Executive functions as a predictor of a person’s performance in information modelling

Danny Romme

April 17, 2014
Contents

1 Introduction 1

2 Current state of affairs in literature 3
  2.1 What is modelling? 3
  2.2 Skills needed for proper modelling 5
  2.3 Mental processes involved in performing any task 6
  2.4 Most important executive functions 7
  2.5 Measuring executive functions 8
  2.6 Intervening factors 9

3 The relation between modelling skills and executive functions 13
  3.1 Hypothesis H1 13
  3.2 Hypothesis H2 15
  3.3 Hypothesis H3 15
  3.4 Hypothesis H4 16
  3.5 Testing the hypotheses 17

4 Observations of modelling sessions 19
  4.1 Materials 19
  4.2 Procedure 19
  4.3 Results and interpretation 20
    4.3.1 Results for H1 21
    4.3.2 Results for H2 23
    4.3.3 Results for H3 25
    4.3.4 Results for H4 27
    4.3.5 Other results from the gathered data 28

5 Conclusions and recommendations 31
  5.1 Conclusive answer to the research question 31
  5.2 Recommendations for further research 32
Chapter 1

Introduction

In Information Sciences it is a key competence to be able to create correct abstract models to tackle complex real-life problems. It is known what makes a model a correct model, but not why some people can become good modellers and others cannot. As a pioneer and leader in the field Ilona Wilmont tries to find the answer in the level of development of an individual’s mental processes. To support her research on this subject I have decided to look into the relationship between modelling skills and a specific mental process: executive functions. Is there a relationship between the level of development of executive functions and modelling skills? And if so, can executive functions be used to predict an individual’s ability to create correct models?

First I will give an overview of the important literature related to this research. Based on this literature I will then state what relationships I expect to find between executive functions and modelling skills. After that an overview of the experimental observations is given, and the gathered data will be discussed. Finally an answer to the main research questions asked above is composed.
CHAPTER 1. INTRODUCTION
Chapter 2

Current state of affairs in literature

The current research is based upon knowledge and theories from literature of different fields of science. On one hand there are the modelling skills known from information science, on the other hand there are the executive functions from psychology with a hint of neuroscience. This chapter is aimed at giving the reader a basic understanding of the concepts and ideas used in this research.

First I will explain what is meant by modelling, and what skills are needed to be a modeller. After that I will explain what mental processes are involved in performing tasks, why executive functions are important, and how to measure the level of development of a person’s executive functions. The last paragraph will discuss intervening factors that can have an effect on modelling skills, executive functions, and more.

2.1 What is modelling?

Modelling is the act of creating an often simplified, or more abstract, representation of something. In the field of Information Technology a model is often called a conceptual model, as it is a model of ones thoughts and ideas or a model of observed behaviour of a system or business process. Usually such models consist of a number of concepts and relations, representing real-world constructs and their supposed interactions with each other. For a more in-depth insight in modelling I have included a summary on this subject from Wilmont et al. (2012).

“[Modelling] requires flexible switching between concrete and abstract representations, and shifts in focus on certain relevant properties. The first type of switching could be considered ‘vertical’: shifting up and down between abstraction levels. The second type of switching is ‘horizontal’, in the sense that even though the focus shifts, the same concept is still being considered. Such abstraction strategies require executive control processes.

Vertical shifting happens in terms of generalisation and instantiation. When viewing abstraction as generalisation, abstraction is a byproduct of concept learning through simple mechanisms of generalisation from examples. When generalising from concrete information, it is shown that elaborate abstract ideas
develop which only partly include the information provided. Experience contributes considerably to the formation of an abstract idea. However, generic properties in modelling are useless if they are not associated with objects in the modelling domain. In order to instantiate, concrete knowledge must be in place. This means a model must be meaningful; it must be clear what the evolving model means in terms of the activities and objects a modeller or user encounters in his daily environment. Sometimes the connection between concrete and abstract goes awry. Usually, people know only partly how things like the complex systems they work with function, making it very hard to make explicit procedural or tacit knowledge. Such concrete knowledge gaps can result in faulty conceptions of a domain, a failure to monitor understanding and progress, and thereby hamper success at problem solving.

Horizontal switching essentially means taking a focused perspective on a certain concept given a certain context. In other words, relations can be made between concepts based on those properties relevant for the situation, otherwise they may be ignored. However, omitted information is never forgotten; at all times it is assumed to be in place, correctly functioning and most importantly, exercising its influence in the process under study. Those concept features considered in combination must be able to generate a coherent description of the concept, but it is highly context-sensitive. In a dynamic environment, features acquire their meaning and mutual relations according to the context. If the context shifts, so do the features that have to be considered and they may acquire totally new meanings. Awareness of the match between context and meaning must be carefully monitored.

In practice, people reason on a certain ‘preferred level of abstraction’, which appears to be mostly context-dependent, fluctuates widely over time and is independent of the capacity to abstract and of general intelligence. The precise mechanisms of why a preferred level of abstraction fluctuates and what drives it are still unclear, but one could argue that this level might be the easiest for working memory to grasp: a balance between the level of detail and the attention needed to see the common and relevant properties of a concept. At whatever level of abstraction, details should be sufficient to assign meaning to the abstraction, without requiring excessive attention to fill in unknown blanks. Deciding whether a preferred level of abstraction is particularly suited to a certain line of reasoning appears to be related to monitoring ability. However, if concepts turn out to be too abstract to understand, unconscious reduction of abstraction level takes place. Several strategies are employed for this: retreating to familiar mental structures, relying on fixed procedures while working, and discussing specific examples rather than a whole set.

The main forms of abstractions we encounter during modelling are concepts and relations. Essentially, every abstraction in modelling can be explained in terms of concepts and the relations between them. However, in order to assign meaning to an abstract concept, a vast body of concrete knowledge is required, both of the actual meaning of the concept, as well as the context(s) in which the concept might occur.

There have been many attempts to formally define a concept, and this definition has been subject to many changing insights over the years. Generally speaking, concepts are what enable us to interpret situations in terms of previous situations that we judge as similar to the present. [1]In earlier times the primary function of concepts was thought to be the identification of items as
belonging to a particular class, but it seems more likely that they participate in the generation of meaning. Concepts acquire meaning in relation to other concepts and activities. This also means that concepts occur only in concrete, specific situations, for without the information contained in the situation, the concept would not be able to acquire any meaning.

Relations describe [how] concepts [... ] influence each other. The analytical nature of relations illustrates that without a thorough, concrete understanding of the concepts, relations do not become truly meaningful. Hence, concepts and relations depend on each other to acquire meaning. In a model, relations determine the position of different concepts. Relations may be of a different complexity, where relational complexity is determined by the number of related dimensions that need to be considered jointly to arrive at the correct solution.”

2.2 Skills needed for proper modelling

Now it is described what modelling is, it is also important to know what skills a person should have to become a decent modeller. In cooperation with an expert in the field of Information Sciences, combined with Wilmont et al.’s view on modelling above and skills identified in the article of Frederiks and Van der Weide (2006), I identified the following valuable skills (numbered for easier reference later on):

1. Abstract and induce. A problem domain is usually full of irrelevant details which should be separated from the essential information that does contribute to the understanding of the problem. This is where abstraction comes in. Induction is used when one uses examples in the abstracted information to form ideas on similar concepts in the problem domain.

2. Structure and classify information. Structuring and classifying information into groups of similar concepts makes it easier to make early statements about these concepts. Differences between the grouped concepts can later be addressed when more is known about these concepts and their role within the problem domain.

3. Having an open mind. Even once the structure of (a part of) the problem domain has been perceived, one should still be receptive to new ideas and information that can change the perceived structure of the problem domain. If not, the resulting description of the problem domain will inherently be flawed.

4. Handle implicit knowledge. On the same not as the previous skill it is important to be able to handle holes in the problem description by the domain expert. Domain experts tend to leave out (unconsciously) those aspects in the application domain which are experienced as generally known and daily routine. In order to detect implicit knowledge, the system analyst should not take things for granted.

5. Keep an overview on progress towards goals. There is no doubt a limitless amount of things can be done with the information in a problem domain,
but only a few lead to a model of said problem domain. To ensure that working towards that model is done efficiently, one should always keep the goal in mind and try to minimise divergence from the path leading to that goal.

6. Keep an eye on the whole instead of only one piece. This does not mean the individual pieces cannot be taken apart from the rest for a more in-depth view, but one should always consider the related pieces from the whole to view the particular piece in the right context.

7. Skilled communicator. The information from the problem domain is mostly received through interaction with a domain specialist. It is key to make sure there are no miscommunications between the modeller and this specialist, as any miscommunication can lead to an incorrect understanding of the problem domain, which in turn will lead to a flawed model.

8. Describe thoughts and ideas. The model of the problem domain is initially formed in one’s mind. It does not matter how good this model is if one is not able to transfer this model from one’s mind onto paper. Only when a model is put on paper can it be useful to others.

2.3 Mental processes involved in performing any task

Besides looking at what skills are needed for modelling, it is also useful to take a look at what mental processes are involved in carrying out any task at all, before anything can be said about the connection between mental processes and skills needed for modelling. Recent literature can be seen agreeing ever more on there being some kind of central control mechanism in the brain that steers one’s actions. This mechanism is commonly referred to as the ‘executive control’ or ‘central executive’ and is assumed to be “an attentional control system of limited processing capacity that has the role of controlling action” (Baddeley and Hitch, 2010). The ‘controlling’ that is referred to here is either the automatic control of routine and habitual behaviour, or the active control involving long-term knowledge to set up and reflect on possible solutions to a problem. Executive control and working memory cooperate closely in the latter case, with working memory being a support to executive control. What sets working memory apart from short term memory is that working memory not only stores information but is also capable of manipulating, or processing, this information. In literature the term working memory is used most frequently to refer to “a limited capacity system that is capable of briefly storing and manipulating information involved in the performance of complex cognitive tasks” (Salthouse and Babcock, 1991; Baddeley and Hitch, 2010).

As can be seen from the definition of executive control there is also some form of attention involved. According to Wilmont et al. (2012), “[a]ttention allows one to concentrate on a certain set of objects, or tasks, in favour of others. It works to translate goals into behaviour, by orienting an individual towards goal-relevant information, and also supervises this process in terms of executive control. Attentional mediation is thought to be a relatively stable personality trait, since it depends on central nervous system characteristics and
functioning. It is shown that selectivity of attention is a significant predictor of capacity to abstract, in situations in which people are faced with multiple competing incoming stimuli."

This means that the main mental processes that can be distinguished are executive control, working memory, and attention.

### 2.4 Most important executive functions

The three mental processes discussed above (executive control, working memory, and attention) closely work together during any given task. For this research, however, I will only focus on executive control. That is because this control mechanism is not some kind of black box, but a collection of individual functions controlling action in their own specific way. Although there is some disagreement about the exact content of these executive functions and therefore also about their names, there is significant overlap between the distinctions made by different authors. These are the most important and well-described executive functions in modern literature:

**Inhibition.**
Inhibition helps to overcome automatic or routine behaviour and action (Shallice and Burgess, 1993; Miyake et al., 2000). It enables people to inhibit prepotent responses, to resist or delay an impulse, and to stop one’s own action or thought (Gioia and Isquith, 2004; Brocki et al., 2010). Also known as Response Inhibition or Motor Inhibition.

**Task Initiation.**
Task initiation is the ability to begin a task or activity, or the process of generating ideas or problem-solving strategies (Gioia and Isquith, 2004) without undue procrastination, in an efficient or timely fashion (Dawson and Guare, 2010). Also known as Initiate.

**Evaluation.**
Evaluation is the ability to stand back and take a bird’s eye view of yourself in a situation. It includes self-monitoring and self-evaluative skills (Dawson and Guare, 2010), like checking one’s own actions during or shortly after finishing a task to assure appropriate attainment of a goal (Gioia and Isquith, 2004). Also known as Self-Monitoring or Metacognition.

**Flexibility.**
Flexibility is the ability to shift between tasks, operations, or mental sets (Miyake et al., 2000), to alter the problem-solving strategy during complex tasks, to switch or alternate attention, and to think flexibly (Gioia and Isquith, 2004). The latter enables people to revise plans in the face of obstacles, setbacks, new information, or mistakes (Dawson and Guare, 2010). Also known as Shift, Shifting, Attention Switching, or Task Switching.
2.5 Measuring executive functions

Although there are numerous tests that claim to be able to measure executive functions, these can not be measured directly. The reason why is aptly formulated by Gioia et al. (2008): “Given the central importance of the executive functions to the direction and control of dynamic ‘real world’ behaviour, reliance on clinic-based test performance measures potentially can yield a limited, incomplete assessment. While performance tests attempt to tap executive functions in explicit and specific ways, multiple confounds can limit their ecological validity and generalizability. Burgess (1997) argues that neuropsychological tests alone are inadequate for assessing executive functions because they artificially and ambiguously fractionate an integrated system. Performance-based measures tap individual components of the executive function system over a short per-time frame and not the integrated, multidimensional, relativistic, priority-based decision making that is often demanded in real-world situations.

The structures and interactive nature of the typical assessment situation may reduce demands on the executive functions, and thereby reduce the opportunities to observe critical processes associated with the executive functions. That is, in many testing situations, the examiner provides the structure, planning, organization and guidance as well as the cueing and monitoring necessary for an individual’s optimal performance. In this manner, executive control is provided by the examiner. As a result, individuals with substantial executive dysfunction can often perform adequately on well-structured tests when the examiner is allowed to cue and probe for more information, relieving the individual of the need to be appropriately inhibited, flexible, strategic in planning, and goal-directed.”

To overcome this minor obstacle of not being able to directly test or measure executive functions I am going to measure them indirectly by observing a person’s behaviour. Executive functions are known to strongly regulate people’s behaviour. Every executive function is responsible for a specific array of behaviour, so by observing the behaviour that is exhibited by someone a lot can be said about the underlying executive functions being active at the moment the behaviour is exhibited. Next the characteristic behaviour is listed for all executive functions.

**Inhibition**

Strong inhibition leads to thinking before acting; people seem to have thought through their actions or words before they do or say them (Dawson and Guare, 2010; Shallice and Burgess, 1993; Miyake et al., 2000).

The inability to inhibit is readily observed in daily activities such as acting without thinking, being easily distracted, and being unable to sit still (Gioia and Isquith, 2004). Also ascribed to the inability to inhibit is impulsivity, which is characterized by rapid, unplanned, inappropriate, and potentially maladaptive conduct (Visser et al., 1996), ranging from, for example, a disinhibited verbal response to severe risk-taking and aggression (Enticott et al., 2006).

**Task initiation**

Caregivers often report that children with initiation difficulties have trouble getting started on homework or chores and that they require prompts or cues to begin. Problems with initiation are not the result of noncompliance or disinter-
2.6. INTERVENING FACTORS

There are some factors that can have an impact on a person’s executive control, working memory, attention, or modeling skills. This could cause deviations in their performance, either positively or negatively.

**Culture**

This factor comprises five dimensions, according to Hofstede et al. (2009). The first dimension is ‘power distance’. In countries or organizations with high power distance people are less inclined to approach their superiors or to contradict their statements. This can lead to undesired interpersonal relations between the modeller and domain expert if one sees the other as superior or inferior to themselves. Discussions of the problem domain can become a one-sided versions of the reality, ultimately leading to incomplete or false information about said problem domain.

The second dimension is ‘individualism’. In individualistic societies people are expected to take care mainly for themselves and their close family mem-
bers, whereas people in collectivistic societies are expected to take care for their extended family for which they receive unconditional protection and loyalty. Individualists, as opposed to collectivists, among other things do not view people as belonging to their group or another group, but show tolerance to all people. They also choose their own friends and friendly relations, instead of relying on the family to decide upon friends. Being a modeller in a collectivist society can be hard if the domain expert finds the modeller not to belong to their own group.

The third dimension is ‘masculinity’. Societies that are masculine expect (especially) men to be assertive and harsh and aimed at material success, where women should be humble and gentle and concerned with quality of life. In feminine societies both men and women should be humble and gentle and concerned with quality of life. Extreme masculinity could be a problem when modelling, if the modeller is solely focussed on the result of finishing the model and thereby loses sight of the match between the model and reality.

The fourth dimension is ‘uncertainty avoidance’. Uncertainty avoidance is the degree to which people feel threatened by uncertain or unknown situations. This feeling can lead to stress and a need for formal or informal rules. Modellers often work in uncertain and unknown situations, but one of their goals is to formulate rules in the problem domain by means of concepts and relations. Still uncertainty avoidance seems not to be a very desired trait for modellers.

The fifth and last dimension is ‘long-term orientation’. Here individuals or societies strive for rewards in the future, mainly by perseverance and frugality. The opposite is short-term orientation, which is aimed at respecting traditions, keeping face, and satisfying social obligations. It looks like modellers should preferably not be short-term oriented if they want have creative freedom in their process of creating a model. They should also be perseverent in getting all information they need to make a complete and sound model.

Experience
This mostly impacts modelling skills. The more experienced someone is, the better they usually perform a task. This also holds for modelling, although being more experienced than someone else is no guarantee for producing better models than they do. Experience does, however, provide knowledge about previous encounters of certain situations, helping the person make decisions on how to deal with these situations, giving them a better chance to arrive at the correct solution.

Education
This intervening factor is twofold. First, the higher-educated a person is the better they can be expected to perform on complex tasks like modelling. Second, people with degrees in natural sciences or formal sciences are used to represent knowledge in the form of models and mathematical formulas, whereas this practice is uncommon in social sciences, behavioral sciences, and cognitive sciences. Both distinctions can have influence on the quality of the produced model.

Age
It is established that with age people’s mental processes start working slower (e.g. Salthouse et al. (1991); Bender and Raz (2012); Lee et al. (2012); Peng et al. (2012)). This does not necessarily mean tasks are performed worse. How-
ever, complex tasks like modelling do call heavily on mental processes, making it harder for older people to do well on such tasks.

**Traumatic brain injury**
While not all brain injuries have the same effect (or any effect at all), a variety of cognitive and behavioural deficits can result from traumatic brain injury, including deficits in attention, memory, processing speed, and executive functions (see Gioia and Isquith (2004) for an overview of research on these deficits).

**Homogenity**
If a group of people is asked to do the same task individually under the same circumstances and this group is very homogenic, differences in the quality of their produced models can be traced to their mental processes. Otherwise differences in modelling quality are likely caused by (one of) the other intervening factors.

On the other hand, if the group of test subjects is heterogenic, modelling results should be adjusted for the intervening factors before differences can be traced back to mental processes.

Although everybody is impacted to some extend by every one of these intervening factors, it is out of scope of this thesis to research the precise effect these have. Any result obtained in this research that deviates from expected results might be explained by the effect of one or more of these factors, but the theoretical grounds on which to base such explanations are not examined in enough depth to justify them. The persons observed in this research are all male, fall in the same age group, all come from The Netherlands, have no brain injury, and all have some previous experience of modelling business processes. This homogenity amongst the test subjects eliminates most of the effect the intervening factors could have on the results.
Chapter 3

The relation between modelling skills and executive functions

The main goal of this research is to find out if peoples’ modelling skills are in any way related to executive functions, depicted in Figure 3.1.

![Figure 3.1: Main research question](image)

To do so, I am going to take specific modelling skills and executive functions, and see if there is any relationship between those. In the previous chapter eight modelling skills and four executive functions have been identified. Based on their descriptions from literature, I have formed an idea how these modelling skills could be related to the executive functions. Next I will discuss these four relationships, why I think these relationships might exist, and how I am going to measure these relationships later on.

3.1 Hypothesis H1

For the first hypothesis I am going to look for a correlation between some of the modelling skills and Inhibition. Based on the literature, my opinion is that Inhibition can be of importance when abstracting and inducing and when structuring and classifying information, which were items 1 and 2 from the previous chapter. I think strong inhibition will help people stop automatic responses to each and every detail in the description of a problem domain, and will help to sift the important information from the unimportant information, which is exactly what is needed for proper abstraction, induction, structuring, and classifying.
On the other hand, I think people with weakly developed inhibition will have problems abstracting, inducing, structuring, and classifying. To be able to make statements about this expected correlation both the executive function and the two modelling skills should be measured for a group of people.

As discussed in the previous chapter executive functions can only be measured indirectly by their behaviour. The way the experiment is set up (see next chapter) means this will be done by observing people during a modelling session. This makes it easy to observe the test subjects’ behaviour, which should in turn make it easy to rate their degree of inhibition.

The behaviour to look for during these sessions according to literature is if people do not seem to think through their actions, are easily distracted, can not sit still, and engage in rapid, unplanned, inappropriate, or maladaptive conduct; all signs of dysfunctioning inhibition. But all these behaviours are very subjective and so the observed degree of inhibitory dysfunction will vary heavily from one observer to another, and may even vary from one moment to another depending on the current mood of the observer. After all, what one person may qualify as distraction, not sitting still, or inappropriate conduct, might not be judged likewise by somebody else. To counter this, objective measurements are required.

Supported by the way BRIEF-A (Roth et al., 2005) measures Inhibition I have decided on some objective measurements to substitute the objective behaviour with. This is no guarantee these measurements correspond one-to-one to the behaviour they are substituted for, but at least these measurements can be quantified and compared to each other, which make them all the more useful and which is good enough for me at this point. The objective measurements I have picked as a measure for Inhibition are the total number of occurrences of (1) the person drumming their fingers, wiggling their legs, or unnecessarily clicking their pen, and (2) the person interrupting others when they are talking. The more occurrences of these actions I notice, the less developed I find their Inhibition to be.

For the same reason I need quantifiable measurements for modelling skills. As a measure for abstraction, induction, structuring, and classifying information I am going to take a closer look at the number of question-answer interactions between the modellers. More specifically I am interested in these interactions that have to do with abstracting, inducing, structuring, or classification of any information regarding the problem they are trying to model. Every time such a question is asked it gets scored. In the end a low score indicates low skills for abstraction, induction, structuring, and classification, and a high score indicates high skills for these. This means persons who score high are seen as the better modellers.

What I hope to find is that people who ask more modelling-related questions do less drumming, wiggling, clicking, and interrupting. This is the idea behind the first hypothesis:

\[ H1: \text{Better developed Inhibition leads to a higher level of modelling.} \]
3.2 Hypothesis H2

The second hypotheses will be about the correlation of certain modelling skills with Task Initiation. Based on the literature I find Task Initiation to be important when communicating, be it verbally or in letter. This is relevant when wanting to be a skilled communicator and when trying to describe thoughts and ideas, which were items 7 and 8 from the previous chapter. I think well-developed task initiation will enable people to speak their minds on their own, where others will need some encouragement to do so. Like for the previous hypothesis both the executive function and the modelling skills should be measured objectively before any statement can be made about their correlation. This means I need a quantitative measure for both.

The behaviour associated with Task Initiation is that people have trouble getting started on tasks and that they require prompts or cues to begin. However, problems with initiation are not the result of noncompliance or disinterest in the task; the person typically has interest in the task and wants to succeed but cannot get started. For an objective measure I want to look at how often somebody needs prompts or cues before starting a task. This is also one of the measurements used in the BRIEF-A (Roth et al., 2005) to determine somebody’s level of Task Initiation. I will assume that counting the number of times somebody needs a reminder to do something they have already said they are going to do, is a valid measure for Task initiation.

Likewise I need an objective measure for a person’s skills of communication and describing thoughts and ideas. In my opinion this can be measured as the number of times a person is the one to begin a conversation, as opposed to a person who only speaks when spoken to. So I will count the number of times somebody starts a conversation (an independent piece of talking about a distinct subject, which has a natural beginning and end) and add it to their score. Since every conversation is started by someone it should be clear at the end of the session who starts conversations more often than others. Subsequently I define that a low score indicates low communication skills and a high score indicates high communication skills. In this respect I regard the people with high communication skills as being the better modellers.

What I hope to find now, is that people who initiate a conversation most often are also the ones who need the least amount of prompts and cues to begin a task. This explains the second hypothesis:

\[
\text{H2: Better developed Task initiation leads to a higher level of modelling.}
\]

3.3 Hypothesis H3

The third hypothesis is about the correlation between some of the modelling skills and Evaluation. Based on the literature from the previous chapter, my expectation is that Evaluation is of importance when trying to keep an overview on the progress towards the goals, and when keeping an eye on the whole instead of only one piece of the problem, which were items 5 and 6. I think well-developed evaluation will help decide whether actions taken are beneficent to the process or not, and will also help decide if an operation on a piece of the problem domain will impact other areas in the problem domain. People with un-
CHAPTER 3. THE RELATION BETWEEN MODELLING SKILLS AND EXECUTIVE FUNCTIONS

derdeveloped Evaluation often rush through tasks without checking their work for mistakes, and may be unaware of how their actions affect others in a social context. This is again a very subjective rating for an executive function, so I will need to objectify it before being able to make statements about a person’s level of evaluation.

For the objective measure for the level of Evaluation of a person I will assume that it is valid to count the number of times it occurs that a mistake of that person is pointed out by somebody else right after the person has made that mistake. This is a similar measure as used in the BRIEF-A (Roth et al., 2005) for their scale called Task Evaluation. I am not interested in corrections that have a longer time in between, because that might mean new insights have emerged or other information has become available, making the mistake only a mistake in retrospect while it was not a mistake at the time. This means that the more times somebody makes a mistake that is noticed immediately afterwards, the less effective their Evaluation is.

The modelling skills associated with Evaluation - keeping an overview on the progress towards the goals and keeping an eye on the whole - deal with the same aspect of making sure actions taken help solve the problem at hand. To try and measure objectively how well somebody is capable of doing this, I am going to count the number of times a person comments on the value of an action. Whether they find the action helpful or not does not matter. I define that the more comments a person makes, the better their modelling skills are.

What I hope to see is that people who make less mistakes are also the ones that comment most frequently on actions taken. This comes down to:

\[H3: \text{Better developed Evaluation leads to better modelling.}\]

3.4 Hypothesis H4

For the fourth and last hypothesis I am going to take a look at the correlation between certain modelling skills and Flexibility. Flexibility is required when transitioning from one situation, activity, or aspect of a problem to another as the situation demands. Lack of flexibility generally makes people unable to adapt to changing conditions. Because of this I think flexibility is needed when trying to have an open mind and when handling implicit knowledge, which are items 3 and 4 from the previous chapter. Here it again applies that deciding if someone is flexible or not based on an impression of how well that person can cope with changing conditions is too subjective to work with. This verdict should be based on an objective measure. To do so, I have chosen to go with one of the measures also present in the BRIEF-A (Roth et al., 2005) of judging the capability of transitioning from one aspect of a problem to another. For this I am going to count the number of times it happens that a person is being told the topic that is being discussed at the moment, in contrast to a topic that has been discussed previously. I regard this as an inability of the person to leave the previous topic and focus on the current one. So the less that happens the better that person’s Flexibility.

As a measure of how well a person does in having an open mind and handling implicit knowledge I have decided to look at the number of occasions when that person notices information on the topic at hand is missing. If the remark is
followed by an assumption of some replacement information the person gets +1 to their score, whereas if the remark is followed by a confession of not being able to continue the person gets -1 for their score. In the end the total score is divided by the total number of remarks about missing information so that the final score will lie between -1 and 1. Scores below 0 will then indicate a low level of open mindedness and scores above 0 will indicate a high level of open mindedness. A high level of open mindedness will signify better modelling skills.

What I hope to see is that people who need to be told the current topic less frequently also assume replacement information more often than not. This brings us to the final hypotheses:

\[ H_4: \text{Better developed Flexibility leads to better modelling.} \]

### 3.5 Testing the hypotheses

![Figure 3.2: Research questions after introducing hypotheses](image)

Since all behaviour that I need to measure has now been objectified, it is time to test the hypotheses. The next chapter explains how the experiment is set up, how measurements were done, and what the meaning is of the data that has been gathered with those measurements.
CHAPTER 3. THE RELATION BETWEEN MODELLING SKILLS AND EXECUTIVE FUNCTIONS
Chapter 4

Observations of modelling sessions

4.1 Materials

The materials used are digital video recordings made by Wilmont et al. (2012) for other research. Over a period of several weeks she recorded multiple meetings and modelling sessions aimed at modelling Everest BV’s business processes and design new ones in order to develop a new automated information system. Everest BV is located in ‘s-Hertogenbosch, The Netherlands, and provides IT business solutions in the fields of government, mortgage and insurance. During the meetings up to eight people cooperated in providing all necessary information about the organisation’s business processes to the modellers. The modelling sessions were used to model all the gathered information into one comprehensive model that covers the entire organisation’s business processes. For my research I am only interested these modelling sessions.

Four of the five recorded modelling sessions were usable for observations; the fifth one’s image quality was too poor. The sessions lasted for about 30 to 90 minutes per sessions and were done in groups of 2 to 3 people. The modellers were all middle-aged males and had moderate to high experience with modelling. They had flip charts at their disposal and all of them could write and draw on it anything they felt was important for the modelling process. Every modeller could make suggestions, ask questions, and give his opinion on anything related to the organisation’s business processes and the modelling task at hand.

4.2 Procedure

To be able to make statements about a person’s executive functions and modelling skills on the basis of their exhibited behaviour, I have related types of behaviour to executive functions and to modelling skills in Chapter 2. From these types of behaviour I picked very specific behaviour in Chapter 3 to concentrate on while observing people. It is these behaviours that I looked for when viewing the four video recordings. I made a table on paper containing rows for every kind of behaviour and separate columns for every modeller. Then whend-
ever I saw or heard anybody exhibit behaviour I was looking for I tallied it in the corresponding table cell. After viewing all four videos the table looked like Table 4.1. Although I know the names of the modellers I have anonymised these into MrX, MrY, and MrZ. The phrases in the column labeled 'Behaviour' correspond to the behaviour I decided on in Chapter 3 to look for.

4.3 Results and interpretation

Table 4.1: Tally of behavioural occurrences

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MrX</td>
<td>MrY</td>
<td>MrX</td>
<td>MrY</td>
</tr>
<tr>
<td>Fingers / legs / pen</td>
<td>9</td>
<td>12</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Interrupting others</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Modelling questions</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Needs prompt / cue</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Starts conversation</td>
<td>7</td>
<td>2</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Mistake pointed out</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Comments on action</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Being told the topic</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Information missing</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

The values in Table 4.1 represent the raw data gathered by observing the video recorded modelling sessions. These modelling sessions did not last for the same amount of time, so to be able to compare the tallied values they should be weighted. I have decided to do this by dividing the values by the number of minutes the sessions lasted. Session 1 lasted 34 minutes, session 2 lasted 38 minutes, session 3 lasted 80 minutes, and session 4 lasted 50 minutes. Below is a table with these weighted values.

Table 4.2: Behavioural occurrences per minute, rounded to two decimals

<table>
<thead>
<tr>
<th>Behaviour</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MrX</td>
<td>MrY</td>
<td>MrX</td>
<td>MrY</td>
</tr>
<tr>
<td>Fingers / legs / pen</td>
<td>.26</td>
<td>.35</td>
<td>.37</td>
<td>.29</td>
</tr>
<tr>
<td>Interrupting others</td>
<td>.15</td>
<td>.03</td>
<td>.18</td>
<td>.03</td>
</tr>
<tr>
<td>Modelling questions</td>
<td>.09</td>
<td>.03</td>
<td>.11</td>
<td>.05</td>
</tr>
<tr>
<td>Needs prompt / cue</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>Starts conversation</td>
<td>.21</td>
<td>.06</td>
<td>.16</td>
<td>.08</td>
</tr>
<tr>
<td>Mistake pointed out</td>
<td>.06</td>
<td>.00</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td>Comments on action</td>
<td>.03</td>
<td>.00</td>
<td>.13</td>
<td>.03</td>
</tr>
<tr>
<td>Being told the topic</td>
<td>.03</td>
<td>.00</td>
<td>.05</td>
<td>.00</td>
</tr>
<tr>
<td>Information missing</td>
<td>.03</td>
<td>.00</td>
<td>.00</td>
<td>.05</td>
</tr>
</tbody>
</table>
4.3. RESULTS AND INTERPRETATION

4.3.1 Results for H1

The first three phrases from the first column of Tables 4.2 refer to the three forms of behaviour observed for hypotheses H1: the number of occurrences of the person drumming their fingers, wiggling their legs, or unnecessarily clicking their pen; the number of occurrences of the person interrupting others when they are talking; and the number of question-answer interactions between the modellers that have to do with abstracting, inducing, structuring, or classification of any information of the problem they are trying to model. For simplicity I will call those phrases A1, A2, and A3 respectively. The hypothesis then becomes ‘when A1+A2 is low, A3 is high’. Taken from Table 4.2 the values will look like this:

<table>
<thead>
<tr>
<th></th>
<th>.06</th>
<th>.11</th>
<th>.05</th>
<th>.21</th>
<th>.38</th>
<th>.08</th>
<th>.02</th>
<th>.32</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1+A2</td>
<td>.41</td>
<td>.38</td>
<td>.55</td>
<td>.32</td>
<td>.25</td>
<td>.46</td>
<td>.26</td>
<td>.24</td>
</tr>
<tr>
<td>A3</td>
<td>.09</td>
<td>.03</td>
<td>.11</td>
<td>.05</td>
<td>.21</td>
<td>.38</td>
<td>.08</td>
<td>.02</td>
</tr>
</tbody>
</table>

Intuitively

It is hard to see at a glance which values for A1+A2 are low, because the values lie close to one another. But looking at the numbers above there seems to be a pattern where low values for A3 belong to high values for A1+A2, although not all high values for A3 belong to low values for A1+A2. This calls for better examination to determine any pattern. There are different ways of determining relationships between multiple value pairs. I will use some of these to examine their use for evaluating data as gathered in the current experiment.

Quantiles

First of all it needs to be defined what ‘low’ is and what ‘high’ is. I will do this based on quantiles, which are subsets that divide a ranked set of data items in a number of equal sets where each group comprises an equal amount of the total data. Data sets can be divided into as many quantiles as wanted, but some divisions are more useful than others. Because there are 9 data points in the total set of data items, I have chosen to use the quantile division called tertiles. This means the data set will be divided into three equal subsets. Any value in the first tertile will then be considered low, values in the second tertile will be considered medium, and all values in the third tertile will likewise be considered high. For A1+A2 it means the tertiles are the subsets \{.06, .24, .25\}, \{.26, .32, .38\}, and \{.41, .46, .55\}. For A3 this means the tertiles are the subsets \{.02, .03, .05\}, \{.08, .09, .11\}, and \{.21, .32, .38\}. If the values are then replaced by low, medium, and high, it looks like this:

<table>
<thead>
<tr>
<th></th>
<th>high</th>
<th>med</th>
<th>low</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1+A2</td>
<td>high</td>
<td>med</td>
<td>low</td>
</tr>
<tr>
<td>A3</td>
<td>med</td>
<td>low</td>
<td>med</td>
</tr>
<tr>
<td></td>
<td>high</td>
<td>med</td>
<td>low</td>
</tr>
</tbody>
</table>

As can be seen two out of the three lows for A1+A2 correspond to a high for A3. The other low corresponds to a low for A3, which is not what was desired. But still this result may be considered as a modest indication that hypothesis H1 could be correct.
CHAPTER 4. OBSERVATIONS OF MODELLING SESSIONS

Cosine similarity

The cosine similarity is a mathematical measure of similarity between the orientation of two vectors. Vectors with the exact same orientation have a cosine similarity of 1, where two vectors with an exactly opposite orientation have a cosine similarity of -1. When there are only positive values, like in this experiment, the outcome of the cosine similarity lies in [0, 1]. If A1+A2 is considered a vector and A3 is considered another vector, the cosine similarity between them would be 0.66 which I would interpret as meaning they are moderately alike.

Correlation

In statistics correlation refers to a broad class of relationships. The most common of these is the ‘Pearson product-moment correlation coefficient’, where outcomes range from -1 to +1. Here +1 means perfect correlation, -1 means perfect anticorrelation, and closer to 0 means more uncorrelatedness. The nature of the hypothesis being investigated means A1+A2 and A3 ideally have a linear relationship and are inversely proportional, so are anticorrelated. However, correlation does not imply causation in either direction, which means that although the outcome might point to evidence for a possible causal relationship it can not be said what that relationship is.

Using Pearson’s correlation coefficient to calculate the correlation between A1+A2 and A3, the outcome is -0.15. This indicates that there is a weak negative correlation. Although the relation is weak, it is still negative as hoped for. Because nothing can be said about the causal relationship underlying this correlation this does not confirm the hypothesis, but in any case it does also not refute it.

Scatter plot

Figure 4.1: Scatter plots for A1+A2 and A3

Besides mathematically defined aspects of similarity based on a decimal number between 0 and 1 or -1 and 1, there is also a way of representing the data as a visualisation. With two variables the most perspicuous visualisation is a two-dimensional scatter plot. A bonus is that such a scatter plot also illustrates the degree of correlation between the two variables plotted. The correlation
calculated in the previous paragraph can so be compared to these plots. To make such scatter plots I have normalised the values of A1+A2 and A3 before plotting them. This also makes it easier to compare these plots with plots in coming sections.

From Figure 4.1a it is clear that the data points do lie in a somewhat negatively correlated distribution as suggested by the outcome of the correlation coefficient calculation. However, it appears that the data point on (0.84,1) causes the correlation coefficient to be less prominent. Omitting this data point in the calculation would lead to a correlation coefficient of -0.58, which would make A1+A2 and A3 reasonably anticorrelated. Unfortunately there are too little data points in this sample to be sure if the aforementioned data point is an outlier or not. Either way, at the moment it looks like there is modest evidence for an inverse relationship between A1+A2 and A3 and so for confirmation of H1.

Conclusion

All the different methods used to interpret the results for A1+A2 and A3 indicate that there is evidence that hypothesis H1 should not be refuted. Nevertheless it is too little to say H1 should be confirmed. To make sure the proposition made in H1 is correct the experiment would have to be carried out with more observations.

This outcome was best predicted by the correlation coefficient and the scatter plot.

4.3.2 Results for H2

The fourth and fifth phrase in the first column of Table 4.2 come from hypotheses H2 and correspond respectively to the number of times somebody needs a reminder to do something they have already said they are going to do, and to the number of times a person is the one to begin a conversation. If I call these B1 and B2 for short, the hypothesis then becomes ‘when B1 is low, B2 is high’. Taken from the table, these are the values for B1 and B2:

| B1 | .00 | .00 | .00 | .00 | .00 | .02 | .00 | .00 |
| B2 | .21 | .06 | .16 | .08 | .13 | .05 | .02 | .00 | .10 |

Intuitively

Seeing as there is little variation in values for B1 I would say there is no statistical relationship to be found between B1 and B2. However, it does look like the statement from hypothesis H2 that people who do not have to be told to do something generally start conversations. Still I am going to do the same analyses for B1 and B2 as I did in the previous section for A1+A2 and A3 to see if there is some kind of statistical relationship.

Quantiles

Let us start again by trying to determine what low is and what high is. As there are nine value pairs for B1 and B2 it would again be the most convenient to divide those into three quantiles. Unfortunately, there is so little variance in the
values for B1 that it is impossible to divide those into three equal subsets. The
best I can do is to call all values of .00 low and the value of .02 high. Things
are different for B2, for which it is possible to make a division into three equal
subsets. These will be {.00, .02, .05} for low, {.06, .08, .10} for medium, and
{.13, .16, .21} for high. Substituting this into the table:

| B1 | low | low | low | low | low | low | high | low | low |
| B2 | high | high | med | high | med | low | low | low | med |

This substitution does not really show other information than what was known
before. Still, it confirms that people that do not have to be told to do some-
thing they already said they were going to do, generally start conversations.
This confirms the hypothesis. Unfortunately there is too little data on people
that do have to be told to do something they already said they were going to
do. This means there is not enough data to say if these people do or do not
often start conversations compared to others.

**Cosine similarity**

If B1 and B2 are assumed to be vectors, the cosine similarity between those two
vectors is 0.06. On a scale from 0 to 1 this is very close to 0. For cosine similarity
a value of 0.06 means the two vectors are almost as different as could be. In
other words, there is no similarity whatsoever between B1 and B2 according to
the cosine similarity measure. I think this means that according to the cosine
similarity measure hypothesis H2 should be refuted.

**Correlation**

Like for hypothesis H1 it would be favourable for the correlation of hypothesis H2
to be as close to -1 as possible, on a scale from -1 to 1. Calculating the correlation
between B1 and B2 using Pearson’s correlation coefficient the outcome is -0.39.
Such a value indicates a modest negative correlation and would mean evidence
that H2 can be confirmed.

**Scatter plot**

Figure 4.2: Scatter plots for B1 and B2

(a) Scatter plot of all value pairs

(b) Scatter plot specified per individual
4.3. RESULTS AND INTERPRETATION

There seems to be disagreement between the outcome of the cosine similarity and the correlation coefficient about whether B1 and B2 are indeed related to one another. A scatter plot might help to see which measure is closer to the truth. Before plotting the value pairs they were normalised so they can be easier compared to the scatter plots in the previous and next sections. These are the resulting plots:

Figure 4.2a shows that there is actually no relationship to be discovered between B1 and B2. The value of -0.39 for correlation that sounded promising is actually skewed considerably by the data point that is plotted on (1,0.1). Without this data point the correlation would have been exactly 0, meaning no relationship at all. However, no relationship still does not mean that H2 should be refuted. Refuting H2 is only necessary when the outcome of the correlation would have been bigger than 0. On the other hand, there is no reason to omit said data point. The only way to increase the conformance between the scatter plot and the calculated correlation coefficient is to have more data points that preferably do not lie on the y-axis. Additional observations would have to be done to achieve this.

Conclusion

It is undecided if H2 should be confirmed or refuted. There is no relationship between the values of B1 and B2, where a negative relationship would be desirable. There being no relationship is caused by the lack of observed people that needed cues or prompts to do something they already said they were going to do. Observing more of such people is needed to have enough data. On the other hand the observed people who do not need cues or prompts show that they are generally able and willing to start conversations. This is partial evidence that H2 can be confirmed with more data, but more studies are needed to substantiate this.

The analyses that were most useful for spotting this non-existent relationship were the quantiles and the scatter plot.

4.3.3 Results for H3

The sixth and seventh phrase in the first column of Table 4.2 are for hypotheses H3 and relate to the number of times it occurs that a mistake of that person is pointed out by somebody else right after the person has made that mistake, and to the number of times a person comments on the value of an action respectively. Calling these C1 and C2 for short, the idea behind the hypothesis becomes ‘when C1 is low, C2 is high’. These values for C1 and C2 are taken from the table:

\[
\begin{array}{cccccccc}
C1 & .06 & .00 & .03 & .03 & .04 & .03 & .02 & .00 \\
C2 & .03 & .00 & .13 & .03 & .11 & .06 & .08 & .10 & .00 \\
\end{array}
\]

Intuitively

At first sight it looks like the higher the value for C1 the higher the corresponding value for C2. This is the opposite of what I hoped to find. Further analysis is needed to confirm such a proportionality.
CHAPTER 4. OBSERVATIONS OF MODELLING SESSIONS

Quantiles

Tertiles will again be used to divide the data set into equal subsets. The nine values of C1 will be divided into \{0.00, 0.02\} for low, \{0.03, 0.03\} for medium, and \{0.04, 0.06\} for high. For C2 there is a small problem, though: dividing the data set into three equal subsets would mean that the subsets for low and medium would hold one identical value. To solve this the identical values must be put in either the subset for low or the subset for medium. The prior would lead to the subsets being \{0.00, 0.03\} for low, \{0.06\} for medium, and \{0.10, 0.11\} for high, with a resulting substitution

\[
\begin{align*}
C1 & \quad \text{high} \quad \text{low} \quad \text{med} \quad \text{med} \quad \text{high} \quad \text{med} \quad \text{high} \quad \text{low} \quad \text{low} \\
C2 & \quad \text{low} \quad \text{low} \quad \text{high} \quad \text{low} \quad \text{high} \quad \text{med} \quad \text{med} \quad \text{high} \quad \text{low}
\end{align*}
\]

The latter leads to the subsets being \{0.00, 0.00\} for low, \{0.03, 0.06, 0.08\} for medium, and \{0.10, 0.11, 0.13\} for high, with a resulting substitution

\[
\begin{align*}
C1 & \quad \text{high} \quad \text{low} \quad \text{med} \quad \text{med} \quad \text{high} \quad \text{med} \quad \text{high} \quad \text{low} \quad \text{low} \\
C2 & \quad \text{med} \quad \text{low} \quad \text{high} \quad \text{med} \quad \text{high} \quad \text{med} \quad \text{med} \quad \text{high} \quad \text{low}
\end{align*}
\]

Especially this last substitution shows a tendency of higher values for C1 resulting in higher values for C2, resulting in refutation of the hypothesis.

Cosine similarity

The cosine similarity between C1 and C2 is 0.80 meaning they are highly similar. This could also be an indication that the hypothesis should be refuted.

Correlation

Using Pearson’s correlation coefficient to calculate the correlation between C1 and C2 the outcome is 0.41. This moderate positive correlation proves that the relationship between them is certainly not inversely proportional. So again it looks like the current hypothesis will be refuted.

Scatter plot

Figure 4.3: Scatter plots for C1 and C2

(a) Scatter plot of all value pairs

(b) Scatter plot specified per individual
The scatter plots for C1 and C2 show that the value pairs lie in a somewhat positively correlated distribution. Note that there are two value pairs on (0,0), which amplifies this positivity. It seems logical from these scatter plots that hypothesis H3 can only be refuted.

Conclusion

All analysis tools confirm my intuitive suspicion that hypothesis H3 should be refuted. This is shown the clearest by the quantile division, the correlation coefficient and the scatter plots.

4.3.4 Results for H4

The last two phrases in the first column of Table 4.2 are used for hypotheses H4 and are derived from the number of times it happens that a person is being told the topic being discussed at the moment in contrast to a topic that has been discussed previously, and the number of occasions when a person notices information missing. In line with the above I will call those phrases D1 and D2 for short, which makes the hypothesis read ‘when D1 is low, D2 is high’. From the table:

| D1 | .03 | .00 | .05 | .00 | .03 | .01 | .00 | .00 | .00 |
| D2 | .03 | .00 | .00 | .05 | .01 | .05 | .04 | .04 | .00 |

Intuitively

Apart from a couple of value pairs where both values for D1 and D2 are the same, it definitely looks like D2 is high where D1 is low and vice versa. Statistical analyses might prove otherwise, but for now I would say there is strong evidence that hypothesis H4 can be confirmed.

Quantiles

To be comparable to the quantile division in the previous sections it would be favourable to use a division into tertiles for D1 and D2. However, there is so little dispersion between the values for both that a division into tertiles would go awry with every possible division. Even worse, there seems to be no usable quantile division for both D1 and D2. The unwelcome conclusion is that in this case quantiles are unusable and so will not be taken into account for the current hypothesis.

Cosine similarity

The calculated cosine similarity for D1 and D2 is 0.27. A low value like this indicates that the vectors are indeed not alike, but if it also means that they are inversely proportional is not discernible. Nevertheless the outcome of the cosine similarity is no reason to refute the hypothesis.

Correlation

The correlation coefficient for D1 and D2 is -0.38, which is almost the same as that for B1 and B2. In that case, however, the outcome was skewed considerably
CHAPTER 4. OBSERVATIONS OF MODELLING SESSIONS

by a possible outlier that emerged from investigating the accompanying scatter plot. If such a distortion is not present for D1 and D2 it might mean that these are really moderately negatively correlated. If so, there is reason to believe that hypothesis H4 can be confirmed.

Scatter plot

Figure 4.4: Scatter plots for D1 and D2

(a) Scatter plot of all value pairs

(b) Scatter plot specified per individual

Although the data points are widely scattered within the plotting domains, I still can see why the correlation coefficient calculated above results in the moderately negative outcome that it does. Besides the two points on (0,0) the points are all plotted close to the diagonal line from (0,1) to (1,0). So I find that the scatter plot confirms the conclusion drawn from the correlation coefficient, although maybe not too evident at first sight.

Conclusion

The analyses done for the data for H4 all seem to point to the same conclusion that there is a moderate inversely proportional relationship between D1 and D2. This means there is evidence that H4 can be confirmed. I find the correlation coefficient the clearest indicator in this case.

4.3.5 Other results from the gathered data

Validating requirements vs processing requirements

When viewing the modelling sessions it became apparent that there was a difference between some of them in the amount of modelling done. The difference was most notable between sessions 1 and 2 on the one hand and sessions 3 and 4 on the other hand. Sessions 1 and 2 were oriented towards validating the system requirements against what had been modelled thus far, and checking if that what had been modelled was in compliance with the requirements. This meant that there were less modelling questions being asked, but more questions about whether or not (a part of) the existing model was correct.
In modelling sessions 3 and 4 the test subjects were eager to process requirements into new parts of the model. This lead to many more questions about how to model these requirements and how to connect them to the existing model.

This difference between orientation during the modelling sessions can be seen in the amount of modelling questions asked per minute and the amount of disruptive behaviour in Table 4.2. While the amount of ticking and interrupting is 0.79 and 0.87 per minute in sessions 1 and 2 respectively, the number of modelling questions asked per minute is just 0.12 and 0.16. Compare this to 0.71 and 0.56 ticks and interruptions per minute in sessions 3 and 4 respectively and 0.59 and 0.42 modelling questions asked per minute.

Such a big difference between sessions that are oriented towards validating requirements and sessions that are oriented towards processing requirements might be an important factor in determining the relationship between disruptive behaviour and modelling question asked. But because there have been only two sessions of either kind, there need to be done more observations to see if such a difference is always present. For now there is not enough reason to suspect it has had a crucial influence on the results for hypothesis H1.

**Influence of a third person**

During the first, second, and third modelling sessions the image was quite similar: MrY was the knowledgable person and MrX wanted his approval for ideas and decisions. This translated in MrY being thoughtful and somewhat reactive to MrX, where MrY would speak his mind on any subject, thus starting lots of conversation, interrupting MrY a lot, and sometimes going off-topic. When in the fourth session MrZ was the knowledgable person, MrX’s behaviour did not change much, but was directed more towards MrZ than to MrY.

MrY’s behaviour, however, changed considerably. He became very passive, both physically and verbally. The former can be seen in Table 4.2 by the considerably lower number of ticking and clicking compared to the first three sessions. The latter is evident from not starting any conversation, but only commenting on actions done by others.

Such a change in behaviour signifies that it could be important for the observations who is put together for the modelling sessions. Behaviour of one person can be influenced by presence (or absence) of another person. Statistically the passivity of MrX during session 4 is not big enough to be of importance for the outcomes of hypotheses H1 to H4, but still it is worth mentioning that such change in behaviour has occured.
Chapter 5

Conclusions and recommendations

5.1 Conclusive answer to the research question

Recapitulating the separate conclusion from the previous sections, it is clear that the results vary reasonably between the four hypotheses. For the first hypothesis, “H1: Better developed Inhibition leads to a higher level of modelling”, only weak evidence is found pointing towards confirmation. However, one of the gathered data points seems to be an outlier compared to the others. Omitting this point leads to a considerably stronger relation between Inhibition and modelling. The question remains, though, if this data point really is an outlier or maybe a presage for more data points in its close proximity when more observations are done.

For the second hypothesis, “H2: Better developed Task initiation leads to a higher level of modelling”, there was no indication of any relationship whatsoever. Therefore no judgment is made on the acceptance of the hypothesis. More observations, especially of people who need prompts or cues, are needed to obtain sufficiently diverse data.

The third hypothesis, “H3: Better developed Evaluation leads to better modelling”, should be refuted according to the evidence found in the data gathered in this experiment. There is a moderate positively correlated relationship between Evaluation and modelling, meaning that in this experiment’s set up people with worse Evaluation were in general regarded as better modellers. The hypothesis should thus be refuted.

For the fourth hypothesis, “H4: Better developed Flexibility leads to better modelling”, evidence indicates that this is indeed true. A moderate relation between these facets has been found, which is sufficient to accept the hypothesis at this moment. Still further research would be helpful to justify such a conclusion.

Looking at these four conclusions the overall conclusion to the research question would be that there seem to be certain personality traits that could be predictive of a person’s modelling skills. As it stands the best candidates to base further research upon are Flexibility and Inhibition.
5.2 Recommendations for further research

Based on the experience I have gained through this research there are a number of things I would like to point out to consider for further research on this topic.

- Choose behaviour wisely. Even though the specific behaviours I choose to look for during the observations were in concurrence with what was described in the BRIEF-A (Roth et al., 2005), some behaviour was more applicable than other. It appears that there are many circumstantial variables that can influence if people do or do not depict certain behaviour. The reason I still continued with the behaviours I choose beforehand is that these seemed best distinguishable against other behaviour, and that they were believed to be indicative for their underlying executive functions. Behaviour should not be chosen based upon what is mostly observed during the modelling sessions, as this will steer the research and cripple the validity of the gathered data.

- Heterogenic test subjects. The test subjects observed for this research were all middle-aged males who had had decent education and had at least some experience in modelling. For more complete data and better understanding of possible relationships between executive functions and modelling skills it is advisable to have a wider range of test subjects. The reason the here observed test subjects were not heterogenic is that it is difficult to set up a real-life(like) problem that has to be modelled, that it is hard to let people model when they have never done so before, that it takes a lot of test subjects to have satisfying heterogenicity, and that it takes an enormous amount of time and arranging to have all those people model the same problem under the same circumstances. This is only feasible when there is interest only in a very specific relationship between executive functions and modelling skills, not when trying to find out if there are relationships at all like in this research.

- More observations. More observations means more data, and more data means better support for potential relationships between executive functions and modelling skills. Also, observing the test subjects in their everyday lives might be a better indicator for the level of development of their executive functions than observing them while modelling. Due to time constraints this was not doable for the current research.
Bibliography


