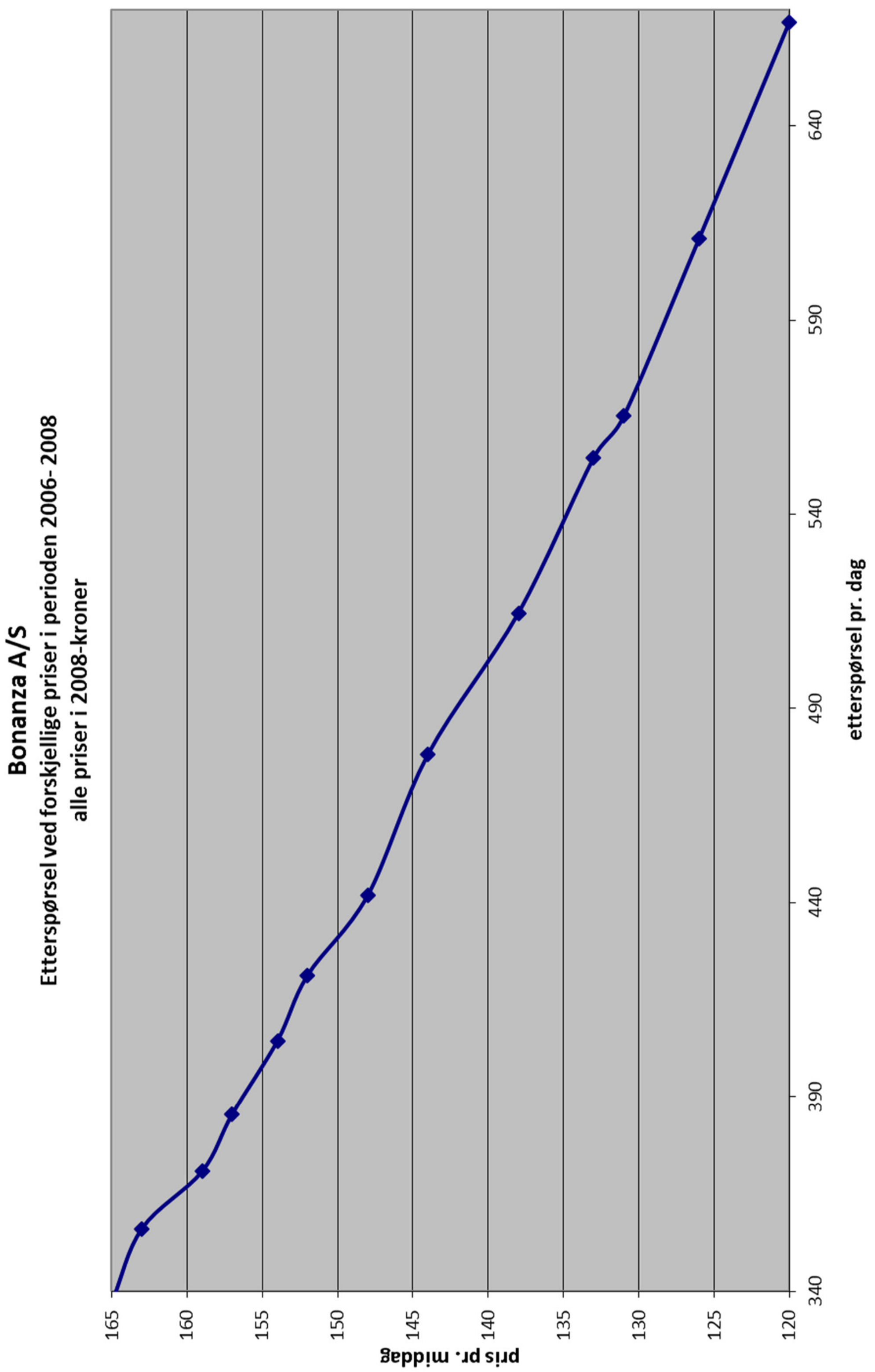
Bonanza A/S - driftsoversikt																					
	mai				juni				juli					august				september			
	uke 19	uke 20	uke 21	uke 22	uke 23	uke 24	uke 25	uke 26	uke 27	uke 28	uke 29	uke 30	uke 31	uke 32	uke 33	uke 34	uke 35	uke 36	uke 37	uke 38	uke 39
Inndata																					
Handlingsvariabler																					
pris pr. middag (kr)	120	126	128	131																	
antall servitører pr. vakt	7	6	6	6																	
antall kjøkkenassistenter pr. vakt	3	3	3	3																	
Ikke-kontrollerbare variabler																					
forventet etterspørsel (måltid/dag)	655	611	611	565																	
Utdata																					
Inntekter																					
faktisk gj.sn. etterspørsel (måltid/dag	664	612	589	575	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
gj.sn. antall måltider solgt (måltid/dag	664	611	589	575	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
omsetning (kr/uke)	557 760	538 902	527 744	527 275	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Variable kostnader																					
ingredienser (kr/uke)	357 199	328 687	316 853	309 321	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
servitører (kr/uke)	93 575	80 207	80 207	80 207	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
kjøkkenassistenter (kr/uke)	37 965	37 965	37 965	37 965	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
totale variable utgifter (kr/uke)	488 738	446 859	435 024	427 493	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dekningsbidrag (kr/uke)	69 022	92 043	92 720	99 782	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

C.1 Historical price-demand data presented as a table

Bonanza A/S - Etterspørsel pr. dag ved forskjellige priser i perioden 2006 - 2008			
alle priser i 2008-kroner			
	pris	etterspørsel	omsetning
	120	667	80005
	126	611	76974
	131	565	74044
	133	554	73745
	138	515	71001
	144	478	68855
	148	442	65395
	152	421	64029
	154	404	62273
	157	385	60516
	159	371	58945
	163	356	58035
	165	337	55611

C.1 Historical price-demand data presented as graphs




Bonanza A/S
Etterspørsel ved forskjellige priser i perioden 2006- 2008
alle priser i 2008-kroner
pris pr. middag
165
160
155
150
145
140
135
130
125
120
etterspørsel pr. dag
340
390
440
490
540
590
640

D Instructions used in the experiments

Instruks til respondenter som deltar i Bonanza-eksperimentet

Introduksjon

Velkommen, takk for at du vil delta i dette eksperimentet.

Din oppgave her i dag er å assistere restaurantsjefen ved Bonanza biffrestaurant, som er en restauranten knyttet til en fornøyelsespark. På grunn av sykdom har restaurantsjefen overtatt ledelsen av et hotell i København på kort varsel, og du er bedt om å overta den delen av restaurantsjefens jobb som angår økonomiske beslutninger.

Restaurantsjefen har høye forventninger til deg, og han regner med at du som økonom kan sørge for at restauranten tjener mest mulig penger hver uke. Det er viktig at du tar oppgaven seriøst og at du gjør jobben så godt du kan. Husk at den som oppnår best resultat for restauranten mottar et gavekort på kr. 1 000,-.

Regnearkmodellen

Når du løser oppgaven skal du benytte deg av denne (viser) regnearkmodellen i Excel. Den viser alle ukene som du skal drive restauranten (mai-september). Du kan anta at etterspørselen etter måltider er den samme gjennom hele perioden.

Som du ser så har modellen en inndatadel og en utdatadel. Handlingsvariablene er de beslutningsvariablene som du skal fastsette hver uke. Resultatet av dine valg vises så i utdatadelen. Restaurantsjefen har allerede drevet restauranten i mai måned, så her ser du resultatet av hans beslutninger.

Når du først har lagt inn verdier for en uke kan du ikke gjøre om valget ditt – du må da gå videre til neste uke hvor du på nytt kan legge inn verdier for beslutningsvariablene. Når du har lagt inn verdier for handlingsvariablene i alle ukene, vil jeg summere resultatet for alle ukene. Den som oppnår høyest totalsum vil motta gavekort pålydende kr. 1 000,-. Dermed er det viktig at du allerede fra starten av finner best mulig pris. Når du har kommet frem til en kombinasjon av pris/antall servitører/assistenter kan du bruke denne kombinasjonen for de resterende ukene.

Du kan selv velge hvilken fremgangsmåte du vil følge for å fastsette inndatastørrelsene. Dersom du trenger penn/papir og kalkulator for å løse oppgaven så har du tilgang til det. Du kan selvsagt også benytte deg av programvaren på datamaskinen for å løse oppgaven, for eksempel dersom du ønsker å sette inn et nytt regneark og gjøre beregninger i Excel. Det er ingen tidsbegrensninger knyttet til oppgaven, du kan bruke så lang tid du ønsker.

Utleveres: Oversikt over etterspørsels- og omsetningstall for de siste to årene. Restaurantsjefen har kommet frem til at det ikke er lønnsomt å selge måltider for mindre enn kr 120,- pr stykk, men ut over det har han ikke noen formening om hva som er optimal løsning.

Tenkte høyt

For at jeg skal forstå hvordan du går frem for å løse oppgaven er det viktig at du hele tiden forteller meg hvordan du tenker for å løse oppgaven. Du trenger ikke å være redd for om det du sier er rett eller galt, eller om du presenterer fullstendige resonnementer. Det er viktig for meg at du hele tiden tenker høyt, slik at jeg kan følge fremgangsmåten for hvordan du løser problemet.

Husk også at det ikke er deg vi tester, men selve opplegget og modellene som vi presenterer for respondentene.

Lyd og bildeopptak

Jeg vil foreta lyd- og bildeopptak mens eksperimentet pågår. Dette er for at jeg i ettertid skal kunne forstå hvordan du gikk frem for å løse oppgaven. Opptakene vil ikke bli gjort tilgjengelig for andre enn de involverte i eksperimentet, verken mens eksperimentet pågår eller etter at det er ferdig.

Registrering av øyeaktivitet (kun aktuelt for Study 2)

Som et ledd i eksperimentet vil vi også registrere dine øyebevegelser. Det vi er interessert i er å se hvordan størrelsen på pupillene endrer seg etter hvert som man utfører en oppgave. Vi ønsker å se om det er noen sammenheng mellom problemløsningsaktivitet og størrelse på pupillene.

For at vi skal kunne registrere øyebevegelsene dine på en korrekt måte, må vi gjennomføre en kalibreringsprosess. Dette for at opptaksutstyret skal tilpasses individuelt til hver deltaker. Deretter er det viktig at du sitter i en viss avstand/vinkel til skjermen når du ser på den. Jeg skal fortelle deg dersom du må endre din avstand/posisjon underveis.

Formål

Vi har utarbeidet en rekke regnearkmodeller som vi er usikre på om fungerer slik vi har forutsatt. Dette ønsker vi å teste i eksperimentet. Vi tester altså ikke deg som person, men snarere om våre modeller fungerer slik vi har tenkt. Basert på våre erfaringer med eksperimentet vil vi foreta justeringer i modellene.

Spørsmål?

E Examples of transcripts (verbal protocols): S1-37 and S2-18

S2-18	*Presentasjonsform*:	**Graf**	*Kompleksitet*:	**Høy**
	Oppnådd resultat:	**1 627 624**	*Prosent av max*:	**68,28 %**

S2-18: Nå skal jeg egentlig snakke til maskinen jeg da?

TV: Ja, tenk høyt. Det spørreskjemaet kan vi ta til slutt. Når du legger inn verdier i regnearket kan du gjerne fortelle hvilken uke du er i og hvilke verdier du legger inn for pris, antall servitører osv.

Skal vi se, nå kan du lukke den der, så får du litt mer ut av skjermen. Da står du fritt til å gjøre akkurat hva du vil, og du har så mye tid som du måtte ønske.

S218: Hvis jeg finner en jeg synes er optimal, så kan jeg kjøre den videre bortover?

TV: Ja, da kan du kopiere og dra bortover.

Her ser du forresten at han har begynt på pris på 120 og så har han gått opp til 165. Det er altså det prisintervallet han har brukt historisk sett. Han har kommet til at han ikke har lyst til å gå lavere enn 120.

...

S218: Jeg kan ikke se noe om antall ansatte her?

TV: Nei, det kan du faktisk ikke.

S218: Nei, så jeg kunne ikke se sammensetningen der.

TV: Nei, det kan du ikke – ikke i de dataene. Det eneste du kan se er her i modellen. De fire ukene han har drevet restauranten. Hvis det kan brukes til noe, det vet jeg ikke men...

S218: Nei, jeg prøver nå jeg, selv om jeg synes det er litt rart å snakke med meg selv.

TV: Forsøk å tenke høyt så godt det lar seg gjøre – og ikke tenk på om det du sier virker fornuftig eller ikke. Og bare spør hvis det er ting du lurer på.

S218: Jada, jeg skal prøve. Jeg er grådig god til prøve og feile metode... Skal vi se...

Tid: 11:12

UKE 23

S218: Ok, jeg prøver med samme pris, men annet forhold, bare se det først... Se om det forholdet som er optimalt... Nei... Men, kan jeg skifte så lenge jeg ikke har trykket Enter, eller er det det når jeg har trykket inn alle variablene?

TV: Når du har tastet inn alle de tre variablene kan du ikke endre. Om du bare har lagt inn f.eks. pris så kan du fortsatt endre.

(P=128, S=7, A=3)

S218: Ja, for jeg har endret nå. Ja… Se det optimale forholdet mellom (?)…

14:16

S218: Nei, det ser ut som det var en dårlig løsning.

TV: Nå må du huske å tenke høyt.

S218: Ja, jeg skal gjøre det. Ja, jeg tenker at jeg er nødt til å se på forholdstallene mellom antall servitører på vakt og kjøkkenassistenter. Jeg ser at i begynnelsen var det mange på vakt, og da var det lave priser, dermed høyere kapasitet, og dermed tjente de mye, men jeg ser at jeg kan ikke ha høy pris og samtidig ha den kapasiteten, det er ikke vits i, da blir den ikke brukt opp. Etterspørselen blir ikke så høy da. Så nå må jeg bare prøve å se hva som er optimal der… (?)…

UKE 24

S218: Sier at det er mer optimalt med forholdet 6 – 3 enn det er med der, likevel, håper at jeg kan ta en høyere pris enn 131 (*dvs 133*) og likevel få… Redusere etterspørselen, men allikevel får jeg inn mer per… Nei, det blir mindre optimalt enn 131… Forholdstallet der…

17:26

S218: Men skal jeg se på DB per uke, eller skal jeg se på… Jeg skjønner ikke, den lilla går jo sakte nedover her.

TV: Skal vi se, hvor er du nå?

S218: Jeg skal ikke se på omsetning sant, jeg skal se på den?

TV: Du skal tjene mest mulig penger, ikke sant.

S218: Nei, jeg bare lurte på hvor, der er den rød, ja, jeg bare så ikke hvor den rød var.

TV: Det går fint.

S218: Ja, jeg bare så ikke at den sammenlignet (?), skulle se på den røde…

Ja, vi får nå se på det. Der var pris der, men må jeg følge de prisene som er her, eller kan jeg ta sånn 129 istedenfor 131?

TV: Ja, det kan du gjøre.

S218: Ja, jeg vet ikke om det har noe for seg.

UKE 25

S218: Skal vi se, 98,99, hvis jeg prøver å ta et tall mellom der i uke 25 siden prisen i uke 22 var… Uke 22 som var 131 var optimalt så langt, så antar jeg at det var riktig forholdstall. Har ikke noe grunnlag for det, men jeg prøver det. Hvis jeg prøver 132 der (*og 6, 3)*, da blir etterspørselen større der… Så avhenger det av kapasiteten min… Da får jeg samme lønnskostnader… Skal vi se… Lønn… Forholdstallene mellom… Skal vi se… Nå får jeg en høyere etterspørsel enn i sted.

Hvis jeg prøver… Ja, den ble høyere. Så det er bedre med 132 enn 131… Eneste jeg kan prøve… Jeg kan endre kapasiteten med forholdstallet her. Det er ikke bedre med 7 enn 6… Hvis jeg reduserer antall servitører så lurer jeg på om det vil ha noe å si for etterspørselen, eller om jeg lekker… Fremdeles har nok kapasitet… Jeg vet ikke hvor mye ledig kapasitet jeg har her allerede.

UKE 26

Skal se om det har noe å si om jeg reduserer antall servitører på vakt. Se om det gir en, om det kan gå…(*132, 5, 3).* Nei. Det optimale er i hvert fall ikke den løsningen. Det eneste jeg kan teste er om det vil være motsatt… Har det noe å si om jeg har motsatt i uke 27…

UKE 27

Eneste jeg kan prøve er om jeg kan ha flere kjøkkenassistenter… Teste, da tar jeg og tester det. (*dvs. 132, 6, 4*)Da antar jeg at det tallet som var, var det optimale der. Antar at det ikke har noe å si om jeg…

UKE 28

Skal vi se hvis jeg øker kapasiteten, om det da vil… Det vil ikke gå der… Gir ikke noe utslag der heller, 6 og 3… (*dvs. 132, 6, 3*)Tror det er mer riktig med 6 og 3 der. 5 og 3 var ikke noe virkningsfullt. Ser at 6 og 3 (?), 5 og 3 gir 90.. Skal vi se, jeg må bare se om omsetningen har vært noe særlig høyere. Skal vi se… Tidligere har den bare vært opp til 90 der… Etterspørselen var oppe i 650, her er etterspørselen… I starten var den mye høyere med 7 og 3… Der er etterspørselen endret, etterspørselen er lik… Avhenger av de som er på jobb, og så høy har jeg hvis jeg… 7 og 3, 612 der…

24:32

S218: Bare ser den ene veien… Selv om prisen virker høy så bare ser jeg etterspørselen… Antar jeg at det er ledig kapasitet, for ellers ville sikkert, her er like mange ansatte og flere, etterspørselen er høyere… Så antar jeg at den utnytter kapasiteten fullt ut, prisen kan ikke være høyere, for da fikk jeg lavere… 131 (?), antar jeg at det er optimalt. Lurer på om det kan være optimalt hver måned. Nå ser jeg at det er lavere enn jeg hadde i uke 22 med 131… Sjekke om det var en tilfeldighet eller om jeg får det en gang til. Vet ikke hvorfor jeg får den, men… Der er jeg oppe igjen i høy. Antar at det bare har noe å si, eller om det er tilfeldigheter som spiller inn. Der er lavere. Hvis det vil si at jeg tar annenhver uke, hvis jeg tar 132 her, 6 og 3… Hvis jeg antar at (?)… Da ble det lavere igjen.

26:12

UKE 31

S218: Skal vi se, nede i 95… Ansetter færre siden etterspørselen er så lav, så jeg at jeg ikke kunne (?)… 5 og 3 der. For nå er jeg nede i 95 og det…

UKE 32

Skal vi se, da prøver jeg med 131 her da… Må jeg ende opp med en sum som er riktig, mest optimalt eller er det…?

TV: Hva mener du?

S218: Nei, jeg har, om jeg kan ha forskjellige, ukentlige…

TV: Ja, du kan prøve i så mange uker du vil, men hvis du nå mener at du har kommer frem den optimale løsningen og ønsker å kjøre på den, så kan du gjøre det.

S218: Ja, jeg synes den er god, men den er ikke optimal i forhold til ting jeg har hatt tidligere, men det trenger ikke ha noe å si…

Der var jeg helt nede i… Sånn som i forhold til den og den og den… Jeg prøvde å se et mønster her da. At det var annenhver, og da tenkte jeg at det kanskje kunne ha noe å si om jeg satte prisen ned annenhver uke og så opp igjen og ned igjen og opp igjen, men det går jo ikke an å holde på å bytte priser annenhver uke.

TV: Det kan du gjøre hvis du ønsker.

S218: Nei, jeg tror ikke jeg tør å gå ned igjen til den. Kan jo være, jeg har jo bare kjørt den en gang…

S218: Jeg prøver å kjøre en gang (?), må bare se om den er høyere generelt. Får med en gang høyere etterspørsel, og det er jo for så vidt positivt, men den hadde blitt høyere om jeg hadde hatt med hele 132, ikke så mye da… Men hvis jeg prøver det forholdet der… 97… Likevel høyere enn den 132en bortsett fra den ene uken jeg hadde 2… Se om det er noen forskjell der eller om det bare er tilfeldigheter. Klarer ikke bli enig helt om den prisen… Den ligger der, hvis den varierer litt… Men jeg har likevel ikke hatt så høy som den jeg hadde der. Jeg tror jeg bare går for hundre og tretti… Det er jevnere… Skal vi se, er egentlig ikke det. 87, 89… (?). Hvis ikke det er en trend hver tredje uke… Skal vi se, høy, høy…

S218: Tar det på tross av at den har vist lavere… Hvis jeg prøver den… Ja, nå gjorde jeg det, dro den bare bort. Mye testing.

TV: Ja men det er greit.

Jeg ønsker at du gir en oppsummering til slutt. Du har jo snakket underveis også, men hvis du kunne oppsummert hva du har gjort i korte trekk – hvordan du gikk fram for å løse oppgaven?

S218: Jeg skulle sikkert regnet litt mer (?)… Ok, jeg har prøvd å se på… Jeg så først på prisene, sjekket bare prisene først, hva som er blitt brukt, og så så jeg på forholdstallene først og fremst da.

TV: Mellom?

S218: Mellom antall servitører og kjøkkenassistenter. Så i begynnelsen var det optimalt med 6 og 3, derfor har jeg prøvd 7 og 3 her da, og så så jeg at det var mindre. Så har jeg testet ut først om, på pris da… I forhold til det forholdstallet som så mest optimalt ut. Og den, (?) Hvor begynte jeg, jeg begynte her, i uke 28… Det var bare for å se forholdstallet da, om det ble dårligere hvis jeg tok en mer. Det ble det, mye dårligere. Så tenkte jeg at det var optimalt forholdstall foreløpig da. Så skulle jeg bare se på pris, og da så jeg at, bare prøvd på 133, da fikk jeg 98, og hvis jeg da prøvde et tall mellom… Samme forholdstall her og her, her og her, og men jeg fikk… Så skulle jeg teste en pris mellom der, hva det ville gi. Det ga meg optimalt i utgangspunktet, pris, at jeg satt igjen med mer penger, høyere DEB. Så skulle jeg bare se, sjekke, så skulle jeg bare se om det hadde noe å si om jeg endret litt på forholdstallene igjen, siden det var mye lavere etterspørsel her, om jeg kunne gå ned i kapasitet med antall servitører og kjøkkenvakter.

TV: I uke 26 nå?

S218: Først i uke 26 sjekket jeg om det hadde noe å si om jeg bare reduserte antall servitører, og så antall kjøkkenvakter om det hadde noe å si, men da ble det mye lavere begge ganger, så da antok jeg hvert fall at dette var det optimale. Så likte jeg egentlig den ganske godt, men så plutselig begynte jeg å tenke her borte, jeg torde ikke trekke den bort med en gang, det kunne, siden det gikk litt nedover plutselig i uke… Det var kjempehøyt i uke 25, men så ble det… Så prøvde jeg her, så var det høyt, men så var det litt varians her. Kanskje 131 summen her borte ble høyere igjen, og kanskje det var mer riktig pris, så testet jeg det noen ganger. Men så så jeg at jeg kom aldri opp på det nivået der. Så tenkte jeg at det kanskje var greiere å ta den. Det var jo for så vidt ikke det da, men hvis du ser 131 der og 132 der, så er den likevel høyere, og da tenkte jeg at kanskje det bare er variasjon generelt, og da bare tok jeg den resten bort, for generelt følte jeg at den var høyere.

TV: OK, det var flott. Du har jo gjort det bedre enn den forrige restaurantsjefen. Han er sikkert fornøyd

S218: Ja, la oss håpe det.

TV: Takk for hjelpen!

S1-37	*Presentasjonsform*: **Tabell**	*Kompleksitet*: **Høy**
	Oppnådd resultat: **2 189 371**	*Prosent av max*: **91,84 %**

S1-37: Etterspørsel pr. dag ja…… Ja, dette blir ikke enkelt… Kan jeg skrive på dette, eller skal jeg bruke et annet ark?

TV: Ja, bare skriv i vei.

S1-37:Det er en klar sammenheng mellom prisen her og antall vakter man trenger.. så setter vi prisen bestemmer vi jo hvor mange kjøkkenassistenter og servitører vi må ha på vakt.. så det viktigste er kanskje å finne prisen.

Kan jeg regne her på dette regnearket?

TV: Ja det må du gjerne. Eller du kan åpne et nytt ark om du vil.

S1-37: Neida, jeg bare gjør det her enkelt, jeg.

TV: Du må huske å tenke høyt.

S1-37: Ja, nå bare regner jeg på et priseksempel her.. en pris på 148 kroner.. for å se hva han har… Skulle hatt et DB her da.. hadde gjort det litt enklere.

Har vi en oversikt over sessongvarisajoner her?

TV: Nei, du kan anta at det ikke er det.. at det er jevnt i hele perioden. Og så har du jo denne oversikten med historiske tall som du kan bruke

S1-37: ja, riktig, riktig.. kan jo bruke denne til å se hva han hadde som tilsvarende tall tidligere i måneden. Om det er stigende med en høyere pris. Kommer forventet etterspørsel?

TV: Nei, det er der du kan skrive inn hva du forventer at den blir..

S1-37: Så da kan jeg bare kjøre på der da.

TV: Nei, den er kun for at du skal kunne legge inn forventningene dine. Etterspørselen blir ikke nødvendigvis det samme.

S1-37: OK.

Burde finne den prisen som har størst DB her da.. så skulle vi hatt en DB her som jeg kunne sett på.. hehe

(lang pause – foretar beregninger i regnearket.)

Ehh.. nå er jo også de kjøkkenassistentene og kelnerne variable kostnader, så man bør jo lage et estimat med antall kelnere/servitører pr kunde for å få et korrekt DB. For kjøkkensjefene og restauranten er vel faste kostnader.. Hvis du har for få ansatte, får du da noen negativ utvikling i modellen.. Du kan tilby dårligere service og få høyere DB.

TV: Det kan jeg ikke si noe om.

S1-37: Hvis vi ikke har med kelnerne så har vi et større DB.

TV: Hva har du satt opp, Du har satt opp en liste over pris, etterpsørsel, DB pr dag.

S1-37: Tallene stemmer jo utrolig bra med modellen da.. med inndata og utdata.. det skal man ha. Her ligger det fast på antall kjøkkenassistenter og antall servitører da.. det har vi vel muligheten til å endre.

Opperer med rimelig lave priser og høy etterspørsel her.. Men hvis vi har en høyere pris.. jeg ser DB er høyere ved en høyere pris. 144 er den prisen jeg har høyest DB på. Og da vil man trenge færre kelnere og assistenter.. så det må jo lønne seg å ha en høyere pris enn de som har vært operert med her. Det er vel den som har best vel.

Er det noen rangeringsverktøy her.. i Excel? Nei, får ta det manuelt.

Skal vi se på disse kelnerne da.. hvor mye det blir..

Så det er ikke noe som tilsier noe om bruk av servitører?

TV: Nei, ikke noe annet enn det som står i teksten.

S1-37: For hvis jeg opperer med en pris på 148 så vil antallet måltider gå betraktelig ned.. så vil kanskje klare meg med færre folk. Ehh.. prøver å finne ut hvor mye det koster å ha servitrisene pr. gjest her. Eller pr. måltid.

Hvor mange dager i uken er det åpent, står det noe om det?

TV: Restauranten holder åpent syv dager i uken.

S1-37: Ikke noe forskjell på dagene?

TV: Nei

S1-37: Da koster det altså, med de tallene som er brukt her, 18,7 kr. Pr. person for å ha kelnere eller servitører.. mens da koster …

For å opprettholde samme servicenivå som i modellen, så forsøker eg å regne ut hvor mye hver kelner jobber med hver kunde. Det var ikke mye..

I modellen har de en servicegrad på 0,009. mens jeg med en pris på 144 som har størt DB, vil ha en servicegrad på 0,01.. så setter ned antall ansatt på jobb. Ja.. Da ble det dårlige service hvis jeg har fire og høyere hvis jeg har fem.

TV: Du må bestemme hvilket servicenivå du vil ha selv.

S1-37: Ja, riktig, det er opp til meg

Jeg går for en høyere pris og burde jo egentlig hatt en høyere service.. Ehh. Den servicen som betyr mest er jo den servicen du får hos servitrisen.. Det på kjøkkenet ser du jo ikke..

Så siden min etterspørsel går ned til 478 fra 656 så tror jeg vi kjører med pris 144, 5 servitriser og to kjøkkenassistenter tror jeg.. da vil vi ha en høyere service på hos personalet hos kunden og litt mindre på kjøkkenet. Og forventet etterspørsel fra tabellen her er 478,.. da knuser jeg dine tall her.. hehe

TV: Dette er første uken i juni

S1-37: Det er bare å kjøre på her hvis jeg fant en modell jeg likte

TV: Ja

S1-37: Bør kanskje forsøke meg litt for å se hvordan det utvikler seg.. Men var det ikke bare en ukes frist på å si i fra hvor mye man vil jobbe.

TV: Det kan du bestemme pr. uke, altså hvor mange du vil ha på vakt.

S1-37: Men jeg har jo ikke noen nye tall å jobbe med.. eller er det det her.. sånn det gikk

TV: Ja, dette viser hvordan det gikk ut fra dine beslutninger

S1-37: OK

Men jeg bestemmer jo selv etterspørselen. Oi, se her da.. det blir jo forskjellig.. eller har jeg tastet feil. Det blir forskjellig..

TV: Ja.

S1-37: Eller er det lagt inn noe lureri her.. hehe.. . Det mangler jo kanskje den viktigste variabelen her.. Korrekt etterspørsel.

Å, der har du faktisk, ja.. den er jo bygget på... den så ikke jeg i det hele tatt. Jeg trodde den var estimert, jeg... OK

Kan jeg spørre hva denne bygger på? Bygger den på servicenivå?

TV: Etterspørselen vil variere litt, og den bygger på flere forhold

S1-37: OK, ja

Har det noe å si hva servicenivået er da?

TV: Det er jo likt for dine uker i juni - så som du ser så har det ikke noe å si egentlig.

S1-37: OK, så servicegraden er ikke lagt inn som en variabel i antallet som kommer, faktisk etterspørsel.?

TV: Jeg kan nok ikke si hvordan modellen er bygget opp!

S1-37: Jeg kan jo se på koden da

TV: Ja, du kan se på formlene i dette regnearket her.

TV: Du begynner på juli nå?

S1-37: JA!

Det er ingen sessongsvingninger?

TV: Neida

05. ok. Jeg er ikke helt fornøyd med utviklingen i DB siste uke. Men det har jo steget... Ehh..

Jeg har jo ikke noen nye tall her da.. Jeg satt opp prisen til 158 kroner.. for å prøve det.. siden det var den prisen som gav størst DB etter.. Opprettholder samme servicenivå, men med en lavere pris. DB ble omtrent det samme - litt lavere enn i de høyeste månedene, litt høyere enn de laveste månedene.

Setter prisen ned til 138 for å prøve å se hvordan det går.. Det beste DB generelt sett.. Det gikk ikke så bra! Går tilbake til 144 kroner, da det gav best DB. Ehh... Estimerer med litt færre kunder enn det som sies i modellen, da antall kunder har vært lavere ved alle observasjoner. 170 ser ut som et bra tall. Prøver igjen med 144 kroner, og får denne gangen et meget godt DB.

Fortsetter med 144 kroner.

Et par ikke så veldig gode uker nå!

Men jeg fortsetter ut måneden med samme pris. Den har gitt best resultater.

TV: Da har du fylt ut alle?

S1-37: Ja, jeg forsøkte meg med litt andre priser som gav et litt annet DB, gikk først opp i pris til 152. Det gav omtrent samme DB, litt dårligere.. men omtrent det samme. Prøvde meg videre på 138, men da fikk jeg et meget skuffende DB.. Dessuten så ble salget da så høyt at det ble for få servitriser og folk på kjøkkenet. Men her er det såpass mye lavere. Når etterspørselen er på 420 så synes jeg det er mer mening å endre på kjøkkenet.. Så må du også se for deg at på kjøkkenet er det en del fast personale som er der hele tiden, så variasjonen blir ikke så stor som det ser ut her.. Den blir ikke 33%.

TV: Hva fant du som optimal pris?

S1-37: 144 gikk jeg tilbake til, for det var den som gav best.

TV: OK. Og den prisen fant du ut i fra beregningene dine her nede i regnearket?

S1-37: ja, men det kan jo være det er feil.

TV: Nei, men dette var greit. Da må jeg bare få lagre her.

Takk for hjelpen!

F.1 Data set for Study 1

No.	Task type	Pres. Format	Electives	Computation	Graph	Analytic	Spatial	Level of info. proc.	Index	Group average
S1_01	L	b	d	3		1		5	0,9993	
S1_02	L	b		2		1		4	0,9698	
S1_03	L	b		3		1		5	0,9998	
S1_04	L	b					1	3	0,9965	0,9914
S1_05	L	g	d	3		1		4	0,9300	
S1_06	L	g	d	3		1		7	0,9954	
S1_07	L	g	d	3	2	1		4	0,9587	
S1_08	L	g					1	3	0,9864	0,9676
S1_09	L	t	d	2		1		4	0,9788	
S1_10	L	t		2	2	1		3	0,9775	
S1_11	L	t		1	2	1		3	0,9801	
S1_12	L	t		3	2	1		5	0,9967	0,9832
S1_13	H	b	d	1		1		4	0,9105	
S1_14	H	b	d	2		1		5	0,9026	
S1_15	H	b	d	1		1		4	0,9743	
S1_16	H	b	d	3		1		7	0,9878	
S1_17	H	b		1		1	1	4	0,7640	
S1_18	H	b					1	3	0,8083	
S1_19	H	b					1	5	0,9240	
S1_20	H	b					1	3	0,8284	
S1_21	H	b					1	4	0,8724	0,8858
S1_22	H	g	d				1	3	0,9028	
S1_23	H	g	d	1		1	1	3	0,5131	
S1_24	H	g	d	2	2	1	1	3	0,7828	
S1_25	H	g	d	2	2	1	1	3	0,8128	
S1_26	H	g	d	3		1		7	0,9358	
S1_27	H	g	d				1	3	0,8530	
S1_28	H	g					1	3	0,9331	
S1_29	H	g					1	2	0,7838	
S1_30	H	g					1	3	0,8061	
S1_31	H	g					1	3	0,8896	
S1_32	H	g					1	2	0,8274	0,8218
S1_33	H	t	d	3	2	1	1	5	0,9843	
S1_34	H	t	d	3	2	1	1	7	0,9663	
S1_35	H	t	d	1		1		3	0,7747	
S1_36	H	t	d	3		1		5	0,9844	
S1_37	H	t	d	2		1		6	0,9184	
S1_38	H	t	d	2		1		4	0,9163	
S1_39	H	t				1		2	0,5918	
S1_40	H	t		1		1		3	0,8394	
S1_41	H	t					1	1	0,6917	
S1_42	H	t		1		1		5	0,8782	0,8546
	H = high	b = both	d = data	1 = simple	1 = line	3 = high	3 = high	1 = low		
	L = low	t = table		2 = compr.	2 = XY	1 = low	1 = low	7 = high		
		g = graph		3 = marg.	3 = XY, compr.					

F.2 Data set for Study 2

No.	Task type	Pres. Format	Electives	Computation	Graph	Analytic	Spatial	Level of info. proc. p.d.	Level of info. proc. p.s.	Index	Group average 1	Group average 2	P.D. min.	Group average P.D. min.	P.S. min.	Group average P.S. min.	P.A. min.	Group average P.A. min.	NFC
S2_04	L	b	d	2	2	1	1	6	6	0,9696			49,00		4,25		53,25		0,8889
S2_26	L	b		1			1	1	2	0,8384			0,97		7,82		8,79		0,2778
S2_32	L	b	d	1		1		3	5	0,9828			5,83		19,85		25,68		0,0000
S2_34	L	b		2		1		6	6	0,9982			26,78		3,92		30,70		0,1667
S2_35	L	b	d	3	3	1	1	6	6	0,9932			12,17		3,47		15,63		-0,2778
S2_42	L	b		2		1		4	4	0,9455	0,9546		10,37	17,52	50,03	14,89	60,40	32,41	-0,0556
S2_06	L	g					1	2	3	0,9381			1,62		58,48		60,10		-0,2778
S2_12	L	g					1	1	3	0,6483			0,87		15,17		16,04		0,1667
S2_31	L	g		1	2	1		4	5	0,9945			33,35		12,15		45,50		-0,1667
S2_36	L	g	d	1		1		5	6	0,9755			24,98		21,63		46,62		-0,1667
S2_37	L	g		3		1		5	5	0,9980			17,68		3,32		21,00		-0,1111
S2_38	L	g		1		1		4	4	0,9922			35,57		7,95		43,52		-0,3333
S2_40	L	g	d	2		1		4	4	0,9419	0,9269		10,10	17,74	8,02	18,10	18,12	35,84	0,0000
S2_30	L	t		1		1		3	3	0,9640			24,18		16,17		40,35		0,1111
S2_33	L	t		2		1		3	4	0,9933			5,50		3,08		8,58		0,2778
S2_39	L	t		3		1		4	4	0,9982			15,08		2,52		17,60		0,1667
S2_41	L	t	d	1		1		3	4	0,9912	0,9867	0,9508	6,20	12,74	17,62	9,85	23,82	22,59	0,2222
S2_02	H	b	d	2		1	1	3	6	0,9017			4,17		46,25		50,42		0,8889
S2_05	H	b	d	3		1	1	5	5	0,9815			10,92		4,00		14,92		-0,3889
S2_07	H	b		1		1	1	1	3	0,8865			2,97		14,18		17,15		-0,1111
S2_10	H	b					1	2	4	0,8635			3,05		14,70		17,75		0,1667
S2_18	H	b					1	1	3	0,6828			0,67		18,33		19,00		0,2222
S2_19	H	b		2		1		5	5	0,8720			20,50		13,00		33,50		-0,1667
S2_23	H	b					1	1	3	0,8216			2,35		19,50		21,85		0,5000
S2_24	H	b					1	1	2	0,5893	0,8249		1,30	5,74	9,07	17,38	10,37	23,12	0,1667
S2_03	H	g		1		1	1	3	4	0,8445			26,00		63,35		89,35		-0,0556
S2_13	H	g	d	2		1		4	5	0,8695			10,20		52,02		62,22		0,5556
S2_15	H	g					1	2	3	0,6721			4,93		3,18		8,12		-0,2222
S2_16	H	g		2		1		4	5	0,8894			6,35		24,00		30,35		0,5000
S2_22	H	g		2	2	1		6	5	0,8320	0,8215		6,05	10,71	58,05	40,12	64,10	50,83	0,2778
S2_01	H	t	d				1	1	3	0,8406			1,83		23,33		25,17		-0,8333
S2_08	H	t	d				1	1	2	0,8387			0,72		17,25		17,97		0,5000
S2_09	H	t		1		1		3	4	0,8909			21,47		3,75		25,22		0,0556
S2_14	H	t		2		1		3	5	0,9766			3,13		12,27		15,40		-0,3889
S2_17	H	t		1		1	1	2	3	0,8063			8,92		17,40		26,32		0,2222
S2_20	H	t		1	2	1	1	3	4	0,8081			30,95		28,38		59,33		0,2222
S2_25	H	t		2	3	1	1	4	5	0,9145	0,8679	0,8391	42,70	15,67	23,88	18,04	66,58	33,71	0,2222
	H = high L= low	b = both t = table g = graph	d = data	1 = simple 2 = compr. 3 = marg.	1 = line 2 = XY 3 = XY, compr.	1= true	1= true	1=low 7=high	1=low 7=high										

F.3 Data set for Study 1 + 2

No.	Task type	Pres. format	Electives	Computation	Graph	Analytic	Spatial	Level of info. proc.	Index	Group average
S1-01	L	b	d	3		1		5	0,9993	
S2-04	L	b	d	2	2	1	1	6	0,9696	
S2-32	L	b	d	1		1		5	0,9828	
S2-35	L	b	d	3	3	1	1	6	0,9932	
S1-02	L	b		2		1		4	0,9698	
S1-03	L	b		3		1		5	0,9998	
S1-04	L	b					1	3	0,9965	
S2-34	L	b		2		1		6	0,9982	
S2-42	L	b		2		1		4	0,9455	0,9839
S1-05	L	g	d	3		1		4	0,9300	
S1-06	L	g	d	3		1		7	0,9954	
S1-07	L	g	d	3	2	1		4	0,9587	
S2-36	L	g	d	1		1		6	0,9755	
S2-40	L	g	d	2		1		4	0,9419	
S1-08	L	g					1	3	0,9864	
S2-06	L	g					1	3	0,9381	
S2-31	L	g		1	2	1		5	0,9945	
S2-37	L	g		3		1		5	0,9980	
S2-38	L	g		1		1		4	0,9922	0,9711
S1-09	L	t	d	2		1		4	0,9788	
S2-41	L	t	d	1		1		4	0,9912	
S1-10	L	t		2	2	1		3	0,9775	
S1-11	L	t		1	2	1		3	0,9801	
S1-12	L	t		3	2	1		5	0,9967	
S2-30	L	t		1		1		*3*	0,9640	
S2-33	L	t		2		1		4	0,9933	
S2-39	L	t		3		1		4	0,9982	0,9850
S1-13	H	b	d	1		1		4	0,9105	
S1-14	H	b	d	2		1		5	0,9026	
S1-15	H	b	d	1		1		4	0,9743	
S1-16	H	b	d	3		1		7	0,9878	
S2-02	H	b	d	2		1	1	6	0,9017	
S2-05	H	b	d	3		1	1	5	0,9815	
S1-17	H	b		1		1	1	4	0,7640	
S1-18	H	b					1	3	0,8083	
S1-19	H	b					1	5	0,9240	
S1-20	H	b					1	3	0,8284	
S1-21	H	b					1	4	0,8724	
S2-07	H	b		1		1	1	3	0,8865	
S2-10	H	b					1	4	0,8635	
S2-18	H	b					1	3	0,6828	
S2-19	H	b		2		1		5	0,8720	
S2-23	H	b					1	3	0,8216	0,8323
	H = high L = low	b = both t = table g = graph	d = data	1 = simple 2 = compr. 3 = marg.	1 = line 2 = XY 3 = XY, compr.			1 = low 7 = high		

No.	Task type	Pres. format	Electives	Computation	Graph	Analytic	Spatial	Level of info. proc.	Index	Group average
S1-22	H	g	d				1	3	0,9028	
S1-24	H	g	d	2	2	1	1	3	0,7828	
S1-25	H	g	d	2	2	1	1	3	0,8128	
S1-26	H	g	d	3		1		7	0,9358	
S1-27	H	g	d				1	3	0,8530	
S2-13	H	g	d	2		1		5	0,8695	
S1-28	H	g					1	3	0,9331	
S1-29	H	g					1	2	0,7838	
S1-30	H	g					1	3	0,8061	
S1-31	H	g					1	3	0,8896	
S1-32	H	g					1	2	0,8274	
S2-03	H	g		1		1	1	4	0,8445	
S2-15	H	g					1	3	0,6721	
S2-16	H	g		2		1		5	0,8894	
S2-22	H	g		2	2	1		5	0,8320	0,8423
S1-33	H	t	d	3	2	1	1	5	0,9843	
S1-34	H	t	d	3	2	1	1	7	0,9663	
S1-35	H	t	d	1		1		3	0,7747	
S1-36	H	t	d	3		1		5	0,9844	
S1-37	H	t	d	2		1		6	0,9184	
S1-38	H	t	d	2		1		4	0,9163	
S2-01	H	t	d				1	3	0,8406	
S2-08	H	t	d				1	2	0,8387	
S1-40	H	t		1		1		3	0,8394	
S1-41	H	t					1	1	0,6917	
S1-42	H	t		1		1		5	0,8782	
S2-09	H	t		1		1		4	0,8909	
S2-14	H	t		2		1		5	0,9766	
S2-17	H	t		1		1	1	3	0,8063	
S2-20	H	t		1	2	1	1	4	0,8081	
S2-25	H	t		2	3	1	1	5	0,9145	0,8768
	H = high	b = both	d = data	1 = simple	1 = line			1 = low		
	L = low	t = table		2 = compr.	2 = XY			7 = high		
		g = graph		3 = marg.	3 = XY, compr.					

G Questionnaire

Generelle opplysninger

RespondentNr: _____ Alder: ____

Kjønn: Mann: ☐ Kvinne: ☐

Til vanlig bruker jeg:

- Briller ☐
- Kontaktlinser ☐
- Verken briller eller kontaktlinser ☐

I dag bruker jeg:

- Briller ☐
- Kontaktlinser ☐
- Verken briller eller kontaktlinser ☐

Kurskode/Tittel for valgfrie fag:

Valgfag I: ____________________ VOA I: ____________________

Valgfag II: ____________________ VOA II: ____________________

Valgfag III: ____________________

Valgfag IV: ____________________

Karakter MET030 skole: _____ Hjemme: _____

Bruk av regneark

		Passer svært dårlig				Passer svært godt
1	Jeg bruker regneark jevnlig til å utføre beregninger	1	2	3	4	5
2	Jeg er vant til å utarbeide grafer og diagrammer i Excel	1	2	3	4	5
3	Jeg er vant til å lese grafer	1	2	3	4	5
4	Jeg er generelt flink til å bruke regneark for å løse økonomiske problemstillinger	1	2	3	4	5

Problemløsningsstil – Need for Cognition

Under finner du 18 utsagn som beskriver forskjellige måter å tenke på når man løser problemer. Noen av utsagnene vil du umiddelbart oppfatte som meget gode beskrivelser av din tenkemåte. Andre vil være meget lite beskrivende og ikke passe i det hele tatt. Atter andre vil oppfattes som moderat gode eller dårlige beskrivelser. Når du skal ta stilling til utsagnene, prøv å tenke deg hvilke tenkemåter du *pleier* å bruke. Du skal altså vurdere om hver av setningene nedenfor *mest typisk* eller *til vanlig* beskriver dine oppfatninger, preferanser og atferd. Sett en sirkel rundt det *ene* tallet som står i den kolonnen som passer *best* for deg. Det er viktig at du besvarer alle spørsmålene.

	Passer svært dårlig				Passer svært godt
1. Jeg foretrekker komplekse fremfor enkle problemer	1	2	3	4	5
2. Jeg liker å ha ansvar for situasjoner som krever mye tenking	1	2	3	4	5
3. Tankevirksomhet er ikke det jeg synes er mest gøy	1	2	3	4	5
4. Jeg gjør heller noe som krever lite tankearbeid, fremfor noe som virkelig utfordrer mine evner til å tenke	1	2	3	4	5
5. Jeg prøver å forutse og unngå situasjoner hvor det er sannsynlig at jeg må tenke grundig omkring noe	1	2	3	4	5
6. Jeg finner det tilfredsstillende å fundere og gruble lenge og grundig på problemer og oppgaver jeg kan løse	1	2	3	4	5
7. Jeg tenker ikke dypere på et problem enn det situasjonen krever	1	2	3	4	5
8. Jeg foretrekker å tenke på mindre, daglige prosjekter framfor langsiktige og større prosjekter	1	2	3	4	5
9. Jeg liker oppgaver som krever lite tankearbeid når en først har lært å gjøre oppgavene	1	2	3	4	5
10. Ideen om å bruke min intellektuelle kapasitet til å komme meg til topps appellerer til meg	1	2	3	4	5
11. Jeg setter stor pris på oppgaver som går ut på å finne nye løsninger på problemer	1	2	3	4	5
12. Å lære seg nye måter å tenke på fascinerer meg ikke i særlig stor grad	1	2	3	4	5
13. Jeg foretrekker at livet mitt er fylt med oppgaver og "puzzles" som jeg må løse	1	2	3	4	5
14. Abstrakt tenking appellerer til meg	1	2	3	4	5
15. Jeg foretrekker en oppgave som er intellektuell, vanskelig og viktig, fremfor en som i noen grad er viktig, men som ikke krever mye tankearbeid	1	2	3	4	5
16. Jeg føler lettelse mer enn tilfredsstillelse etter at jeg har løst en oppgave som krever mye mental innsats	1	2	3	4	5
17. For meg er det nok at noe fører til at jobben blir gjort; jeg bryr meg ikke om hvordan og hvorfor det virker	1	2	3	4	5
18. Jeg ender ofte opp med å fundere og gruble over ting, selv om de ikke angår meg personlig	1	2	3	4	5

H Observation form

OBSERVASJONSSKJEMA

RespondentNr:

Modellbetegnelse:

Tidsbruk:

Problemdefinering	
Problemløsing	

Presentasjonsbehov
Respondenten indikerer behov for:

Visualisering/graf	
Tall/tabell	

Grafer:

Linjegraf	
XY-enkel	
XY-utvidet	

Beregninger:

Enkel	
Utvidet	
Omfattende	

Økonomiforståelse:

Behersking av regneark:

Fremgangsmåte for løsing:

Prøv/feil vilkårlig	
Prøv/feil fra erfaring	
Optimaliserer på kapasitetsbegrensninger	
Annet	

Progresjon:

Syn:

Bruker linser	
Bruker briller	
Verken briller eller linser	

Dominant øye:

Venstre	
Høyre	

Generelle observasjoner:

Forklaring til innholdet i skjemaet:

Presentasjonsbehov

- Brukes dersom respondenten uttrykker behov for et annet presentasjonsformat enn det han/hun har tilgjengelig.

Modellbetegnelse

- Angir hvilken modelltype som benyttes i eksperimentet: BEV=Enkel verbal, BEG=Enkel graf, BEB=Enkel graf og tabell, BKV=Kompleks verbal, BKG=Kompleks graf, BKB=Kompleks graf og tabell

Tidsbruk

- Problemdefinering – tiden respondenten bruker på å forstå/sette seg inn i modellen, hva det spørres om osv.
- Problemløsing – Tiden respondenten bruker for å kjøre modellen (fatte ukentlige beslutninger)

Grafer

Dersom respondenten utarbeider egne grafer, indikeres dette her. Detaljgraden på grafene er rangert på en skala fra 1-3.

- Verdien 1 er brukt ved svært enkle linjegrafer, uten detaljopplysninger
- Verdien 2 er brukt dersom respondenten har utarbeidet en XY-graf, der enkelte verdier for X og Y er plottet inn, for så å tegne en linje mellom disse.
- Verdien 3 er benyttet ved svært detaljerte XY-grafer. Her har respondenten utarbeidet skalaer på X- og Y-aksen, og brukt grafen til å lese av verdier bestemte punkter.

Beregninger

Dersom respondenten har utført beregninger, enten i regnearket, på papir eller ved hjelp av kalkulator, er dette angitt i denne kolonnen. Omfanget av beregningene er klassifisert på en skala fra 1 til 3.

- Verdien 1 indikerer at respondenten har utført svært enkle beregninger (for eksempel å legge sammen to tall).
- Verdien 2 er brukt dersom respondenten har gjort utstrakt bruk av beregninger i sitt arbeid for å komme fram til sin optimale løsning. Her har respondenten gjerne satt opp en tabell med ulike priser, for så å beregne dekningsbidrag for disse prisene.
- Verdien 3 er brukt for å markere at respondenten har utført til dels avanserte beregninger. Her er det gjennomført grensebetraktninger med hensyn til antall servitører (og kjøkkenassistenter) som er nødvendig å ha på vakt ved de ulike etterspørselsalternativene. Videre kan det være utført en marginalbetraktning for å se hvordan dekningsbidraget blir påvirket av en enhetsøkning i prisen.

Økonomiforståelse

- Gi din subjektive vurdering av respondentens økonomiforståelse. Maksimeres dekningsbidraget? Sees sammenhenger mellom kostnader og inntekter og endringer i dekningsbidrag?

Behersking av regneark

- Gi din subjektive vurdering av respondentens bruk av regnearket. Evne til å utvikle formler, modeller, grafer etc.

Fremgangsmåte for løsing

- Prøv/feil vilkårlig – Brukes dersom respondenten ikke ser trender i sine forsøk
- Prøv/feil erfaring – Brukes når respondenten lærer av tidligere erfaringer
- Optimaliserer på kapasitetsbegrensninger – Finner kapasitetsbegrensninger for servitører (og evt. Kjøkkenassistenter). Setter pris slik at etterspørselen møter kapasitetsbegrensningene
- Annet – Ved annen fremgangsmåte beskrives denne her.

Progresjon

- Her indikeres det om respondentens forståelse for løsing av oppgaven utvikles i løpet av perioden.

Syn

- Kryss av for alternativet som passer for respondenten på tidspunktet eksperimentet ble utført.

Dominant øye

- Kryss av for dominant øye. Dominant øye kan finnes ved følgende fremgangsmåte: Peke på et punkt et stykke unna. Lukk venstre øye. Dersom du må flytte fingeren for fortsatt å peke på punktet når venstre øye lukkes, er ditt venstre øye dominant. Dersom synsbildet ikke forandrer seg, har du et dominant høyreøye.