Arguments, Values and Weight Assessment.

Holistic versus analytical argumentation and cultural values in weight assessment processes.

Roelof van Dalen
S0068853
Master Thesis
University of Twente, Enschede, The Netherlands
14th of August 2013

Supervisors:
Dr. J.M.G. Heerkens
M.R. Stienstra MSc

Summary

This thesis is centred on two types of cultural differences in argumentation in weight assessment processes. One part on the process of argumentation in weight assessment and one part on the content of arguments by looking at values.

Weight assessment processes concerns determining the importance of (sub) attributes, in this experimental setting these attributes were safety and comfort when buying a new vehicle for a fictitious taxi company. These attributes can be split up into sub attributes, for example you can split up the attribute comfort into the sub attributes air-conditioning and comfortable seats. The importance of (sub) attributes is called weights. You can weigh in a relative way, you link two or more (sub) attributes together. For example: safety is more important than comfort. You can also weigh in an absolute way, stating one (sub) attribute is important. For example: safety is important.

This research was originally done by Hans Heerkens in 2003 among Dutch students. Later on this research was expanded by Floor Richters and Julia Doednik on respectively Australian and Russian students. This thesis is a secondary analysis of the raw data collected in the previous research on the subject of cultural differences.

Culture distinguishes a group from other groups and it influences the way we think and act (Hall, 1976), this is not directly visible but manifested in behaviour (Root, 1994). Researching a cultural difference regarding argumentation in weight assessment processes would be interesting, because if you are aware of cultural differences you are able to anticipate on the behaviour someone else would portrait in certain situations. As said, in this thesis we research cultural differences on process and cultural differences on content.

Researching a cultural difference on the process of argumentation was done by analysing results in the light of the contradiction between holistic versus analytical ways of thinking. From the work of Nisbett (2001) we've derived two indicators on the basis of argumentation in weight assessment processes:
- Attribute versus sub attribute weighing
  - Holistic argumentation supports weights on the attribute level.
  - Analytical argumentation supports weights on the sub attribute level.
- Relative versus absolute weighing
  - Holistic argumentation supports relative weighing.
  - Analytical argumentation supports absolute weighing.

Both indictors point in the same direction. The Russian group is most holistic of all the three groups, the Australian group is most analytical of all the three groups and the Dutch group falls between the Russian and Australian group.

Researching a cultural difference on the content of argumentation was done by the usage of the 10 value-types devised by Schwartz (1994): hedonism, power, achievement, stimulation, self-direction, universalism, benevolence, conformity, tradition and security. These values are present in every culture. The difference between cultures is made up by the relative importance of these value types, not every values is equally important. According to the theory of Schwartz & Bilsky (1987) values guide behaviour to the socially preferable road and therefore choices. In our case the choices for

certain weights. So it would not be exceptionally if we would find values. However the reality of this research is far from this. Values are almost not present in weight assessment processes, more precisely: I almost didn't find any. When values are measured, a large portion is not even aimed on importance assessments at all. Due to the low amount of values found a more in depth analysis on value types and the analysing a difference in relative importance is not viable. The amounts of values are too low to investigate any correlations between culture, value types and their relative importance.

The low amount of values could have a number of causes:

- In contradiction to the theory the presence of values in weight assessment processes is low.
    - This would contradict the theory on the guiding effect of values.
    - The experimental subjects did not feel any involvement with the fictitious assignment. The experimental subjects were aware of this fact and the lack of consequences of their weight assessment could mean the socially preferable road doesn't exist at all.
- The choice of being conservative with coding
    - To enhance intercoder reliability we were conservative with coding, doubtful 'values' were not included.
- The suitability of the think aloud method.
    - With this method we can only measure what experimental subjects say. Therefore they have to be conscious of their thoughts in order to verbalize them. Values are not directly visible. (Hofstede, 1994) They are at the core of culture and guide behaviour in an unconscious way. If experimental are not conscious of the influence of values we cannot measure them using the think aloud method.

On the basis of the research on values we cannot prove any difference between experimental groups, the relative importance of values and their influence on weights. We can draw conclusions on the process of argumentation in the light of holistic versus analytical argumentation. There is a big difference in the use of argumentation between the groups. The Russian group is the most holistic of all the three groups, the Australian group is the most analytical of all the three groups and the Dutch group can be found between the Russian and Australian group.

Preface

This thesis is written as the final assignment for the Master of Business Administration, track International Management. International Management is a concept which consists of two words, these two words are both covered in this thesis. The 'International' part by looking at different cultures and the 'Management' part by looking at weight assessment.

In the light of globalization the international management will become even more important. Decisions nowadays have to be made with regard to the global marketplace, this means business has to be conducted in and between other countries. This global marketplace also implies dealing with cultural differences. The awareness of a cultural difference would be a step forward. Managers could use this knowledge to make better informed decisions when it comes to dealing with other cultures. This should lead to better understanding between colleagues and business partners from other cultures and more efficient and effective decision making in the international context.

This Master thesis is a secondary analysis based on previous research done by a number of people associated with the University of Twente, either as a dissertation or as a Master or Bachelor thesis. I would like to thank Hans Heerkens, Floor Richters and Julia Doednik for the usage of the raw data from their research.

Furthermore I would like to thank my supervisors Hans Heerkens and Martin Stienstra for their input in this thesis. Their critical input and the corresponding discussions were vital for the quality of this thesis.

Roelof van Dalen

August 2013, Wapserveen

Table of contents

Chaper 1: Introduction

The sub title of this thesis is: 'Holistic versus analytical argumentation and cultural values in weight assessment processes.'. Right away we can see a division into two parts. One part of this division consists of an indirect analysis of culture in weight assessment. This is done by looking at the process of argumentation and the differences between holistic and analytical argumentation. The other part consists of a direct analysis of culture. This analysis is based on the content of the weight assessment by looking at cultural values.

As a bonus, but somewhat outside of the core of this paper, we can look at the amount of arguments and the possible positive relationship with importance. A proposition on this matter was formulated on the basis of exploratory research done by Anne Veltman among Dutch students (2010). He states: 'The more important an attribute the more arguments are given to support the importance of this attribute.'

In these few sentences we've already used a lot of concepts: culture, holistic, analytical, argumentation, weight assessment. All relevant concepts will be discussed in detail in the theoretical framework. In order to avoid any misconceptions and to fully understand the sections up to the theoretical framework I will introduce these concepts in the next section.

1.1 Background

Within organisations decision making is a very important part of management tasks. Decision making is choosing between one or more available alternatives. On all levels managers constantly have to take on this task which influences the process and the future of an organization. In international management these decisions have to be made between different cultures.

When a choice has to be made between alternatives (for instance goods and courses or actions), the attractiveness of each alternative can be expressed in terms of the scores of each alternative on a number of attributes and the weight (importance) of each of these attributes (Keeney & Raiffa, 1976). Weight assessment is a part of decision making. The alternative from which to choose from are described by attributes. These attributes are the characteristics that describe the alternatives. The importance assessment of these attributes is called the weight assessment process. (Heerkens, 2003)

The weight assessment process can be described by the Weight Assessment Model (WAM) devised by Heerkens (2003). His model is to my knowledge the only way in which to describe a weight assessment process. Other authors have devised methods for measuring importance that decision-makers attach to various attributes (Harte & Koele, 1995, Jaccard, Brinberg & Ackerman, 1986), but they only show the result of the process. In this paper we are particularly interested in the process to this result which can be described using the WAM. The WAM will be one of the foundations of this research. This paper will be a secondary analysis of Hans Heerkens research material, this material consists of 'think aloud' protocols. In these protocols the research subjects had to think aloud while making a weight assessment.

Typically arguments are used to justify the importance of certain attributes. In decision-making theory importance are the weights and weight combinations. Arguments are necessary not only to convince your boss you've made an acceptable assessment, but also to convince other stakeholders and yourself. This means weight assessment and the use of arguments are an integral part of decision making.

This thesis will address two kinds of cultural differences. The difference in the use of argumentation will be guided by the contradiction between holistic and analytical ways of thinking. This will be addressed in section 2.7.

Researching values and their presence in arguments should give us an insight on if people from other cultures use values in weight assessment. Schwartz & Bilsky (1987) listed five elements of values. These will be discussed in section 2.6 detail. For now I've devised a short definition of values from these elements: Values are implicit concepts or beliefs in a cultural group and they've got a guiding effect to socially preferable behaviour.

1.2: Problem statement

This thesis is aimed at two types of cultural differences: one part the difference in the usage of arguments in weight assessment (process) and the other part the presence of values in weight assessment (content).

There is no clear cut way to identify values in weight assessment. This means unambiguous identification rules have to be compiled. We will do this using extensive theoretical research and we will apply these rules to the raw data. Furthermore we have to extract indicators of holistic and analytical argumentation, we will do this using an article of Nisbett (2001).

There is little knowledge about values in weight assessment and there are no procedures to identify values in the weight assessment process. As stated in the introduction, the process of weight assessment has been described in detail by Heerkens (2003). He does not focus on values and on how to identify them. Other authors address weight assessment, but they focus primarily on the result of the process. We are interested in the content of weight assessment and the existence of values in this process. The procedures to identify values have to be made and tested.

In the first part of this thesis we're going to give a literature overview of the relevant concepts: weight assessment, attributes, importance, WAM, arguments, culture, holistic versus analytical and values. To fully comprehend and solve the knowledge problem regarding values in weight assessment a solid theoretical framework is needed. The next thing we are going to do is to make rules to identify arguments and values in 'think aloud protocols', these are based on the theoretical framework. These rules have to be tested and eventual be amended and/or changed.

In short there are six main goals in this research:
1. To create a solid theoretical framework regarding relevant concepts and their relationships: weight assessment, attributes, importance, WAM, culture, arguments, holistic versus analytical ways of thinking and values.
2. To develop a method to measure arguments in weight assessment processes.
3. To analyse the data on arguments in the light of the contradiction between holistic and analytical ways of thinking.
4. To develop a method to measure values in weight assessment processes.
5. To research a cultural difference regarding the presence of values as a basis for arguments.
6. To research a potential relationship between the amount of arguments and importance.

1.3 Research questions

This thesis is based on process and content of weight assessment. To research a cultural difference in these areas we've devised a main question which covers arguments and values in weight assessment.

Main question

To what extend is there a cultural difference in the usage of arguments and the presence of values as a basis for arguments in weight assessment processes?

To get a complete grasp of the subject at hand a solid theoretical foundation is needed. This means doing a literature study regarding: 'weight assessment processes', 'attributes', 'importance', 'the WAM' and 'culture.

- Weight assessment processes, in this part the aim is to introduce the concept weight assessment and the link with decision making.
- Attributes, full understanding of this concept is needed to grasp the WAM. This means knowing what attributes are and what their role in the weight assessment process is.
- Importance, the concept of importance is the key concept of weight assessment.
- WAM, as stated earlier the WAM is a model which captures the weight assessment process. Consequently the WAM helps to understand the weight assessment process, this is needed fully grasp the main question.
- An introduction of culture is needed to understand the possibility of differences in the process of argumentation and the presence of values in weight assessment.

With this foundation the theoretical framework is not yet completed, there are more relevant concepts. These concepts will be covered by the sub questions. The sub questions of this research are designed to pave the road to answering the main question. In the following chapter we discuss every sub question and we describe which steps are needed to answer the sub question.

1.4 Sub questions

The main question consists of the presence of values as a basis for arguments. This will be handled in het next three sub questions.

Sub question 1:
Is there a cultural difference in the usage of arguments in the light of holistic versus analytical argumentation?

- To do a literature study regarding arguments and the contradiction between holistic and analytical.
- To make criteria of indicators regarding arguments and holistic versus analytical argumentation.
- To operationalize the concept of arguments to the research situation: make arguments measurable in the think aloud protocols.
- To operationalize holistic versus analytical argumentation to the research situation: make indicators of holistic versus analytical argumentation.
- Apply the measurement rules about arguments to all the think aloud protocols and check if there is a difference between the experimental groups.
- Analyse the results in the light of the indicators of holistic versus analytical ways of thinking.

Sub question 2:
Is there a cultural difference in the presence of values as a basis for arguments in weight assessment processes?

- To do a literature study regarding values.
- Make criteria regarding values.
- Operationalize the concept of values to the research situation, make values measurable in the think aloud protocols.
- Apply the measurement rules about values to all the think aloud protocols. In practice this means counting the amount of values used as a basis for arguments in the think aloud protocols.
- Check if there is a cultural difference between the cultural groups in the light of the presence of values as a basis for arguments.

Sub question 3:
Researching the proposition: 'The more important an attribute the more arguments are given to support the importance of this attribute.'

- Analysing the results in the light of importance of attributes and the amount of values present in the procotols.

We now know what we want to research and how we are going to do this. The next section concerns the research population of this thesis.

1.5 Research method

This thesis is a secondary analysis based on the raw data gathered from research done by Hans Heerkens and later on Floor Richters and Julia Doednik.

Their research was aimed at a non-routine weight assessment by laymen. In this case the experimental subjects had to make a weight assessment between the main attributes 'safety' and 'comfort' of minibuses. This weight assessment was part of the acquisition of a new fleet of minibuses, more precise the experimental subjects were asked to give advice to the management of a fictional travel company that transported passengers to Amsterdam Shiphol Airport. The advice was aimed at the importance judgement of safety versus passenger comfort. In addition they would have to be able to explain their judgements to the management of the company. It was stressed that the subjects were allowed to perform the assessment in any way they liked and that there were no limits as to the kind of reasoning that was allowed. (Heerkens, 2003) More on the think aloud method in chapter 3: Measuring and Identifying.

1.6 Subjects for study

Three groups of students which comprise the research population:
- Australian students
- Russian students
- Dutch students

I've chosen these groups on the basis of their availability, the 'think aloud protocols' of these groups are available to me from earlier experiments and the experiment has been done in the native languages of the experimental subjects.

This research will be a secondary analysis of data collected for prior research. This research was done by different researchers, but with the use of the same experiment. The experiment was done in the experimental subjects native languages and this data is available for me to use.

For me the analysis of the Dutch think aloud protocols will not be a problem, because this is also my native language. The protocols derived from the Australian group are English and the analysis of these protocols will not pose a problem. The protocols regarding the Russians students are translated into Dutch by a native speaker of the Russian language. This means overall there will not be any big language barriers regarding the analysis of the think aloud protocols.

1.7 Limitations

- The research is aimed at individual assessments, not on group assessments.
    - It is not possible to use the think aloud method on group assessments.
    - Individual assessments are often the foundation of what happens in a group. You first have to understand the individual assessment before you can understand group assessments.
- The weight assessment is non-routine.
    - This means the weight assessment is unique, complex and the experimental subjects have to make an explicit weight assessment. In routine decision making there is a big change the decision-maker does not make an explicit weight assessment. Like buying new printer paper. In that case the decision-maker does not have to make an explicit weight assessment. He or she just orders the same kind he/she always does.
- The weight assessment is made in an organizational context.
    - In this context the weights are made explicit. This leads to a more natural experiment, it is not odd for the experimental subjects to make an explicit weight.
    - This context provides relevance for Business Administration, my area of study.
- The experimental subjects are laymen. This means the experimental subjects have no knowledge on the area of the attributes and have no knowledge or experience on assessment processes.
    - A weight assessment is always non-routine for a layman.
    - Due to the lack of experience a laymen has no prejudice regarding the contents of the weight assessment.
- The experiment takes place in the native languages of the experimental subjects.
    - In this way you'll avoid any problems regarding a language barrier concerning the experimental subjects. This could be an important concern especially with reference to the use of think aloud protocols.
- The weight assessment process is being described by the Weight Assessment Model (WAM) devised by Heerkens (2003)
    - This model is to my knowledge the only way to describe a weight assessment process.
    - This model is needed to identify which arguments are aimed at justifying importance.
- Values are measured using the value-type classification made by Schwartz (1994).
    - He uses explicit values which are divided into 10 value-types.
    - By using a concept of values in which they are already explicit you'll avoid any problems with the operalization of (implicit) values into explicit and measurable values in think aloud protocols.

1.8 Structure of this thesis

Chapter 2: 'Theoretical Framework' is the elaboration of the relevant concepts needed to answer the questions from chapter 1. Chapter 3: 'Measuring and Identifying' will use the concepts from chapter 2 and operationalize them into measurable constructs. Furthermore this chapter will handle the research method in more detail. Chapter 4: 'Results' is aimed at the results of the research by answering the sub questions. Chapter 5: 'Conclusion and Discussion' will cover the proposition as stated earlier, but the main goal of this chapter is to answer the main question of the research along with a critical discussion of this thesis.

1.8 Structure of this thesis

Chapter 2: Theoretical Framework

Chapter 2 of this thesis is aimed at defining the concepts at hand. In order to fully comprehend the main question and the sub questions a literature study of concepts is needed. Subsequently this chapter will serve as a solid theoretical basis regarding: weight assessment, attributes, importance, WAM, arguments and culture. Furthermore we will go into holistic versus analytical thinking and values, to answer the sub questions.

2.1 Weight assessment processes

Weight assessment and decision making are closely interlinked. Decision making is making a choice between two or more alternatives. The decision making process consists of the identification of problems and chances for every available alternative and the weighing of the different alternatives. (Howard, 1988)

For the weighing of the different alternatives you'll have to use the attributes that describe the alternatives. Determining the importance of attributes is called weight assessment. (Heerkens, 2003)

For example, when you're buying a new car, the available cars at the dealership are the alternatives from which you can choose. These alternatives can be described using the attributes of these cars. For example: 'The maximum speed of car A is 180 km/h?'. These alternatives can also be described using the importance of attributes. For example: 'The maximum speed of a car is of great importance'.

In short, the weight assessment process is the process of determining the importance of all attributes. The weight assessment can be structured using the Weight Assessment Model (WAM), this model consists of 7 main phases and can be grouped into three categories.

We will discuss the 7 phases of the WAM. This will cover a few important concepts: weight assessment processes, attributes, sub attributes, importance and weight. For the complete understanding of weight assessment processes we will first define importance. These concepts are vital to the WAM. For the understanding of the concept they need to be explored before the WAM itself.

2.2 Importance and weight

Importance is a term which is used very much, but if you ask someone to define importance they would most probably have a difficult time explaining the concept. This is not strange, there are multiple different definitions of importance. Some definitions are rather general and others are very precise.

A very general definition of importance is stated by Jaccard, Brinberg & Ackerman, (1986):
An attribute is said to be important if a change in the individual's perception of that product attribute leads to a change in the attitude towards that product. With this definition there is still some vagueness around the concept importance. Another definition of importance is given by Alpert (1971), important attributes are attributes that are essential for an object but are shared by all alternatives and thus do not determine choice. There is much confusion between importance and determinance, which is the influence on choice. We can see that there is no real consensus about the concept importance. In this thesis we use a more precise definition which is used in many experiments in which subjects have to judge the importance of attributes. This definition links importance with attractiveness for the decision maker.

Importance is the relative influence of the attitude concerned on the attractiveness (in the eyes of the decision maker) of each of the alternatives to be chosen from. (Heerkens, 2003)

This definition of importance is the same as the definition of 'weight' in decision theory. There are reasons to believe that this definition of importance is used by subjects when working with experiments. (Heerkens, 2003)

This means if an attribute is very important and a certain alternative scores relatively high on that particular attribute, this alternative will be relatively attractive in the eyes of the decision maker. For example when you're buying a new car. If a decision maker finds 'safety' to be an important attribute and a car X scores high on safety tests, car X will be relatively attractive in the eyes of the decision maker. The more important the attribute is and therefore the higher the weight that is given to a certain attribute the faster the attractiveness increases when the score on the attribute increases.

2.3 Weight Assessment Model (WAM)

In the previous parts we've discussed importance. This is needed to comprehend the WAM. In this section we will further discuss this model introduced by Heerkens. (2003). To understand the differences between the phases a distinction between attributes and sub attributes is needed. This distinction will be covered first.

2.3.1 Attributes and sub attributes

In the first section of this chapter we've seen an example about buying a new car and choosing between alternatives. The question is what makes the alternatives differ from each other. To show how alternatives differ from each other you would use the characteristics of the different alternatives.

For instance when you're buying a new car. Car A might have a higher top speed, has ergonomic and adjustable seats or is just simply cheaper. These characteristic of the car (top speed, comfortable seats and price) are called 'attributes' in decision theory and describe the alternatives to choose from. (Heerkens, 2003)

The term sub attributes is used when an attributes is split into smaller (sub) attributes. You could split an attribute like safety into multiple sub attributes. For example: the presence of seatbelts, airbags and anti-lock braking systems.

The distinction between attributes and sub attributes will serve as an indicator for holistic versus analytical ways of thinking in section 2.7.

2.3.2 Phases of the WAM

Heerkens (2003) identified 7 phases of the WAM. They are presented in a sequential way in table 1. In reality subjects often go back and forth between phases and most often address phases more than once. The model is divided into a structuring cluster (phases 1 and 2), a weighing cluster (phases 3 to 6) and an evaluation cluster (phase 7). In the structuring phase the problem is formulated and the attributes are processed in such a way that they can be weighted. The weighing takes place in the weighing cluster. After the weighing the subject reflects on their process and/or results in the evaluation phase. (Heerkens, 2003)

Phase	Category	Phase name
1.	Structuring cluster	Problem identification
2.	Structuring cluster	(Sub-) attribute processing
3.	Weighing cluster	Absolute sub-attribute weighing
4.	Weighing cluster	Homogeneous sub-attribute weighing
5.	Weighing cluster	Heterogeneous sub-attribute weighing
6.	Weighing cluster	Attribute weighing
7.	Evaluation cluster	Evaluation

Table 1: The Phases of the WAM (Heerkens, 2003)

Phase 1: Problem identification
In this phase the subject is going to get a better grasp of the problem at hand. This could consist of activities like reading the assignment or reformulating the assignment for a better understanding of the situation. When reformulating subject may say what they are going to do, but they could also mention things they are not going to do. In this phase the subjects are going to define the task lying ahead.

Phase 2: (Sub-) attribute processing
When you want to weight attributes you first have to know what you are weighing. You have to process the attribute in such a way that is becomes suitable for weighing. For instance making an attribute more concrete by defining it in such a way that is has a measuring unit. For example you could measure the attribute safety by using Euro NCAP (European New Car Assessment) crash test scores.

There are five ways of processing an attribute:
- Decomposing. Splitting up an (sub) attribute into several sub-attributes.
- Re-formulating. The (sub) attribute is given a different name while meaning the same attribute with a similar, not necessarily identical, measurement unit.
- Concretising a (sub-) attribute.
- Integrating two or more (sub) attributes into a new 'sub-attribute.
- Making an attribute more abstract. This is the complement of concretising.

Phase 3: Absolute (sub) attributes weighing
A statement about the importance and weight of a (sub) attribute is made without making a reference to other (sub) attribute. For example stating that: 'Safety is very important'.

Phase 4: Homogeneous sub-attribute weighing
In contrast to phase 3, the weighing in phase 4 does reference to other sub-attributes. In this phase two or more sub attributes belonging to the same attribute are weighted against each other and arguments for the weighing is given. For example: 'Anti-lock braking systems are more important than airbags. Both anti-lock braking and airbags belong to the same main attribute, namely safety.

Phase 5: Heterogeneous sub attribute weighing
This phase contains the same sort of activities as phase 4, only the sub attributes to be weighted belong to different attributes. For example: 'Airbags are more important than comfortable seats.' Airbags belong to the main attribute safety and comfortable seats belong to the main attribute comfort.

Phase 6: Attribute weighing
In this phase the main attributes are weighted against each other. For example: 'Safety is more important than comfort'.

Phase 7: Evaluation

In this phase the subjects reflect on their actions and the results of their actions. For instance they could reflect on the chosen weights, the arguments for choosing these particular weights and on the way they have worked on the problem.

In this thesis we are interested in arguments, whether this is on the process of argumentation or on the content of argumentation. For this thesis we are interested in one particular type of argument: the argument that is aimed at justifying weights. These arguments can be found in the weighing cluster of the WAM (phases 3 up to and including phase 6). The next section will elaborate on these types of arguments.

2.4 Arguments

In section 2.1 we've seen the connection between weight assessment and decision making. Weight assessment is not directly aimed at choosing between alternatives, but at determining the importance of attributes that make up the different alternatives.

In decision making theory national culture is seen as an influential factor in decision making. Briley et al. (2000) see cultural differences in decision making while asking for arguments to support decisions. In this light a cultural difference can be expected. Briley et al. (2000) are focussed on the choices which are made on the bases of the arguments and not on the arguments themselves. This thesis is aimed at arguments in weight assessment processes and not on the choices made based on this weight assessment.

The influence of culture through values on decision making is has been shown by many studies. *'It has been suggested that cultural values have direct influence on personal work related values, and that they are essential in understanding general attitudes to work in modern organizational settings'* (Hofstede, pp. 225, 1998)

To research a cultural difference in arguments used in weight assessment we are going to look at different approaches in decision making. There are two approaches in decision making theory which are basically very different, the value based approach and the reason based approach. We will discuss these first before we go on to the role of argumentation.

2.4.1 Value based versus reason based approaches

Decision making theory distinguishes two different approaches: value based and reason based. Value based decision making is commonly used in economics, management science and decision research. This approach is characterized by assigning numerical values with each alternative and the ultimate choice is on the basis of maximization of this numerical value. (Shafir, Simonson en Tversky, 1993). In this light the distinction between weight assessment and decision making becomes important.

The value approach suggests choosing the alternative with the highest numerical value. In weight assessment the choice is not between different alternatives. The choice is the assessment of importance of different (sub) attributes. In the experiment the subjects had to assess the importance of safety and comfort and their sub attributes. Looking at the value based approach you'd expect experimental subjects to assign numerical values to the (sub) attributes in order to describe the importance of (sub) attributes. These numerical values can later on be used to choose the alternative with the highest numerical value. In practice this does not happen, experimental subject barely assign numerical values to (sub) attributes. Besides this, this value based theory is aimed at choosing between alternatives and not aimed at determining importance.

The value based approach, while interesting, is not suitable for this research. The other approach is the reason based approach. This approach is characterized by the identification of various reasons and arguments that influence the decision. These reasons and arguments explain decisions in terms of reasons for and against the various alternatives. The reason based approach is not as formal as the value based approach, highest numerical value prevails. The reason based approach is essentially quantitative and rather vague. But, there are two major advantages connected to the reason based approach. This approach is more suitable than the value based approach in complex and real world decisions. Furthermore a focus on reasons is closer to the way we normally think about decision making. (Shafir, Simonson en Tversky, 1993)

The reason based approach is natural way of looking at decision making and is a good method to analyse complex situations. The experiment in this research is complex. Even though there are only two main attributes. These attributes can be divided into an infinite number of sub attributes with an infinite number of possible weights. Using the reason based approach also seems logical in the light of the other advantage. This approach represents a way of thinking which is close to the way we normally think about decision making. The subjects in this experiment are laymen, they have no knowledge or experience on decision making. An analysis based on 'how you'd normally think about decision making' seems logically.

The value based approach in this section should not be confused with the (cultural) values which are the subject of section 2.6. The value based approach is an approach to decision making, which comes down to choosing the alternative with the highest numerical value. Values (cultural) are the core of culture and guide behaviour. They are two fundamentally different concepts.

The reason based approach is focussed on arguments, more on this in the next section.

2.4.2 Arguments

In general an argument is the stated reason why the subject made a certain choice, the justification of a certain choice by a subject to themselves or to others. (Shafir, Simonson en Tversky, 1993). As we've seen in section 2.4.1 these reasons can be for and against alternatives. This means they can be positive, they encourage the choice of a certain alternative. They could also be negative, they discourage the choice of a certain alternative.

In the light of the current research we can make the previously mentioned choice more specific. In this thesis we are interested in choices associated with importance assessment. In the specific situation of weight assessments an argument would be used by subjects to justify certain weights and weight combinations. (Heerkens, 2003) This definition is especially aimed at the WAM. In the light of importance the arguments we are interested in are only present in the weighing cluster, phases 3 through 6. These arguments are aimed at the content of the weight assessment and are directly linked to weights.

You could also find arguments in the other phases of the WAM. For instance in phase 1, the justification of certain reformulations the experimental subject has made. Or, in phase 2 of the WAM where you would for example find arguments justifying the decomposing of (sub) attributes. This might have an effect on importance, only this has not been proven empirically at this time. In light of this observation we assume that only arguments present in the weighing cluster of the WAM are directly linked with importance. Arguments from other clusters of the WAM will not be taken into account in this research.

In this thesis we are mostly interested in unique arguments and arguments which are directly linked to weights. With chains of arguments there could be issues on both areas. The chain of arguments

does not consist of a repetition of the original argument in different words. Each argument is new and unique. The arguments which form a chain are aimed at supporting each other, but they are all linked to the same weight assessment.

Example:

- *'When buying a minibus good seats are very important'. The first order argument could be: 'As a taxi driver you're driving long distances and you need to sit correctly'. The second order argument could be: 'If you do not sit correctly you'll likely get back problems.' The third order argument could be: 'If you've got back problems you're not able to do your job properly.'*

Al these arguments are aimed at the importance of good seats in a minibus; but not all do so directly. The second order argument is aimed at the first order argument and the third order argument is there to support the second order argument. In this way you'll get a chain of arguments.

It might be the case that a 3rd order argument cannot be linked to a certain weight without the other arguments in the chain, because these arguments are there the experimental subject explicitly makes this link. The key point of these chains is that even though the arguments in the chain are all aimed to justify the same weight or weight combination they are all individual and unique arguments and are treated as such.

Concluding, there are a number of things to analyse on the field of arguments. The main point is the justification of weights and not choices in general. The next point is the uniqueness of the arguments and the usage of chains. The third part of the analysis is on positive and negative argumentation.

This thesis is focussed on cultural differences in weight assessment, the next section is aimed at this subject.

2.5 Culture

The thesis is aimed at researching a possible cultural difference in argumentation between the three experimental groups. Communication between people of different cultures can only be successful if both parties are aware of the difference in the perceptual field. A message falling outside the receiver's perceptual field, therefore, cannot transmit the sender's meaning, although it may well transmit a meaning not intended by the sender. (Root, 1994).

We're researching a possible difference in the process of argumentation by looking at holistic versus analytical approaches (see section 2.7). We're also looking at the content of argumentation by looking at the presence of values in arguments (see section 2.6). Before going into depth on holistic versus analytical approaches and values we are going to look at the broad concept of culture. Culture can explain differences in the way people think (process) and what people think (content).

Culture is a concept which is used a lot, in some cases even misused. Earlier on in this thesis we've adopted the definition of Hofstede as a basis for culture. According to Hofstede (2004) culture is: *"The collective programming of the mind that distinguishes one group or category of people from another."*

In the introduction this was a very practical definition, short and clear. For a better understanding of the concept the definition of Hofstede will be expanded into a more elaborate framework. This will be done on the basis of work done by different researchers.

There is widespread consensus among anthropologists about four elements of culture (Hall, 1976).

- Culture is learned by individuals as they grow up in their society.
- The elements of culture are an integrated whole.

- Culture consists of behaviour traits that are shared by members of a social group.
- Culture distinguishes a group from other groups with different cultures.

A culture provides rules that enables members of the same social group to cope with the physical and social environment in that social group. The rules include existential values that shape perceptions of the world and norms that guide social behaviour. These rules do not only guide our own behaviour, but also raises an expectation on how we expect others to behave. (Root, 1994) In short culture influences the way we think and act, this is not directly visible but manifested in behaviour.

The preceding mentions about culture and the influence of culture on behaviour could raise the expectation of enormous cultural influences on daily life and that all individuals of a cultural group would portrait the same behaviour. In practice this is not the case. Behaviour may to a certain extent be determined by collective programming, but there are also other forces which play a role in determining behaviour.

Culture is one of three levels of mental programming. (Hofstede, 1991) This is captured in figure 1, this triangle represents the three levels of human mental programming.
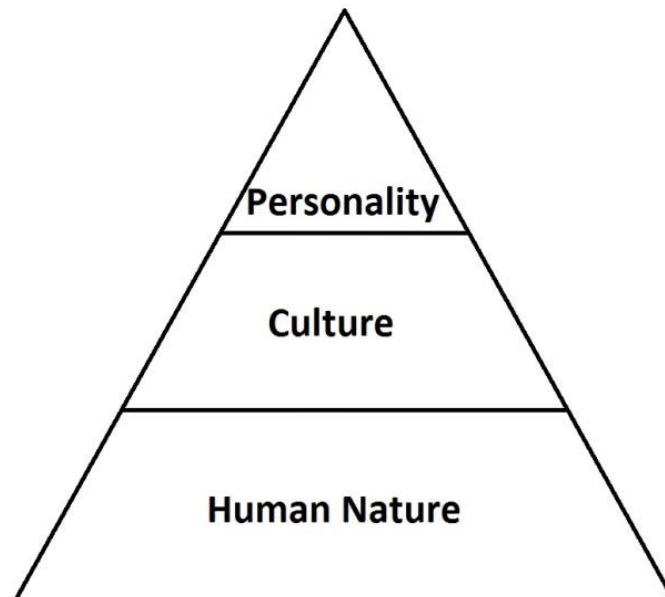
Figure 1: Three levels of uniqueness in human mental programming

Human nature relates to universally shared basic characteristics. This comes down to behaviour aimed at fulfilling basic human needs. For example the need for food and water. The characteristics of human nature are not a distinguishing factor, all of mankind share these basic human needs. Personality is the opposite of human nature and is very specific. This means it differs from individual to individual. The middle part of this triangle of mental programming is the cultural level, this level determines the collective human programming. (Hofstede, 1991)

All these three levels determine human behaviour, but in this thesis we are only interested in the cultural part of mental programming. With this background information on human mental programming we are going to look more closely into culture. Hofstede (1991) pictures culture as an onion, see figure 2.
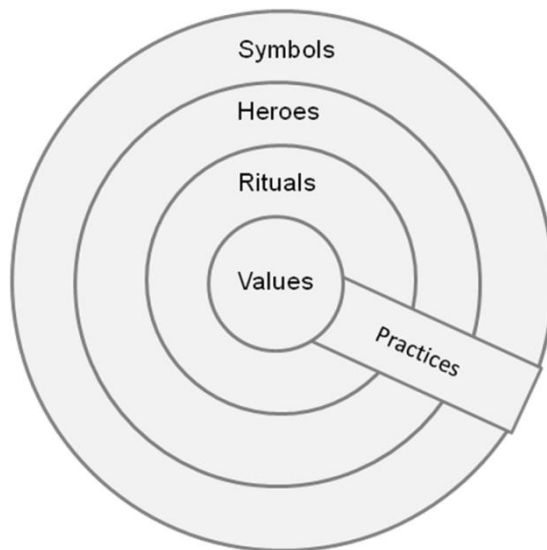
Figure 2: Layers of culture.

The outer rings of this onion are most visible. Symbols include languages, flags and particular ways to dress or not to dress. In the case of the Netherlands you would immediately think of windmills and wooden shoes.

Heroes are persons who serve as a model for preferred behaviour, these could be real or imaginary and dead or alive. A good example of this would be the king or queen as head of state in the Netherlands.

Rituals are collective activities, like social and religious ceremonies. But, also the way we do simple things. For instance the way we greet each other when we meet, the Dutch shake hands while the Chinese greet each other with a nod of the head.

At the heart of the onion we can find values, these values are really the core of culture and they affect the outer rings of the onion. The outer layers of this model are visible through practices. Values are, unlike the outer layers of culture, not visible. (Hofstede, 2004) We will discuss the problems concerning measurability connected to this invisibility in section 2.6.

Culture also creates differences in the systems of thought. These differences have been around for a long time and they still exist between China and other East Asian societies and America and other European-influenced societies. Section 2.7 will handle the contradiction between holistic an analytical systems of thought and the influence on argumentation.

2.6 Values

We've now got a clear picture of weight assessment, attributes, importance, WAM, arguments and culture, but in what way can this be linked to the content of argumentation?

As we could see in 'culture as an onion' in section 2.5, values are at the core of culture and therefore you can expect a difference in the use of values between different cultures. This makes the concept of values relevant for this thesis on cultural differences in weight assessment. But the statement about values raises a lot of new questions which should be answered. In this section we're going to address the question: 'what are values?'.

We stated earlier that culture is a frequently (mis)used concept. Values is also a concept on which there is a lot of ambiguity. Hitlin and Piliavin (2004, p. 359) also note this observation: *When employed, the term 'values' has been used variously to refer to interests, pleasures, likes, preferences, duties, moral obligations, desires, wants, goals, needs, aversions and attractions, and many other kinds of selective orientations.* With this notion about values it becomes clear that values is a label which can be placed on a number of things. Furthermore the conceptualization and measurement of values is very incoherent and differs from one scientific discipline to another. *Most surprising is the almost complete lack of reference that the major empirical researchers on values make to relevant social theory and vice versa.* (Hitlin and Piliavin, 2004, p. 360) This remark gives a clear picture the ambiguity of values, the lack of consensus on the matter and the misuse of the concept in general.

With the lack of consensus it is wise to explore multiple definitions of the concept 'values' and to search for dominant or recurring element in those different approaches to values. Hitlin and Piliavin

(1994) made a review of the concept values. They discussed multiple definitions and ways of thinking about the concept. The next part will be largely based on that paper.

*Perhaps the most influential definition of values traces back to Kluckhohn (1951, p. 395): "A value is a conception, explicit or implicit, distinctive of an individual or characteristic of a group, of the desirable, which influences the selection from available modes, means, and ends of action. (*Kluckhohn as cited in Hitlin and Piliavin, 1994, p. 362) Values are shared by a (cultural) group and will guide people to take the desirable action. This might not be in the best interest of the person itself. The desirable is not the personal property of the person, but it relates to the desirable for the social group.

From the perspective of Kluckhohn, values will always lead to the socially desired action. According to Lesthaeghe and Moors as cited in Hitlin and Piliavin (1994, p. 362) he overstates the influence of values. They argued that Kluckhohn '*takes a functionalist, deterministic view in which values are cultural imperatives that necessarily lead to certain actions.'.* Values are enduring beliefs that a specific action is socially preferable to an alternative action. Knowing what is socially preferable is does not imply that this road is automatically chosen. People also have a mind of their own and they could choose to ignore the socially preferable road. Even if they share (some of) the involved values and agree with this socially preferable road. For instance this could be the case with conflicting values and conflicting personal interests. This concurs with the notion of Hofstede (1991), not only culture and therefore values determine human behaviour. Human behaviour is also guided by other influences, see figure 1 in section 2.5. These other influences are not the focus of this thesis. Human nature relates to basic human needs. In this category you'd expect no differences between the experimental groups. Human nature is not a distinguishing factor. Personality is a distinguishing factor on individual level. Every human being is different. This cannot be used to research group differences. Looking at personality does not provide characteristics on which you can identify groups and group differences. Culture does distinguish groups and is shared by the individuals in the group. The core of this thesis is cultural differences in argumentation. This section on values is on the content of argumentation. This is done by looking at cultural differences in the presence of values.

We've now got a broad idea of what values are, but what are the commonalities in all this theory on values. Schwartz & Bilsky (1987, p. 551) listed five elements that are dominant in the literature. *According to the literature, values:*
- *Are concepts or beliefs.*
- *Are about desirable end states or behaviours.*
- *Transcend specific situations.*
- *Guide selection or evaluation of behaviour and events.*
- *Are ordered by relative importance, not all values have equal influence.*

The first notion of the cultural foundation of values is certainly a recurring theme in the literature. It is safe to say that values are at the core of culture. Therefore we can expect difference in the relative importance of values between cultures and a different use of values among different cultures. Furthermore the values that exist in a certain social group or culture are shared beliefs which are implicitly or explicitly taught from generation to generation.

The next element of values is about the guiding aspect of values. Values serve as a guiding mechanism to the socially preferred action or lead to the socially preferred end state. But I agree with the previously mentioned notion of Lesthaeghe and Moors about the deterministic view of Kluckhohn and the possibility of overstating the influence of values. Values are indicators of what is socially preferable, but it will not automatically lead to that specific choice. People also have a mind of their own and they could also deviate from the socially preferable road. This concurs with the

concept of culture and the other two levels of human mental programming devised by Hofstede (1991).

In the section about culture as an onion (Hofstede, 1991) we've seen that values are not visible and in the light of the previous section they can only be identified in their guiding effect on behaviour. In short: you can only observe values through their influence on socially preferred end states or behaviour. Schwartz (1994) identified ten value-types based on research done within 70 cultures around the world. All of these values are aimed at specific behaviour or end states. Instead of using a definition of values as being implicit. We are going to use specific and explicit values divided into 10 value-types.

To observe values in this thesis we use these ten values by Schwartz, each value is defined in terms of the broad goal it expresses:

- **Hedonism**
  Defining goal: pleasure or self-centred sensual gratification. (pleasure, enjoying life and self-indulgent)
- **Power**
  Defining goal: status and prestige, control over people and resources, emphasize the attainment or preservation of a dominant position within the more general social system. (authority, wealth and social power)
- **Achievement**
  Defining goal: personal success through demonstrating competence according to social standards, thereby obtaining social approval and recognition. (Ambitious, successful, capable and influential)
- **Stimulation:**
  Defining goal: excitement, novelty and challenge in life. This value-type stimulates maintaining an optimal and positive, rather than a threatening level of activation. (a varied life and an exciting life)
- **Self-direction**
  Defining goal: autonomous thought and action, which means a certain degree of independent choosing, creating and exploring. (self-respect, freedom and privacy)
- **Universalism**
  Defining goal: understanding, appreciation, tolerance, and protection for the welfare of all people and for nature. Universalism combines two subtypes of concern: for the welfare of those in the larger society and world and for nature (social justice, equality, world at peace, unity with nature, wisdom and protecting the environment)
- **Benevolence**
  Defining goal: to preserve and enhance welfare of those with whom one is in frequent personal contact. Most critical are relations within the family and other primary groups. Benevolence values emphasize voluntary concern for the welfare of others. (helpful, honest, responsible, forgiving, loyal, true friendship and mature love)
- **Conformity**
  Defining goal: restraint of actions, inclinations and impulses likely tom upset or harm others and violate social expectations or norms. Conformity entails complying with social expectations or norms, they promote cooperation in order to avoid negative outcomes for self.(obedient, self-discipline, politeness, honouring elders and responsibility)
- **Tradition**
  Defining goal: respect, commitment, and acceptance of the customs and ideas that one's culture or religion provides. Which means engaging in traditional and religious activities. Tradition entails subordination to more abstract objects, like religious and cultural customs and ideas. (respect for tradition, devout and accepting my portion in life)

- ***Security***
  Defining goal: safety, harmony and stability of society, of relationships and
  of self. There are two subtypes of security: some serve primarily individual interests, others
  might serve wider group interests. (social order, family security and national security)

All these values are empirically derived from research done in 70 countries. Among them are Australia, the Netherlands and Russia; which are the countries our subjects of study are from. So is the view on values of Schwartz fundamentally different from other researchers on this area? The answer is no. Schwartz follows the tradition of seeing values *'as the criteria people use to evaluate actions, people, and events.'* (Schwartz, 2006, p. 1)

In one culture power might be more dominant and in another culture this could be tradition, this is the relative importance of the different values within a culture. In different cultures you can see a different value mix. This difference in the relative importance of the different value types makes up the differences in culture regarding values.

The content of argumentation which leads to a certain behaviour can be measured. In this thesis we're going to identify the value types devisid by Schwartz (1994). Doing an analysis of the presence of the different value types gives the relative importance of these value types.

2.7 Holistic versus analytical

In the previous section we've seen a theoretical difference in the content of argumentation between cultures. Looking at experimental groups by analysing the presence of values is one way of looking at culture in weight assessment. In this thesis we're also going to look at the influence of culture though the holistic versus analytical systems of thought. The influence of culture on the process of weight assessment.

Holism is emphasizing the importance of the whole and the interdependence of its parts and this way of thinking is concerned with wholes rather than analysis or separation into parts. While the analytical approach is focussed on the central object and its attributes, independent of the context. (Nisbett, 2001)

We are researching the process of argumentation in the light of this contradiction for two reasons:
- The level on which arguments are used (sub attribute versus attribute level) is an indicator of this contradiction. Analysing the level in which arguments are used we can see if one of the experimental groups is more holistic or analytical than the other groups.
- The usage of arguments on relative weighing is also an indicator of this contradiction. Analysing usage in which arguments to justify relative weights we can see if one of the experimental groups is more holistic or analytical than the other groups.

Nisbett (2001) wrote an article about the contradiction between the two concepts In the article he looks at historic differences between 'East' and 'West' and the contemporary impact on the way of thinking in the US and China. He uses the US and China to give really striking examples of this contradiction, because of the huge differences in culture between the US and China. He describes this difference as : '*China and other East Asian societies remain collectivist and oriented toward the group, whereas America and other European-influenced societies are more individualist in orientation*.' (Nisbett ,2001, p. 295)

To elaborate on holistic versus analytical systems of thought only the paper of Nisbett (2001) was used. To be critical it is necessary to review multiple sources in order to come to a theoretical

framework. In the case of the contradiction between holistic and analytic systems of thought there wasn't much more theory other than articles of Nisbett himself. See appendix 2: Searching for theory, for more detail on this search for theory.

On the basis of Nisbett (2001) we've seen that there is in fact a cultural difference between nations in ways of thinking. The think aloud experiment which is the basis of this thesis results in a verbal representation of the thought process linked with weight assessment. In this process we can identify arguments (see section 3.2: Identifying arguments). The theoretical differences between holistic and analytical systems of thought and the possibility to identify arguments allows us to see if there is a difference between the groups in this experiment. To research this possible difference we have to look more closely at holistic and analytical ways of thinking and devise indicators of both ways of thinking.

'The authors find East Asians to be holistic, attending to the entire field and assigning causality to it, making relatively little use of categories and formal logic, and relying on "dialectical" reasoning, whereas Westerners are more analytic, paying attention primarily to the object and the categories to which it belongs and using rules, including formal logic, to understand its behavior.' (Nisbett, 2001, p. 291)

The most interesting element in the light of the process of argumentation is the contradiction between the need for context.

The holistic approach is characterized by emphasizing the importance of the whole, the interdependence of its parts and the relationship between objects. (Nisbett, 2001) In the light of weight assessment and argumentation you would expect a holistic experimental subject to spend more time on arguments concerning relative weighing. This way he/she justifies the relationship between attributes. The relationship between attributes is a vital part of seeing the interdependence of the different attributes for the final weighing. If you do not use relative weighing you cannot see the interdependence of its parts and the relationship between objects. So the amount of arguments aimed at relative weighing is good indicator of a more holistic approach to argumentation. Regarding the level of argumentation, if you go really deep into sub attributes you lose sight of the whole. Holistic argumentation would take place on the attribute level in order not to forget the bigger picture involved with holism.

The analytical system of thought is characterized by the tendency to focus primarily on the central object and its attributes, independent of the context. (Nisbett, 2001) Argumentation is not focussed on relationships between attributes, but at one attribute at the time independent of the context. In the light of weight assessment and argumentation you would expect an analytical experimental subject to spend more time on arguments concerning absolute weighing on sub attribute level. This way he/she can really focus on the different sub attributes without looking at the context. If you use absolute weighing you focus on one (sub) attribute at the time and you can see this (sub) attribute independent of the context. So the amount of arguments aimed at absolute weighing is a good indicator of a more analytical approach to argumentation. Regarding the level of argumentation. If you go really deep into sub attributes you can also see the separate elements independent of the context. Analytical argumentation would take place on the sub attribute level in order really focus on the separate parts of the weight assessment.

Concluding: the amount of arguments aimed at relative versus absolute weighing and the amount of arguments aimed at the attribute level versus the sub attribute level are good indicators of a holistic versus an analytical approach to argumentation. With these indicators on holistic and analytical argumentation we can make a good analysis on the cultural differences in the process of argumentation.

The next chapter is about measuring and identifying arguments and values in think aloud protocols. First we will discuss the think aloud method after this we will operationalize the concepts arguments and values form the theoretical framework into measurable constructs.

Chapter 3: Measuring and identifying

This chapter, about measuring and identifying, builds upon the previous section. In the previous section we've discussed the relevant concepts. We now know what they are and what their characteristics are. In this chapter we are going to use the theory from the previous section in order to make arguments and values measurable in think aloud protocols. This means we have operationalize these concept. Concretely we have to make decision-making-rules regarding identifying argument and values. These decision-making-rules should answer to the questions: 'What terms should be met in order to classify expressions in the think aloud protocol arguments?' and 'What terms should be met in order to classify arguments as being value based?' Before we answer these questions we're going to explore measuring in the 'think aloud' method.

3.1 Think aloud method

This thesis is a secundary analysis of the raw data from previous work done by Hans Heerkens, Floor Richters and Julia Doednik. Their research was based on the 'think aloud' method. The principle of this method is simple: let an experimental subject do an exercise, in this case a weight assessment, in which they have to verbalize all off their thoughts.

3.1.1 The method

Think aloud protocol analysis is a way of studying reasoning processes. The experimental subjects are asked to do an exercise, in this case assessing the importance of (sub) attributes when buying a new taxi bus. The use of the think aloud method stimulates the experimental subjects to go through the whole process explicitly, to verbalize all of their thoughts. This verbalization can then be transferred into writing, the written verbalization is called a 'think aloud protocol'. The analysis of these protocols is a good way of information gathering on cognitive activities (Ericsson and Simon, 1993) Using this method we can get a clear picture of the thought process of the experimental subjects. This thesis is aimed at the process and content of these thoughts and therefore this method is perfectly applicable. Furthermore analysing these think aloud protocols is a common method while studying clinical reasoning. (Banning, 2008) Clinical reasoning entails applying knowledge to a certain situation to develop a solution. Banning did her research in the field of nursing. She studied the reasoning process in problem solving using the think aloud method. In this thesis the experimental subjects have to come to a solution in terms of importance assessment.

Practically speaking this method is very suitable for a secondary analysis. The think aloud method produces very rich but rather unstructured data. This means the variables do not have to be specified in advance. Experimental subjects have got unlimited freedom in their thought process. Therefore we get the rich and valid data which represents the thoughts of the experimental subjects. (Ericsson and Simon, 1993) Due to the unstructured nature of this data the analysis of the data is difficult and there are a lot of pages with this unstructured data. Making sense of this data requires solid theory building and clear decision making rules. The usage of the WAM (Heerkens, 2003) helps to structure the data and gives us insights in the process of weight assessment. For instance, the phases of the WAM are explicitly used to seperate arguments linked the actual weighing from arguments that are not.

3.1.2 Intercoder reliability

This thesis is an individual assignment. In this case there was no possibility of an independent coder. Which means the person analysing the protocol is the same person as the researcher who has constructed the coding scheme. According to Van Someren et al. (1994) this could cause a problem,

the researcher is usually too attached to a certain research hypothesis to do the coding with an objective mind. The solution is simple: find coders who are not involved in the research project and who have no specific interest in the outcome. This will give the best guarantee for objective (reproducible) coding, which of course is the ultimate goal.

In this research it was not possible to use independent coders, to overcome this difficulty I took a number of steps. These steps are taken to prevented problems like Van Someren et al. (1994). By discussing the results and adding the examples from that discussion into the coding rules I feel confident about the correctness of the coding.

The steps to ensure objective coding:
- Made clear coding rules.
- Applied the coding rules to three protocols.
- Adjusted the coding rules and clarified them with real examples from the protocols.
- Discussed the new coding rules with my examiners.
- Adjusted the coding rules according to the result of the discussion.
- Applied the coding rules to all the protocols and discussed the results with my examiners. Before this discussion I'd prepared a short document which contained examples of coding segments which would fall in some sort of 'grey area'.
- After discussing these grey areas I've decided to apply the coding rules in a very strict manner. These grey areas, including the discussion, are added to the already existing examples to clarify the coding rules.

What is the result of all this: clear coding rules enhance intercoder reliability. When analysing the same protocols different coders should come to the same conclusions. Because there is only one coder I've set up a general rule: 'When there is doubt about a segment being an argument or value be conservative and not mark that segment as an argument or value.'

3.2 Identifying arguments

In chapter three we've discussed the concept of arguments in theory. In this part we have to operationalize this theory into decision-making-rules about when a phrase can be classified as an argument, so called coding rules.

For this we're going to use criteria of arguments and apply these criteria to the 'think aloud protocols'. The concepts we've abstracted from the theoretical framework are: 'justification of choice, link with importance assessment, chain of arguments and repetition'.

To do measurements in the think aloud protocols these concepts should be made into decision-making-rules. This is easier said than done, for each of the criteria there will be some grey areas whether or not a phrase fully satisfies the demands of the criteria at hand. To overcome this problem these grey areas have to be explored and decisions about black or white have to be made. To justify the choices made in the grey areas these choices will be accompanied by examples, not only to justify but also to clarify.

As a general rule, all of the criteria have to be met in order to label a phrase as an argument.

3.2.1 Justification

Arguments are used to justify choices to themselves or to others. The choices which are made are not part of the argument, only the justification is.

Example

- *A good car stereo is important, as a driver you'll not always have customers in the taxi bus, playing some music during that time would be a good way to relax.*

The choice in this example is the importance of a good car stereo this is not part of the argument, only the second part regarding relaxing while waiting on new customers is.

Arguments should have a direct link to choices. This link should be clear and this link should be made by the experimental subject, not by the coder.

Example

- *I think comfortable seats are very important, as a taxi driver you're driving long distances and you'll need to sit correctly.*

The choice in this example is the importance of comfortable seats. The argument which follows is directly and clearly linked to this choice. There is no need for creativity from the coder to artificially link these two segments. Not only is the content of the argument rationally linked to the choice, but also the proximity of the two segments is an indicator of a direct link.

Example

- *I think comfortable seats are very important, imagine a customer who has had a long flight, they probably want to get home as soon as possible.*

In this example there is no direct link between the choice of the importance of comfortable seats and the following segment. Having comfortable seats does not imply a quicker trip from the airport. The second part of this example could be an argument regarding the importance of a high top speed, but it is not up to the coder to artificially make this link.

Example

- *If comfort should be equal to a normal car, the taxi bus should have individual seats like a normal car.*

In this example you could say that the choice for individual seats is justified by the idea of having comfort equal to a normal car. In this case the notion of individual seats is not a choice, but a concretization of comfort equal to a normal car. In this example the experimental subject has not made any choice and there is no argument.

Example

- *Personally I think engine noise is always something which can totally ruin comfort.*

In this example the experimental subject does not explicitly links engine noise with an importance assessment, but he/she makes very clear it could totally ruin comfort. This is an implicit importance assessment for having low engine noise. The argument in this case is: the personal notion of ruined comfort. This is typically an example of an argument which is in the grey area, but there is no need for great creativity to link low engine noise and comfort together. This will be coded as an argument, the experimental subject makes the link itself (although implicitly).

Example

- *Comfort is very important to me personally.*

In this example the experimental subject gives a notion of what he/she thinks is important, in this case comfort. The trigger word here is 'personally', this makes the example an argument. The

argument is that he/she thinks comfort is important. Therefore the importance assessment of comfort is implicitly made clear.

The most important criteria of an argument is: it should be used as a justification for a certain choice. Without the link to a choice a statement made by experimental subjects is not an argument. Argument and choice should be logically and directly linked. It is not up to the coder to fabricate this link artificially. If the link is not in the think aloud protocol there is no way to empirically indicate an expression as an argument.

3.2.2 Weighing

In the specific situation of weight assessments an argument would be used by subjects to justify certain weights and weight combinations. The arguments we are interested in are linked to weights. Arguments used in phases one, two and seven of the WAM are not taken into account because they have no direct link to weights.

Example

- *I think a high top speed is not very important, in the Netherlands you'll never get the chance to drive that fast.*

In this example the argument regarding driving fast is linked to the importance assessment of top speed.

Example

- *I'm going to read the assignment one more time, because things are not entirely clear to me yet.*

In this example the experimental subject gives a justification regarding reading the assignment one more time. The justification is not aimed at the weighing cluster, but the justification is linked to phase 1 (problem identification) and does not have any connection with the actual weighing.

Example

- *I'm going to split safety into passive and active safety, maybe this will give me more new ideas for sub attributes.*

In this example the phrase regarding getting new ideas for sub attributes is aimed at splitting up the attribute safety. This is not part of the weighing cluster, but this concerns phase two ((sub-) attribute processing) and does not have any connection with the actual weighing.

Example

- *I'm very satisfied about the way I handled the assignment, my assessments are clear and easy to justify.*

In this example the phrase regarding clear assessments and easy justification is not linked to weights or weight combination, but to the way the experimental subject grades his/her own performance. This is part of phase seven (evaluation) of the WAM and does not have any connection with the actual weighing.

This thesis is aimed at researching (among other things) a possible link between importance and arguments. Therefore only arguments directly linked with importance assessment are taken into account. In some cases there is an argument, but the importance assessment itself is missing. As a general rule we do not mark this as an argument. Only if it is evident that there is only one conclusion possible we would mark this as an argument.

Example:

- *It would not be a good newspaper article if one of your taxi busses is involved in an accident and all your passengers are dead.*

This example would be a good argument for the importance of safety, in fact the only logical conclusion of this argument is: safety is important. Therefore we mark this as an argument, even though the importance assessment of safety itself is missing.

Example:

- If you really focus on comfort the taxi will be more modern and probably the safety will also be higher.

This example could be an argument for the importance of comfort, because it also improves safety. It could also be an argument for the in importance of safety. If the taxi bus is comfortable it will also be safe, so safety is not important. In this example the experimental subject does not give an importance assessment and there are multiple conclusions possible. In this thesis this would not be marked as an argument.

3.2.3 Chain of arguments

To justify weight assessments experimental subjects could use a chain of arguments that support each other. First, second, third (and so on) order arguments are treated separately. They are all unique. Even though they are aimed at the same weight assessment. For example: a chain of three arguments aimed at the same weight assessment are treated the same way as three separate and totally different arguments. In both cases they will account for an increase of three arguments in the amount of values.

Example

- *When buying a minibus good seats are very important. As a taxi driver you're driving long distances and you need to sit correctly, if you do not sit correctly you'll likely get back problems. With back problems you're not able to do your job properly.*

This example holds one importance assessment regarding good seats and multiple arguments which are all linked to each other and to the choice. The first order argument is about driving long distances and the need correct sitting. This argument is directly linked to the importance of good seats. The second order argument concerns incurring back problems. This is aimed at strengthening the first order argument, but can also be directly linked to the importance of good seats. The third order argument regarding not being able to do your job properly cannot be directly linked to the importance of good seats without the first two arguments. So without the chain, in the case of the absence of the first two arguments, the third order argument would not classify as an argument due to the fact that it is not up to the coder to fabricate this link. In the case of this example this link is explicitly thought aloud by the experimental subject. In this way all the arguments in the chain are classified as unique arguments.

These are all individual arguments which form a chain to strengthen each other, they are all aimed at the justification of the same weight assessment, but they are all unique and therefore counted as separate arguments.

A chain of arguments contains multiple arguments that are linked together by the experimental subjects, but they are not the same. Higher order arguments are aimed at strengthening lower order arguments. Sometimes the higher order arguments have no direct link to the importance assessment without the lower order arguments. So we only take these arguments into account if the chain of arguments is made explicit by the experimental subjects. In the case of a chain all the arguments in the chain are counted as separate arguments.

3.2.4 Repetition

With repetition there is no question of first, second, third order arguments the presented arguments are not new or unique. They only serve as a clarification of the original argument. This is not a justification, but a clarification or rephrasing of the original argument.

The presented arguments are not aimed at presenting new justification for the importance assessment, or to introduce new justification that strengthens the original argument. With repetition there is no case of novel information that would justify the importance assessment or to strengthen the original argument. The arguments have got the same content, they are only different in phrasing. To avoid double counting these clarifications are not taken into account as a new argument, but as part of the original argument.

Example
- *Safety is not very important in the choice between comfort and safety, in Holland we've got a APK. Every car and bus is checked every year, this is mandatory. Every vehicle is safe because of the safety check every year.*

In this example you could identify three arguments for safety of every car: '*in Holland we've got a APK*'; '*Every car and bus is checked; this is mandatory*'; and '*the safety check every year*'. The last two phrases are not aimed at something new, they are there to clarify and specify the original argument about the APK. The APK has got all the information of the latter two expressions, namely a yearly mandatory safety check. The clarification or rephrasing of this argument does not contain any new information. Therefore these phrases are not treated as separate arguments, but all as one argument.

The key distinction between repetition and a separate argument is the novelty of the information that makes up the argument. Is the information new, we will count the argument as a separate argument. Is the information in the argument not new, we will not count the argument as a new argument. Repetition will be coded separately, It is still an argument, only not a new one.

3.3 Identifying values in arguments

In chapter three we've discussed the concept of values in theory. In this part we have to operationalize this theory into decision-making-rules about when a phrase can be classified as a value, so called coding rules.

In the theoretical framework there has been a separation between general characteristics and more precise characteristics in the form of value types.

These more general characteristics of values would be very hard to measure. For instance measuring that values are the core of culture would theoretically be possible. In this research setting this is impossible. For example: '*In my culture you are responsible for the safety of the customers and therefore have to make sure you've got safe taxi busses.'* This example would be a great way you could encounter culture and values in the think aloud protocols. The experimental subject explicitly expresses the need for safety because of the way people in his culture think about responsibility. Theoretically it is possible to find expressions with similar content in the think aloud protocols, but it will highly unlikely if we would actually find expressions like this.

A more explicit view on values are the ten value-types by Schwartz (1994). To measure values in the think aloud protocols we're going to use these ten different types of values by Schwartz (1994)'. Furthermore we are going to make criteria of these values and apply these criteria to the think aloud protocols. These value-types are more specific than the general characteristics of values listed in the theoretical framework and are more practical to measure in the think aloud protocols

To do measurements in the think aloud protocols these concepts should be made into decision-making-rules. This is easier said than done. For each of the criteria there will be some grey areas whether or not a phrase fully satisfies the demands of the criteria at hand. To overcome this problem these grey areas have to be explored and decisions about black or white have to be made. To justify the choices made in the grey areas these choices will be accompanied by examples, not only to justify but also to clarify.

3.3.1 Universalism

Defining goal: understanding, appreciation, tolerance, and protection for the welfare of all people and for nature. Universalism combines two subtypes of concern: for the welfare of those in the larger society and world and for nature.

Example:

- *I think the bus should be fuel efficient, it should use as little fuel as possible. This is better for the environment.*

In this example the experimental subject directly links the low fuel usage to protecting the environment. This relates to the protection of nature in the value-type universalism, in this case the experimental subject explicitly refers to a value.

3.3.2 Benevolence

Defining goal: to preserve and enhance welfare of those with whom one is in frequent personal contact. Most critical are relations within the family and other primary groups. Benevolence values emphasize voluntary concern for the welfare of others.

Example:

- *To quantify all of the sub attributes of comfort and safety is not really possible, it will never be objective, but subjective. It would be better to have everything quantifiable, you'll get a more honest comparison.*

In this example you could make a very logical assessment about the experimental subject: 'the experimental subject is really concerned with the honesty of the comparison of alternatives on sub attributes, because these sub attributes cannot be objectively measured. In this case he/she is concerned about the honesty of his/her way of working. This could jeopardize the relationship with the (fictional) management of the company.' In spite of the logic, this way of coding requires creativity from the coder. Earlier on I've decided to be strict with coding rules, not to get creative and therefore in this thesis these kinds of sections will not be seen as values.

Example:

- *It's probably ehm, it, it wouldn't be fair that some people get more of the ehh, climate control than other people, they all pay the same, so it should be evenly distributed throughout.*

In this example the question at hand is the even distribution of cold or hot air from climate control. The experimental subject is worrying about honesty. All the passengers are paying the same price, so all of these should get the same kind of luxury. To come to this there is no need for creativity and this will be marked as the value type benevolence.

3.3.3 Tradition

Defining goal: respect, commitment and acceptance of the customs and ideas that one's culture or religion provides. Which means engaging in traditional and religious activities. Tradition entails subordination to more abstract objects, like religious and cultural customs and ideas.

Example:

- *Safety is very important. That's my opinion and I can't really explain it… Sjeremetjevo has the most dangerous highway!*

In this experimental group the experimental subjects are frequently saying these sorts of things as an argument for the importance of the attribute safety. You could rule this as the value tradition, apparently there is something in the minds of these experimental subjects to have a high regard to safety due to unsafe road conditions in their country. In this thesis this is not a value, it is just taking environmental factors into consideration.

Example:

- *I pray to God nothing happens to the passengers when they are driven around on our dangerous highways.*

In this example the last part is not part of a value. In the previous example we've seen this is just taking environmental factors into consideration. These factors drive the experimental subject to a prayer to God, he or she is being devout. This is part of the value tradition.

3.3.4 Conformity

Defining goal: restraint of actions, inclinations and impulses likely to upset or harm others and violate social expectations or norms. Conformity entails complying with social expectations or norms, they promote cooperation in order to avoid negative outcomes for self.

Example:

- *My first thought is that it's very strange to be using safety as a way of weighing two things,*
- *because we have safety standards*
- *and as consumers we assume that when a government or a body sets a safety standard that… and that things comply with it then they, they should be good enough.*

The first segment is neither an argument nor a value. The latter segments could be both. They are certainly arguments to assess the attribute safety as not important. In this case this could be based on the value tradition. In this case you could also see this as the value conformity, complying with social expectations or norms. There is even the word comply in this segment, but simply ruling segments as values because they share a common word is stretching it. In this case there might not even be a value, because it is simply conforming to minimum legal requirements. This is a sort of thing which you could see a lot in the think aloud protocols. Experimental subjects use minimum legal requirements as a basis for assessments. In this thesis these segments are seen as using or interpreting knowledge about local conditions and this will not be marked as a value.

Example:

- *If I've cleared up everything about the possibilities, it is easier to see what customers want. That's most important, you should give customers what they want.*

  Due to the terms used in the first segment you could see this as the value conformity. Conformity entails complying with social expectations or norms. In some way this is what the experimental subject is doing in these cases. You could also see this as an economical way of thinking. From a marketing point of view it's a very good idea to give customers what they want. In this thesis segments about giving customers what they want are not marked as values.

Example:

- *If the taxi is not safe the customers will not only be very dissatisfied, but also very mad, because customers expect a safe journey.*

In this example the experimental subject is trying to avoid mad customers by providing safe journeys. In this case he/she is not only conforming to legal requirements or simply giving the customer what

they want, but he/she is conforming to a social expectation of safe transportation. In this case I've marked this as a value.

3.3.5 Security

Defining goal: safety, harmony and stability of society, of relationships and of self. There are two subtypes of security, some serve primarily individual interests, others might serve wider group interests.

Example:

- *High safety has to be standard.*
- *The car should be ready.,,,, I mean for dangerous situations on the road.*

In these example the experimental subject makes a remark about high safety without coupling this to an argument or any particular situation or person. This statement potentially serves a wide group of people, with the desired end state of keeping them safe. So there is a potential value and with a little creativity you could mark this as a value. Especially with the value security these kinds of situations often occur, you could suspect a value and with some creativity you could fit these segments within the definition of security. This is not the way to go, the rules should be applied in a strict manner. This is not marked as a value.

Example:

- *Safety is important, I have a duty of care to make sure that they're safe while we're travelling.*

In this example the experimental subject is stating safety is importan and he/she wants to ensure safety while traveling. There is no need for creativity to mark this as a value, but there could be some discussion about the value type. I marked this as security, it is aimed at providing safe travelling. This could also be benevolence. The experimental subject is talking about duty of care which could be linked to responsibility as a part of benevolence. The latter way of thinking takes some creativity, therefore I marked this as security.

3.3.6 Power

Defining goal: status and prestige, control over people and resources, emphasize the attainment or preservation of a dominant position within the more general social system.

Example:

- *I can give as many reasons as I want, but the management will not listen to me anyway.*

In this example the experimental subject complains about the lack of power they have over the management, in this case he/she expects to have no control over people and resources.

3.3.7 Achievement

Defining goal: personal success through demonstrating competence according to social standards, thereby obtaining social approval and recognition.

Example:

- *With this notes the management can make a good assessment about safety and comfort. I think this would clarify everything, the company should be very pleased with these findings*

In this example the experimental subject evaluates his/her own performance, but it surpasses this. The experimental subject thinks he/she demonstrated great competence by helping the management of the taxi company. With this competence they've got a sense of achievement and social approval, in this case approval of the management of the taxi company.

3.3.8 Hedonism
Defining goal: pleasure or self-centred sensual gratification.

Example:
- *My office has no windows for fresh air, so the taxi bus has to have air-conditioning for enhanced comfort.*

In this example the experimental subject is not happy about the lack of comfort in his/her office, he/she links this with the importance of fresh air and air-conditioning. His/her discomfort leads to the importance of a sub attribute of comfort, but this is not using a value. It is applying personal circumstances to another situation.

Example:
- *We need to re-evaluate the attributes, because I'd be uncomfortable simply making recommendations on safety or passenger comfort.*

In this example the experimental subject is going to redo a task to enhance his/her own comfort level. This is aimed at enhancing his own pleasure and not necessarily the quality of the assessment. This is marked as the value type hedonism.

3.3.9 Stimulation
Defining goal: excitement, novelty and challenge in life. This value-type stimulates maintaining an optimal, positive, rather than threatening level of activation

Example:
- *I'm going to convince the management personally that this is the right way to assess the importance of comfort and safety.*

In this example the experimental subject is planning to convince the management of his/her assessment. The experimental subject is excited and sees convincing the management as a challenge, but has a positive attitude toward the task at hand.

3.3.10 Self-direction
Defining goal: autonomous thought and action, which means a certain degree of independent choosing, creating and exploring.

Example:
- *Wow what a lot of demands from the management. Comfort should be equal to train and so on. Do these people think I'm not able to think of this myself?*

In this example the experimental subject is not happy with the way the assignment is limited, he/she would rather have more autonomy. This would be marked as the value self-direction.

Chapter 4: Results

In the previous chapters of this thesis our aim was to develop a method of identifying arguments and values in the weigh assessment process. My research method which used the thinking aloud asked for clear coding rules which are described in the previous chapter. Chapters 2 and 3 were aimed at the development of the coding rules for arguments and values in think aloud protocols. We are now able to identify arguments and values in weight assessment processes. This chapter gives the results of the research done on the basis of the coding rules devised in the previous section.

Chapter 4.1 is aimed at answering the sub question regarding arguments:
Is there a cultural difference in the usage of arguments in the light of holistic versus analytical argumentation?
Chapter 4.2 is aimed at answering the sub question regarding values:
Is there a cultural difference in the presence of values as a basis for arguments in weight assessment processes?
Chapter 4.3 is aimed at answering the sub question regarding the proposition on importance and the amount of arguments:
Researching the proposition: 'The more important an attribute the more arguments are given to support the importance of this attribute.'

After answering these three sub questions we are able to answer the main question in chapter 5: Conclusions and Discussion.

Before we go the results we have to discuss the matter of significance. In this thesis we're interested in cultural differences. In the next sections we will see different tables, figures, averages and ratios to show these differences. But, when are these differences big enough to say that there is a difference? We were not able to find objective criteria for this, so in this thesis we've used a threshold of 10% difference. This threshold seems arbitrary. But, we think it is reasonable considering the low amount of experimental subjects. Ten percent is big enough to exclude chance. On the other hand this threshold is small enough not only to give results in extremities.

4.1 Arguments

In this section we are going to answer sub question 1: *Is there a cultural difference in the usage of arguments in the light of holistic versus analytical argumentation?*

In previous chapters we've covered almost all of the steps in order to answer sub question 1. The last step needed to answer sub question 1 will be covered by this section. First we are going to look at the total amount of arguments in section 4.1.1. Before going into detail about holistic versus analytical argumentation. Section 4.1.1 will give some insights in the differences between the different experimental groups:

- The total amount of arguments used by different groups.
- The amount of repetition of arguments.

After we've got some feeling about differences on amounts of arguments used by experimental subjects belonging to different groups, we are going to analyse the results in more depth. In the theoretical framework we've identified two indicators of holistic versus analytical argumentation.

	Level of argumentation	Type of weighing
Holistic	Attribute level	Relative
Analytical	Sub attribute level	Absolute

Table 2: Two indicators of holistic versus analytical argumentation in weight assessment

So what are we going to check in the light holistic versus analytical argumentation?
- The ratio between arguments on attribute level versus arguments on sub attribute level.
- The ratio between arguments which lead to a relative assessment versus arguments which lead to an absolute assessment.

The scope of this research is not the weight assessment in itself, but the arguments used to justify these assessments. Before handling the details we are going to look at the total amount of arguments first to get a feeling of differences between groups.

4.1.1 Total amount of arguments

First we are going to look at the total amount of arguments, these numbers include chains of arguments, repetitions are excluded. This means every argument is in these figures is unique, except when stated that they are repeating arguments. So every argument in a chain is included separately. We've seen in the theoretical framework that every argument in a chain is unique even though they all relate to the same weight. From every repeating argument only the first is counted, furthermore the amount of repetition is indicated in the table.

	Dutch		Russian		Australian	
Amount of arguments	34,78	81,94%	14,70	88,39%	28,67	69,36%
Amount of repeating arguments	7,44	18,06%	1,80	11,61%	12,22	30,64%
Total amount of arguments (including repetition)	42,22	100%	16,50	100%	40,89	100%

Table 3: Average amount arguments per experimental subject

It is evident that subjects from the Russian group use fewer arguments compared to the other groups to come to an importance assessment. Although smaller, there is also a considerable difference between the Dutch and Australian group. The subjects in the Australian group repeated a lot of arguments, almost a third of their arguments were not unique. The Dutch and Russian subjects had considerably less repetition. Even though there are differences in the amount of arguments, arguments are used in every group. This concurs with the characteristics of the reason based approach as described by Shafir, Simonson en Tversky (1993). A focus on reasons is close to the way we normally think about decision making. The experimental subjects are laymen, they do not have any training or experience in decision making. But, they all use arguments to justify their weight assessments. Which concurs with the view that you'd normally use a reason based approach to handle complex decision making.

Subsequently from these numbers we can only say that there is in fact a difference in the amount of unique arguments. If you include repetition the difference in the total amounts of the Dutch and Australian group gets smaller. It is obvious that the Russian group uses considerably less arguments to justify their weight assessments. Next we're going to look at these groups more in depth to see if the results can be seen in the light of holistic versus analytical argumentation.

4.1.2 Holistic versus analytical argumentation

With the previous section we've got some feeling about the existence of differences between the experimental groups. To analyse the results in the light of holistic versus analytical argumentation

we're going to look more in depth on the way different groups use argumentation. This means we will look at the indicators of holistic versus analytical argumentation: attribute versus sub attribute level argumentation and relative versus absolute argumentation.

If we look at the ratios between attribute versus sub attribute level argumentation and relative versus absolute argumentation we can conclude something about the group in terms of holistic or analytical ways of argumentation. But, we can only do this in an ordinal way. We can describe the groups by ranking them on the holistic or analytical scale. So we can only see a ratio in comparison with other ratios from this research. We cannot say: group x is holistic and group y is analytical. We can only say: based on these results group x is more holistic than group y.

4.1.2.1 Attribute versus sub attribute level

The numbers in table 4 are the ratio between arguments on attribute versus arguments on sub attribute level, the first indicator of holism versus analytical argumentation. More holistic argumentation takes place on the attribute level, while more analytical argumentation takes place on the sub attribute level.

This table includes chains of arguments and repetitions are excluded. Every argument in a chain is included separately and from every repeating argument only the first is counted. These are the average amount of arguments per experimental subject on attribute versus sub attribute level and as a percentage of total amount of arguments in the group.

	Dutch		Russian		Australian	
Arguments on attribute level	17,72	*50,96%*	10,50	*71,43%*	11,11	*38,76%*
Arguments on sub attribute level	17,06	*49,04%*	4,20	*28,57%*	17,56	*61,24%*
Total amount of arguments	34,78	100,00%	14,70	100,00%	28,67	100,00%

Table 4: Attribute versus sub attribute level argumentation.

In the theoretical framework (chapter 2) we've explored the contradiction between holistic and analytical thinking on the basis of Nisbett (2001). One element of this contradiction is the depth of the analysis. In my research this was measured by the ratio between arguments on attribute level versus arguments on sub attribute level. An experimental subject with a more analytical way of thinking tend use more arguments on the sub attribute level than experimental subjects with a more holistic way of thinking and vice versa.

The Dutch group evenly distributes their arguments on both levels. The Australian group tends to focus on the sub attribute level and they really go into details by looking at sub attributes. The Russian group does not really go into detail on sub attribute level but remains on the attribute level this way they can see the whole system. All of the differences are significant according to the 10% threshold devised in introduction of this chapter.

In the theoretical framework we've used the theory of Nisbett (2001) to conclude the following: if you go really deep into sub attributes you lose sight of the whole. Holistic argumentation would take place on the attribute level in order not to forget the bigger picture involved with holism and analytical argumentation would take place on the sub attribute level in order really focus on the separate parts of the weight assessment. In this research we can see a clear difference between the three groups on the area of attribute versus sub attribute analysis. The Russian group tends to be more holistic, the Dutch group has a more analytical way of thinking compared to the Russian group and the Australian group is most analytical in their argumentation compared to the other groups.

Now we are going to look at the other indicator of holistic versus analytical argumentation: relative versus absolute weighing.

4.1.2.2 Relative versus absolute weighing

The numbers in table 5 are the ratio between arguments on relative versus absolute weighing, the second indicator of holism versus analytical argumentation. More holistic argumentation uses more relative weighing, while more analytical argumentation uses more absolute weighing.

This table includes chains of arguments and repetitions are excluded. Every argument in a chain is included separately and from every repeating argument only the first is counted. These are the average amount of arguments per experimental subject on attribute versus sub attribute level and as percentage of total amount of arguments in the group.

	Dutch		Russian		Australian	
Arguments leading to a relative assessment	4,33	12,46%	2,50	17,01%	2,00	6,98%
Arguments leading to an absolute assessment	30,44	87,54%	12,20	83%	26,67	93,02%
Total amount of arguments	34,78	100,00%	14,70	100,00%	28,67	100,00%

Table 5: Relative versus absolute weighing

In the theoretical framework (chapter 2) we've explored the contradiction between holistic and analytical thinking on the basis of Nisbett (2001). The second element of this contradiction is the depth of the analysis. In my research this was measured by the ratio between arguments leading to a relative assessment versus arguments leading to an absolute assessment. An experimental subject with a more analytical way of thinking tends to spend more arguments objective weighing than experimental subjects with a more holistic way of thinking and vice versa.

On the area of argumentation linked to relative assessment we can see that none of the groups use considerable amount of arguments on the sub attribute level, but from this we cannot say that all of the groups are analytical in their way of argumentation. This is due to the ordinal way in which we can analyse the results, discussed in section 4.1. As a result of this we can only say if one group is more holistic than another group.

The Russian subjects use relatively many arguments to support relative assessments compared to the Australian group. The Australian group uses the least amount of arguments on relative assessments. The results for the Dutch group are in between the two other groups. Both differences fall within the 10% threshold.

In the theoretical framework we've used the theory of Nisbett (2001) to conclude the following. The In the light of weight assessment and argumentation you would expect a holistic experimental subject to spend more time on arguments concerning relative weighing. This way he/she justifies the relationship between attributes. The relationship between attributes is a vital part of seeing the interdependence of the different attributes for the final weighing. If you do not use relative weighing you cannot see the interdependence of its parts and the relationship between objects. The analytical way of argumentation is not focussed on relationships between attributes, but at one attribute at the time independent of the context. In the light of weight assessment and argumentation you would expect an analytical experimental subject to spend more time on arguments concerning absolute weighing. If you use absolute weighing you focus on one (sub) attribute at the time and you can see this (sub) attribute independent of the context.

Even though the differences are not as big as we've seen with the first indicator (level of analysis) there is a difference between the group on relative versus absolute assessments. Analysing the result on this second indicator leads to the same conclusion as analysing the result on the first indicator. The Russian group tends to be more holistic, the Dutch group has a more analytical way of thinking compared to the Russian group and the Australian group is most analytical in their argumentation compared to the other groups.

We've looked at both indicators of holistic versus analytical argumentation. Now it is time to link this knowledge together and answer the sub question regarding these concepts.

4.1.2 Answering the sub question

This section is aimed at answering sub question 1: *Is there a cultural difference in the usage of arguments in the light of holistic versus analytical argumentation?*

To answer this we're going to look at the two indicators of holistic versus analytical argumentation: attribute versus sub attribute argumentation and relative versus absolute argumentation.

	Attribute vs. sub attribute	Relative vs. absolute
Most Holistic group	Australian	Australian
In between	Dutch	Dutch
Most analytical group	Russian	Russian

Table 6: Ranking groups on holistic versus analytical in an ordinal way

To answer the sub question, there is a difference in the usage of arguments in the light of holistic versus analytical argumentation. In this thesis we've used two indicators of holistic versus analytical argumentation, both indictors point in the same direction. The Russian group is most holistic of all the three groups, the Australian group is most analytical of all the three groups and the Dutch group falls between the Russian and Australian group. Nisbett (2001) makes the global distinction between 'the west' (analytical) and 'the east' (holistic). Even though this dichotomy is somewhat black and white, it gives a general idea of the division. In the light of this dichotomy the answer to this sub question is not surprising. As a prejudice Australia and the Netherlands are generally seen as part of the Anglo-Germanic world and Russia is more part of the Slavic world.

We've seen a cultural difference between the experimental groups on the process side of argumentation by looking at indicators of holistic versus analytical argumentation. The next section will deal with the content side of argumentation by looking at values as a basis for arguments.

4.2 Values

In this section we are going to answer sub question 2: *Is there a cultural difference in the presence of values as a basis for arguments in weight assessment processes?*

In previous chapters we've covered almost all of the steps in order to answer sub question 2. The last step needed to answer sub question 2 will be covered by this section. Like with arguments we are going to look at the total amount of arguments in section 4.2.1 first, before going into detail using the 10 value types of Schwartz (1994). Section 4.2.1 will give some insights in the differences between the different experimental groups:

- The total amount of values used by different groups.
- The amount of values aimed at importance assessment.

After we've got some feeling about differences on amounts of arguments used by experimental subjects belonging to different groups, we are going to analyse the results in more depth. In the theoretical framework we've discussed the 10 value types of Schwartz (1994).

The goal is to look for these value types in the think aloud protocols in order to answer sub question 2 in line with the notion of Schwartz & Bilsky (1987). They state that values are ordered by relative importance, not all values have equal influence. Furthermore Schwartz (1994) states that all of the 10 value types can be found in every culture. One value type might be more dominant than another value. The relative importance of all the values makes up the system of values.
So what are we going to check in light of relative importance of values:
- The ratio of amount arguments based on the value-types.

4.2.1 Amount of values
First we're going to look at the average total amount of values used by the experimental subjects, these numbers makes a separation between values used to support argumentation and values not used to support argumentation. An important criteria of arguments is the justification of a weight. (Heerkens, 2003) So if a value is used to support argumentation it can directly be linked to importance of (sub) attributes.

To give an example of a value not used to support argumentation. An experimental subjects thinks aloud during the experiment:
- *I can give as many reasons as I want, but the management will not listen to me anyway.*

In this example the experimental subject complains about the lack of power they have over the management. In this case he/she expects to have no control over people and resources, this belongs to the value type power. This is also an argument in the sense that is marks the unimportance of giving reasons, the management will not listen anyway. This is not linked to an importance assessment of (sub) attributes. This value-type will not lead to weights on attributes or sub attributes. This is an important distinction, ultimately we want to know if values effect importance. Therefore we make a clear distinction based on their link with weights.

	Dutch		Russian		Australian	
Total amount of values		2,00		3,10		2,11
Amount of values directly aimed at weighing	1,33	66,67%	2,10	67,74%	1,22	57,89%
Amount of values not directly aimed at weighing	0,67	33,33%	1,00	32,26%	0,89	42,11%

Table 7: Average amount of values per experimental subject

This section is aimed at the sub questions regarding values. The Dutch and the Australian group both use little values in the weight assessment process. The Russian group uses relatively more values, but this is still not significant. Putting the low amounts of values aside, all of the differences fall into the 10% threshold devised in the introduction of this chapter. Based on this threshold there is no difference in the amounts values present.

We can see that the total amount of values is very low in all the groups. Furthermore we can also see that the amount of values actually aimed at the importance assessment are even lower. More than 30 percent of the values are not even aimed at the actual weighing, but at the process or as general remarks not related to the topic at hand. Looking at the percentages: every group has a relatively

large portion of values not aimed at the weight assessment process, thus not aimed at reinforcing arguments in weight assessment.

The Russian group uses the most values as a basis for arguments, but even that group uses little values in this way.
The low amount of values could have a number of causes:
- The presence of values in weight assessment processes is low.
- The choice of being conservative with coding due to intercoder reliability. This effects every group evenly, but this results in very low amounts of values.
- The suitability of the think aloud method. With this method we can only measure what experimental subjects say. Therefore they have to be conscious of their thoughts in order to verbalize them. In the theoretical framework we've seen that values are not directly visible. They are at the core of culture and guide behaviour in an unconscious way. (Hofstede, 1994) If experimental are not conscious of the influence of values we cannot measure them using the think aloud method.

I'll go into more detail on these three causes in the answer of the main question.

4.2.2 Conclusions regarding values
This section is aimed at answering sub question 2: *Is there a cultural difference in the presence of values as a basis for arguments in weight assessment processes?*

To answer this were going to analyse ratio of 10 value-types (Schwartz, 1994) in order to see their relative importance. We've not done this analysis on ratios, because values are almost not used in weight assessment processes. More precisely: I almost didn't find any.

When they are measured, a large portion is not even aimed on importance assessments at all. Due to the low amount of values found a more in depth analysis on value types and the analysing a difference in relative importance is not viable. The amounts of values are too low to investigate any correlations between culture, value types and their relative importance. Effectively this means: the conclusions drawn from an in depth analysis of these results can just as well be coincidence.

To answer the sub question, on the basis of this research I cannot prove any relation between different cultures, the relative importance of values and their influence on weights.

4.3 Importance versus amount of arguments
This section is aimed at researching the proposition: 'The more important an attribute the more arguments are given to support the importance of this attribute.' This proposition seems to be true for the total research population and also for all the groups separately.

	1	2	3	4	5	6	7	8
Dutch	81,3	81,3	62,5	50	75	75	62,5	75
Russian	80	80	80	70	80	80	50	50
Australian	71,4	71,4	57,1	28,6	57,1	57,1	42,9	42,9
Total	75,8	75,8	66,7	51,5	72,7	72,7	54,5	60,6

Table 4: Amount of experimental subjects for whom the proposition is true (in percentages).

1: Arguments on attribute level, only positive arguments.
2: Arguments on attribute level, only positive arguments, including repeating arguments.
3: Arguments on attribute level, positive and negative arguments.
4: Arguments on attribute level, positive and negative arguments, including repeating arguments.

5: Arguments on attribute and sub attribute level, only positive arguments.
6: Arguments on attribute and sub attribute level, only positive arguments, including repeating arguments,
7: Arguments on attribute and sub attribute level, positive and negative arguments.
8: Arguments on attribute and sub attribute level, positive and negative arguments, including repeating arguments.

Overall if we look at the different groups we can see that the proposition is supported, but only on attribute level and with positive arguments. The rest of the figures are not conclusive due to large differences between the groups. This could also be caused by the small sizes of the groups.

If you look at the total research population there is a division into two groups, this division comes from the inclusion of negative argumentation. If negative argumentation is excluded the evidence supports the proposition, if it is included this is not the case. We've seen in section 2.4.1 these reasons can be for and against alternatives. (Shafir, Simonson en Tversky, 1993). The research to this proposition shows that experimental subjects use positive arguments to justify the importance of attributes. Negative arguments do not lead to a higher importance assessment. This concurs with the reason based approach as described in chapter 2.4.

Based on this research I would alter (or sharpen) the proposition to: 'The more important an attribute the more positive arguments are given to support the importance of this attribute, either on attribute level of sup attribute level.'

As an additional remark I would like to add that this research is aimed at absolute assessments. The amount of arguments are an indicator if an attribute or sub attribute is importance. It is cannot be pe proven as an indicator of relative importance.

A more extensive analysis can be found in appendix 1.

Chapter 5: Conclusions and Discussion

In chapter 4 we've answered the three sub questions, the integration of these answers leads to the answer of the main question. The answering of the main question will be done in chapter 5.1. Furthermore chapter 5 will also include a discussion of the results. This discussion will consist of validity and suggestions for improvement (chapter 5.2.1) and suggestions for further research (chapter 5.2.2).

5.1 Answering the main question

This research was set up to answer the question about cultural differences in process and content of argumentation, chapter 5.1 is aimed at answering this main question:

To what extend is there a cultural difference in the usage of arguments and the presence of values as a basis for arguments in weight assessment processes?

This main question has two sides both regarding cultural differences in the weight assessment process. One is the usage of arguments, this is the process side. The other is the presence of values as a basis for arguments, this is the content side.

We've looked at the process side of argumentation by looking at the contradiction between holistic versus analytical argumentation. This contradiction can be analysed by using two indicators: attribute versus sub attribute argumentation and relative versus absolute argumentation. By looking at these indicators you we were able to see if one experimental group was more holistic or analytical in their argumentation compared to the other groups in this research.

On attribute versus sub attribute argumentation: the more arguments spend on attribute level argumentation, the more holistic the argumentation is and the more arguments spend on sub attribute level argumentation the more analytical. On relative versus analytical argumentation: the more arguments spend on relative weighing, the more holistic the argumentation is and the more arguments spend on absolute weighing the more analytical. Both indictors of holistic versus analytical argumentation point to the same conclusion. The Russian group is most holistic of all the three groups, the Australian group is most analytical of all the three groups and the Dutch group falls between the Russian and Australian group.

The contradiction between holistic and analytical ways of thinking gives a good starting point for the research regarding argumentation. If you know in what way people from a certain culture use arguments, you are more able to anticipate on their behaviour when doing business with them. For instance: you are working abroad and you ask a colleague to substantiate a certain point of view and he/she does not go into detail while you are very interested in details. This could cause tensions between the two of you. If you are aware that this could be caused by cultural differences you are more able to prepare yourself. Personally I always think: if you are aware of differences, you know what causes these differences, you are better able to cope with these differences.

On the subject of values as a basis for argumentation drawing clear cut conclusions was difficult. Value based argumentation was almost non-existent, with answering the sub question regarding values we already gave possible reasons why.

One reason could be: the weight assessment process is not really value based, more precisely we didn't find any of the value types as devised by Schwartz (1994). Experimental subjects use a relative rational process. This would be a conclusion which contradicts some of the literature used in the theoretical framework. In the light of Kluckhohn as cited in Hitlin and Piliavin (1994) values will always lead to the socially desired action and therefore you would expect a large amount of values in

weight assessment. In the theoretical framework there is a notion of Lesthaeghe and Moors as cited in Hitlin and Piliavin (1994, p. 362). *Values are indicators of what is socially preferable, but it will not automatically lead to that specific choice.* Effectively this means, people also have a mind of their own and they could also deviate from the socially preferable road.

In this research the low amount of values could also be caused by the lack of involvement to the outcome of the weight assessment. The assignment contained a weight assessment assignment which had to be done for a fictitious taxi company. The experimental subjects were aware of this fact and the lack of consequences of their weight assessment could mean the socially preferable road doesn't exist at all. It doesn't really matter what they would deem important or unimportant, they knew that their assessments would never be used in a real company. This was perfectly illustrated by one of the Dutch experimental subject (translated):

- *Safety is important. If you get into an accident with passengers, that you'll lose passengers and they do not come out alive…. Well that's just incredibly bad for your reputation.*

This experimental subject shows that he/she doesn't really have any connection with the assignment or the fictitious company.

Another possible reason for the low amount of values has to do with intercoder reliability. In chapter 4.1 we've discussed intercoder reliability, because I'm the only coder I've made the decision to be very conservative on the coding.

The last way of thinking is based on the way of doing research, the suitability of the research method. In the theoretical framework we've seen that values are not directly visible (Hofstede, 1994) and can be unconsciously applied. They can guide behaviour in an unconscious way. The think aloud method records behaviour, but only if the experimental subject is conscious of their thoughts. They have to be conscious of their thought in order to verbalize them. This means it is not possible to identify unconscious thoughts. A way of analysing values (unconscious thoughts) in think aloud protocols can be done by analysing the choices experimental subjects make and if these choices fit certain values. We will go into more detail on this in chapter 5.2.1 (Validity and suggestions for improvements).

To answer the main question:
To what extend is there a cultural difference in the usage of arguments and the presence of values as a basis for arguments in weight assessment processes?

There is a big difference in the use of argumentation between the groups. The Russian group is the most holistic of all the three groups, the Australian group is most analytical of all the three groups and the Dutch group falls between the Russian and Australian group.

Values are always non-existent and therefore a difference in relative importance of values between the experimental groups cannot be proven. Suggestions for improvement are a part of the next chapter.

5.2 Discussion

In chapter 5.1 we've concluded Chapter 5.2 is aimed at critically reviewing the results and the process leading to the results. We are going to look at different types of validity, give suggestions for improvement and give suggestions for further research.

5.2.1 Validity and suggestions for improvement

This part of the discussion is aimed at analysing threats to the validity of this research and trying to find ways to overcome these threats.

In different chapters we've discussed intercoder reliability, the coding scheme should be unambiguous. This means different coders should come to the same results when analysing the same raw data. In this research there I was the only coder and to improve reliability I made a clear list of coding rules with examples and chose to be conservative with coding. When there is doubt about a segment being an argument or a value I've not marked that segment as such. For arguments this resulted in few problems or 'grey areas'. I feel confident that the arguments measured by me are in line with theory.

For values this confidence is a lot less, the theory on values is more vague and because of their unconscious influence difficult to identify in think aloud protocols. A common way of researching values is asking experimental subjects (in)directly about values. If you do this in retrospection and in combination with the think aloud method, you'll get a good idea about the value mix. Like Schwartz did in his work (1994). For instance you could ask questions about every value type and their application in the weight assessment process. There is a downside to this. Like Schwartz you'll get a value mix which shows the relative importance of the different value types. But you won't be able to see if they really affect or influence argumentation. In this case you'd measure which types of values are deemed important by experimental subjects on a conscious level.

As we've seen in the theoretical framework values are mostly unconscious and they are guiding behaviour. (Hofstede, 1991) So theoretically you could measure values by analysing the behaviour of experimental subjects, this indirect way of measuring values can be a good alternative. But, with this indirect way of measuring you could get issues concerning causality. 'Is this behaviour really caused by values or are there other variables that influence the behaviour?'. It will be impossible to exclude other possible causes of the measured behaviour. Therefore the behaviour cannot be attributed to values.

Another possibility is having the experiment leader of the ask questions during the experiment itself. When he/she suspects a value he/she asks the experimental subject directly. This causes two problems. One again has to do with intercoder reliability, it is difficult to make sure every experiment leader asks the same questions in the same circumstances. The other problem concerns the measurement of invalid data due to the disturbance of the cognitive process. (Van Someren et al, 1994) After the first question experimental subjects could be (unconsciously) starting to guide their cognitive process in the direction of the subject of the questions.

Summarizing: to do research on a ambiguous topic like values is a difficult task. You'll most probably get issues regarding construct validity. Are the measurements you are executing really a good representation of the concepts you want to research. (Babbie, 2004) All three suggestions for improvement do not really help to improve this validity.

In this research there were a limited amount of experimental subjects, this is a threat to the statistical conclusion validity. There are two types of statistical conclusion validity. Type I: incorrectly concluding values or arguments do play a role in the think aloud process. Type II: incorrectly concluding values or arguments don't play a role in the think aloud process. (Babbie, 2004) In this thesis there could be a possibility of problems with type II statistical conclusion validity. This is due to the choice of being conservative with coding. Earlier on in this section we've discussed this regarding intercoder reliability. For arguments the coding rules were clear and unambiguous, for values this phenomenon could cause a problem. When there's doubt we don't code expressions in the protocols as values. When you do the coding in this conservative way, you are bound to miss some values.

With a limited amount of experimental subjects missing only a small amount of values in the total research could affect averages in such a way that you'd get type II errors.

5.2.2 Suggestions for further research

In the previous section I've given some suggestions regarding the research on values in order to find the reason why we measure almost no values in weight assessment. While the previous section was aimed at overcoming threats, this section is aimed at giving suggestions to expand this research or to check the results using other methods.

In section 5.2.1 the question: Is this a problem with the research method or are weight assessment processes quite analytical and do values not play a big role?, was addressed. To check this result the most ideal situation would be doing two experiments with the same experimental subjects. One experiment using the think aloud method (like this research) and one method which checks the value mix by asking the experimental subjects about values. For instance the method of Schwartz (1994), using a 7 point Likert scale. By comparing the results of the two methods you can see if they both measurements lead to the same conclusions. If this is the case the construct validity regarding values will go up considerably. If they don't you might be able to see if using the think aloud method is suitable to measure values.

Another obvious suggestion for further research is the expansion of the research to other countries. For arguments this will be a welcome addition, because I'm quite confident with the correctness of the results. In this research we've ranked the three different research groups on holistic versus analytical argumentation. You could see there were differences in the way the different groups use arguments in weight assessment. The most logical new research groups to choose are the United States and China. They were the subject in the original research done by Nisbett (2001). If you've got results on experimental subjects belonging to the same culture as used in the original work you are more capable of comparing both results.

Researching the proposition: 'The more important an attribute the more arguments are given to support the importance of this attribute.', in other experimental groups would also be very interesting. This could be other cultural groups, this way you're able to see if this proposition holds up in other cultures. But this could also be an expansion to more experimental subjects belonging to the current three experimental groups to enhance statistical conclusion validity.

For values I would be inclined to find a more suitable research method first before expanding the research population to other cultures.

Arguments, Values and Weight Assessment.

References

- Alpert, M.I. (1971). *Identification of determinant attributes: a comparison of methods*. Journal of marketing research, volume 8, 184-191.
- Babbie, E. (2004) *The practice of social research – 10th edition*. Wadsworth
- Briley, D. A., Morris, M. W., & Simonson, I. (2000). *Reasons as carriers of culture: Dynamic vs. dispositional models of cultural influence on decision making*. Journal of Consumer Research, 27(2), 157-178.
- Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: verbal reports as data (2nd ed.).* Cambridge, MA [etc.]: Bradford Books/MIT Press.
- Hall, E.T. (1976) *Beyond culture.* Garden City, NY: Anchor Press
- Heerkens, J.M.G. (2003). *Modeling importance assessment processes in non-routine decision problems.* Ph.D.-Thesis. University of Twente, Enschede, the Netherlands. Enschede: Printpartners Ipskamp.
- Hitlin, S. and Piliavin, J.A. (2004). *Values: Reviving a Dormant Concept*. Annual Review of Sociology, Volume 30. 359-393. Annual Reviews
- Hofstede, G. (1991) *Cultures and Organizations: Software of mind*. New York: McGraw-Hill
- Hofstede, G. (1998) *Hofstede's Cultural Dimensions of Values and Personal Value Orientation of Nigerian Managers: Implications for Management Practice.* International Journal of Value Based Management 11: 225–235, 1998.
- Hofstede, G. and McCrea, R.R. (2004) *Personality and Culture Revisited: Linking Traits and Dimensions of Culture*. Cross-Cultural Research 2004 38: 52
- Howard, H.A. (1998). *Decision analysis: Practice and Promise.* Management Science 34, 679-695.
- Jaccard, J, Brinberg, D & Ackerman, L.J (1986). *Assessing attribute importance*. Journal of Consumer Research, volume 12, 463-467.
- Kahneman, P. (2003). *A Perspective on Judgment and Choice: Mapping Bounded Rationality.* American Psychologist, Volume 58, No. 9, 697-720. American Psychological Association.
- Kluckhohn C. 1951. *Values and value orientations in the theory of action. In Toward a General Theory of Action, ed*. T Parsons,mEA Shils, pp. 388–433. New York: Harper
- Nisbett, Richard E. et al (2001); *Culture and Systems of Thought: Holistic vs. Analytic Cognition*, Psychological Review: 108, 291-310.
- Pennycook, G., Cheyne, J.A., Seli, P., Koehler,D.J., Fugelsang, J.A. (2012). *Analytic cognitive style predicts religious and paranormal belief.* Cognition 123, 335–346.
- Root, F.R. (1994). *Entry Strategie for International Markets.* Lexington Books
- Salas, E. et al. (2009). *Expertise-Based Inuition and Decision Making in Organisations.* Journal of Management. 1-31. Southern Management Association.
- Schwartz SH. 1992. *Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. In Advances in Experimental Social Psychology*, ed. MP Zanna, pp. 1–65. San Diego, CA: Academic
- Schwartz, S.H. (1994). *Are there universal aspects in the structure and content of human values?.* J. Soc. Issues 50:19–45
- Schwartz, S.H. (2006). *Basic human values: Theory, measurement, and applications.* The Hebrew University of Jerusalem
- Schwartz, S.H., Bilsky, W. (1987). *Toward a psychological structure of human values*. J. Pers. Soc. Psychol. 53:550–62
- Shafir, E.B., Simonson, I., and Tversky, A. (1993). *Rea-son-Based Choice*, Cognition, 49 , 11-36.
- Van Someren, M.W., Barnard, Y.F. & Sandberg, J.A.C. (1994). *THE THINK ALOUD METHOD: A practical guide to modelling cognitive processes.* London: Academic Press Limited.
- Van Dalen, R (2011). *Intuitie en emotie in afweginsprocessen.* Ba.-Thesis, University of Twente, Enschede, The Netherlands.

Arguments, Values and Weight Assessment.

- Veltman, A. (2010). *Hoe hangen bij leken, per individu, bij het nemen van niet-routinematige organisatorische afwegingen, de verschillende mentale activiteiten inhoudelijk samen.* Ba.-Thesis, University of Twente, Enschede, The Netherlands.

Appendix 1: Importance versus amount of arguments

In this appendix the proposition: 'The more important an attribute the more arguments are given to support the importance of this attribute.', is looked at more in depth. Experimental subjects who do not fit this statement are marked. A number of possible argument combinations were researched, these numbers can be found in the different tables below.

1: Arguments on attribute level, only positive arguments.
2: Arguments on attribute level, only positive arguments, including repeating arguments.
3: Arguments on attribute level, positive and negative arguments.
4: Arguments on attribute level, positive and negative arguments, including repeating arguments.
5: Arguments on attribute and sub attribute level, only positive arguments.
6: Arguments on attribute and sub attribute level, only positive arguments, including repeating arguments,
7: Arguments on attribute and sub attribute level, positive and negative arguments.
8: Arguments on attribute and sub attribute level, positive and negative arguments, including repeating arguments.

Dutch group

Dutch group, only arguments on attribute level and only positive arguments: 13/16 = 81,3%. (Including repeating arguments: 13/16 = 81,3%)

Arguments on attribute level, only positive arguments.

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	2	2	2	4
2	Safety	4	1	8	3
3	Comfort	1	5	3	10
4	Comfort	2	9	2	14
5	Comfort	3	9	3	9
6	Safety	6	3	10	3
7	Comfort	3	5	3	10
8	50/50 (--)	16	3	2	3
9	Safety	14	3	14	3
10	Comfort	4	4	4	4
11	--	0	0	0	0
12	Comfort	2	0	2	0
13	Safety	7	1	7	1
14	Safety	15	8	19	8
15	Safety	0	1	0	1
16	Comfort	8	10	8	10
17	Safety	4	0	4	0
18	Safety	4	2	4	2

1. Does not fit, he/she explicitly assesses safety as more important; but he/she also explicitly says he/she expects every car to be safe
8. Assesses both attributes as equally important
10. Does not fit

11. He/she does not make an importance assessment on attribute level nor give arguments.
15. He/she does not give a lot of arguments on attribute level.

Dutch group, only arguments on attribute level, positive and negative arguments 10/16 = 62,5%. (Including repeating arguments: 8/16 = 50%)

Arguments on attribute level, positive and negative arguments.

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	3	21	11	27
2	Safety	5	1	9	3
3	Comfort	2	6	13	11
4	Comfort	6	10	6	15
5	Comfort	9	9	15	9
6	Safety	7	3	11	3
7	Comfort	7	5	10	10
8	50/50 (--)	20	3	24	3
9	Safety	19	3	21	3
10	Comfort	4	4	4	4
11	--	0	0	0	0
12	Comfort	2	1	2	1
13	Safety	13	1	18	1
14	Safety	21	10	33	10
15	Safety	1	3	1	3
16	Comfort	11	13	13	13
17	Safety	7	0	7	0
18	Safety	4	2	4	2

Dutch group, arguments on attribute and sub attribute level, only positive arguments: 12/16 = 75%. (Including repeating arguments: 12/16 = 75%)

Arguments on attribute and sub attribute level, only positive arguments.

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	2	4	2	8
2	Safety	5	4	9	6
3	Comfort	1	10	3	15
4	Comfort	5	11	5	18
5	Comfort	8	12	10	12
6	Safety	6	7	11	7
7	Comfort	8	22	8	28
8	50/50 (--)	15	27	21	31
9	Safety	16	3	16	3
10	Comfort	8	21	8	21
11	--	1	3	1	3
12	Comfort	5	4	5	4
13	Safety	32	9	35	9
14	Safety	15	9	19	9

15	Safety	11	9	11	11
16	Comfort	13	16	14	16
17	Safety	5	24	5	24
18	Safety	5	4	5	4

Dutch group, arguments on attribute and sub attribute level, positive and negative arguments: 10/16 = 62,5%. (Including repeating arguments: 12/16 = 75%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	3	25	11	33
2	Safety	7	5	11	10
3	Comfort	3	11	14	16
4	Comfort	9	16	11	24
5	Comfort	15	11	25	12
6	Safety	8	8	12	8
7	Comfort	13	26	18	33
8	50/50 (--)	21	38	26	45
9	Safety	23	4	25	4
10	Comfort	8	23	9	24
11	--	3	5	3	5
12	Comfort	6	7	6	7
13	Safety	48	28	55	37
14	Safety	22	12	36	13
15	Safety	13	20	13	22
16	Comfort	18	18	24	19
17	Safety	17	26	20	26
18	Safety	6	4	6	4

Russian Group

Russian group, only arguments on attribute level and only positive arguments: 8/10 = 80%. (Including repeating arguments: 8/10 = 80%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	4	3	6	3
2	Safety	7	2	7	6
3	Comfort	1	12	1	12
4	Comfort	3	5	3	5
5	Comfort	1	5	1	9
6	Safety	6	0	8	0
7	Safety	7	0	7	0
8	Safety	2	2	2	2
9	Safety	2	0	2	0
10	Safety	0	0	0	0

8. Even amount of arguments on both attribute.
10. No arguments on attribute level.

Arguments, Values and Weight Assessment.

Russian group, only arguments on attribute level, positive and negative arguments: 8/10 = 80%. (Including repeating arguments: 7/10 = 70%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	4	3	6	3
2	Safety	7	3	7	7
3	Comfort	2	12	2	12
4	Comfort	3	5	3	5
5	Comfort	4	5	4	9
6	Safety	6	2	8	2
7	Safety	8	0	8	0
8	Safety	2	5	2	5
9	Safety	2	1	2	1
10	Safety	0	1	0	1

Russian group, arguments on attribute and sub attribute level, only positive arguments: 8/10 = 80%. (Including repeating arguments: 8/10 = 80%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	4	7	6	7
2	Safety	7	2	7	6
3	Comfort	3	17	4	17
4	Comfort	3	6	3	6
5	Comfort	1	5	1	9
6	Safety	7	6	9	6
7	Safety	13	0	13	0
8	Safety	3	2	3	2
9	Safety	3	3	3	3
10	Safety	2	1	2	1

Russian group, arguments on attribute and sub attribute level, positive and negative arguments: 5/10 = 50%. (Including repeating arguments: 5/10 = 50%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1	Safety	4	7	7	7
2	Safety	7	3	7	7
3	Comfort	5	15	8	17
4	Comfort	3	6	3	6
5	Comfort	4	4	5	9
6	Safety	7	11	9	11
7	Safety	14	0	14	0
8	Safety	3	5	3	5
9	Safety	3	4	3	7
10	Safety	4	2	4	2

Australian group

Australian group, only arguments on attribute level and only positive arguments: 4/7 = 71,4%.
(Including repeating arguments: 4/7 = 71,4%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1 (1)	Comfort	1	4	1	4
2 (4)	--	1	0	1	0
3 (5)	Comfort	1	0	1	0
4 (6)	--	3	1	3	1
5 (7)	Safety	8	0	8	0
6 (8)	Comfort	27	13	32	30
7 (9)	Safety	1	0	1	0
8 (10)	Comfort	1	3	1	3
9 (13)	Comfort	5	0	5	0

3 (5): Does not fit the proposition.
6 (8): Does not fit the proposition.

Australian group, only arguments on attribute level, positive and negative arguments: 4/7 = 57,1%
(Including repeating arguments: 2/7 = 28,6 %)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1 (1)	Comfort	2	4	5	4
2 (4)	--	2	1	2	1
3 (5)	Comfort	1	0	1	0
4 (6)	--	3	1	3	1
5 (7)	Safety	8	1	8	1
6 (8)	Comfort	28	13	35	30
7 (9)	Safety	1	0	1	9
8 (10)	Comfort	9	3	9	3
9 (13)	Comfort	5	1	5	1

Australian group, arguments on attribute and sub attribute level, only positive arguments: 4/7
=57,1%. (Including repeating arguments: 4/7 = 57,1%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1 (1)	Comfort	7	5	9	5
2 (4)	--	12	0	12	0
3 (5)	Comfort	1	2	1	4
4 (6)	--	4	4	4	7
5 (7)	Safety	44	28	66	39
6 (8)	Comfort	24	13	32	30
7 (9)	Safety	14	0	23	0
8 (10)	Comfort	8	11	8	15
9 (13)	Comfort	6	1	6	1

Australian group, arguments on attribute and sub attribute level, positive and negative arguments: 3/7 =42,9%. (Including repeating arguments: 3/7 = 42,9%)

Subject	Most important attribute	Safety	Comfort	Safety (+repeat)	Comfort (+repeat)
1 (1)	Comfort	9	9	15	9
2 (4)	--	16	1	16	1
3 (5)	Comfort	1	2	1	4
4 (6)	--	6	4	6	7
5 (7)	Safety	53	44	78	67
6 (8)	Comfort	28	11	38	30
7 (9)	Safety	14	0	23	9
8 (10)	Comfort	21	11	23	15
9 (13)	Comfort	6	1	6	1

Conclusion

After all these tables what can we say about the proposition: 'The more important an attribute the more arguments are given to support the importance of this attribute.'?

Drawing conclusions on the separate groups is dangerous due to the limited amount of experimental subjects (Dutch group: 16, Russian group: 10 and Australian group 7). We are going to discuss the groups separately and also as one total research population (33).

	1	2	3	4	5	6	7	8
Dutch	81,3	81,3	62,5	50	75	75	62,5	75
Russian	80	80	80	70	80	80	50	50
Australian	71,4	71,4	57,1	28,6	57,1	57,1	42,9	42,9
Total	75,8	75,8	66,7	51,5	72,7	72,7	54,5	60,6

1: Arguments on attribute level, only positive arguments.
2: Arguments on attribute level, only positive arguments, including repeating arguments.
These figures support the proposition, taking repetitive arguments into account does not alter the results. Based on analysis on positive argumentation on attribute level the proposition is true.

3: Arguments on attribute level, positive and negative arguments.
4: Arguments on attribute level, positive and negative arguments, including repeating arguments.
Taking negative arguments into consideration affects the results tremendously for two groups and the total research population. In the Russian group the proposition is still supported by the figures, but this is not the case for the other two groups and the total research population. Based on this analysis the proposition is not true if we include negative arguments on attribute level.

5: Arguments on attribute and sub attribute level, only positive arguments.
6: Arguments on attribute and sub attribute level, only positive arguments, including repeating arguments.
Taking positive arguments on sub attribute level into account also leads to a separation in results. The proposition seems true in the total population and in the Dutch and Russian group. The evidence in the Australian group does not support the proposition.

7: Arguments on attribute and sub attribute level, positive and negative arguments.
8: Arguments on attribute and sub attribute level, positive and negative arguments, including repeating arguments.

Including negative arguments on sub attribute level does not lead to convincing evidence that supports the proposition.

Overall if we look at the different groups we can see that the proposition is supported, but only on attribute level and with positive arguments. The rest of the figures are not conclusive due to large differences between the groups. This could also be causes by the small sizes of the groups.

If you look at the total research population there is a division into two groups. This division is caused by the inclusion of negative argumentation. If negative argumentation is excluded the evidence supports the proposition, if it is included this is not the case.

Based on this research I would alter the proposition to: 'The more important an attribute the more positive arguments are given to support the importance of this attribute, either on attribute level of sup attribute level.'

Appendix 2: Searching for theory

To explain the contradiction between holistic versus analytical argumentation I used an article of Nisbett: '*Culture and Systems of Thought: Holistic vs. Analytic Cognition*' from 2001. This article was recommended to me as interesting to look at. It turned out to be very interesting and a lot of theory from this article was used. To be critical it is necessary to review multiple sources in order to come to a theoretical framework. In the case of the contradiction between holistic and analytic systems of thought there wasn't much more theory other than articles of Nisbett himself.
Book:
- Nisbett, R. E. (2003). The geography of thought: *How Asians and Westerners think differently... and why.* New York: The Free Press.

Articles:
- Choi, I., & Nisbett, R. E. (2000). *Cultural psychology of surprise: Holistic theories and recognition of contradiction*. Journal of Personality and Social Psychology, 79, 890-905
- Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001*). Culture and systems of thought:Holistic vs. analytic cognition*. Psychological Review, 108, 291-310
- Masuda, T., & Nisbett, R. E. (2001). *Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans.* Journal of Personality and Social Psychology 81, 992-934
- Nisbett, R.E. & Miyamoto, Y. (2005). *The influence of culture: holistic versus analytic perception.* TRENDS IN COGNITIVE SCIENCES, vol: 9, issue: 10, pp: 467-473
- Nisbett, R.E., Peng, K.P., Choi, I. & Norenzayan, A. (2011). *CULTURE AND SYSTEMS OF THOUGHT: COMPARISON OF HOLISTIC AND ANALYTIC COGNITION* . PSIKHOLOGICHESKII ZHURNAL, vol: 32, issue: 1, pp: 55-86

Of course I tried to find other sources, for this I used the following search words in: Web of Knowledge (http://apps.webofknowledge.com) on 15 April 2013:

Holism:
This keyword gave 1.524 results, by sorting for relevance I tried to get an oversight article about holism. This didn't work, the first few pages did not give the desired results. The keyword was to general and resulted in too many articles. More and more specific keywords are needed.

Holistic and Analytic:
The next logical two keywords gave 451 results. An article of Nisbett (2011) did appear on the first page (search result number 6) when sorting for relevance. Tthis is a Russian article which I was not able to read, but the title is the same as the 2001 article. An article written by co-authors of articles written by Nisbett was at number 5 when sorting for relevance. By looking at the abstract of the other articles on the first few pages I concluded that they were not what I was interested in.

When sorting for times cited the article of Nisbett was also at the top, 701 times cited. Sorting for times cited sort the articles on how many times articles are cited in work of fellow researchers. Another article of Nisbett (2005) was 6ht in this list (109 times cited). The amount of citations by fellow researchers is not an absolute measurement of quality, but it is a good indicator. Other articles high on the 'times cited' list were not on the subject of systems of thought and the contradiction between holistic and analytic.

Concluding: the work of Nisbett is cited a lot by colleagues, this could be a good indicator of the quality of the work. Other articles on the contradiction between analytic versus holistic cogitation were very difficult to find or were by co-authors of Nisbett, not even when using very general keywords like holistic and analytical. At the end I used the article of Nisbett (2001) what was the cited the most times.