INCREASING CONSISTENCY

EVALUATION OF THE RELIABILITY MEASUREMENTS

EMPLOYED IN THE BRANDING CONSTELLATION THEME

STUDENT	C.C.A.J.M. VAN REIJ
STUDENTNUMBER	840.467.163
UNIVERSITY	OPEN UNIVERSITEIT OF THE NETHERLANDS
COURSE	THESIS B9131B
STUDY	ACCOUNTING & FINANCE
THESIS THEME	BRANDING CONSTELLATION
EXAMINER	DRS. W.E. JURG
SECOND REVIEWER	DR. J.M.C. SCHIJNS
DATE	31 AUGUSTUS 2010

SUMMARY

Branding Constellations

Branding constellations is a problem identification technique studied within one of the master thesis themes of the Department of Management Science of the Open University in The Netherlands. The innovative assumption of these branding constellations is that the elements of a branding system set up by branders in a personified representation are capable of perceiving and expressing the unconscious relationships between these elements in the branders' minds. This helps them in identifying their branding problems so they can develop more insights in feedback loops between these elements. The personified representations - who are ignorant about branding problems - are called *stand-ins*.

Case Studies

The branding constellation theme distinguishes three kinds of comparative case studies: first-person, second-person, and third-person based on research of Bradbury and Bergmann Lichtenstein (2000). A *first-person case study* examines the usefulness (relevance, validity, reliability, and/or precision) of a branding constellation on a branding problem that is the student's responsibility as a brander. The *second-person case study* focuses on the usefulness of a branding constellation on a branding problem of a brander who is an acquaintance or a colleague of the student. A *third-person case study* generally attempts to falsify a hypothesis regarding branding constellations.

Research Problem

A significant problem within the branding constellation theme is the inconsistency that exists in the applied reliability measurements. This hinders profound comparison between the findings of different case studies focused on the above mentioned theme, according to Karel (2009). Therefore, this study is aimed at evaluating the consistency of the reliability measurements used in several of these studies. In addition, this study develops two standard excel sheets to be able to measure the reliability of branding constellations consistently in the future. Thus, the *research objective* of this thesis is:
The evaluation of the reliability measurements used for branding constellations within the branding constellations theme, including the presentation of two reliability measurements to be used within the theme in the future. This leads to the following five research questions:

1. Which reliability measurements haven been employed in scientific studies on projections?
2. Which reliability measurements can be derived from the reliability definitions in the definition list used by the branding constellation theme?
3. What are the most important reliability measurements that were used by the students in the branding constellation theme?
4. How can these most important measurements be standardised?
5. What are the similarities and differences between the findings on the most relevant reliability measurements of the previous students and the standardised measurements?

Reliability Measurements in Projection Literature
The reliability measurements of projections were found in Schultheiss (2008), Donoghue (2000), Lilienfeld et al. (2000), Perry (1998), Lundy (1985), Wagner (1985), and Jensen (1959). The authors performed different test-retest studies with each one population with a time interval. The test results varied, depending on the time interval of the retest. The overall conclusion was that the longer the time interval was, the lower the reliability score that was measured. However, the reliability measurements were employed without indicating what the 'r' stood for and how it was calculated.

Reliability Measurement Categories Derived from Definition List
Within the branding constellation theme a list of definitions is being used. The list contains numerous different definitions of the concept *reliability*. These definitions were classified in five different interpretation clusters based on the interpretations of these definitions: (1) different scorers come to similar conclusions; (2) similar conclusions are presented at different times; (3) different measurements result in similar conclusions; (4) the actual truth of the conclusions; and (5) the precision of the conclusions that are drawn.

Most Important Reliability Measurements
Karel (2009) distinguishes three categories of reliability measurements (1) test-retest reliability measurements, (2) triangular reliability measurements, and (3) precision measurements. The *test-retest reliability measurements* compare the findings of two branding constellations. The most important measurement in the test-retest category of reliability measurements is the *bilateral test-retest reliability.* This reliability is measured as the consistency in the relationships between the stand-ins of the elements in two branding constellations on the same branding problem. The *triangular reliability measurements* compare the findings of branding constellations with traditional problem identification techniques such as a literature study, historical analysis, and product/market analysis. *Precision measurements* assess the degree to which statements of the stand-ins have the possibility to be tested empirically. This thesis focuses on both bilateral test-retest measurements and the precision measurements. The category of triangular reliability measurements was not taken into account due to the fact that triangular measurements are only suitable for first-person case studies according to Karel (2009), while this thesis is a third-person case study.

Standardized Reliability Measurements
The measurements of the bilateral test-retest reliability and the precision of all theme students were analysed on the exactness of their excel-calculations. The most exact excel-calculation on the bilateral test-retest reliability was conducted by Halters (2009) and the most exact excel-calculation on precision was conducted by Jasper (2009). These 'best in class' excel calculations were taken as a starting point for the standardised excel measurements. The reliability measurements were subsequently standardised and their calculations were automated in excel as much as possible.

Similarities and Differences between Findings of Previous Students and Standardised Reliability Measurements
The theme students designed similar excel sheets for the measurement of the bilateral test-retest, but the measurements were very different. The most

significant difference was found between the measurements used in the earlier theses and the later theses. The earlier theses focused on interpreting the relationship of a stand-in with the other stand-ins directly. The later theses took all statements of the stand-ins into account in interpreting the relationship between the stand-ins. In the precision measurement the design and the approach of the precision measurements by the theme student was much more similar. The difference between the precision measurements was in the, the calculations, and the scored statements. The earlier theses focused on measuring the statements of the stand-ins of the personal elements, such as the director. Most of the later theses also included the statements of the stand-ins of the abstract elements, such as the brand. The scores on bilateral test-retest reliability of the branding constellations are higher based on the standardised version. While the theme students concluded that branding constellations are moderately bilateral test-retest reliable (+0.6 on a -2:+2-scale), the standardised version concludes that they are very bilateral test-retest reliable (+1.6 on a -2:+2-scale). The scores on precision are more similar: the average of the precision measurement of the theme students was (+0.52 on a -2:+2-scale)and the average of the standardised precision measurement was (+0.57 on a -2:+2-scale).

Implications

It is expected that the standardised reliability measurements will improve the consistency of the future bilateral test-retest measurements and the precision measurements within the branding constellation theme. The standardised measurement simplifies these measurements for future theme students, while leaving little room for researchers' errors. The great number of researcher errors shows how necessary such a standardisation is. Thus, the other future theme measurements should also include further measurement standardisations.

PREFACE

The title *Increasing Consistency* indicates the objective of this thesis: increasing consistency of the reliability measurements used within research on the topic of branding constellation. The beads in the picture on the cover are the symbols for the consistency this thesis aims for. If you change the colour the shape will be the same, even if there were thousands of them.

My name is Celine van Reij. I was born in Weert in 1980. In 2000, I moved from Roermond to Amsterdam to study *Bedrijfskunde voor de Financiële Sector* ('Finance and Business') at the Amsterdamse Academie, and later at the Vrije Universiteit. Due to the introduction of the Bachelor Master System in the Netherlands, the study was not continued. The Open Universiteit of the Netherlands offered me the opportunity to finish my study. I decided to grasp this opportunity and I enrolled in the Master of Science, Accounting and Finance, of the School of Management of the Open University in The Netherlands. I would like to express my gratefulness for giving me this opportunity, especially to Ms. Verstappen.

The combination of the Master Accounting and Finance and the branding constellation theme is an unusual combination, but not for me. During my time at the Vrije Universiteit my favourite course was *Dienstenmarketing*. During my professional career I combined the best of both topics; by working as a financial controller for a Marketing & Brand department.

A special thanks to Wim Jurg for giving me the opportunity to pursue my thesis within the branding constellation theme. I am especially grateful for his support, commitment, and encouragement during the writing of my thesis. Wim Jurg shall maintain the CD-ROM with the confidential additional information to this thesis, this CD-ROM can be withdraw at Wim Jurg.

I also would like to express my gratitude to Dr. J.M.C. Schijns for his time and willingness to be the second reviewer of this thesis.

Amsterdam, August 31 2010,

Celine van Reij.

TABLE OF CONTENTS

1. INTRODUCTION

The *branding constellation theme* is a student research group within the School of Management at the Open Universiteit in The Netherlands (OUNL). It focuses on the usefulness of branding constellations.[1] The *branding constellation* is a technique to identify brand related problems such as whether a new taste of the branded product will be valued by the customers. *Usefulness* covers four dimensions within the theme: (1) relevance, (2) validity, (3) reliability, and (4) precision. This thesis evaluates the reliability and precision dimensions of the measurements used within former studied branding constellations within the research group. This introduction covers the research problem and the outline of this thesis.

1.1 Research Problem

This section describes the research problem. Subsection 1.1.1 presents the research objective and subsection 1.1.2 contains the research questions.

1.1.1 Research Objective

A major problem within the branding constellation theme is the inconsistency in the reliability and precision measurements, due to different approach of the measurements methods, which hinders a good comparison between the case study findings within the theme according to Karel (2009: 23). Therefore, this study evaluates these reliability measurements and develops standards to be able to measure the reliability and precision of branding constellations consistently within the theme in the future. Thus, the *research objective* of this thesis is:

The evaluation of the reliability measurements used for branding constellations within the branding constellations theme, including the presentation of reliability measurements to be used within the theme in the future.

1.1.2 Research Questions

To fulfil the research objective five steps were taken. First of all, a literature review on the reliability of projection techniques was executed, since branding constellations are perceived by branders and marketing experts as projections (Jurg, 2010: 7). Second, a study was conducted to investigate the definitions concerning reliability definitions used in the branding constellation theme's definition list, this in order to derive potential measurements connected to the term 'reliability'. Third, the reliability measurements used by students within the theme were described, analysed, and compared. Fourth, the reliability measurements to be used within the theme in the future were standardised. As a fifth and final step, the newly suggested reliability measurements were compared to the old measurements.

Thus, the study consists of five research questions:

1. Which reliability measurements have been employed in scientific studies on projections?
2. Which reliability measurements can be derived from the reliability definitions in the definition list used by the branding constellation theme?
3. What are the most important reliability measurements that were used by the students in the branding constellation theme?
4. How can these most important measurements be standardised?
5. What are the similarities and differences between the findings on the most relevant reliability measurements of the previous students and the standardised measurements?

1.2 Thesis Outline

The introduction to the branding constellation theme is presented in chapter 2 and the methodology in chapter 3. The thesis chapters then devoted to answering the research questions;

- Chapter 4 deals with the reliability measurements used in scientific studies on projections.
- Chapter 5 describes the reliability definitions in the definition list used by the branding constellation theme and derives potential reliability measurements.
- Chapter 6 focuses on the similarities and differences between the most important reliability measurements that were used by students in the branding constellation theme.
- Chapter 7 deduces the reliability measurement that should be used within the branding constellation theme in the future.
- In chapter 8, the findings on the new reliability measurements are compared with the previous reliability measurements.
- Chapter 9 closes the main text with theoretical implications for the branding constellation theme, a discussion, and a reflection.

2. BRANDING CONSTELLATION THEME

The purpose of this thesis is to contribute to the branding constellation theme by evaluating and improving the reliability measurements of branding constellations used within the theme. This chapter introduces the branding constellation theme. Section 2.1 focuses on the origin and development of the branding constellation theme. Section 2.2 presents the positioning of branding constellations in marketing.

2.1 Origin and Development of the Branding Constellation Theme

The branding constellation theme originated in 2002 as the Dean of the School of Management at the Open Universiteit in the Netherlands, Prof. Dr. Herman van den Bosch, asked Drs. Wim Jurg to start a research group on branding constellations parallel to his own PhD-thesis on the perceived usefulness of branding constellations. This dissertation is recently finished and will be defended November 1, 2010.

The branding constellation theme tries to describe and analyse branding constellations through 'objective' methodologies. Jurg (2008: 10) argues that a follow-up study on his thesis' findings should, among others, focus on the reliability of branding constellations by systematically comparing independently performed constellations by different members of brand teams on the statements of the stand-ins, the generated insights, and their intuitive truth. These reliability measurements generally have been named *test-retest reliability measurements* within the branding constellation theme (Karel 2009: 17). In addition, he argues that the branding constellation findings should be contrasted with the findings of scientifically accepted techniques, for instance, the relationships between the branding elements in branding constellations should be compared to outcomes of questionnaires and/or interviews. These reliability measurements generally have been named *triangular reliability measurements* within the theme. The third category of general reliability measurements within the theme is the category of precision measurements: the degree to which falsifiable and/or verifiable statements can be deduced from branding constellations.

The branding constellation theme focuses on comparative case studies. Three kinds of comparative case studies[2] can be distinguished: first-person, second-person, and third-person based on research of Bradbury and Bergmann Lichtenstein (2000: 551-564), these different case studies are explained in more detail below:

First-person case study: A case study in which the usefulness (relevance, validity, reliability, and/or precision) of a branding constellation is studied on a branding problem that is the student's responsibility as a brander. The first-person case study is primary research. In *primary research* the students collect their own data. This kind of study was performed by Gomersbach (2004), Davidse (2005), De Velde Harsenhorst (2006), De Heij (2006), Holwerda (2006), Vertregt (2007), and Karel (2009).

Second-person case study: A case study in which the usefulness of a branding constellation is studied on a branding problem of a brander who is an

acquaintance or a colleague of the student. The second-person case study is also a primary study. This kind of study was performed by Van Geel (2004), Mathijssen (2005), Ten Have (2007), and Halters (2008).

Third-person case study: A study of the falsification of a hypothesis regarding branding constellations, for instance

- Branding constellations are metaphors (Van Zwienen, 2005)
- Branding constellations are a form of lateral marketing (Van Mechelen, 2005)
- Branding constellations fit the Soft System Methodology (Simons, 2005)
- Branding constellations are a brainstorming technique (Harrewijn, 2006)
- Branding constellation are a form of action research (Labots, 2006)
- Branding constellations require emotional intelligence (Stroo, 2006)
- Quantum theory might explain the working of branding constellations (Blootens, 2006)
- Emotions of stand-ins correlate with their distances and directions (Schuurman, 2006), and
- Branders employing branding constellations score similar on the MBTI personality test than regular branders (Claus, 2008).

All hypotheses could not be falsified, except for the quantum theory, the sociometrics, and the personality hypotheses. Blootens concluded that quantum theory seems useful as a metaphor rather than as a description of reality fitting branding constellations. Schuurman concluded that the emotions of stand-ins do not correlate with their mutual distances and directions and Claus concluded that the MBTI-personality of branders employing branding constellations is significantly different on all four MBTI-dimensions. In total 23 students finished their master theses on branding constellations. Figure 2.1 presents all these theses with the brand/company studied and the kind of comparative case studies that was applied.

	Students	Company	Kind of case study		
			First-person	Second-person	Third-person
1	Van Geel (2004)	Legermuseum – 2004		X	
2	Siezen (2004)	Stork Fokker – 2003		X	
3	Gomersbach (2004)	Rabobank – 2004	X		
4	Mathijssen (2005)	RSM – 2004		X	
5	Davidse (2005)	DE&SP – 2005	X		
6	Van Zwienen (2005)	Blooming – 2004			X
7	Van Meer (2005)	Expert cases – 2004			X
8	Van Mechelen (2005)	KPN Mobile – 2002			X
9	Simons (2005)	MultiCopy – 2002			X
10	De Velde Harsenhorst (2006)	EODD – 2004	X		
11	De Heij (2006)	SKBA – 2004	X		
12	Holwerda (2006)	Philips-LG – 2004	X		
13	Stroo (2006)	Sigma – 2003			X
14	Harrewijn (2006)	Friso – 2002			X
15	Schuurman (2006)	Alex March/October – 2003			X
16	Labots (2006)	Expert cases 2003 – 04			X
17	Blootens (2006)	Hooghoudt - 2003 – 05			X
18	Ten Have (2007)	Lipton – 2007		X	
19	Vertregt (2007)	GTI – 2005	X		
20	Meijer (2008)	Local Rabobank – 2005	X		
21	Claus (2008)	Brander's Myers-Briggs – 2007			X
22	Halters (2008)	IDS Scheer – 2007			X
23	Karel (2009)	Comfort in Living – 2008	X		
		Total	8	4	11

Figure 2.1: Thesis, company at which the research was performed and the kind of comparative case studies applied

From Figure 2.1 can be concluded that third-person case studies is conducted most often and the second-person case study least often. This thesis can be categorized as a third-person case study.

2.2 Positioning of Branding Constellations in Marketing

According to Jurg (2010: 17) the identification of branding problems is an important subject in the marketing of a brand. Jurg (: 26) defines *brands* as self-organising systems that need energy from the environment to stay alive and that structural changes develop from positive feedback loops. He further argues that today virtually any entity on the planet with an ability to sustain an attraction is treated as a brand, such as celebrities, churches, cities, companies, countries, football teams, social movements, and political parties (: 25). Jurg (: 27) defines *branding* as conducting operations that make a positive contribution to the branding system; a *positive contribution* refers to marketing programme decisions made to improve the unique brand position in the consumers' minds in order to increase the brand's value.

Jurg (: 18) further argues that identification of branding problems could benefit from a systems perspective. The systems perspective studied within the theme is *branding constellations.* This is a new application of systems constellations employed to identify branding problems. Systems constellations cover by name two variations (: 33) family constellations and organisation constellations, applied to personal and organisational problems, respectively. The innovative assumption of branding constellations is that people who are set up by a brander as personified representations of elements of a branding system are able to express the implicit relationships between these elements in the brander's mind (: 107).

Jurg (: 15-16) argues that problem identification is the first and most important stage of marketing research. Marketing research starts with problem identification and is generally followed by research design, data collection, data analysis, and a research report. The problem identification is considered not only the most important, but also the most difficult stage, in the marketing research process. Problem identification has received little attention in the in the marketing field. Although marketers know that accurate marketing research should start with proper identification of a particular problem, and that its success depends on the quality of this problem identification, they fail to act in line with this knowledge. Problem identification processes in marketing practice were generally ad hoc and do not follow a systematic procedure. When marketing problems are not well-identified, it is very likely that the research will be a waste of time and money.

As a consequence of the limited attention granted to proper problem identification, there is also limited attention given to problem identification techniques and their validation, according to Jurg (:49). There is no generally accepted way to validate new problem identification techniques, and there is a substantial lack of valid research focused on the usefulness of problem identification techniques, according to Jurg (:49). Figure 2.2 presents Jurg's (: 49) comparison to other problem identification techniques, such as brainstorming, lateral marketing, the Soft Systems Methodology (SSM), psychodrama, projections, and the Zaltman Metaphorical Elicitation Technique (ZMET).

Dimensions	Holistic perspective	Reductionism perspective
Emotional approach	**Branding constellations**	Brainstorming, psychodrama, projections, and ZMET
Rational approach	Cognitive mapping and SSM	Lateral marketing

Figure 2.2: Positioning of branding constellations to other problem identification techniques (Jurg 2010: 49)

Jurg (: 49) argues that branding constellations differ from other problem identification techniques, since they combine a holistic perspective and an emotional approach. A *holistic perspective* focuses on the elements and relationships emerging from the whole rather than decomposing problems into the basic elements that form the core of the problem (the latter is classified as the reductionism perspective). An *emotional approach* includes bodily experiences, feelings, and intentions as well as spontaneous verbal 'outbursts'

based on these emotions rather than a logical verbalisation and an encouragement to employ grounded arguments (which is classified as a rational approach).

The key message of Jurg's PhD-thesis (2010: 7) is that branding constellations are perceived by marketing experts as a kind of systems projection. While other authors on systems constellations argue that it does not matter who is chosen by a brander as a stand-in for a branding element, the marketing experts think that this projection is the heart of branding constellations. A detailed description of the branding constellation procedure can be found in appendix A.

3. METHODOLOGY

This chapter describes the data design of the research that was used to fulfil the objective of this thesis. Section 3.1 deals with the collection and analysis of measurements used in scientific studies on the reliability of projection techniques. Thereafter section 3.2 describes the research procedure regarding the reliability definitions within the branding constellation theme's definition list to derive reliability measurements. Section 3.3 portrays way in which the most important reliability measurements were deduced. Section 3.4 describes the way in which the standardised reliability measurement was produced. Section 3.5 describes the way in which the reliability measurement that should be used in the branding constellation in the future was inferred.

3.1 Reliability Measurements in Literature on Projection Techniques

This section concentrates on the design to find the most relevant reliability studies on projections techniques in all their different sorts.

The thesis website of the branding constellation theme of the School of Management of the Open Universiteit provided five journal articles on projection techniques, by name Boddy (2005), Fassenbender (1997), Haire (1950), Hauser (1979), and Yoell (1974). In his reference list Boddy (2005) includes Anderson (1978), Levy (1994), and Lilienfeld et al. (2000).

A search with Google Scholar on the words 'projection technique' led to an article from Donoghue (2000). This article refers to Wagner (1985) and Catterall et al. (2000). During the search for the article of Wagner, a relevant article of Lundy (1985) was found, as both their articles were published in the same journal issue.

From these articles could be concluded that many authors also use the term 'projection test' instead of 'projection techniques'. A search on the website of google scholar using the search criterion 'projection test', led to one new article by Greenberg (1959). A further search on the criteria 'projection' and 'reliability' led to the articles of Perry (1998), Stein (1998), Thorndike (1985), and Yeager (2003).

Analysing these articles, made clear that the word 'consumer' could further refine the search criteria. By searching on the website of google scholar using the search criteria 'projection technique', and 'consumer' the article of Doherty (2010) was found. The article of Doherty (2010) refers to one new article, by name Steinmann (2009).

A search on the website of google scholar using the search criteria 'projection technique' and 'test-retest' led to the article of Jensen (1959), while a search on the website of ebook.com[3] using the search criteria 'projection' and 'test-retest' led to an article of Blatt (1975). Further, a search on the website of google scholar using the criteria 'reliability' and 'precision' led to the article of Thompson (2010). A search on the website of google scholar using the criteria 'reliability' and 'test-retest' led to the article of Schultheiss et al. (2008). After the search to

journals and e-book, the second search opportunity was in library of the Vrije Universiteit Amsterdam. Using the search criteria 'reliability' led to the books by Litwin (1995) and by Thorndike and Thorndike (2010). Using the search criteria 'projection technique' led to books written by Harrower et al. (1960) and Murstein (1965). In the library next to the book by Murstein (1965), books by Mc Donald (1999) and by Murstein (1963) were found.

Figure 3.1 presents the articles that included reliability measurements of projection techniques. The order of the articles is presented on publishing date, starting with the youngest, because the youngest shows the most current studies.

Author	Year of publication
Schultheiss et al.	2008
Donoghue	2000
Lilienfeld et al.	2000
Perry	1998
Lundy	1985
Wagner	1985
Jensen	1959

Figure 3.1: Articles studied on measurements of projections

Chapter 4 presents the essence of projection techniques and a methodological and integrative overview of the reliability findings, employed as in the premaster study of Schapendonk and Arts (2010: 7). The *methodological overview* presents the methods found in the literature to measure the reliability of projection techniques. The *integrative overview* presents the author, the studies referred to, projection technique, respondents, and findings. When authors present different findings, only the findings on the Thematic Apperception Test (TAT) were taken into account to keep the analysis transparent. *TAT* is a projection technique to reveal repressed aspects of personality, motives, and needs for achievement, power, and intimacy, and for problem-solving abilities (Yoell 1974: 33). The choice for TAT was based on the fact that branding constellations and TAT are both based on visual stimuli.

3.2 Reliability in Branding Constellation Theme Definition List

The definition list of the branding constellation theme was used to derive potential reliability measurements. This section explains the classification of these definitions.

Within the branding constellation theme a list of definitions has been in use (see CD-ROM, file: *Definitions Excel Files*). The definition list has been developed by Jurg during his PhD study. The list contained numerous different definitions of the concept 'reliability'. These definitions were classified in different interpretation clusters based on the interpretations of the definitions.

The interpretation procedure started with the file (file on CD-ROM: *1. Definitions Branding Constellation Theme 20100621*) containing all the definitions used within the theme at that time. The second step selected the term 'betrouw…' in

column B, because the definition list was formed from the Dutch terms for reliability (file on CD-ROM: *2. Definitions on Betrouw…*). The third step separated the definitions per author and was numbered (file on CD-ROM: *3. Definitions per Author on Betrouw…*, column B). The fourth step selected from the total list all English definitions containing *reliability* (by selecting *reliab…*). Only these definitions were taken into account while other versions and synonyms of the word reliability were not taken into account (file on CD-ROM: *4. English Definitions per Author on Reliab…*). The file containing the definitions of 'reliability' is called: *5. English Definitions per Author on Reliability (*file on CD-ROM, column G*)*. This file was used to analyse the definitions of reliability. These definitions were divided into interpretation clusters based on the selections of key-words from these definitions. During the selections all synonyms or words with similar meaning were grouped, these groups of synonyms or similar words were presented in chapter 5.

3.3 Most Important Reliability Measurements

In this section the measurements of the students within the branding constellation theme are described. The branding constellation theme has been using a wide variety of measurements to analyse the reliability of branding constellations. The first step to compare the different measurements was making an overview of all the reliability measurements used in the branding constellation theses: accuracy, equivalent reliability, falsification, inter brander reliability, inter element reliability, inter rater reliability, precision, test-retest reliability, and triangular reliability. A detailed overview of these measurements is presented in section 6.1, Figure 6.1. The second step was to match similar measurements. These two steps were based on the index and on the measurements perform and enclosed on the CD-ROM's of the theses. It turned out that accuracy, falsification, and precision denoted similar measurements. The same was true for equivalent and triangular reliability, and for the insight test-retest, inter brander test-retest, and inter element test-retest reliability. The reliability of the inter rater *reliability* and these similar measurements are currently studied by Van Elshout, and were therefore not taken into account in this thesis. As the triangular measurements only fit first person case studies according to Halters (2009: 58), these measurements were not taken into account in this thesis either.

The two most important measurements were operationalised as the measurements used by the highest number of students. The measurements that were taken into account in this thesis are (1) the bilateral test-retest reliability and (2) the precision measurements as these are the measurements that are used most often. A detailed presentation of these numbers is presented in section 6.1. Figure 3.2 presents the numbers of the findings: the bilateral test-retest reliability and the precision measurements.

Measures	Number of students
Bilateral test-retest reliability	9
Precision	14

Figure 3.2: Quantity bilateral and precision measures preformed by theme students

3.4 Standardised Reliability Measurements

This section describes the methodology resulting in standardised excel working files for the bilateral test-retest reliability measurement and for the precision measurement. In subsection 3.4.1 the research on the performed bilateral test-retest is described. In subsection 3.4.2 the research on the performed precision measurements is described. The development of the excel file is will enable researchers in future to perform a standardised bilateral test-retest reliability measurement. In subsection 3.4.3, the development of the excel file is described, this file can be used to perform a standardised bilateral test-retest measurement. In subsection 3.4.4, the development of the excel file to perform a standardised precision measurement is described.

3.4.1 Bilateral Test-Retest Reliability

The theses regarding the branding constellation theme also employed a variety of bilateral test-retest measurement procedures. The similar procedures were filtered out based on the measurements presented in the theses and enclosed on their CD-ROM's. The overview of the bilateral test-retest measurements is presented in section 6.2 Test-retest reliability, Figure 6.2. To improve the understanding of the bilateral test-retest measurements, all measurements of the theses were repeated. Detailed information of the study on the bilateral test-retest measurements is presented in appendix C. The study of the bilateral test-retest analyses was done in each (student's) bilateral test-retest file (see the CD-ROM *Support files/Excel standardised bilateral test-retest measurement*).

3.4.2 Precision Measurement

As reported in section 3.3.1, it turned out that 'precision' was also named 'falsification' and 'accuracy' by earlier theme students. First, an overview was made of all the 'precision' measurement procedures employed in the theses. This overview is presented in section 6.3, in Figure 6.5. The next step was to match similar measurement procedures. The filtering of the 'precision' measurements and all the similar tests was based on the measurements in the theses and on the enclosed CD-ROMs. To improve the understanding of the precision measurements the measurements of the theses were also repeated. Detailed information of the study to the theses is given in appendix D (see the CD-ROM *Support files/Excel standardised precision measurement*).

3.4.3 Standardised Bilateral Test-Retest Measurement

The design of the standardised Excel sheets was the first step to take. The second step was to develop the fill-out sheets for *stand-ins, and transcription.* The third step was to develop the first stand-in sheet (named *S1*). In this sheet all the statements of this stand-in were shown. This was possible by linking the *S1* to the *stand-in* sheet and the *transcription* sheet, by using a formula named sum-product (count-if). Errors in de the name of the stand-ins will be recognised by the sum-product (count-if) formula. Therefore a check on the stand-in names is included in the *transcription* sheet. If an incorrect (or not filled in the *stand-ins* sheet) stand-in name would be used, than the *transcription* sheet in column J would turn red.

The phase in the *transcription* was filled by using the abbreviation shown in the top of the sheet. The cell J changed colour automatically. In the stand-in sheets, column J to column AA represents the stand-ins. The correct stand-in name is shown after filling out the sheet *stand-ins*. These columns allow future students to fill out the scores for the bilateral measurements. The sheet *S1* was copied to *S2-S20, facilitator, brander and public*, including the formats and formulas. The formulas referring to the different stand-ins were adjusted per stand-in file. If the *stand-in* name is equal to the '*who'* in the sheet *transcription*, the input from the *transcription* is copied into the cells in the stand-in sheet. In all to rows where the *stand-in* is not equal to the '*who',* a "0" appears, when this row does not contain a statement of this particular stand-in. Pushing the grey square button on the top of the sheet will make these "0" rows disappear.

The fourth step was to develop the calculation of the bilateral test re-test scores. For the bilateral test-retest score all statements were scored. Due to the use of formulas, the calculations of the scores were done automatically to prevent calculation errors by future students. The formula employed was the sum-product (count-if) formula. A total overview of all the average scores is presented in the sheet *Total bilateral test-retest.* The rows that are not used due to less than the possible 18 stand-ins stay empty. This sheet also contains a button to delete the empty rows.

The fifth step for the bilateral test-retest measurement was to develop a formula to compare the results of the first branding constellation to the results from the second branding constellation. This was done by making an extra sheet containing a table with the possible results from the first constellation on the top of the sheet and the possible results of the second constellation on the right. By using and combining the formulas V-lookup and H-lookup the total score of the two constellations is presented in the sheet Total bilateral scores. The formula that combines the two results tables find the combination results in the result table, the result table is filed with figures with only one figure behind the comma (1.0, 1.3, 0.3, etc). Therefore in the total average table an extra formula was inserted to round off the results, otherwise the combination table would have given an error. The rows that were not employed due to less than the possible 18 stand-ins stay empty. Also this sheet contains a button to delete the empty rows.

The final step was creating legend sheets for the scores and scoring, an introduction for the student and other readers, and a manual for the students.

The program Excel gets upgraded every few years, therefore this study employed 'all around' formulas. These formulas were longer and more difficult to understand, but were suitable for all the excel versions. An extra sheet has been included in the file to explain the formulas employed, named *Formulas*.

3.4.4 Standardised Precision Measurement

A similar design to the standardised bilateral test-retest measurement was employed to develop a standardised precision measurement excel file. The first five steps from the development of the bilateral test-retest measurements file were repeated. Only the stand-in sheets needed another adjustment.

The sixth step for the precision score was including the sheet *Overview prec. sc. per phase* to present the scoring per phase. This was possible by employing the sum-product (count-if) formula, which is included in every stand-in sheet (*S1* t/m *S20*, see the cells P2:T10). The formulas selected the phase and calculated the frequency and average of the scores. The rows that are not used due to less than the possible 18 stand-ins give the amount “0”. This sheet also contains a button to delete al the “0” rows.

The seventh step was including a sheet where a selection can be made by future students on abstract or personal elements. By using the formula sum-product (count-if) the scoring of the abstract stand-ins or person stand-ins will be shown (see sheet *Overview prec. sc. abstract* and *Overview prec. sc. person*). This is possible due to the selection the future student have to make in the sheet *Stand-ins*, in column D, selection ‘person’ or ‘abstract’. The rows that were not used due to less, than the possible 18, stand-in give the amount “0”. Also this sheet contains a button to delete the “0” rows.

One of the last steps was creating legend sheets for the scores and scoring, an introduction for the student and other readers, and a manual for the students.

3.5 Previous versus Standardised Reliability Measurements

The development of the standardised reliability measurement caused for some changes in the reliability measurements. The impact of the changes in the reliability measurements of the theme students were measured. The outcome of the previous theme reliability measurement were compared to the standardized reliability measurements.

4 RELIABILITY MEASUREMENTS OF PROJECTION TECHNIQUES

In this chapter an answer is presented to the question *"Which reliability measurements have been employed in scientific studies on projections?"* Section 4.1 describes the essence of projection techniques. Section 4.2 presents the findings on the literature studies regarding the reliability of projection techniques.

4.1 Projection Techniques

Projection (or projection) techniques in qualitative marketing consumer research involve the presentation of ambiguous emotional stimuli, and asking stand-ins to make sense of them (Jurg 2010: 45-46; based on Boddy, 2005). Projections are considered useful when stand-ins have difficulty expressing some brand-related emotions, and researchers need some way of accessing these from within the respondents' minds. But it is considered vitally important that the respondents themselves make the interpretations in marketing research. Jurg argues that projections in marketing research are applied differently than in psychology. In psychoanalysis, a *projection technique* is a defence mechanism in which people subconsciously attribute to others than their own unacceptable thoughts or emotions. In contrast to the modest role of the facilitator in marketing research, the role of the facilitator in psychoanalysis is to confront clients with sensitive *interpretations*: verbal interventions through which the analysts make clients consciously aware of their subconscious projections. Jurg further argues that the dominant psychoanalytic thought in the UK is the *Object Relations School*. Here, *objects* were thought of as subconscious memory traces of experiences rather than subconscious physical drives as in psychoanalysis. The Object Relations School refers to *projection identification* rather than projection: the subconscious process of splitting off some parts of the self, projecting them on to people.

4.2 Literature Overview of Reliability Measurements

Several researchers studied the reliability of the projection techniques. The reliability studies are ordered in reliability measurements, starting with the test-retest reliability measurements, followed by the split-half, odd-even split, maximum split-half, inter rater, and triangular reliability measurements, respectively. The figures are split based on the author, researcher, projection method, respondents, and findings.

The authors described different projection techniques, the studies regarding the reliability of the TAT projection technique was examined further in this thesis. If the TAT was not discussed another performed reliability study of the projection technique was selected. The TAT is known as: "the picture interpretation technique because it uses a standard series of provocative yet ambiguous pictures about which the subject is asked to tell a story".[4]

Figure 4.1 presents the projection technique studies on the test-retest measurements. The relevant parts of the articles were included in the Appendix C. The studies are ranged on the authors' last name.

Author	Study	Projection variation	Methodology	Findings
Schultheiss et al.(2008)		PSE Picture Story Exercise 8 picture PSE	2 Weeks interval 24 male and 63 female participants (mean age 20 years), 5 different ethnic groups	r=0.39 (count-corrected power), r=0.37 (achievement), r=0.61 (affiliation)
Lilienfeld et al. (2000)	Kraiger (1984)	Thematic Apperception Test	Graduates. Test: write stories that are unique. Retest: write stories as similar as possible to test 1.	Test r=0.52 Retest r=0.38
	Winter (1973)	Thematic Apperception Test	Graduates. Tests write stories that are unique. Retest write stories as similar as possible to test 1	Test r=0.27 Retest r=0.61
Lundy (1985)		Thematic Apperception Test	102 Students Year 1 half test (n=93) Year 2 half test (n=87)	Test r=0.48 (affiliation) Retest r=0.56 (intimacy)
Jensen (1959)	Lindzey and Hermar (1955)	Thematic Apperception Test	20 People, two months interval four TAT cards	Average r=0.51
	Tomkins (1947)	Thematic Apperception Test	45 Women, interval two to ten months	2 Month r=0.80 6 month r=0.60, 10 month r=0.50

Figure 4.1: Test-retest reliability studies on projection techniques

Jensen (1959) describes the study of Tomkins, where 45 women were shown pictures, which represent stories to memories. After a period of time they were asked what they had seen. The results of the intervals (2 months 0.80, 6 months 0.60, and 10 months 0.50) suggest the curve of forgetting. The larger the interval of time is the lower the correlation result was. Lindzey and Herman (1955) tried to get around this problem by requiring making up a different story on retest if they recalled their first story; only three scored reliability coefficients were significantly larger than zero.

Lilienfeld (2000) argues that the tests of Kraiger (1984) and Winter (1973), Winter (1973) result gained higher scores than those of Kraiger (1984). In the retest Kraiger (1984) asked the respondents to write a different story, but Krager (1984) concluded that the respondent feel obliged to create similar stories. Winter (1973) gave the instruction to create either a unique story or a similar story as possible to their earlier story; the correlations of the retest of Winter (1973) were significantly higher. Conclusion was that the reliability for TAT was unsolved and needed more research regarding test-retest reliability.

The scores of Lundy were slightly higher than the predicted (0.40-0.50), the results were consistent across males and females.

Schultheiss et al. (2008) tested the internal consistency Cronbach's alpha (0.02 to 0.43), ipsative correlation coefficient (between 0.20 – 0.40). The *Cronbach's alpha* measures the internal consistency reliability among a group of items combined to form a single scale (Litwin, 1995:24). The ipsative correlation coefficient is if an individual's total score on the test is always equal to a given constant for each individual.

The findings on the reliability of projection techniques vary between the test 0.48 and 0.80, for the retest 0.27 and 0.61. The test result varied depending on the time interval of the retest; generally speaking it could be argued that longer time interval leads to a lower reliability score.

The reliability of projection techniques is also tested using the split-half method, which is similar to the Spearman-Brown odd-even split. Split-half method is a single test splits into two subtests. Items can be assigned at random, or odd-numbered items may be assigned to one subtest and an even-numbered item assigned to the other, in other words the Spearman-Brown odd-even split. The correlation between the half-test scores is the reliability of either (Mc Donald 1999: 95). The results of the entire test in the table vary between for the test 0.31 and 0.68, for the retest 0.62 and 0.91. By using special selected pictures the internal consistency can be considerably increased according to McClelland (Jensen 1959: 123). In the Figure 4.2 the split-half and odd-even split tests

Author	Researcher	Projections	Reliability method	Conducted research	Research findings
Perry (1998)	Cramer (1988)	Thematic Apperception Test	Split-Half	School aged children	r=.68 to r=71(three defences)
Wagner (1985)		Hand Test	Spearman-Brown Odd-even split	200 Individuals, randomly divide 2 equal size samples	Sample 1: r=0.64 sample 2: r=0.62
Jensen (1959)	Child et al.	Thematic Apperception Test	Split-half Reliability	183 Students, 6 month interval	Test r=.39, retest r=.91
Jensen (1959)	Child et al.	Thematic Apperception Test	Split-half Reliability	200 Personality questionnaires	High reliability scores (mean .73) and positive correlation

Figure 4.2: Reliability of projection techniques split-half

The reliability of projection techniques is tested using the maximum split-half method, see following Figure 4.3. Wagner (1985: 579) obtains maximum stable reliability scores substantially superior to the odd-even method. He concludes that projection tests were probably more internally reliable than had been reported before.

Author	Researcher	Projection technique	Reliability method	Conducted research	Research findings
Wagner (1985)		Hand Test	Maximum split-half	200 Individuals, randomly divide 2 equal size samples	sample 1: r=0.80 sample 2: r=0.77

Figure 4.3: Reliability of projection techniques maximum split-half

The reliability of projection techniques is tested using the inter rater method. The results vary between 0.61 and 0.81, depending on the chosen variable denial and projection, identification, and defence variable. Wagner (1985) compared the two test odd-even reliability and the maximum split-half, in both samples the maximum split-half was greater than the odd-even. The both reliability measurements proved to be comparable, in both test the test scored higher than the retests. The scores of the test variables were also comparable in both test results.

In Figure 4.4 the inter rater reliability is shown.

Author	Researcher	Projection Technique	Reliability Method	Research	Findings
Perry (1998)	Cramer (1991)	Thematic Apperception Test	Inter rater	Inexperienced to experienced raters, School aged children	r=0.81 (denial and projection) r=0.64 (identification)
Perry (1998)	Cramer & Gaul (1988)	Thematic Apperception Test	Inter rater	School aged children	Inter rater reliability 0.81 (defence)

Figure 4.4: Reliability of projection techniques inter rater

Perry concluded that children exposed to a failure condition reported more negative affect in their stories, and used more defences, particular denial and projection.

According to Jensen (1959) a projection test is clearly not the same test in de hands of different examiners or under different conditions of administration. The inter rater reliability represents a reliability measurement of one measurement by two or more researchers.

The reliability of projection techniques is tested using the triangulation method, by combining two or more methods of data collection. The projection technique is combined with informal interviewing and individual interviews to enhance the outcome, see following Figure 4.5. Levy (1994: 5) states another way of examination the results is comparing the data to normative date. However, such norms do not usually exist in market research. The projection techniques were usually employed in combination with other quantitative of qualitative research techniques (Catterall et al. 2000: 248), for example by adding a questionnaire.

Author	Researcher	Projection Technique	Reliability Method	Research	Findings
Donoghue (2000)	Solomon (1994)		Triangulation	Combining project technique with informal interviewing	Enhance the value of the data

Figure 4.5: Reliability of projection techniques triangulation

Donoghue (2000) states that projection techniques are often used in conjunction with individual interviews and focus-groups moderators also use projection stimuli to enhance focus discussions. Projection technique combining with informal interviewing will enhance their value.

In Germany the projection techniques have been used in relation to answering organisational questions. In this study German managers were asked to response to a word association task, requiring visualizing their companies as animals. The words given, needed to be rated on high or low corporate success. Due to the experience of the managers in their companies the word association did not lead to an unconscious process, but it provided data for an interview or for a survey. Fassenbender (1997: 174) concluded that the word association approach is a good technique for organisational consulting.

There were several advantages to using projection techniques, including the amount, richness, and precision of the information that is collected (Donoghue 2000; Steinman 2009: 42). Projection techniques, when used properly, enable the researcher to access presumably unreachable beliefs, attitudes, values, motivations, personality, cognitions, and behaviours (Donoghue 2000; Steinman 2009: 42). Steinman (2009:42) states that the nature of projection techniques is that the true purpose of the instrument is well disguised and, in most instances, the subjects were not aware of the purpose of the exercise.

There were other researchers who feel uncomfortable about using projection techniques, due to ethical concerns issues of validity, reliability, interpretation of data, and the choice and design of the projection techniques (Catterall 2000: 250). Catterall et al. (2000: 252), state that there is more research needed, users agree that designs should be kept simple, avoiding too much detail or stylisation. Jensen (1959) agrees on this part by arguing that the most satisfactory solution, to the high standard error of a projection technique, is to eliminate or reduce as many of the sources of error variance as possible by making the administration, scoring, and interpretation of the projection test more standardized and objective.

From the researched reliability measurements the test-retest reliability and the triangulation are the most important reliability measurements.

5 RELIABILITY MEASUREMENTS DERIVED FROM BRANDING CONSTELLATION THEME DEFINITION LIST

In this chapter the answer to the question *"Which reliability measurements can be derived from the reliability definitions in the definition list used by the branding constellation theme?"* is presented. Section 5.1 presents the reliability findings from the definition list and section 5.2 presents the sub definition of reliability from the definition list. Section 5.3 closes this section by answering the question.

5.1 Analyses of the Definition List

The objective of this section is to categorise the definitions of reliability from the definition list into interpretation clusters. The first step was to select keywords in each definition of the authors (selected from column D in the excel file: *6.English Definitions per Author on Reliability, Different*). For instance, the first definition in this list was the definition of *administrative reliability:* "estimates unreliability due to having different scorers" (Peter, 1977: 397). The selected keyword in this definition was 'different'.

Words with the same meaning, found during the analysis, were also selected in the interpretation cluster 'different'. To be specific, the words that were considered synonyms or related to 'different' were: 'alternative', 'more', 'other', 'same', 'second', 'separate', 'subset', and 'two'. If two or more words were found in a definition, only the first word was taken into account and this definition was positioned in the interpretation cluster including the first word. For example: "Extent to which different items intended to measure the same thing correlate with each other" (Easterby-Smith, Thorpe, & Lowe, 2004: 135): this definition was taken into account for the word "different" and therefore not for the word "other".

The following keywords were interpreted as equal to 'scorers': 'author(s)', 'individual(s)', 'judge(s)', 'observer(s)', and 'researcher'. For instance, the brander is the author of the formulation of the branding problem; he is the individual performing the constellation; he judges the findings of the constellation; he is the observer of the branding constellation; and finally, he is a researcher of his branding problem.

Thus, 'scorers' is the key term of the key notion of the first interpretation cluster. *Reliability* means within this interpretation cluster that two or more different scorers have the same interpretation.

This interpretation, however, does not show what should be interpreted in the same way. For instance, the second definition in the definition list (the excel file: *7. English Definitions per Author on Reliability, Different Scorers Cluster,* column D, row 37) includes the term "same conclusions": "Whether a different researcher would have reached the same conclusions" (Noordegraaf in Schijns and De Oude-De Wolff, 2003b: 23). The following keywords were interpreted as corresponding to 'conclusions': 'category', 'finding', 'measurement', 'observation', 'response', 'result', 'scale', and 'score' (or their plurals). Similarly, the following keywords were interpreted as corresponding to 'same': 'equivalent' and 'similar'.

'Same' and 'equivalent' were the highest degrees of similarity. To conclude the first interpretation cluster was called 'scorers', which included that 'different scorers came to similar conclusions'.

A second interpretation cluster was derived from the definition of Gill and Johnson (2005: 119): "Extent to which a measuring device will produce the same results when applied more than once to the same person under similar conditions". The following keywords were interpreted as corresponding to "more than once to the same person under similar conditions": 'occasion', 'variables', and 'time'. A different 'occasion' included a different 'time'. This interpretation cluster was named 'time'.

A third interpretation cluster was derived from the definition of Kerlinger and Lee (2000: 653): "Develop two equivalent forms of the measurement instrument. [...] Each person would then have two scores, and again, the pairs of scores would be used in a correlation formula to compute the correlation". The identification of a branding problem by branding constellation can be compared with the identification of the problem by other problem identification techniques.

The following keywords were interpreted as corresponding to 'instrument': 'method', and 'test'. The third interpretation cluster was named different 'measurements'. Key words with a similar meaning found during the analysis were also taken into account in the interpretation cluster 'measurements', by name 'repeat…', 'consisten…', 'replica…', 'separate', and 'correlat…'.

The keyword 'repeat…' was based on the definition of *reliability* as "To be admitted as scientific knowledge, observations must occur under the prescribed circumstances not once but repeatedly" (Neale and Liebert, 1986: 8). The keyword 'consistency' was based on the definition of *reliability* as "Consistency of performance from each purchase to the next" (Aaker, 1991: 92/3). This keyword 'consistent' was based on the definition of *reliability* as "The extent to which measures were free from random error and give consistent results" (Proctor, 2004: 529). The keyword 'replica…' was based on the definitions of *reliability* as "refers to how replicable the finding is, or how dependably it will be repeated on another occasion, or by another searcher" led us to not to a new definition of *reliability,* because the definition was already taken into account at the selection of the word other. Their other two definitions containing the word 'replica..' were taken into account with the selection of the words different or consistent. The keyword 'separate' was based on the definition of *reliability* as "Separate sets of measures were employed which were designed to be as similar as possible" (Peter, 1977: 395). The keyword 'correlat…' was based on the definition of *reliability* as "Characteristic of a test possessed by virtue of the positive inter correlations of the items composing it" (De Groot, 288; after Kuder and Richardson, 1937: 159).

A fourth interpretation cluster was derived from the definition by Neale and Liebert (1986: 37) of *reliability* as indicated for instance by "The degree that a particular observation has yielded a 'true' score". The selection on 'true' presented definitions that cannot be placed in previous interpretation clusters, therefore the new interpretation cluster is called 'truth'. The keyword 'error variance' was interpreted as similar to 'truth'. Therefore, the definition of

reliability as "Proportion of error variance of the total variance yielded by a measuring instrument subtracted from 1.00, the index indicating perfect reliability" Kerlinger and Lee, 2000: 648) did not lead to a new interpretation cluster.

A fifth interpretation cluster was derived from the definition of *reliability* by Kerlinger and Lee (2000: 643): "Lack of distortion or precision of a measuring instrument". Precision was interpreted as a new interpretation cluster.

Keywords with a similar meaning found during the analysis were also taken into account in the interpretation cluster 'precision', by name 'accura..', and 'tru…' as these terms did not lead to a new interpretation cluster. The keyword 'accura…' was based on the definition of *reliability* as "Making accurate choices concerning suitable network partners and establishing open minds within communication processes" (Wilhelmer & Wagner-Luptacik, 2008: 3). This did not led to a new the interpretation cluster.

Not all reliability definitions were included in these keyword selections. The definitions that did not contain any of the previous keywords, were denoted with the keyword 'blank'. For example: "Stability of observations over time" (Miles & Huberman, 1994: 278; after Kirk and Miller, 1986) and "Stability if an observation through time" (Silverman, 2005a: 225; after Kirk & Miller, 1986). Stability is considered similar to consistency. Thus, these definitions did not lead to new interpretation clusters.

To conclude, five interpretation clusters were derived from all reliability definition from the branding constellation definition list:

1. Scorers (different or same)
2. Times (different or same)
3. Measurements (different or same)
4. Truth
5. Precision.

Figure 5.1 presents the five interpretation clusters and the number of definitions of reliability referring to this interpretation cluster.

Definitions	Interpretation cluster
37	Times
34	Measurements
33	Scorers
8	Truth
2	Precision

Figure 5.1: Overview interpretation clusters

Most definitions refer to the first three interpretation clusters: times, measurements, and scorers (different or same). However, in their definitions many authors combine these two of the three clusters. Figure 5.2 presents a matrix containing combinations of the first three interpretation clusters. The fields show how the authors combined the first three interpretation clusters in their reliability definitions.

		Different			Same		
		Scorers	Times	Measurements	Scorer	Time	Measurement
Different	Scorers	(Peter, 1977: 397)(Noordegraaf in Schijns & de Oude-de Wolff, 2003b: 23) (Peter, 1977: 399) (Riege, 2003: 81)(Easterby-Smith, Thorpe, & Lowe, 2004: 53)	(Aaker, 1991: 92/3) (Ereaut, 2003: 147) (Mariampolski, 2001: 57) (Proctor, 2004: 194) (Riege, 2003: 81) (Perreault & Leigh, 1989: 140)	(Silverman, 2005a: 123) (Silverman, 2005a: 225; Silverman, 2005b: 210, 224, 380)		(Miles & Huberman, 1994: 278)	(Gummeson, 2000: 91) (Gummeson, 2000: 185) (Cohen, 2008: 90) (Yin, 1994: 33; Yin, 2003: 34) (Kerlinger & Lee, 2000: 642) (Neale & Liebert, 1986: 43) (Perreault & Leigh, 1989: 13) (Silverman, 2005a: 229)(Kerlinger & Lee, 2000: 653)
	Times	(Aaker, 1991: 92/3) (Ereaut, 2003: 147) (Mariampolski, 2001: 57) (Proctor, 2004: 194) (Riege, 2003: 81) (Perreault & Leigh, 1989: 140)	(Proctor, 2004: 186) (Robson, 2002: 101) (Robson, 2002: 551) (Easterby-Smith, Thorpe, & Lowe, 2004: 53)(Kerlinger & Lee, 2000: 653)(Neale & Liebert, 1986: 40)	(Silverman, 2005a: 33) (Peter, 1977: 395) (Peter, 1977: 395)	(Proctor, 2004: 529) (Gill & Johnson, 2005: 119) (Remenyi, Williams, Money & Swartz, 2005: 288) (Franzen & Moriarti, 2009: 365) (Peter, 1979: 8) (Peter, 1979: 9)(Peter, 1979: 8)		(Chandler & Owen, 2003: 61) (Mason, 2002: 187) (Kerlinger & Lee, 2000: 663) (Neale & Liebert, 1986: 8) (Peter, 1977: 394) (Peter, 1977: 397) (Miles & Huberman, 1994: 278)(Silverman, 2005a: 225)(Peter, 1977: 395) (Peter, 1979: 8) (Peter, 1977: 395)
	Measerements	(Silverman, 2005a: 123) (Silverman, 2005a: 225; Silverman, 2005b: 210, 224, 380)	(Silverman, 2005a: 33) (Peter, 1977: 395) (Peter, 1977: 395)			(Peter, 1979: 6)	
Same	scorer		(Proctor, 2004: 529) (Gill & Johnson, 2005: 119) (Remenyi, Williams, Money & Swartz, 2005: 288) (Franzen & Moriarti, 2009: 365) (Peter, 1979: 8) (Peter, 1979: 9)(Peter, 1979: 8)				(Peter, 1981: 136)(Kerlinger & Lee, 2000: 653)(Neale & Liebert, 1986: 42)
	Time	(Miles & Huberman, 1994: 278)		(Peter, 1979: 6)			(Peter, 1977: 395)(Silverman, 2005a: 226)
	Measurements	(Gummeson, 2000: 91) (Gummeson, 2000: 185) (Cohen, 2008: 90) (Yin, 1994: 33; Yin, 2003: 34) (Kerlinger & Lee, 2000: 642) (Neale & Liebert, 1986: 43) (Perreault & Leigh, 1989: 13) (Silverman, 2005a: 229)(Kerlinger & Lee, 2000: 653)	(Chandler & Owen, 2003: 61) (Mason, 2002: 187) (Kerlinger & Lee, 2000: 663) (Neale & Liebert, 1986: 8) (Peter, 1977: 394) (Peter, 1977: 397) (Miles & Huberman, 1994: 278)(Silverman, 2005a: 225)(Peter, 1977: 395) (Peter, 1979: 8) (Peter, 1977: 395)		(Peter, 1981: 136)(Kerlinger & Lee, 2000: 653)(Neale & Liebert, 1986: 42)	(Peter, 1977: 395)(Silverman, 2005a: 226)	(De Groot, 288)(Neale & Liebert, 1986: 39) (Easterby-Smith, Thorpe, & Lowe, 2004: 135)

Figure 5.2: Authors of the definitions divided over the interpretation clusters

Each field represents a combination of the interpretation clusters. For example ‘different scorers’ (horizontal) and ‘same time’ (vertical), could also be found as ‘same time’ (horizontal) and ‘different scorers’ (vertical), the order of the words in the definitions of the reliability were not taken into account for the analyses on the interpretation clusters. Several fields were empty; these combinations were not employed by authors in a reliability definition.

Figure 5.3 presents an overview of all the stated 13 combinations of the interpretation clusters, with the quantity reliability definitions that refer to this combination.

Definitions	Interpretation cluster	
11	Different times	Same measurements
9	Different scorers	Same measurements
7	Different times	Same scorer
6	Different scorers	Different times
6	Different times	
5	Different scorers	
3	Same measurements	
3	Same scorer	Same measurements
3	Different times	Different measurements
2	Different scorers	Different measurements
2	Same time	Same measurement
1	Different scorers	Same time
1	Different measurements	Same time
58	*Total*	

Figure 5.3: Number of definitions per combination of interpretation clusters

Of the total 58 definitions in 11 definitions the authors of these definitions refer to ‘different times’ (or occasion) and the use of the ‘same method’ for the reliability of a study. In only 1 definition the author refers to use ‘different measurements’ at the ‘same time’ for the reliability of a study.

The authors of the reliability definitions regarding the interpretation cluster ‘truth’ were presented in Figure 5.4.

Truth
Kerlinger and Lee (2000: 643)
Kerlinger and Lee (2000: 648)
Kerlinger and Lee (2000: 648)
Kerlinger and Lee (2000: 657)
Neale and Liebert (1986: 37)
Peter (1977: 395)
Peter(1977: 395/6)
Peter (1979: 7)

Figure 5.4: Authors of the truth interpretation cluster

Figure 5.4 shows that the interpretation cluster ‘truth’ is derived from only three authors, whereas the first three interpretation clusters are based on 25 authors in 28 different articles.

The authors of the reliability definitions regarding the interpretation cluster precision are presented in Figure 5.5.

Precision
Kerlinger and Lee (2000: 643)
Wilhelmer and Wagner-Luptacik (2008: 3)

Figure 5.5: Authors of the precision interpretation cluster

Figure 5.5 shows that the interpretation cluster ‘precision’ is even derived from only two authors, whereas the first three interpretation clusters are based on 25 authors.

5.2 Reliability Sub Definitions

The definition list employed by the branding constellation theme also includes sub forms of reliability, named *reliability sub definitions*: reliability … or … reliability. The reliability sub definitions were included in the interpretation cluster allocation presented in section 5.1. Figure 5.8 presents an overview of the reliability sub definitions found in the reliability definition list and the category involved.

Sub definition	Description	Interpretation cluster
Administrative reliability	"estimates unreliability due to having different scorers" (Peter 1977: 397)	Different scorers
Alternative forms [reliability]	"Separate sets of measures are employed which are designed to be as similar as possible" (Peter 1977: 395)	Different measurements
Diachronic reliability	"Stability of observations over time" (Miles & Huberman, 1994: 278; after Kirk & Miller 1986)	Different times
Equivalent reliability	"Extent to which different items intended to measure the same thing correlate with each other" (Easterby-Smith, Thorpe, & Lowe 2004: 135)	Different measurements
Index of reliability	"Correlation of the true score with the observed score" (Kerlinger & Lee 2000: 653)	Truth
Inter item reliability	"Involves correlating scores on two or more subsets of items on a test" (Neale & Liebert 1986: 42)	Different measurements
Interjudge reliability	"Involves having the same individual rated, examined, or observed by two or more individuals" (Neale & Liebert 1986: 43)	Different scorers
Internal consistency of reliability	"Parts of a multi-item scale can be correlated with other parts to obtain a very close approximation of the correlation of a measure with itself" (Peter 1981: 136)	Different measurements
Inter-rater reliability	"Giving the same data to a number of analysts (or raters) and asking them to analyze it according to an agreed set of categories" (Silverman 2005a: 229)	Different scorers Same measurement
Reliability coefficient	"Proportion of the 'true' variance to the total variance of a measurement instrument" (Peter 1977: 395)	Truth
Reliability heuristic	"Consumers choose the brand that seems most reliable or has a reputation of reliability" (Franzen & Moriarti 2009: 365)	Not applicable
Split-half reliability	"Based on two halves that are usually considered as equivalent or parallel. If we develop this concept further by considering each item as a separate test, we can derive some of the measures that are commonly found in the psychological and educational literature" (Kerlinger & Lee 2000: 653; after Kuder & Richardson 1937) (KR-20, KR21)	Different measurements
Synchronic reliability	"Similarity of observations within the same time period" (Silverman, 2005a: 226; after Kirk & Miller 1986: 42)	Different times Same scorers
Test-retest reliability	"Identical set of measures is given to the same sample on two separate occasions" (Peter 1977: 395)	Different times Same measurement

Figure 5.8: Reliability sub definitions

Figure 5.8 shows that four of the five interpretation categories would have been found by analysing the sub definitions of reliability. Only the category 'precision' would not have been found this way.

Within the branding constellation theme a list of definitions is in use. The list contains numerous different definitions of *reliability*. These definitions were classified in five different interpretation clusters based on the interpretations of the definitions: different scorers come to similar conclusions; similar conclusions at different times; different measurements result in similar conclusions; the truth of the conclusions; and the precision of the conclusions.

The most definition refer to ‘different times’ (or occasion) and ‘same method’ regarding reliability. The interpretation clusters ‘different time’ and ‘same method’ are in combination similar to test-retest. Definition of test-retest reliability according to Peter (1977: 395) “Identical set of measures is given to the same sample on two separate occasions”

6 MOST IMPORTANT RELIABILITY MEASUREMENTS

In this chapter the answer to the question "*What are the most important reliability measurements that were used by the students in the branding constellation theme?"* is presented. Section 6.1 presents and analyses the reliability forms employed in the theme theses. Section 6.2 presents and analyses the test-retest reliability measurements employed in the theses. Section 6.3 presents an analysis of the precision measurements.

6.1 Overview of Reliability Measurements in the Theses

In this section an outline is given of the different reliability measurements employed in the theses: inter brander, inter element, inter rater, precision (falsification and accuracy), test-retest, and triangular (equivalent)reliability,. Figure 6.1 presents their definitions and the other names within the theme used to denote these measurements. Figure 6.1 presents only those theme theses that included a reliability measurement as part of their study.

Measurement	Description	Other theme names
Accuracy	The assessment to what extent the statements and movement of the stand-ins in the branding constellation-as-executed were verifiable and unambiguous (Vertregt 2007: 9)	Falsification, precision
Equivalent reliability	Implies a design in which two equivalent forms of measurement are developed. The theme has developed a parallel form of the branding constellation to test the opinions on the equivalent reliability of the branding constellation (Gomersbach 2004: 5)	Triangular reliability
Falsification	According to Popper science is about discovering and correcting mistakes. He argues precision is no truth criterion, but vagueness and contradictions can be a clue for mistakes (Van Zwienen 2005: 15)	Accuracy, precision
Inter brander reliability	Considers the degree in which the constellation is rated the same by two or more branders in equal settings (Van Geel 2004: 5)	Inter rater reliability
Inter element reliability	Considers the fit of the behaviour of the representatives in the branding constellation: "Do their behaviours harmonize?" (Van Geel 2004: 5)	Inter rater reliability
Inter rater reliability	Involves the degree in which the constellation is rated the same by two or more individuals (Gomersbach 2004: 5)	Inter brander, inter element, and insight reliability,
Precision	What degree statements of the stand-ins have the possibility of being untrue (Ten Have 2007: 29)	Accuracy, falsification
Test-retest reliability	The question here is, whether the same or similar results occur if the minds of the branders are 'measured' again with the branding constellation (Siezen 2004:13)	Presented in subsection 6.2
Triangular reliability	Outcomes of the branding constellation proceedings are compared with the findings of traditional problem identification data sources, namely a literature study, historical analysis and product/market analysis (Ten Have 2007: 7)	Equivalent reliability

Figure 6.1: Reliability measurements in the theses

Figure 6.1 shows that six terms were changed during the development of the theme. This made the analysis more complicated. For instance, within the branding constellation theme the comparison between branders of a brand team was called *inter rater reliability* by thesis number 20 (Meijer, 2008: 35) and *insights test-retest reliability* by thesis number 23 (Karel, 2009: 18). *Inter brander reliability, and inter element reliability,* were also interpreted as a comparison between the insights of branders and marketing experts by thesis number 19 (Vertregt, 2007: 25). The reliability of the inter rater reliability measurements is studied by theses number 25 (Van Elshout).

An overview of the reliability measurements performed by the theme students is given in Figure 6.2.

			Reliability					
	Students	Company	Inter brander	Inter element	Inter rater (Insight*)	Precision, accuracy falsification	Test-retest	Triangulation (equivalent*)
1	Van Geel (2004)	Legermuseum - 2004	X	X	X	F	X	X*
2	Siezen (2004)	Stork Fokker - 2003	X	X	X	F		X*
3	Gomersbach (2004)	Rabobank - 2004	X	X	X	F		X*
4	Mathijssen (2005)	RSM - 2004				F	X	
5	Davidse (2005)	DE&SP - 2005	X	X	X	F		X*
6	Van Zwienen (2005)	Blooming - 2004				F	X	X
9	Simons (2005)	MultiCopy - 2002					X	X*
10	De Velde Harsenhorst (2006)	EODD, - 2004				P	X	
11	De Heij (2006)	SKBA - 2004				F	X	
12	Holwerda (2006)	Philips-LG - 2004	X	X	X	A		X*
15	Schuurman (2006)	Alex March/October - 2003					X	
17	Blootens (2006)	Hooghoudt - 2003 - 05					X	X
18	Ten Have (2007)	Lipton - 2007			X*	P	X	X
19	Vertregt (2007)	GTI - 2005			X	A		
20	Meijer (2008)	Local Rabobank - 2005	X	X	X	A		
22	Halters (2008)	IDS Scheer - 2007			X*	P	X	X
23	Karel (2008)	Comfort in Living - 2008			X*	P	X	X
		Total	*6*	*6*	*10*	*14*	*11*	*11*

Figure 6.2: Reliability measurements in the branding constellation theme theses

Figure 6.2 shows that precision, test retest, and triangular reliability were studied most often. *Precision* is defined as the assessment to what degree statements of the stand-ins have the possibility of being untrue; the precision in the branding constellation is the verification of the statements of stand-ins. In other words the measurable elements were compared. “Measurable” means that the elements can be checked: the brander herself, and the stand-ins (Van Geel 2004: 40). Within the theme the precision measurement was called *falsification* by thesis number 1 Van Geel (2004: 40), *accuracy* by Vertregt (2007: 30) and *precision*

by Karel (2009: 19). The precision measurements of the theme students will be discussed in this thesis.

6.2 Outline of Test-Retest Reliability Measurements

The students of the theme theses used a variety of test-retest reliability measurements. In this thesis only the *most important* test-retest measurements were taken into account, operationalised as 'used by 5 or more students'. For an enhanced impression were the different test-retest reliability measurements found in the theses described in Figure 6.5. The definitions derived from the theses that performed the test-retest reliability measurement.

Test-retest	Description	Other theme name
Analytic	The relationships between branding elements from the first and the second branding constellation (Blootens 2007: 26)	Bilateral
Harmony	Compares the relationships between elements of the May branding constellation with those of the October one (De Heij 2006:32)	Bilateral
Bilateral	Examines the similarities and dissimilarities between the relationships of combinations of the constellated elements in both branding constellations (Halters 2008: 47)	Bilateral
Insight	Compares the gained insights by the two branders for the three sources (transcriptions, workshop questionnaire and email questionnaire) in total (Ten Have 2007: 29)	Insight
Messages	Compares the messages of the two branding constellations (Halters 2009: 32)	Messages
Most comparable moment	By contrasting the most comparable moments in the two constellations (Ten Have 2007: 38)	Storyline
Proceedings	Compared the positions, directions and statements of the stand-ins in the initial projection constellation and in the final vision constellation (Halters 2009: 33)	Proceedings
Qualitative analytic	Evaluated the happenings per element in the two constellations (Ten Have 2007: 42)	Qualitative Analytic
Quantitative analytic	Regarded to what degree the relationships of similar elements were alike (Ten Have 2007: 45)	Bilateral
Relations	Compares the relationships between elements of the first branding constellation with those of the second one (Van Zwienen 2005: 21)	Bilateral
Storyline	Two branding constellations' outlines were compared (Ten Have 2007: 41)	Messages
Systemic	The storylines of the two branding constellations were compared (Schuurman 2006:55)	Messages

Figure 6.3: Test-retest definitions in theme theses

Figure 6.3 shows that many terms changed during the development of the theme. This made the analysis more complicated. For instance, synonyms for the bilateral test-retest reliability were relation test-retest reliability (Van Zwienen, 2005 and Simons, 2005), harmony test-retest reliability (De Heij, 2006), analytical test-retest reliability (Blootens, 2006), and quantitative analytical test-retest reliability (Ten Have, 2007).

Figure 6.4 presents the test-retest reliability measurements performed by the theme students.

	Students	Company	Test-retest										
			Analytic	Harmony	Bilateral	Insight	Messages (storyline*)	Most Comparable Moment	Proceedings	Qualitative analytic	Quantitative analytic	Relations	Systemic
1	Van Geel (2004)	Legermuseum-2004							X				
4	Mathijssen (2005)	RSM - 2004			X								
6	Van Zwienen (2005)	Blooming - 2004			X							X	
10	De Velde Harsenhorst (2006)	EODD - 2004			X				X				
11	De Heij (2006)	SKBA - 2004		X	X								
15	Schuurman (2006)	Alex - 2003	X		X								X
17	Blootens (2006)	Hooghoudt-2003-05	X		X								X
18	Ten Have (2007)	Lipton - 2007			X	X	X	X		X	X		
22	Halters (2008)	IDS Scheer-2007			X	X	X		X				
23	Karel (2008)	Comfort in Living – 2008			X	X	X		X				
		Total	*2*	*1*	*9*	*3*	*3*	*1*	*4*	*1*	*1*	*2*	*2*

Figure 6.4: Outline of test-retest in theses

Figure 6.4 shows that the bilateral test retest was studied most often within the category of the test-retest reliability measurements. Therefore, this thesis focuses on the bilateral test-retest reliability. Van Zwienen (2005) and Mathijssen (2005) could not be taken into account in the bilateral test-retest measurements, because no two similar elements were involved in both constellations.

From all the reliability measurements the test-retest measurement and the precision measurement are the most performed measurements. From all the test-retest reliability measurements the bilateral test-retest was the most performed test-retest measurement. The two most important reliability measurements of the branding constellation theme are the bilateral test-retest measurements and the precision measurement.

7 STANDARDISED RELIABILITY MEASUREMENTS

This chapter provides an answer to the question *"How can these most important measurements be standardised?"* Section 7.1 describes the similarities and differences in the reliability measurements by the theme students. In section 7.2 the standardised reliability measurements by the theme students.

7.1 Similarities and Difference in the Reliability Measurements

In this section the similarities and dissimilarities in the reliability measurements performed by the theme students are described. Subsection 7.1.1 described the similarities in the bilateral test-retest measurements and subsection 7.1.2 the differences in the bilateral test-retest measurements. Subsection 7.1.3 presents the similarities in the precision measurements and subsection 7.1.4 the differences in the precision measurements.

7.1.1 Similarities in the Bilateral Test-Retest Measurements

Similar measurements of the bilateral test-retest were named analytic, harmony, quantitative analytic, and relation test-retest measurements. These measurements were all based on the transcription of the constellation. The statements of the stand-in were each study scored based on the interpreted relationship with other stand-ins/elements. A 5-point scale [-2; +2] was used to indicate the relationship between the stand-ins. The value of the scores is presented in Figure 7.1.

Score	Motivation
2	The relationship between elements in the branding constellation is very positive
1	The relationship between elements in the branding constellation is rather positive
0	The relationship between elements in the branding constellation is neutral
-1	The relationship between elements in the branding constellation is rather negative
-1	The relationship between elements in the branding constellation is very negative

Figure 7.1: Legend stand-in relationship scores

7.1.2 Differences in the Bilateral Test-Retest Measurements

The earlier theses measured if the stand-in had a relationship with the other stand-ins and if this relationship was comparable with the same stand-in in the second constellation. The theme students of first group selected only a few statements that illustrated their scores of these relationships. The later theses selected all statements of the stand-ins and scored these one by one (see Figure 7.2).

	Students	Company	Bilateral test-retest	
			Total relationships	All relationships
10	De Velde Harsenhorst (2006)	EODD - 2004	X	
11	De Heij (2006)	SKBA - 2004	X	
15	Schuurman (2006)	Alex - 2003	X	
17	Blootens (2006)	Hooghoudt - 2003-05	X	
18	Ten Have (2007)	Lipton - 2007	X	
22	Halters (2008)	IDS Scheer-2007		X
23	Karel (2008)	Comfort in Living - 2008		X

Figure 7.2: Bilateral test-retest the two groups

The measurements of the earlier theses were very similar. The only difference was the scoring. Some theme students used scoring with numbers while others used '+/-' to score the statements regarding the relationship between elements. The measurement of the second group was even more similar. A difference was that some theme students measured the average of all scores to reach bilateral test-retest, while other students based their results on a legend of the bilateral test-retest scores, presented in Figure 7.3.

Score	Legend bilateral test-retest scores
2	The relationships scores are very similar: the scores are both higher than 1.5 or lower than -1.5; or when the scores are both between 1,5 and 0.5; or when the scores are both between -1.5 and -0.5; or when the scores are both between -0.5 and 0.5
1	The relationships scores are rather similar: one of the relationship scores is higher than 1.5 and the other is between 0.5 and 1.5; or when one of the relationship scores is lower than -1.5 and the other is between- 0.5 and -1.5; or when the difference between the relationship scores is smaller than 0.5, but are not very similar
0	The relationships scores are neither similar nor dissimilar: the relationships scores differ between 0.5 and 1.0; for instance the relationship score on constellation 1 is 0.3 and on constellation 2 is 0.8
-1	The relationships score are rather dissimilar. One of the relationships is higher than 1.5 and the other is lower than 0.5 but higher than -0.5; or one of the relationships is lower than -1.5 and the other is lower than 0.5 but higher than -0.5; or one is 1 and the other -1 and the other '0'
-2	One of the relationships is higher than 1.5 and the other is lower than -0.5; or one of the relationships is lower than -1.5 and the other is higher than 0.5.

Figure 7.3: Legend bilateral test-retest scores (Halters 2009:35)

7.1.3 Similarities in the Precision Measurements

Similar measurements of the precision measurements were called 'accuracy' and 'falsification' as indicated in section 6.1.

The total conversations during the branding constellation were transcribed for evaluation of the statements. All the statements were sorted per stand-in. Some of the statements were not taken into account during the scoring, for example,

- Repeated statements
- Statements of the facilitator and the brander
- Comments of the camera people.

The theme students used a similar format and approach to measure precision. The statements or the applicable statements were scored on a five-point [-2:+2]-scale by all theme students, see figure 7.4

Score	Motivation
2	The statement of a stand-in of a brand element is directly measurable, and therefore can be checked on its truth immediately
1	The statement of a stand-in of a brand element can be operationalised using definitions that were / might be found in literature
0	The statement of a stand-in of a brand element cannot be operationalised using definitions that were / might be found in literature
-1	The statement of a stand-in of a brand element is multi interpretable
-2	The statement of a stand-in of a brand element is in contradiction with another statement made by this stand-in.

Figure 7.4: Precision scores 5 point scale

The average score of all the scores of the statements of a stand-in reported, and analysed in the precision interpretation 10-points scale, see Figure 7.5.

Score	Precision interpretation
1.6 – 2.0	The precision is based on a scale of -2 until 2 is perfect. A lot of the statements are falsifiable.
1.2 – 1.6	The precision is based on a scale of -2 until 2 is very good. Some of the statements are falsifiable.
0.8 – 1.2	The precision is based on a scale of -2 until 2 is good. Most of the statements are falsifiable with use of operationalisation.
0.4 – 0.8	The precision is based on a scale of -2 until 2 is pretty good. Most of the statements are falsifiable after operationalisation.
0.0 – 0.4	The precision is based on a scale of -2 until 2 is moderate. A few of the statements are falsifiable after operationalisation.
-0.4 – 0.0	The precision is based on a scale of -2 until 2 is weak. Less than a few of the statements are falsifiable.
- 0.8 - -0.4	The precision is based on a scale of -2 until 2 is poorly. Most of the statements are multi interpretable.
- 1.2 - -0.8	The precision is based on a scale of -2 until 2 poor. Most of the statements are multi interpretable.
- 1.6 - -1.2	The precision is based on a scale of -2 until 2 bad. A few statements are in contradiction with other statements made by stand-ins.
-2.0 - -1.6	The precision is based on a scale of -2 until 2 very bad. Most of the statements are in contradiction with other statements made by stand-ins.

Figure 7.5: Precision interpretation scores, 10-point scale

7.1.4 Differences in the Precision Measurements

A major difference between the theme students for measuring 'precision' was the use of the name for this measurement. 'Accuracy', 'falsification', and 'precision' were different names for equal measurements. In addition, most students used different methods to calculate the average of the precision scores (average formula, sum divided by count, or calculated by hand). These methods also caused some unseen researcher errors.

Van Geel (2004) and some other students used different colours in their measurements as a classification for the phase or stage the branding constellation was in at that time.

Not all stand-ins were included in the precision measurements by some of the students. For example Siezen (2004) scored only the stand-ins that represented personal elements and he did not include any abstract elements. Figure 7.6 presents an overview of the students' total stand-in present during the constellation and the quantity of stand-ins the theme students analyzed.

	Students	Company	Stand-ins	
			Analysed	Present
1	Van Geel (2004)	Legermuseum - 2004	6	10
2	Siezen (2004)	Stork Fokker - 2003	6	10
3	Gomersbach (2004)	Rabobank - 2004	6	9
4	Mathijssen (2005)	RSM – 2004	10	10
5	Davidse (2005)	DE&SP - 2005	1	12
6	Van Zwienen (2005)	Blooming - 2004	4	10
10	De Velde Harsenhorst (2006)	EODD - 2004	8	9
11	De Heij (2006)	SKBA - 2004	6	11
12	Holwerda	Philips-LG - 2004	6	6
18	Ten Have (2007)	Lipton - 2007	8	8
19	Vertregt (2007)	GTI – 2005	9	9
20	Meijer (2008)	Rabobank - 2005	6	6
22	Halters (2008)	IDS Scheer - 2007	9	10
23	Karel (2008)	Comfort in Living - 2008	9	9
		Average	*7*	*9*

Figure 7.6: Stand-ins present and analysed

The explanation for not including all the stand-ins in the precision measurement according to Mathijssen (2005: 135): "For the falsification score, only the representatives from concrete elements have been taken into account: the statements of these representatives can be verified and thus falsified in real life. Representatives of abstract elements – such as the brand and the product or program - have not been taken into account for this score because these were not verifiable and measurable in real life".

7.2 Standardised Reliability Measurement

In this section the standardised reliability measurements are described. The standardised reliability measurements are based on the findings of the retested of the performed reliability measurements. In subsection 7.2.1 the standardised reliability measurement are described. In subsection 7.2.2 the update of the

bilateral test-retest measurements is described, based on the findings of the standardised bilateral test-retest measurements. In subsection 7.2.3 the standardised precision measurement will be described, based on the finding of the retest of performed precision measurements.

7.2.1 Standardised Reliability Measurement

In this subsection the findings from the retested reliability measurement regarding the standardized reliability measurement are described. The objective of the standardised bilateral test-retest measurement and standardised precision measurement is to equalize measurements and to exclude researcher errors in the measurements.

The measurements of the reliability measurements are based on the transcriptions of the branding constellations, and the results of the measurements are based on the calculations of scores. The transcriptions of the branding constellations were employed to transport the statements of each stand-in to the particular stand-in excel sheets. In the previous measurements many researcher errors were discovered in this selection. In the standardised measurements a function is build-in that sort and transports the statements of each stand-in in a separate sheet. This function transports all the information regarding this statement, the time, phase, and additional information. For this function the student needs to fill out the elements/name/roles of the stand-in in a stand-in sheet. The theme students' fills in the transcription in the standardised excel sheet. In the column 'who' presents the sender of the statement needs to be filled in. The column 'who' is linked with the function to transport the statements to the stand-in sheet (who). If the 'who' is not comparable with the filled in stand-in names a red alert will appear.

The transcription contains the statements of the stand-ins. Every statement contains the name of the sender and receiver, time, phase, and additional information regarding the statement. To prevent missing information, the function transports all the information automatically to the relevant work sheets.

Previous students used different phases. In addition, the phases were most of the time not indicated or explained. The reliability measurements per phase were not comparable due to the use of the different phases by the theme students. The phases of standardised reliability measurements are: the introduction phase, projection phase, intervention phase, vision phase, debriefing, and evaluation based on Jurg (2010). The legend of the phases is presented on every stand-in sheet as a reminder.

The calculation of the previous theme students caused for some unforeseen researcher errors in the measurements. Therefore all calculations of the reliability measurements are build-in. Specific analyses for each reliability measurement are pre-build and are calculated automatically to prevent researcher errors.

The repeated statements by the stand-in were generally not taken into account; even if the repeated statements were made later on. The theme students assumed that all repeated statements were repeated due to a lack of clarity. In

the standardised reliability measurements all statements are included. The repeated statements due to misunderstanding needs to be placed in [..] in the transcription to be left out of the analysis.

The previous reliability measurements within the theme only included the statements of the personal stand-ins leaving the abstract elements out of the analysis. In the standardised reliability measurements all the stand-in are included in the measurements both the personal and the abstract elements. Also the brander and facilitator are included and the reliability of their statements can be measured too.

7.2.2 Standardised Bilateral Test-Retest Measurement

For the development of the branding constellation it is important to compare different studies. To achieve this objective the performed measurements have to be standardised. Therefore one format for the measurement of the bilateral test-retest has been introduced.

In the standardised bilateral test-retest all stand-ins are included in the measurements. The statement of a stand-in that indicates a relationship with another stand-in receives a score based on the scoring table. At the end, the stand-ins that were not mentioned to have a relationship with this stand-in, receive the score “0”; assuming that their relationship is neutral. The build-in function in the standardised bilateral test-retest measurement calculates the relationships scores to the bilateral test-retest scores as Murstein (1963: 149) argues that the scoring system should be as simple as possible.

To simplify the bilateral test-retest measurements, compared to the other reliability measurements, the previous 5-point scale of Halters has been adjusted. This scale caused confusion and errors. The standardised scale prevents calculation researcher errors. The standardised bilateral test-retest overall scoring scale is presented in Figure 7.7.

Score	Motivation
2	The relationships scores are very similar: the relationships scores difference is smaller than 0.8
1	The relationships scores are rather similar: the relationships scores difference is between 0.9 and 1.6
0	The relationships scores are neither similar nor dissimilar: the relationships scores difference is between 1.7 and 2.4
-1	The relationships score are rather dissimilar: the relationships scores difference is between 2.5 and 3.2
-2	The relationships score are dissimilar: the relationships scores difference is higher than 3.3

Figure 7.7: Standardised bilateral test-retest scores

The results of the relationship scores from the first branding constellation and the second branding constellation are compared. If the difference of the average scores is smaller than 0.8 the scoring is “2”. If the difference is bigger than 0.9, but smaller than 1.6 the scoring is “1”, etc.

7.2.3 Standardised Precision Measurement

The analyses in the previous theme precision measurements were only based on the average scores of one stand-in or of the all stand-ins in one constellation. For the ongoing development of the branding constellation development it is needed to analyse more factors of the precision measurements. All stand-ins are now included in the standardised precision measurement. The build-in function in the standardized precision measurement provides the scores of the personal elements and the abstract elements separately. Another build-in function in the standardised precision measurement provides the scores per phase in the constellation (per stand-in).

To prevent incorrect scores to the statements of the stand-ins by the theme students, a build-in function provides for the explanation of the given score next to the given score.

In the standardised precision measurement only the applicable statements of the stand-ins are taken into account. The statement 'yes' is not taken into account due to the not falsifiable state of this statement. An *applicable statement* means a full sentence.

To simplify the precision measurement and to bring them in line with the other scales within the theme, the previous 7-point scale has been adjusted to a 5-point scale. The 7-point scale is presented in Figure 7.5. The standardised 5-point scale is presented in Figure 7.8.

Score	Motivation
1.2 – 2.0	The precision is based on a scale of -2 until 2 is prefect. A lot of the statements are falsifiable
0.4 – 1.2	The precision is based on a scale of -2 until 2 is good. Most of the statements are falsifiable with use of operationalisation (literature research)
-0.4 – 0.4	The precision is based on a scale of -2 until 2 is moderate. A few of the of the statements are falsifiable after operationalisation and/or statements are multi-interpretable
-1.2 - -0.4	The precision is based on a scale of -2 until 2 prefect. Most of the statements are multi interpretable
-2.0 - -1.2	The precision is based on a scale of -2 until 2 very bad. Most of the statements are in contradiction with other statements made by stand-ins.

Figure 7.8: Standardised precision measurement scores

Similarly, the 7-point scale of the variation has also been redesigned to a 5-point scale, see Figure 7.9.

Score	Motivation
4.0 – 3.2	The variation of the precision scores are based on a scale of 0 until 4, a score between 4.0 - 3.2 is very big: the scores are very different
3.2 - 2.4	The variation of the precision scores are based on a scale of 0 until 4, a score between 3.2 - 2.4 is big: the scores are different
2.4 - 1.6	The variation of the precision scores are based on a scale of 0 until 4, a score between 2.4 - 1.6 is moderated: the are scores equally different as they are similar
1.6 – 0.8	The variation of the precision scores are based on a scale of 0 until 4, a score between 1.6 - 0.8 is small: the scores are equally
0.8 – 0.0	The variation of the precision scores are based on a scale of 0 until 4, a score between 0.8 - 0.0 is very small: the scores are equally different

Figure 7.9: Standardised variation precision measurement scores

To increase the objective of the precision and bilateral test-retest measurements tool an introduction and manual has been added to the file for the theme students as well for the second reviewer or second researcher. The introduction and the manual contain buttons that calculated and buttons that find the correct information/part of the file. As stated earlier, all the calculations and formulas were fixed in the measurement sheets; to encourage the development of the file an extra sheet has been added explaining the used formulas and calculations.

The bilateral test-retest measurement and the precision measurement were standardised after some experience with the previous measurements. The success in the use of the standardised reliability measurements is in the added introduction, manual, and formula sheet.

8 RESCORING RELIABILITY MEASUREMENTS

This chapter answers the fifth and final research question: *"What are the similarities and differences between the findings on the most relevant reliability measurements of the previous students and the standardised measurements?"* In section 8.1 the bilateral test-retest measurements of the theme students are compared to the standardised bilateral test-retest measurements. In section 8.2 the precision measurements of the theme students are compared to the standardised precision measurements.

Detailed information of the retest of the reliability measurements are described in appendix C for the bilateral test-retest measurements and in appendix D for the precision measurements.

8.1 Bilateral Test-Retest Measurement Findings

Comparing the bilateral test-retest measurement of the theme students to the standardised measurement gain the following scores, see Figure 8.1. The theme measurements versus the standardised bilateral test-retest measurements are given.

	Students	Company	Theme students scores	Standardised bilateral test-retest scores
10	De Velde Harsenhorst (2006)	EODD – 2004	1.0	1.6
11	De Heij (2006)	SKBA – 2004	0.0	0.2
15	Schuurman (2006)	Alex – 2003	0.9	1.6
17	Blootens (2006)	Hooghoudt-2003-05	0.6	1.9
18	Ten Have (2007)	Lipton – 2007	0.5	1.6
22	Halters (2009)	IDS Scheer-2007	0.3	1.2
23	Karel (2009)	Comfort in Living – 2008	0.7	1.6
		Total average	*0.6*	*1.4*
		Average (excl. De Heij)	*0.7*	*1.6*
		Variation (excl. De Heij)	*0.7*	*0.7*

Figure 8.1: Theme student versus standardised bilateral test-retest measurements

The scores on the test-retest reliability are higher using the standardised version. While the theme students concluded that branding constellations are moderately bilateral test-retest reliable, the standardised version concludes that they are very bilateral test-retest reliable. The most prominent change is found in the case study of Blootens (2006), where the conclusion changed from moderately reliable (0.6) to very reliable (+1.9); including even the highest score. The lowest score is found in the case study by Halters (+0.3, +1.2 respectively). The positions in the scoring overview have, however, remained similar: Halters (2009) still scores lower than average and the score of Karel (2009) is still just above the average of all scores.

The bilateral test-retest measurements of De Heij were impossible to repeat due to missing information (see Appendix C.2). Thus, the scores of De Heij are rather incomparable. Thus, the average was also counted without taken the study De Hey into account.

The scores of the other theme students are not comparable due to their different measurements. The variation of the scores of the theme students is large: 1.6 (excluding the score of De Heij). The variation of the scores of standardised bilateral test-retest measurements is clearly lower 0.7 (excluding the scores of De Heij).

Evaluating the two 'best of the class' bilateral test-retest measurement Halters (2009) and Karel (2009) the measurements of these two students, it must be noted that they had fewer adjustments than the other students in their bilateral test-retest measurements.

The standardised bilateral test-retest shows a standardised method for measuring the bilateral test-retest, which is suitable for all the bilateral test-retests in the branding constellation. The given scores by the theme students are calculated automatically. This excludes any calculation researcher errors in the scores. The relationships can be evaluated by stand-in and in total. Due to the use of one format and the build-in calculation the scores of branding constellations are comparable. The standardised bilateral test-retest measurement is convenient and repeatable for future theme students.

8.2 Precision Measurement Findings

Figure 8.2 presents the theme measurements, their corrected scores for researcher errors, and the standardised precision measurements.

The column *students presented* show the scores presented in the theses of the theme students. The column *students corrected* show the scores of theme students, after correcting their calculation errors and including all stand-ins. The average scores of the included stand-ins were scored with "0". Figure 7.7 presented an overview of the total stand-in present and analysed. Appendix D, shows a detailed measurement per theme student.

	Students	Company	Scores		
			Students presented	Students corrected	Standardised
1	Van Geel (2004)	Legermuseum - 2004	0.5	0.5	0.9
2	Siezen (2004)	Stork Fokker - 2003	0.3	0.2	0.4
3	Gomersbach (2004)	Rabobank - 2004	0.2	0.01	0.4
4	Mathijssen (2005)	RSM - 2004	0.4	0.6	0.6
5	Davidse (2005)	DE&SP - 2005	0.3	0.3	0.5
6	Van Zwienen (2005)	Blooming - 2004	0.6	0.2	0.6
10	De Velde Harsenhorst (2006)	EODD - 2004	0.5	0.5	0.5
11	De Heij (2006)	SKBA - 2004	0.5	0.3	0.4
12	Holwerda (2006)	Philips-LG - 2004	0.4	0.4	0.3
18	Ten Have (2007)	Lipton - 2007	0.8	0.8	0.7
19	Vertregt (2007)	GTI - 2005	0.7	0.7	0.5
20	Meijer (2008)	Rabobank - 2005	0.4	0.4	0.5
22	Halters (2008)	IDS Scheer - 2007	0.8	0.8	0.6
23	Karel (2008)	Comfort in Living - 2008	0.7	0.7	0.8
		Average	*0.52*	*0.48*	*0.57*

Figure 8.2: Theme student scores vs standardised precision measurement

The differences between the original scores and the standardised ones are limited. The more 'extreme' scores have become more average, except for the score of Holwerda (2006). Evaluating the 'best of the class' precision measurement Karel (2009), the difference with the standard is limited due to a correction of only a few scores.

In the stacking comparables of the theme students Karel (2009) is above average, in the standardised precision measurement Karel (2009) is still above the average. Siezen (2004), Van Zwienen (2006), and Holwerda (2006) were not included in the stacking comparables of the latter theme students.

9. DISCUSSION

The first section 9.1 presents the conclusion and section 9.2 the implication of the research findings for the branding constellation theme. Section 9.2 examines the findings on the internal and external validity, as well as on their internal and external reliability. Section 9.3 presents the recommendations. Section 9.4 closes this chapter with a personal reflection on this thesis regarding the branding constellation theme.

9.1 Conclusion

In this section the conclusion and answers to research questions will be given. In subsection 9.1.1 the reliability measurements of projections were given. In subsection 9.1.2 the reliability measurements derived from the branding constellation theme definition list findings. In subsection 9.1.3 the most important reliability measurements findings. In subsection 9.1.4 the standardised reliability measurements, and in subsection 9.1.5 the rescoring of the reliability measurement.

9.1.1 Reliability measurements of the projection techniques

The reliability measurements employed in the scientific studies on projections were the test-retest, inter rater, and triangulation. These measurements were the most founded measurement of the reliability on projection techniques in the scientific studies. The conclusion of Jensen (1959) is that in a test-retest the time interval is an important subject. Jensen (1959) argues that longer time interval leads to a lower reliability score. Based on the test result of a maximum split-half test (form of test-retest) Wagner (1985) stated that projection tests are probably more internally reliable than reported before. Catterall et al. (2000: 252), state that there is more research needed, users agree that designs should be kept simple, avoiding too much detail or stylisation. Jensen agrees on this part by arguing that the most satisfactory solution, to the high standard error of a projection technique, is to eliminate or reduce as many of the sources of error variance as possible by making the administration, scoring, and interpretation of the projection test more standardized and objective. However, the reliability measurements were employed without even indicating what the ‘r’ stood for and how it was calculated.

9.1.2 Reliability measurements derived from branding constellation theme definition list

The branding constellation theme list contains numerous different definitions of *reliability*. These definitions were classified in five different interpretation clusters based on the interpretations of the definitions: different scorers come to similar conclusions; similar conclusions at different times; different measurements result in similar conclusions; the truth of the conclusions; and the precision of the conclusions. Of the total 58 definitions, the most refer to ‘different times’ (or occasion) and ‘same method’ regarding reliability. The interpretation clusters ‘different time’ and ‘same method’ are in combination similar to the test-retest reliability concept. The definition of test-retest reliability according to Peter (1977: 395) is “Identical set of measures is given to the same sample on two

separate occasions". To increase the reliability of the data and study the truth measurement and precision measurement need to be included.

9.1.3 Most important reliability measurements

The most important measurements of the reliability measurements in the branding constellation theme were the bilateral test-retest measurement and the precision measurement as they were the most often performed measurements.

9.1.4 Standardised reliability measurements

The reliability measurements were analysed on the similarities and differences in the measurements. The similarities are the measurement approaches and the use of the scales. The differences between the reliability measurements are caused by the quantity of stand-ins included, quantity of statements included, and researcher errors in the measurements. The most important measurements can be standardised by developing an Excel tool for the measurements. For the standardised bilateral test-retest measurement and the standardised precision measurement these standardised measurements' tools have been developed. The standardised reliability measurements include all stand-ins and all applicable statements. The functions in the standardised reliability measurements exclude any calculations errors in the measurements.

9.1.5 Rescoring of the reliability measurements

The scores on bilateral test-retest reliability of the branding constellations are higher based on the standardised version. While the theme students concluded that branding constellations are moderately bilateral test-retest reliable (+0.6 on a -2:+2-scale), the standardised version concludes that they are very bilateral test-retest reliable (+1.6 on a -2:+2-scale). The scores on the precision are more similar: the average of the precision measurement of the theme students was (+0.52 on a -2:+2-scale)and the average of the standardised precision measurement was (+0.57 on a -2:+2-scale).

9.2 Implication

The objective of this thesis was to study the most important reliability measurements of the branding constellation theme. This thesis presents overviews of all the reliability measurements performed by the theme students. The most important reliability measurements are the bilateral test-retest measurement and the precision measurement. In the previous performed bilateral test-retest measurements and the precision measurement many measurement differences were found, which made these scores not comparable. The previous bilateral test-retest and precision measurements were re-measured based on the standardised reliability measurement. Therefore these scores are now comparable for follow-up theme students.

9.3 Validity and Reliability of the Study

Subsection 9.2.1 describes the internal validity and subsection 9.2.2 the external validity of the study. Subsection 9.2.3 discusses the internal reliability and subsection 9.2.4 the external reliability of the study.

9.3.1 Internal Validity

According to Yin (1975: 376), “internal validity raises the question of whether a study’s research design is adequate to support the study’s conclusion”.

The analysis from the definition list was based on the keywords in the definitions. Only the keywords in the definition were taken into account, not an interpretation of the definition. Writing errors in the definition list may have led to missing or misplacing definitions, searching on a word. All the missing definitions were taken in account by using the selection ‘Blank’.

The search in the scientific research with the search criteria ‘projections’ or ‘projection techniques’ was not successful. Therefore the focus was on the search criterion ‘projective techniques’. The word ‘projection’ is too commonly used to be employed to find journal articles on projection techniques. Due too little experience on the subject of reliability, the research on literature was very difficult. The articles were clearly written for experts, which made the analysis an enormous struggle.

In the analysis on the reliability definitions the order of the keywords was not taken into account. Therefore, Figure 5.2 shows some authors double. If an author referred to ‘different times’ and ‘same method’ the author will also be found in the figure in ‘same method’ and ‘different times’.

The included phases in the standardised precision measurements are based on the phases of the PhD dissertations of Jurg (2010: 23) (appendix A). The previous theme students used a variance of phase combinations. This allows for many new measurements and interpretations that were not employed in this thesis.

Unfortunately, during the finishing of this thesis none of the theme students was engaged with the measurements of the bilateral test-retest and the precision measurements. Therefore the feedback received was only from rather uninvolved theme students.

9.3.2 External Validity

“External validity raises the question of whether a study’s conclusion can be generalized to other situations” (Yin, 1975: 376).

In this thesis the objective was to evaluate and standardised the reliability measurements employed within the branding constellation theme. This does not allow for an analysis on external validity.

9.3.3 Internal Reliability

"Internal reliability is about whether the researcher would generate the same outcomes if he would do the same research again" (Yin, 1975: 375). If I would do the same study again, I would arrive at similar standardised reliability measurements. I am not aware of any other errors or biases in the calculations.

9.3.4 External Reliability

External reliability is about whether or not other researchers would generate the same outcomes using the same methodology, setting, and time (Yin 1975: 375).

The external reliability was secured and adapted by the checking and testing of three follow-up theme students and the theme coordinator. The feedback of the theme students can be found on the CD-ROM in the directory *Support files/Feedback*.

9.3 Recommendations

For the development of the branding constellation theme more research is needed on several subjects. To stimulate the research and the development an online database is recommend. In the last years, 24 theme students collected a enormous quantity of data and information, like articles, text, video's, studies and theses. The quantity data asks for an online database so the data is better attainable for future development of the branding constellation theme.

Another recommendation for future research regarding branding constellation theme is to research the usefulness of the inter rater reliability and the triangular in a first person case study. In another follow-up thesis regarding branding constellation the subject could be the validity forms used in the previous theses.

The standardised reliability measurement for the measuring of the reliability in branding constellation is functional in the program Excel. The tool is developed for theme students with minimal experience with Excel. Adjustments should only be made by excel-experienced theme students. In a few years the standardised reliability measurement will need some adjustments to keep up with the development of the branding constellation. Or in a few years another program is more suitable for measuring the bilateral test-retest or precision measurements, for example Access. The advantage of Access is the reliability; the disadvantage is the difficulty to use and adjust.

A final recommendation includes the precision measurement. Rather than limiting the focus on the average, the future measurement may focus on the '+2' score-statements and study their truth as is also done in the introspective validity measurements.

9.6 Reflection

To obtain better insight in the branding constellation, I first analysed the branding constellation DVDs and read the theses from Karel (2009), Halters (2009), Van Zwienen (2006), and later some others. The insights gained on

measuring the bilateral test-retest reliability and precision evolved by retesting the measurements of other theme students.

The retesting of the theme students measurement was a learnable experience. Most students did not use the maximum of Excel. It is a pity to see the effort the several student had to accomplish get their results. They calculated all scores by hand or retyped data in to another file. The standardised reliability measurements developed in time. It was for me pleasurable to see the result of an effective formula, although some took me hours/days to develop and install.

A hassle during writing this thesis was my English. I needed some help to accomplish a readable thesis. In past months my English improved, but it is still not of good quality.

One of the best parts I learned from the thesis is communication. At first I hesitated to asked questions. As time went by, the question started to become a obstacle. This experience was a good example for me to learn, due to the time pressure by waiting to ask my question I got myself in some trouble. I learned to pay more attention formulating the right question and ask it then as soon as possible.

The Open Universiteit estimates 600 hours for producing a final thesis. While working on this thesis I kept an overview of how much time I actually spent on it. Due to the time pressure I experienced some stressful moments in the last few months. Thanks to KLM and Martinair it was possible to claim all my holidays in a sort period of time. The best days for me were those with a good start in the morning, some easier subjects in the afternoon and a good finish in the evening. The evenings, after work, were not as appalling as I expected.

During the period of the water damage in our house too less attention was paid to the thesis. It took me 2 weeks to realize I could not combine writing a thesis and the restoration the apartment. So, the apartment was made liveable, the restoration will soon follow.

Figure 9.3 shows the planned versus the actual time spent on the accomplishment of this thesis.

Task	Planned hours	Actual hours
Introduction, preparation	20	30
Literature study	50	45
Analysis of definition list	40	25
Preparation retesting	15	10
Retesting and analyses	75	130
Building measurement tool (Excel)	50	60
Writing thesis and appendices	300	410
Finalisation	50	30
Total	*600*	*740*

Figure 9.3: Planned versus actual time spent

I am thankful for the opportunity to gain a lot of knowledge of the branding constellation, research, and writing it's was a lifetime experience. I feel I have succeeded in taking the branding constellation theme one step further on the road to reliable reliability measurements and also to other usefulness measurements.

LITERATURE OVERVIEW

Articles on projection techniques

- Donoghue (2000), S., Projective Techniques in Consumer Research, *Journal of Family Ecology and Consumer Sciences,* 2000, Vol 28: 47-53
- Jensen (1959), A.R., The reliability of projective techniques: review of the literature, *Acta Psychologica,* 1959, Vol 16: 108-136
- Lilienfield (2000), S.O., Wood, J.M., Garb. H.N., The Scientific Status of Projective Techniques, *Psychological Science in the Public Interest,* 2000, Vol 1, Iss 2: 27-66
- Lundy (1985), A., The Reliability of the Thematic Apperception Test, *Journal of Personality Assessment,* 1985, Vol 49, Iss 2: 141-145
- Perry (1998), J.C., Lanni, F.F., Observer –Rater Measures of Defense Mechanisms, *Journal of Personality,* 1998, Vol 66, Iss 6: 993- 1024
- Schultheiss (2008), O.C., and Leining, S.H., Schad, D., The Reliability of the Picture Story Exercise Measure of Implicit Motives: Estimates of Internal Consistency, Retest Reliability, and Ipsative Stability, *Journal of Research in Personality,* 2008. Vol 42: 1560-1571
- Wagner (1985), E.E., Alexander, R.A., Roos. G., Prospero, M.K., Maximizing Split-Half reliability Estimates for Projective Techniques, *Journal of Personality Assessment,* 1985, Vol 49, Iss 6:579-581

Background articles of the branding constellation theme

- Anderson (1978), J.C., The Validity of Haire's Shopping List Projective Technique, *Journal of Marketing*, 1978, Vol 15: 644-649
- Catterall (2000), M., and Ibbotson, P., Using Projective Techniques in Education Research, *British Educational Research Journal,* 2000, Vol 26, Iss 2: 245-256
- Blatt (1975), The Validity of Projective Techniques and their Research and Clinical Contribution, *Journal of Personality Assessment,* 1975, Vol 39, Iss 4: 327-343
- Boddy (2005), C, Projective techniques in market research: valueless subjectivity or insightful reality?, *International Journal of Market Research*, 2005, Vol 47, Iss 3: 239-254
- Bradbury (2000), H., Bergmann, B.M., Relationality in Organizational Research: Exploring the Space Between, *Organization Science*, 2000, Vol 11, Iss 5: 551-564
- Doherty (2010), S., and Nelson, R., Using Projective Technique to Tap into Consumers' Feelings, Preceptions and Attitudes…Getting an Honest Opinion, *International Journal of Consumer Studies*, Vol 34: 400-404
- Fassenbender (1997), P., Implications for Use of Projective Techniques in Organisational Consulting, *Perceptual and Motor Skills,* 1997, Vol 84: 171-175
- Greenberg (1959), A., Pictorial Stereotypes in a Projective Test, *Journal of Marketing,* 1959:72-74
- Haire (1950), M., Projective Techniques in Marketing Research, *Journal of Marketing,* 1950, Vol 14, Iss 5: 649-656
- Hauser (1979), J.R., and Koppelman, F.S., Alternative Perceptual Mapping Techniques: Relative Accuracy and Usefulness, *Journal of Marketing Research,* 1979, Vol 16: 495-506

- Jurg (2008), W.E., *Constellations to identify branding problems*, 2007
- Jurg (2007), W.E., Constellations to identify Branding Problems, The Knowing Field - International Constellations Journal, London, Vol 9, Iss 1: 61-67
- Jurg (2006), W.E. and Bloemer, M., *Improving predictability using branding constellations,* Submitted paper for the 36th EMAC conference, Reykjavik University, 2006
- Jurg (2006), W.E., Bloemer, J., Doorewaard, H., Validity of Branding Constellations, 35th EMAC paper
- Jurg (2003), W.E., Bloemer, J., Doorewaard, H., Het Merk en zijn Marketeer op de Sofa, *Tijdschrift voor Marketing*, Vol 37, Iss 6: 48-50
- Levy (1994), S., Interpreting Consumer Mythology, *Marketing Management,* 1994, Vol 2, Iss 4: 2-9
- Stein (1998), M., Projective Identification in Management Education, *Journal of Managerial Psychology*, 1998, Vol, 13, Iss 8: 558-566
- Steinman, R.B., Projective Techniques in Consumer Research, *International Bulletin of Business Administration,* 2009, Iss 5: 37-45
- Thompson, B.L., and Green, S.B., and Yang, Y., Assessment of the Maximal Split-Half Coefficient to Estimate Reliability, *Educational and Psychological Measurement SAGE,* 2010, Vol 70, Iss 2: 232-251
- Thorndike (1985), R., Testing the test, *Journal of counselling and development*, Vol 63: 528-530
- Yadav (1985), S.B., Korukonda, A. 'Management of Type III Error in Problem Identification'*, Interfaces,* 1985, Vol 55, Iss 7: 55-61
- Yeager, (2003), J., Innovative Motivational Profiling: Comparing Marketing Projective Technique Versus Linguistic Forensic Techniques, *The Qualitative Report,* 2003, Vol 8, Iss 1: 1-24
- Yin (1975), R.K., Helad, K.A., Using the case survey method to analyze policy studies, *Administrative Science Quarterly,* 1975, Vol 20: 371-381
- Yoell (1974), W.A., The fallacy of projective techniques, *Journal of advertising,*1975, Vol 3, Iss 1: 33-36

Background books on projection techniques

- Harrower (1960), et al., *Creative Variations in the Projective Techniques*, C.C. Thomas, Springfield, 1960
- McDonald (1999), R.P., *Test Theory*, *A Unified Treatment,* University of Illinois at Urbana-Champaign, Lawrence Erlbaum Associates Publishers London, 1999
- Murstein (1963), B.I., *Theory and Research in Projective Techniques (Emphasizing the TAT),* John Wiley and Sons Inc, New York, 1963
- Kotler (2000), P., Marketing Management, the millennium edition, Prentice Hall International, Inc, Upper Saddle River, New Jersey, 2000
- Litwin (1995), M.S., *How to Measure Survey Reliability and Validity*, SAGE publications, 1995
- Riel (2001), van, C.B.M., Corporate communication, managen van reputatie, Kluwer, Alphen aan den Rijn, 2001

- Thorndike (2010), R.M. and Thorndike-Christ, T, *Measurement and Evaluation in Psychology and Education*, Pearson Education, Boston, 2010, 8th edition
- Vries (2005), de, W. Jr, Helsdingen, Van, P.J.C., Dienstenmarketingmanagement, Wolters-Noordhoff Groningen, Houten, 4th edition, 2005

PhD Dissertations

- Jurg (2010), W.E. The perceived usefulness of branding constellations, Open Universiteit the Netherlands, version May 5 2010

Master theses in the branding constellation theme

1. Van Geel, van (2004), The usefulness of the branding constellation - A Museum Branding Case
2. Siezen (2004), The usefulness of the branding constellation - A Brand Endorsement Case
3. Gomersbach (2004), The Usefulness of the Branding Constellations - An Introspective Account Management Case
4. Mathijssen (2005), The Usefulness of the Branding Constellation - An International Marketing Advise Case
5. Davidse (2005), The Usefulness of the Branding Constellations - A Web Branding Case
6. Zwienen, van (2005), The Usefulness of the Branding Constellation - Two Brand Name Cases
7. Meer, van (2005), The Usefulness of the Branding Constellation - The Two 2004 Expert Seminar Cases
8. Mechelen, van (2005), The Usefulness of the Branding Constellation - A Brand Extension Case Study
9. Simons (2005), The Usefulness of the Branding Constellation - A Brand Name Logo Case
10. Velde Harsenhorst, de (2006), The Usefulness of the branding constellation - A Merger Case Study
11. Heij, de (2006), The Usefulness of the Branding Constellations - A Second Brand Endorsement Case
12. Holwerda (2006), The Usefulness of the Branding Constellations - A Competency Case Study
13. Stroo (2006), Appendices A brand Distribution Case Study
14. Harrewijn (2006), The Usefulness of What is Lying Behind - An Advertisement Campaign Branding Constellation Study
15. Schuurman (2006), The Usefulness of the Branding Constellation - Making the Invisible, Visible
16. Labots (2006), The Usefulness of the Branding Constellation - Beyond the Tip of the Iceberg
17. Blootens (2007), A Threefold Branding Constellation Study - Beyond Fields
18. Ten Have (2007), The Reliability of Branding Constellations

19. Vertregt (2007), The Usefulness of the Branding Constellation - A Third Brand Endorsement Case
20. Meijer (2008), The Usefulness of the Branding Constellation - Local Faces for Local Customers
21. Claus (2008), Birds of a Feather Flock Together
22. Halters (2009), A Branding Constellation Reliability Study - Trust is a Belief in Reliability
23. Karel (2009), A Branding Constellation Reliability Case Study - Identifying a New Opportunity

Premaster reports

- Schapendonk, J., and Arts, C., Groot Worden Door Klein te Blijven, Open Universiteit the Netherlands, 2010

FIGURES OVERVIEW

APPENDIX OVERVIEW

CD-ROM FILE OVERVIEW

Thesis van Reij 2010
- Thesis van Reij 2010
- Thesis van Reij 2010 – appendices

Thesis van Reij 2010 (support files)
- Excel - standardised bilateral test-retest measurement
- Excel - standardised precision measurement

- Excel measurement - bilateral test-retest
- Excel measurement - definition list
- Excel measurement - precision

- Feedback

- Figure - pictures
- Figure - tables

- Literature - articles (incl background)
- Literature - Jurg (PhD dissertations & papers)
- Literature - these branding constellation
- Literature - these premaster

- PowerPoint – presentation thesis van Reij 2010

Thesis van Reij 2010 (zip file)

END NOTES

[1] Open Universiteit
Intranet, 2 augustus 2010
http://studienet.ou.nl/webapps/portal/frameset.jsp?tab=community&url=%2Fbin%2Fcommon%2Fcourse.pl%3Fcourse_id%3D_3319_1
Het thema *branding* focust op de identificatie van merk- en interventieproblemen. De probleemidentificatie-technieken die daarbij centraal staan zijn de merkopstelling en de interventie-opstelling, de toepassing van de systeemopstelling op merkproblemen respectievelijk interventieproblemen. Het 'gouden' artikel over de merkopstelling vind je in de bijlage. Dit artikel is gebaseerd op het afstudeeronderzoek van Ruud Simons. Voor verdere informatie over de merkopstelling kun je terecht op www.merkopstelling.nl. Daarnaast zijn ook andere onderzoeken naar branding mogelijk in onderling overleg.

[2] Open Universiteit
Intranet, 2 augustus 2010
http://studienet.ou.nl/webapps/portal/frameset.jsp?tab=community&url=%2Fbin%2Fcommon%2Fcourse.pl%3Fcourse_id%3D_3319_1
Het thema *branding* focust op de identificatie van merk- en interventieproblemen. De probleemidentificatie-technieken die daarbij centraal staan zijn de merkopstelling en de interventie-opstelling, de toepassing van de systeemopstelling op merkproblemen respectievelijk interventieproblemen. Het 'gouden' artikel over de merkopstelling vind je in de bijlage. Dit artikel is gebaseerd op het afstudeeronderzoek van Ruud Simons. Voor verdere informatie over de merkopstelling kun je terecht op www.merkopstelling.nl. Daarnaast zijn ook andere onderzoeken naar branding mogelijk in onderling overleg.

[3] ebook.com
Website, 15 July 2010
http://e-bookmarket.com/index.php?keyword=projective+technique+precision&page=results&filetype=pdf

[4] Wikipedia.org
Website, 20 July 2010
http://en.wikipedia.org/wiki/Thematic_Apperception_Test