E-commerce web site evaluation

Developing a framework and method for the systematic evaluation of e-commerce web sites and using Correspondence Analysis to represent the results graphically per industry

Rian van der Merwe

Department of Industrial Engineering, University of Stellenbosch

Student number: 12530859

rianvdm@hotmail.com

Thesis presented in partial fulfilment of the requirements for the degree of

M.Sc.Eng in Industrial Engineering

at the University of Stellenbosch

Study leaders

Mr James Bekker, Department of Industrial Engineering, University of Stellenbosch

Prof. Leyland Pitt, School of Marketing, Curtin University of Technology, Perth, Australia

December 2001

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it to any other university for the purpose of attaining an academic qualification.

Signature Date

What does a life of total dedication to truth mean? It means, first of all, a life of continuous and never-ending stringent self-examination. We know the world only through our relationship to it. Therefore, to know the world, we must not only examine it but we must simultaneously examine the examiner.

M. Scott Peck, 'The Road Less Travelled'

Acknowledgements

Many people made the successful completion of this thesis possible. I would like to thank the following people sincerely for their help and support:

Mr James Bekker, who provided me with just the right mix of freedom and 'the right word at the right time' to create an environment where I could learn as fast and as much as possible.

Prof. Leyland Pitt, who envisioned the thesis and also gave me the opportunity of a lifetime by inviting me to study at his university in Perth, Australia. I am in great debt to you. I will not forget our deal about the drink I owe you when I arrive in Perth!

All the staff at the Department of Industrial Engineering, University of Stellenbosch, who taught me so much and who supported me throughout the 5½ years I studied there.

My father, still my first line of defence in all my academic efforts, who had to read many first tries and answer questions all the way. Thank you for never getting tired.

The rest of my family, who were always near with a supporting word—especially during the gruelling final stages.

My countless friends, who were always interested in what I'm doing and in whose company I immediately relax and feel at home.

Abstract

The corporate web site is essential to companies who use the Internet for e-commerce purposes. For these companies, the web site is the platform used to communicate with customers and facilitate business transactions. Internet companies will not be able to do business successfully with an ineffective web site, because this implies that the only contact point that the company has with customers is not functioning properly. It is, however, extremely difficult to identify what an effective e-commerce web site constitutes of. A great need therefore exists for a comprehensive and accurate method to evaluate the performance of the web sites of Internet companies, not only individually but also in comparison with the web sites of other companies in the same industry. Managers of Internet companies would certainly like to know how their web sites perform, what they can do to increase their performance, and which web sites in their industry can be used as a benchmark in certain areas. This thesis aims to address these needs by fulfilling three objectives:

- To develop a framework and criteria for the comprehensive evaluation of e-commerce web sites.
- To use this framework and sound statistical reasoning to develop a method that can be used to evaluate e-commerce web sites quantitively, and represent the results graphically per industry.
- To implement this method by developing computer software that enables users to evaluate web sites and plot the results.

To accomplish these objectives, the following methodology was followed:

- Review the research done in the field of web site evaluation for both general and e-commerce web sites.
- Review the research on different techniques in the field of Multidimensional Scaling, and identify an appropriate technique for developing two-dimensional plots of web site evaluation data.
- Expand the web site evaluation research and develop a framework and objective criteria for the evaluation of e-commerce web sites, based on solid business principles.
- Develop a method to gather web site evaluation data that is grouped within industries, and to represent the results graphically using an appropriate Multidimensional Scaling technique.
- Implement the method by developing computer software to automate the process.

This document describes the course of the methodology in detail. It reports on the e-commerce web site evaluation framework that was developed; Correspondence Analysis as the Multidimensional Scaling technique used to analyse the evaluation data; the development of the e-commerce web site evaluation method; and the software that was developed in Microsoft Visual Basic to implement the evaluation method.

All three objectives were fulfilled in this thesis, in spite of some concerns that are also discussed. The evaluation framework and accompanying software can be used to evaluate all aspects of e-commerce web sites, and the output can be used to draw meaningful conclusions about how these sites can be improved.

Opsomming

Die korporatiewe webwerf is onontbeerlik vir maatskappye wat die Internet vir e-handel doeleindes gebruik. Die webwerf is vir hierdie maatskappye die platform wat gebruik word om met kliënte te kommunikeer en om saketransaksies te fasiliteer. Internetmaatskappye sal nie in staat wees om suksesvol sake te doen as hulle webwerwe oneffektief is nie, omdat dit sal impliseer dat die enigste raakpunt wat die maatskappy met kliënte het, nie behoorlik funksioneer nie. Tog is dit moeilik om te identifiseer waaruit 'n effektiewe e-handel webwerf bestaan. Daar is dus 'n groot behoefte aan 'n omvattende en akkurate metode waarvolgens die werkverrigting van die webwerwe van Internetmaatskappye geëvalueer kan word, nie net individueel nie, maar ook in vergelyking met die webwerwe van ander maatskappye in dieselfde industrie. Bestuurders van Internetmaatskappye sal beslis wil weet hoe goed hulle webwerwe funksioneer, wat hulle kan doen om die werkverrigting van webwerwe te verbeter, en watter webwerwe in hulle industrie as uitstaande voorbeelde in sekere areas kan dien. Hierdie tesis spreek die bogenoemde behoeftes aan deur drie doelstellings uit te voer:

- Om 'n raamwerk en kriteria vir die omvattende evaluasie van e-handel webwerwe te ontwikkel.
- Om hierdie raamwerk en gegronde statistiese beredenering te gebruik ten einde 'n metode te ontwikkel wat gebruik kan word om e-handel webwerwe kwantitatief te evalueer, en om die resultate grafies per industrie uit te beeld.
- Om hierdie metode te implementeer deur rekenaarprogrammatuur te ontwikkel wat gebruikers in staat stel om webwerwe te evalueer en die resultate te plot.

Die volgende metodologie is gevolg om hierdie doelstellings te laat slaag:

- Bestudeer die navorsing gedoen in die veld van webwerf evaluasie van sowel algemene as e-handel webwerwe.
- Bestudeer die navorsing oor verskillende tegnieke in die veld van Multidimensionele Gradering (*Multidimensional Scaling*), en identifiseer 'n toepaslike tegniek vir die ontwikkeling van tweedimensionele grafiese voorstellings van webwerf evaluasiedata.
- Brei die webwerf evaluasienavorsing uit en ontwikkel 'n raamwerk en objektiewe kriteria vir die evaluering van e-handel webwerwe, gebaseer op stewige sakebeginsels.
- Ontwikkel 'n metode om webwerf evaluasiedata te versamel wat in industrieë gegroepeer is, en stel hierdie resultate grafies voor deur die gebruik van 'n toepaslike Multidimensionele Graderingstegniek.
- Implementeer die metode deur die ontwikkeling van rekenaarprogrammatuur om die proses te outomatiseer.

Hierdie dokument beskryf die verloop van die metodologie in detail. Dit lewer verslag oor die e-handel webwerf evaluasieraamwerk wat ontwikkel is; Assosiasie-analise (*Correspondence Analysis*) as die Multidimensionele Graderingstegniek wat gebruik is om die evaluasie-data te analiseer; die ontwikkeling van die e-handel webwerf evaluasiemetode; en die programmatuur wat ontwikkel is in Microsoft Visual Basic om die evaluasiemetode te implementeer.

Al drie doelstellings is in hierdie tesis bereik, ten spyte van sommige probleme—wat ook bespreek word. Die evaluasieraamwerk en meegaande programmatuur kan gebruik word om alle aspekte van e-handel webwerwe te evalueer, en die resultate kan gebruik word om betekenisvolle gevolgtrekkings te maak oor die wyse waarop hierdie webwerwe verbeter kan word.

Table of contents

List of tables

List of figures

Glossary of terms

This list contains terms that are used in the thesis. The glossary aims to define these terms in the context that they are used in this thesis as clearly as possible.

Banner ads. Small, rectangular advertisements on web sites, which sometimes include simple animations. It is possible to click on these advertisements to get more information about the product/service that is being advertised.

Business-to-Consumer. Refers to e-commerce where a business entity provides products or services to individual customers and not to other business entities or organisations.

Cache. High-speed memory that gives a computer rapid access to data that are used often.

Correspondence Analysis. A Multidimensional Scaling technique that displays the rows and columns of a data matrix as points in dual low-dimensional vector spaces (paraphrased from [Greenacre, 1984:54]).

Customer Relationship Management (CRM). Methodologies, software and usually Internet capabilities that help companies to manage customer relationships in an organized way. It provides seamless integration of every area of business that is relevant to the customer.

Data entry point. Data entry points refer to the devices through which data enter the e-commerce information system. These entry points include the Internet, call centres and CRM software.

E-commerce. A collective term describing the complete process of buying and selling products and service using Internet technologies and the World Wide Web.

E-commerce web site evaluation framework. The framework developed in this thesis to serve as a basis for the evaluation of e-commerce web sites using specific criteria. The framework consists of five criteria categories, with several criteria groups and individual criteria in each category.

E-commerce web site evaluation method. The complete process of e-commerce web site evaluation developed in this thesis—from the gathering of evaluation data to the interpretation of output results.

Front-end vs. Back-end. *Front-end* refers to the part of a web site that the user sees on his computer screen, and *back-end* refers to the unseen operations ensuring the smooth functioning of the web site and the business.

Hard coded. A term used in computer programming to refer to cases where data that could be stored in variables using some form of input to determine their values, are rather stored as constants using explicit program code.

Hits. The amount of visits a web site receives during a certain period of time.

Industry. A term describing a specific type of business, for example the banking industry or the computer industry.

Internet. The complete communication infrastructure that links computer networks worldwide, irrespective of the communication protocol.

Internet company. A company using the Internet for the purpose of e-commerce. Not necessarily a company using the Internet as its only means of doing business.

IP address. IP stands for Internet Protocol, and refers to the procedures used to send data like web site content and e-mail to the correct computer requesting the data. The IP address refers to the unique address that identifies a computer when data is transferred to and from the computer over the Internet.

Multidimensional Scaling. A group of statistical mapping techniques that seek to find an understandable visual representation of the similarities, dissimilarities or ordered rankings between different data points.

Path. The directory on a computer's hard drive or portable drive where certain files are located, relative to a root reference. For example, when browsing the Internet with Microsoft Internet Explorer, temporary Internet files are automatically stored in the *C:\Windows\Temporary Internet Files* directory.

Search engine. A web site (or part of a web site) that can be used to search the World Wide Web (or the specific site that is visited) for a specific subject or word.

Statistical mapping. A group of statistical analysis techniques that aim to facilitate data analysis by displaying information in a spatial format.

Web log. Web logs record general information about users' behaviour on sites, e.g. the number of total hits per day, hits per individual page and the order in which pages are visited.

Web page. Web pages are documents on the World Wide Web that may include for example text, photographs, illustrations, video clips, music files or computer programs.

Web site. A web site is a collection of web pages located on a server connected to the World Wide Web.

Word Wide Web. A computer-based network of web sites that a user can move through by using links from one site to another. The information is stored on computers around the world, and the Hypertext Transfer Protocol (HTTP) is used for information exchange.

1. Introduction

The Internet is quickly becoming the most important avenue for organizations in many industries to interact with their customers and other stakeholders. As a result, many organizations have made considerable investments in their corporate web sites. Many managers now see the web site as a major strategic asset, and for Internet companies it has become essential to the way they do business. Web sites have developed a great deal since the Internet became commercially accessible, and they are becoming more and more powerful. It seems, however, that many customers are still not satisfied with them, and therefore choose not to do business online. Forrester Research estimate that "poor web design will result in a loss of 50 per cent of potential sales due to users being unable to find what they want, and a loss of 40 per cent of potential repeat visits due to initial negative experience" [Cunliffe 2000:297]. It seems that there is still much room for Internet companies to improve their web sites. In order for companies to find out how their sites can be improved, progress needs to be made in the assessment of the performance of web sites, especially e-commerce web sites. While there have been a number of advances made in the area of assessing a web site's communication performance—for example in the counting of hits—the comprehensive evaluation of web sites providing e-commerce solutions have received much less attention. Managers of Internet companies would undoubtedly like to know how the companies' web sites perform, not only in isolation but also compared to those of its competitors.

The purpose of this thesis is to address important needs in the field of web site evaluation as applied to e-commerce in particular.

This introduction aims to give the reader a clear understanding of the thesis structure by focusing on two aspects:

- A discussion of the objectives and the significance of the research.
- The methodology followed to fulfil the objectives, and the structure of the document.

1.1 Objectives and significance of research

Considering what has been said in the opening paragraphs above, the researcher, in conjunction with Prof. Leyland Pitt of the University of Curtin in Perth, Australia, identified the following **limitations and needs in the current e-commerce field**:

- No framework for the comprehensive evaluation of e-commerce web sites currently exists in common literature. Internet companies have no way of assessing the performance of their web sites in a thorough manner.

- Internet companies have no way of assessing how their sites perform compared to others in the same industry.
- What is needed is an evaluation framework that can be used by a third party to evaluate web sites as objectively as possible.
- The need further exists to compare the evaluation results of different web sites within the same industry with each other, based on sound statistical analysis. A graphical representation of this comparison is needed to facilitate easy understanding of the differences between sites. Companies will then be able to see how well their web sites perform—individually and compared to others.

An important distinction must be made between *e-commerce* web sites and *general* web sites. **E-commerce web sites** are defined as web sites that provide a means for customers to do business on the Internet. This definition includes both web sites that *sell products* (for example Amazon.com) and web sites that *provide services* (for example ABSA online banking, [www.absa.co.za]). In this thesis the focus is specifically on Business-to-Consumer e-commerce sites. **General web sites** are defined as sites that provide information or fulfil an entertainment purpose, but have no commerce component. Examples include sites that provide the latest news (for example [www.news24.co.za]) or sites that provide information on the movie industry (for example the *Internet Movie Database*, [www.imdb.com]).

It is also important to realise that the research focuses strictly on e-commerce *web site* evaluation. The general operations of the company and the effectiveness of its e-commerce strategy are not of significance here, even though these are very important factors that also have to be in place for an Internet company to be successful. In this thesis it is the company's web site and the specific operations related to it that are the primary target for evaluation. It is important to isolate the web site of an Internet company for evaluation at some stage, because that is often the only part of the company that the customer actually 'sees'. It is the first and only method of communication with the potential customer. The operational systems of the company can be perfect, but an ineffective web site will result in certain 'dot-com death'.

The need for a graphical representation of web site comparison data led Prof. Pitt and I to a statistical analysis concept called *statistical mapping*, which aims to facilitate data analysis by displaying information in a spatial format. Based on prior knowledge, Prof. Pitt then suggested Multidimensional Scaling (a specific group of statistical mapping techniques[1]) as a suitable tool for the graphical representation of comparison data.

---

[1] Multidimensional Scaling is defined and discussed in full in Chapter 3.

Considering the above discussion, **three objectives** were set for the thesis:

1. **To develop a framework and criteria for the comprehensive evaluation of e-commerce web sites.**
2. **To use this framework and sound statistical reasoning to develop a method[2] that can be used to evaluate e-commerce web sites quantitively. This includes the use of a Multidimensional Scaling technique to plot the evaluation results of different web sites within the same industry on a two-dimensional scale.**
3. **To implement this method by developing computer software that enables users to evaluate e-commerce web sites and plot the results.**

It is important to note the specific **significance of this research**. As mentioned earlier, at present Internet companies have no way to measure their online performance[3] based on objective criteria. Furthermore, they cannot compare their performance with other companies in the same industry. This study is significant because it gives companies an opportunity to evaluate the as-is situation of their web sites with respect to other companies. They will be able to see in which areas they perform well, and in which areas there are room for improvement. It will also be possible to determine which sites within the industry perform better than they do in certain areas, and why. This objective information can be used to make strategic decisions on the paths that should be followed to improve the efficiency and performance of their web sites.

1.2 Methodology and structure of document

This document is structured to follow the methodology used to fulfil the objectives set in the previous section. Table 1 explains the structure of the document and implicitly also outlines the methodology. It is clear that the document is divided into two main parts, the first being a literature review, and the second focusing on the development of the evaluation framework and method.

---

[2] The word *method* as it is used in this context is defined as all the procedures followed to evaluate e-commerce web sites—from the gathering of evaluation data to the interpretation of the output results.

[3] For the purpose of this thesis, *performance* is defined the degree to which a web site fulfils its set (or likely) objectives.

Part 1: Literature review	
Chapter 2	▸ Review the research done in the field of web site evaluation for both general web sites and e-commerce web sites. ▸ Identify the best way to gather and represent web site evaluation data in this thesis.
Chapter 3	▸ Review the research done on different Multidimensional Scaling techniques and understand the progress already made in this area. ▸ Choose and understand an appropriate Multidimensional Scaling technique for developing two-dimensional plots of web site evaluation data, based on the decision made in chapter 2.
Part 2: Development of computerised e-commerce web site evaluation method	
Chapter 4	▸ Expand the web site evaluation research and develop a framework for the evaluation of e-commerce web sites based on solid business principles. ▸ Use this framework to develop individual criteria for the comprehensive evaluation of e-commerce web sites. These criteria should be as objective as possible. ▸ Link the evaluation framework and criteria with the appropriate Multidimensional Scaling technique and develop a method for gathering evaluation data for web sites, grouped within industries, and representing the results graphically.
Chapter 5	▸ Implement the method by developing computer software to automate the process.
Chapter 6	▸ Conclusions and recommendations

Table 1 - Structure of the document

A simplified graphical representation of the thesis methodology is shown in Figure 1.

The ability of customers to understand and use e-commerce web sites effectively is of extreme importance if Internet companies want to excel in their business. This thesis thus also aims to help companies and customers to understand the components of an effective e-commerce web site, and to develop a solid method to determine what changes should be made to improve these web sites.

Part 1 of the thesis follows, which provides an in-depth review of the research conducted on the concepts and tools that are used to fulfil the various objectives set in this introduction.

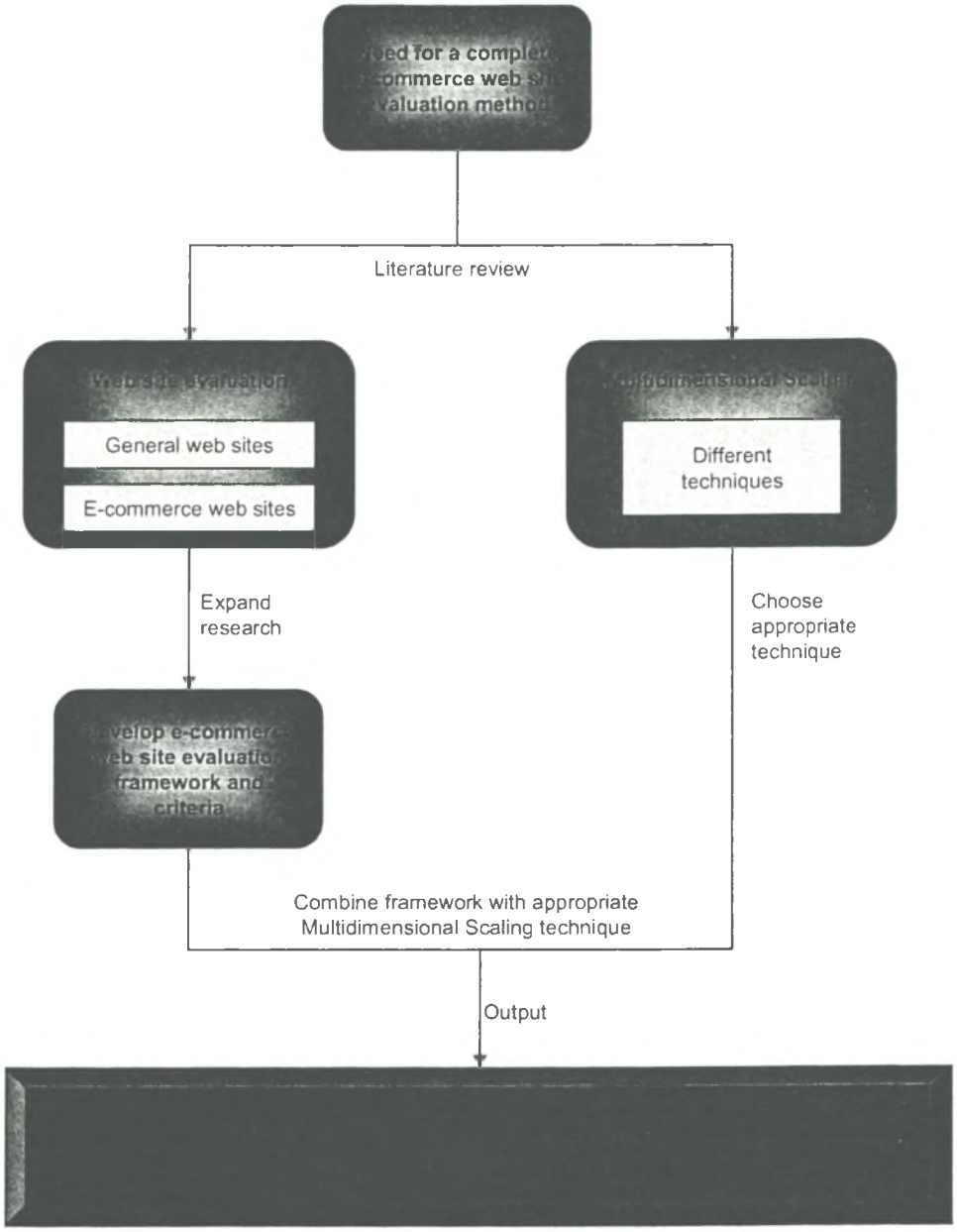


Figure 1 - Graphical representation of thesis methodology

Part 1

Literature Review

2. Evaluating web sites

The excitement of seeing a beautiful colour photograph of Mt. Everest on the Internet for the first time has worn off over the years, even for users who grew up using 286-processor computers with black and white screens and no real use beyond simple word processing. Since the start of commercial use of the Internet, a world of new business applications for the medium—beyond mere communication with text and images—became a reality. The Internet and the interface of the World Wide Web have evolved to such an extent that it is now possible to dazzle users with impressive, multimedia-rich web sites. Unfortunately, for consumers using the Internet for more serious activities than watching the latest movie trailers, this is becoming a huge problem. The hype of impressive front-ends is over, and consumers doing business on the Internet now demand web sites that are easy to use, provide relevant information where and when they want it, and is reliable every time a transaction is made. In order to provide users with such web sites, it becomes necessary to find out what the characteristics of a good web site are and how to evaluate whether these criteria have been met on a site or not.

It is only recently that researchers began to realise the importance of finding criteria to evaluate web sites effectively. This chapter gives an overview of available literature on web site evaluation. It discusses both reliable and less reliable sources to illustrate how web site evaluation research has evolved. It also shows that the same marketing rules do not apply for web sites as for printed media. The chapter is divided into **three sections**. The first gives an overview of literature on the *evaluation of general web sites*. The second section broadens the horizon to *complete e-commerce web sites*. Section three discusses an important aspect of e-commerce web sites, namely *database quality*.

2.1 General web site evaluation

When people started to use the Internet for commercial purposes, it was basically a way to communicate static information with connected users. Back in 1995, the Internet was all about what users could see on their screens, and it was not utilised to do business yet. It is therefore understandable that the first approaches to web site evaluation considered only its basic appearance. The first formal approach to the evaluation of web sites was developed by Boyd Collins in late 1995. He founded the *Infofilter project*, a model intended

for librarians who needed to evaluate the quality of information on the Internet[4]. The model was based on six criteria Collins developed by combining evaluation criteria for printed media with what he believed was relevant for web sites. These criteria are shown in Table 2.

Criterion	Some common attributes
Content	Information accuracy Value of information Currency of information
Authority	Author of the site Author's qualifications Legitimacy of site's sponsor
Organisation	Ease of navigation Clear transition between pages Logical and effective layout
Searchability	Quality and accuracy of search engine Search engine's ease of use
Graphic design	Effective use of images Effective use of other graphics and media Effective use of colour
Innovative use	Attributes that make the site stand out from others

Table 2 - Boyd Collins criteria for web site evaluation

Although innovative at the time, it is clear that the Internet and its uses have changed to such an extent that these criteria are not sufficient to evaluate web sites any more. The Infofilter project ceased operation in July 1997 because its participants realised that it was becoming obsolete. It seems, however, that many authors have not realised that these criteria have lost their relevancy to a large extent.

Using common Internet search engines to find literature on web site evaluation reveals hundreds of web sites claiming to provide web site evaluation criteria. When studied closely it becomes clear that most of these sites use the Boyd Collins criteria as a platform to which they add their own criteria arbitrarily. There is mostly no sign of research on which these criteria are based, which makes it unreliable to a large extent. For example, one article [Alexander & Tate, 1996] gives the following five evaluation criteria: *Authority*, *Accuracy*, *Objectivity*, *Currency* and *Coverage*. Another web site [*Evaluating Websites*, no date] gives even more related criteria: *Purpose*, *Content*, *Authority*, *Scope*, *Audience*, *Currency*, *Special Features*, *Organization*, *Accessibility*, *Reliability* and *Quality of Site Search Engine*. When these criteria are compared to the Boyd Collins criteria it becomes clear that they are still used today—with some minor alterations as each author sees fit.

The biggest problem with these criteria is that they are based on traditional marketing strategies for printed media before the Internet became such a huge force in business. The

[4] For more information on the Infofilter project, visit [http://www.usc.edu/users/help/flick/Infofilter], last accessed on 2 May 2001.

traditional ways of evaluating marketing strategies cannot be used to evaluate web sites, mainly for two reasons:

- As Schubert & Selz (1999) states, "Underlying Internet technology forces marketing activities to be different from the ones applied to traditional sales channels". The Internet makes marketing interactive and it gives the customer the power to choose what he wants to see and for how long he wants to see it. Internet technology opens the door to many new marketing opportunities, but over-use of new technologies on sites, making them extremely difficult to use and navigate, can also scare away customers very easily. It is extremely difficult to find the perfect balance between technological innovation and ease of use.
- The Internet is not nearly only a marketing tool anymore. The way business is currently integrated with the Internet means that web sites can no longer be evaluated only by looking at their front-ends—the back-end operations supporting them are just as important to customers. Whether or not a site can deliver on its promises is what makes customers decide if they want to use the web site again or never come back to it. It is therefore very important to evaluate a site's reliability[5] in this regard.

Another problem with the criteria is that they are very subjective and hard to quantify. Say, for example, an evaluator wants to evaluate a site's *Searchability*. He has to have a good idea of what makes a site easily searchable, and then he needs to quantify the site's *Searchability*. If the evaluator's idea of a searchable site differs from how other evaluators see it, the evaluation becomes unreliable and cannot be used to compare sites with each other.

Moving away from Internet resources on evaluation, the past few years have also seen an increase in the number of research articles on web site evaluation published in academic journals. The advice given in these publications range from extremely rigid ("Use no more than three images per page" [D'Angelo & Little, 1998:74]) to very vague ("Design the web site for content, not appearance", [D'Angelo & Little, 1998:75]. Generally, however, these publications are much more reliable than the multitude of web sites claiming to provide criteria for web site evaluation. They sometimes even move a step beyond mere front-end evaluation. An overview of some of the more reliable sources on the subject follows.

[5] *Reliability* in the context of this thesis is properly defined in Section 4.1 on page 47.

2.1.1 The static quaiity/dynamic quality approach by Day (1997)

Day (1997) says the following about web site design: "Offering prescriptive statements such as 'too many/too few icons are bad' is a little like saying 'blue cars are bad'. They are not bad or good, they either work or they do not. But why they work is not a function of the icons, the graphics, the content, the colour, the length of the page, the number of links, the speed of the modem or any other technological variable. A web site works because the people it serves likes it" [Day, 1997:109]. Day's research is based on this notion, and aims to find what she calls 'likeability factors' for web sites.

She divides the quality of web sites into two broad categories, namely **static quality** and **dynamic quality**. Static quality refers to the aspects of the site that do not change and that would definitely be noticed if it stopped working, like hyperlinks or the order process. Dynamic quality is what makes the site unique—"we notice it because it attends to our individual needs in a unique and completely relevant way" [Day, 1997:110]. It is the features that make customers come back to the site, like special offers and good after-sales support. She builds her evaluation model on the idea that "we need to establish the basics of static quality and then be responsive enough to offer dynamic quality as well" [Day, 1997:111]. Day proposes 7 criteria by which the static quality of web sites can be evaluated:

- **Customer focus** involves evaluating whether or not a company knows who their customers are and if they target that specific group.
- **Purpose** measures the site's ability to inform users "why the web site is there and what it will do for them" [Day, 1997:111].
- **Content** refers to the quality of the information on the site.
- **Structure** measures the ease of navigation and the logical structure of the site and individual pages.
- **Housestyle** refers to the consistency and effective use of the visual elements (graphics and multimedia) on the site.
- **Action** evaluates how easy it is for a customer to do what the site wants him to do. For example, in the case of an online bookstore the effectiveness of the order process will be evaluated.
- **Promises delivered** measure the company's ability to fulfil the promises it makes to customers, for example on-time delivery or fast after-sales service.

After ensuring static quality, Day believes it is then up to each company individually to decide how dynamic quality of their sites can be achieved.

2.1.2 User-defined criteria developed by Misic & Johnson (1999)

Misic & Johnson (1999) used important components of web sites as identified by typical users, as well as important features identified by visitors to the authors' own web site to develop a set of evaluation criteria. These criteria were grouped into three categories, namely **functional/navigational metrics**, **content and style metrics** and **contact information**. Table 3 outlines the specific metrics developed in each category group to evaluate web sites:

Criteria Category Group	Metric used
Functional/navigation metrics	Ease of finding e-mail/phone numbers Ease of finding main page Speed of finding main page Subpage loading speed Uniqueness of functionality Ease of overall navigation Ease of returning to main page
Content and style metrics	Hit counter provided Currency of information (date last updated) Effective use of colour Effective use of graphics Colour consistency Style consistency Wording
Contact information metrics	E-mail contact provided Site maintainer identified Phone numbers available Mailing address available Ease of locating people on the site

Table 3 - Metrics used in Misic & Johnson (1999) web site evaluation study

Although these metrics are by no means extensive, it addresses some important web site issues that should be kept in mind when the evaluation framework is designed.

2.1.3 Positive and negative web site features by Abels et al. (1999)

Abels et al. (1999) gathered information on positive and negative web site features through a group session involving selected faculty from four different business schools in America. The group identified 33 positive features and 18 negative features that were ranked and clustered into six major criteria groups. In order of importance, these groups are **usability**, **content**, **linkage**, **structure**, **special features** and **appearance**. Table 4 outlines some of the metrics developed for each group.

Criteria group	Specific metrics
Usability	Ease of use Overview of site provided Effective navigational structure
Content	Useful information Current information Concise, non-repetitive information
Linkage	Pages provide links that integrate relevant information on the site and at other sites Links provide access to related information All links are functional
Structure	Intelligible, straightforward organising scheme Text broken into appropriate, well-labelled subsections Large blocks of text are minimised
Special features	Search facility provided
Appearance	Site visually attractive on screen Graphics not essential to site—can be turned off for a text only version of the site Printouts of pages attractive without large dark areas

Table 4 - User criteria for web site evaluation developed by Abels et al., 1999

On review of these criteria it becomes clear that the specific metrics are very rigid at times, and their relevance is debatable. The criteria groups as a whole, though, can be very useful in further research.

2.1.4 The web site Quality Evaluation Method (QEM) by Olsina et al. (1999)

Olsina et al. (1999) used the main indicators of software quality (namely usability, functionality, reliability, efficiency, portability and maintainability) as a platform for their research on the evaluation of primarily academic web sites. They decided to use four of these criteria categories for their web site Quality Evaluation Method (QEM), namely **usability**, **functionality**, **site reliability** and **efficiency**. They believe that "these characteristics give evaluators a conceptual framework of quality requirements and provide a baseline for further decomposition" [Olsina et al. 1999:3]. The authors developed more than 120 quality characteristics for web sites, in up to six levels of detail. For the purpose of this section, only three of these levels are shown in Table 5. The full list of quality characteristics can be found in Appendix A.

Level 1 Quality characteristics	Level 2 Quality characteristics	Level 3 Quality characteristics
Usability	Global site understandability	Global organisation scheme Quality of labelling system User-oriented guided tour Image map
	On-line feedback and help features	Quality of help features Web site last update indicator Addresses directory FAQ feature (Frequently Asked Questions) On-line feedback
	Interface and aesthetic features	Cohesiveness by grouping main control objects Presentation permanence Stability of main controls Style Issues Aesthetic preference
	Miscellaneous features	Foreign language support What's New feature Screen resolution indicator
Functionality	Searching and retrieving issues	Web-site Search Mechanisms Retrieve Mechanisms
	Navigation and browsing issues	Navigability Navigational Control Objects Navigational Prediction
	User-oriented domain-related features	Content Relevancy On-line Services
Site reliability	Non-deficiency	Link Errors Miscellaneous Errors or Drawbacks
Efficiency	Performance	Static Page Size
	Accessibility	Information Accessibility Window Accessibility

Table 5 - Quality characteristics for academic web sites (Olsina et al., 1999)

Although these characteristics were developed for academic web sites, they can be used as guidelines for the development of general web site evaluation criteria, as many of them are of a universal nature. Another benefit of this approach is that many of the lower level characteristics are objective and quantifiable.

2.1.5 Good web site design elements by Cunliffe (2000)

Cunliffe (2000) views good web site design as a blend of three elements, namely **content**, **visual appearance** and **usability**. He holds that there are too little experienced web site designers, a problem that is currently giving rise to a definite usability crisis with regards to e-commerce web sites. He blames this problem on factors such as development teams with limited skills and experience, and limited resources available to these teams.

Cunliffe's other concern is that even if a site follows certain design guidelines, it is not necessarily a usable site as these guidelines are typically based on practical experience and not justifiable research. He proposes several methods to test the usability of the site, including techniques such as **competitive analysis** ("analysing existing web sites provided by similar organisations with similar business objectives" [Cunliffe 2000:301]), **scenarios** (using employees to simulate the actions of real users) and **inspection methods** (evaluating sites with a predetermined set of criteria).

2.1.6 Automatic classification and evaluation by Bauer & Scharl (2000)

Bauer and Scharl (2000) developed a method for automatic classification and evaluation of web sites. Their belief is that "the utilisation of dedicated software agents to examine web sites is more efficient and immune against intra- and interpersonal variances than human evaluation" [Bauer & Scharl, 2000:31]. This approach makes it possible to evaluate thousands of web sites at a time, which is not possible with manual evaluation. The drawback is, however, that certain non-quantifiable information like *ease of navigation* becomes unobtainable. Their automatic process uses no subjective data and only gathers quantitative data such as the number of images on a page and the number of links out of a page to classify and evaluate the site. The criteria developed for classification are grouped into three categories: **content**, **interactivity** and **navigation**. Table 6 details the information that is gathered automatically:

Criterion	Variable
Content	Number of documents on site Total amount of kilobytes downloaded Number of different file types (extensions) Number of images
Interactivity	Number of forms Number of documents with JavaScript Number of Java applets Number of MailTo links
Navigation	Frames (Yes/No) Number of internal links Number of external links Number of anchors Number of links to anchors

Table 6 - Criteria for automated classification of web sites (Bauer & Scharl, 2000)

Web sites are classified using these criteria and then evaluated by comparing the data with other sites of the same type. Although the idea of an automatic approach to web site evaluation is very good, the problem with this approach is that so much evaluation information is lost because it would need some form of human opinion to gather it.

2.1.7 Automatic site evaluation methods in general

Except for Bauer & Scharl's work on automatic web site evaluation methods, Spiliopoulou (2000) is also doing research in the field of automatic data mining of web usage logs to evaluate sites. The benefit of these groups of methods is that they use totally objective and easily quantifiable data to perform the evaluation. There are, however, some drawbacks and analysts should therefore be very cautious when applying them:

- **Visitors cannot be uniquely identified**. Log information can only record a user's IP address, and there is no assurance that more than one user does not use the same computer. This means that different people can access the same site from the same

computer, while the usage log will only record one user. Accurate data on users are therefore difficult to assimilate.

- **The use of caching**. Frequently used pages on a site are stored in the computer's memory and are therefore not retrieved directly from the server it originates from. This implies that log data sometimes do not contain a full record of a user's session on a site.
- **Datasets must be of high quality**. The data in a web server log normally needs extensive pre-processing before the datasets can be used for analysis.
- **Reliability of data**. It is difficult to assess whether all recorded data are in fact reliable. For example, some complex search engines, called *automated spiders*, browse through a site and do not follow normal user behaviour—it just goes through all the pages systematically to find the search terms requested by the user. Collecting behavioural data for such cases is obviously not relevant to an automated analysis of user behaviour.
- **Data combination**. In order to make the analysis meaningful, the web log should be combined with other relevant information on the site, such as page description, number of images and links per page, etc. This is time-consuming and can confuse the data further.

Although it is the ideal to evaluate web sites on purely objective data, it is not entirely possible, because of the reasons mentioned above. The solution for this problem is to break down subjective evaluation criteria to the lowest level of complexity possible, so that there is no confusion about exactly what should be evaluated and how the evaluator should do it.

2.1.8 Summary

This section provided an overview of some of the approaches to web site evaluation found in recent literature, and showed the diversity of ways proposed to evaluate web sites. The interested reader is referred to D'Angelo & Little (1998), who provide a comprehensive overview of web site design literature published in the past few years.

Although many of the approaches outlined above are valuable, they have two major limitations that greatly reduce their applicability:

- The criteria are **rarely based on in-depth research**, and **mostly rely heavily on the intuitive feeling of the evaluator**. Even when research is apparent, because of the multitude of different techniques one cannot help but wonder whether every base is

covered or if there is still some part of the site that will not be evaluated when only one of the techniques are used.

- All the **evaluation criteria deal with individual sites, and no method is proposed to scale the results and compare them to other sites**. As mentioned in the introduction, when companies would like to know not only how their web sites perform, but also how they perform with regards to other sites in the same industry, they have no way of doing that in a formalised and statistically correct manner. Of course it is possible to add the results obtained from the evaluation of different sites and compare them among each other, but the results are difficult to interpret and the reliability of the results cannot be determined.

These two limitations are addressed in this thesis. While this section provided an overview of *general web site* evaluation techniques, the next section goes beyond the front-end and explores ways to evaluate *e-commerce web sites*. As mentioned in the introduction, general web sites basically provide information or entertainment while e-commerce web sites facilitate business on the Internet. Customers use these sites to perform business activities on the Internet, without the need to go to a physical store. E-commerce always involves a transaction that takes place over the Internet, in other words money is exchanged between a customer and a supplier/provider. This definition includes both shopping for products (physical or digital—anything from groceries to camping gear to the latest software) and buying some sort of service (for example online banking or booking hotel reservations). The company providing the product or service is referred to as an *Internet Company* in this thesis. In some ways, e-commerce web sites can be seen as the 'store' of the Internet Company. The same rules do not apply for these sites as for general web sites. E-commerce web sites entail not only the information on the site, but other very important aspects such as security considerations, customer profiles and order processes. The next section focuses on research done in this area of web site evaluation.

2.2 E-commerce web site evaluation

It is clear that the methods described in the previous section cannot be applied directly to e-commerce web sites. The reason is simply that successful e-commerce sites do not depend on appearance alone. Much more important is the ability of the site to fulfil the customers' needs reliably and provide efficient customer service. People shop or do business online because of the benefits it involves—benefits that traditional commerce does not provide. It is therefore important to provide users with enough innovative online benefits so that they will become loyal customers. The purpose of e-commerce web site evaluation is to measure not only how attractive a company's web site is, but also how effectively the site is providing benefits to customers and fulfilling their promises reliably.

Literature on e-commerce web site evaluation is hard to find, as it is a fairly new field of research that has not received much attention. Research is still in its infant stages, and is often only involved with one or two important factors of e-commerce web site design. For example, many current studies are concerned with the security aspects of e-commerce sites (Hawkins et al., 2000; Labuschagne & Eloff, 2000). Good security is definitely a necessity, but is not a sufficient condition for a successful e-commerce web site. There are many other factors that also play a role and they must be identified and evaluated.

One of the interesting research projects in this field is the **Web Assessment Model**, a project started by Petra Schubert and Dorian Selz in 1997. The history and development of this model is outlined in their article *Measuring the effectiveness of electronic commerce web sites with the web assessment model*, published in a book by Barnes & Hunt (2000). Another interesting approach is the one followed by Liu et al. (2000), who used the **inputs of webmasters from Fortune 1000 companies** to assess which criteria is important for e-commerce sites. The rest of this section will provide an overview of these two approaches, as it is believed that they currently make the two most valuable contributions to e-commerce web site evaluation.

2.2.1 The Web Assessment Model

The original *Web Assessment model* (WA model) was developed in 1997 at the Competence Centre for Electronic Markets (CCEM) at the University of St. Gallen in Switzerland. The model used a set of criteria to measure the success of e-commerce applications. Empirical studies were conducted over the next few years to test the validity and usability of the model. Customers were asked to evaluate several existing e-commerce web sites using the WA model. More than 70 questionnaires were collected over the duration of the test phase. The results of these studies, combined with more research, resulted in the fundamental

revision of the WA model at the beginning of 2000. The *Extended Web Assessment Method* (EWAM) was the result, and facilitates a comprehensive evaluation of e-commerce sites from the customer's point of view. Both models are significant and will be discussed in this section.

The models are based on the belief that e-commerce web sites include both marketing and technical issues, and criteria should be developed to evaluate both issues. The original WA model was based on the traditional three phases of market transactions—the *information phase*, the *agreement phase* and the *settlement phase*. It also went one step further and added a *community component* because of the Internet's unique ability to facilitate virtual communities.

- In the **information phase** "customers collect information on potential products and services" (Schubert & Selz, 1999). The customer uses this phase to gather information on ways to fulfil a need or solve a problem he may have.
- In the **agreement phase** a link is established between buyer and supplier. Details such as product specifications, form of payment and delivery are communicated between the two parties.
- In the **settlement phase** the actual payment for and delivery of the product takes place. After-sales interaction is also included in this phase.
- The **community component** is a further degree of interaction between customers and companies, and customers among themselves. A good example of this is the Amazon.com community, which allows users to publish their own reviews on the products they buy on the site. These communities are believed to build a high level of trust among members, and between members and the company (Figallo, 1998).

The following four tables, taken from Schubert & Selz (1999), give an overview of the criteria that was developed in the original WA model for each of the phases described above. Although the community component is very important for e-commerce, it is not applicable to all e-commerce web sites. Evaluating virtual communities definitely form a part of the evaluation criteria developed in this thesis, but not in the detail described in Schubert & Selz (1999). The criteria they developed in that phase are therefore only discussed briefly in this document.

Criterion	Explanation
Good user interface	The *user interface* assesses ease of use for frequent users as well as for first time visitors. This does also comprise loading times of pages and guidance in the interaction process with the Web site when completing a transaction.
Good structure of content	The *structure of content* measures ease of access as well as first and second impression of the logical structure of the content. Tables of contents, navigational frames or image maps are typical features to facilitate navigation.
Reasonable information quantity	The *quantity of information* focuses on the range of information on the company, its products and its services.
Apparent benefits from stored customer profile	Most Web sites require customers to register or at least to supply some basic personal information. Good Web marketers should remunerate their customers for revealing this kind of information. This could be either by · Directly crediting money or services, examples are www.bonusmail.com and www.cybergold.com · Granting discounts for product sales
Good products/ service combination possibilities (cross-selling: combine products and/or services)	*Combination possibility* examines the breadth of the product range and the possibility to combine various product offerings (either to the company's own products, or third-party goods/services) online. It measures the amount of cross-selling, i.e. the combination of various goods/services (such as an airline ticket and a hotel reservation).
Good availability/ performance of the system	*Availability/performance* (in respect to geographical aspects) measures the global availability of the system. It judges the availability to customers regardless of their geographic location (e.g. different language versions). Special mirror sites could e.g. improve global performance. Since this aspect is one of the main advantages of the Internet it gains special consideration. *Availability/performance* (in respect to time) measures the loading times, which are of great importance for user comfort.
Cost benefits passed on to the client	The use of electronic sales channels often reduces transaction costs. Provided that margins remain unchanged vendors should be able to offer products on their Web site at a lower price than in a comparable physical store.

Table 7 - Web Assessment criteria in the Information phase (Schubert & Selz, 1999)

Criterion	Explanation
Adjustable customer profile (e.g. payment information)	Business transactions usually require customers to reveal some basic personal information, e.g. payment information. For a greater comfort this kind of information can be safely stored for reuse in a subsequent session.
Guided ordering according to profile (personalized services)	In order to receive a higher degree of personalized services customers could be willing to reveal additional information. Besides, the system might track user activity. A detailed user profile containing personal information such as age, gender, hobbies, preferences, etc. helps to treat each customer differently. This could result into guiding mechanisms, enable the system to come up with suggestions, or even to grant special client discounts.
Possibility of customized products	Some customers might be interested in buying combinations of products (product systems) or only fragments of a product (e.g. only parts of a magazine or newspaper). The Web site could support the customization of user-designed products.
Transparent, interactive integration of business rules	The underlying business rules should be transparent to the user. Business rules are general terms and conditions, guarantees, possibility for returning products, etc. Click buttons to accept terms and conditions and guided interaction are helpful in this context.
Good implementation of security issues (digital signature, secure server, TTPs)	Good Web sites should offer reliable security features (such as SSL, digital certificates, etc.) or implement accepted standards (e.g. SET, TRUSTe, P3P).
Good contact possibilities with vendor (help desk for problems during order process)	*Contact possibility* examines the various ways to establish communication with the vendor. It may comprise the implementation of a help desk or a call center. The Web site could offer · the opportunity to write and read questions of common interest (FAQs) · a feedback possibility via E-Mail or Web forms (i.e. via the electronic medium) The feedback response times must be adequate to the medium used.

Table 8 - Web Assessment criteria in the Agreement phase (Schubert & Selz, 1999)

Criterion	Explanation
Easy selection of generic services	*Generic services* are software modules that are available on the entire Web platform and always present themselves in a uniform interface. Generic services support an electronic transaction (such as the purchase of a book online). Examples are electronic payment systems, logistics services, electronic contracting modules, etc. Their brick-and-mortar counterparts are power sockets, telephone hooks, water taps, and the postal system that should be the same wherever you are (this applies at least within the same country). An easy selection of such services means that they are integrated into the settlement process. Examples are the selection of different choices, e.g. payment systems (E-cash, credit card, SET, check, bill, etc.) or logistics services (UPS, FedEx, US Post, etc.). The tracking of order information should also be supported.
Good integration of generic services	A good integration of such services means that they are sensibly used wherever necessary, comforting the user by their common interface and their routine operation. Typical generic services in Electronic Commerce applications are payments, electronic contracting (dealing with prices and conditions) and logistics.
EC-application makes effective use of customer profile (e.g. payment and logistic information)	During the settlement of a business transaction some basic personal information need to be revealed (e.g. payment information or delivery address for physical goods). For greater comfort this kind of information could be safely stored for reuse in a subsequent session.
Good tracing and tracking (e.g. direct access to personal order information)	A good example for an integration of a logistics service (in this case of a third party) can be found at www.amazon.com. After ordering customers are provided with information on how to trace their order at the UPS tracking site.
Good IT-integration (connection with customer's infrastructure)	Especially for small and medium-sized businesses (SMEs) an export filter (a link to their accounting system) for financial data could be of great value (e.g. information can be exported into MS Money).
Convenient after-sales support	The Web site should also support the handling of after-sales services (e.g. guarantee form, feedback form).

Table 9 - Web Assessment criteria in the Settlement phase (Schubert & Selz, 1999)

Criterion	Explanation
Good access to community	Good link between product offer and community component of the Web site
Uniqueness/originality of information	The information provided on the community is difficult to obtain from other sources
Adequate number of members	The value of the community is its members. More members imply more information and a higher perceived value.
Well-implemented personalization and collaborative filtering mechanisms	Ability of community to link members with people with similar interests
Member may choose his/her appearance within the community	Some Web sites offer visual representations of people in the form of cartoon animals, comic strip characters, etc.
Privacy is sufficiently protected	Level of anonymity
Perceived real added value from membership	Evaluates the value of membership
Good push mechanisms	Information is automatically sent to members
Good pull mechanisms	Members can ask for information updates

Table 10 - Web Assessment criteria for the Community component (adapted from Schubert & Selz, 1999)

After the original WA model was developed and tested, further research was conducted. EWAM combined new research findings in the Internet marketing field with a model designed for the acceptance of information systems called the *Technology Acceptance Model* (TAM) as developed by Davis (1985).

EWAM differs from the WA model in two main respects:

- Firstly, each EWAM criterion is assigned to a specific *criteria group*. These main groups resulted from research conducted in the field of user needs, and are *ease of use (**EOU**)*,

usefulness (***USEF***) and *trust* (***TRUST***). These groups have also to some extent been identified in the previous section as being very important in web site evaluation.

- Secondly, EWAM expands the existing market transaction phases as used in the WA model, and consists of six possible category groups to which criteria can be assigned, as opposed to four in the WA model. These are the *information phase*, *agreement phase*, *settlement phase*, *after-sales phase*, the *community component* and *criteria that concern all phases*.

A graphical representation of the EWAM is shown in Figure 2 on the next page. Table 11 on page 23 shows a list of the criteria—and the groups they were assigned to—as they were developed for EWAM. For a more detailed description of EWAM, the reader is referred to Schubert & Selz (2000).

What makes the criteria in both the WA model and the EWAM so useful is that it provides an overview of all possible category groups that should be evaluated when designing evaluation criteria. EWAM's ability to combine market transaction phases with three very important criteria groups (*ease of use*, *usefulness* and *trust*) makes it a very important model on which further research can be based.

There is, however, a significant problem with the WA and EWAM approaches in the context of this thesis. The criteria in these approaches are on a very high level, still open for much interpretation. For example, the criterion *satisfying customer support* demands of the evaluator to form a very subjective opinion on customer support and, in fact, what he feels good customer support is. Most of these criteria rely on too much subjective interpretation by the evaluator, which makes the data somewhat unreliable. It can still be used as an indication of the web site's performance, but the use of statistical techniques to quantify an objective evaluation of a web site is not recommended. The challenge is to use these criteria and expand it to a lower level of complexity so that all ambiguous meanings are eliminated, and easy objective evaluation becomes possible.

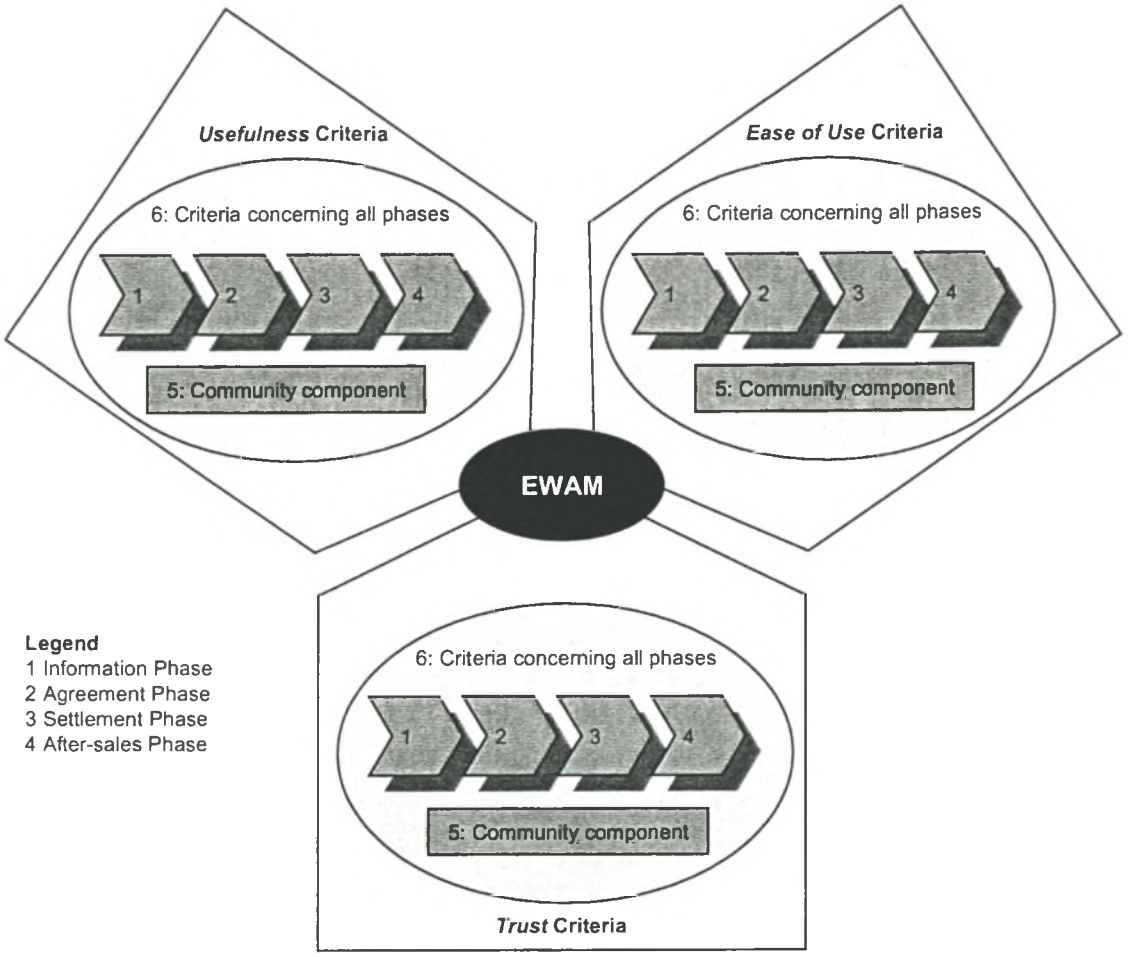


Figure 2 - Diagram of the Extended Web Assessment Method (Schubert & Selz, 2000)

Category Group	Criteria	Criteria Group
1. Information Phase	Web page and specific offers easily found	EOU
	Good structure of content	EOU
	Reasonable Information Quantity	EOU
	Quality of content meets user expectations	USEF
	Cost benefits passed on to the client	USEF
	Bundling: good combination possibilities for products/service	USEF
	Good recommendation system	USEF
	Adequate application of hypermedia	USEF
2. Agreement Phase	Transparent and interactive design of ordering process	EOU
	Fair and individual prices	USEF
3. Settlement Phase	Easy selection and good integration of generic services	EOU
	Good integration in customer's IT-infrastructure	USEF
	Good tracking and tracing	USEF
4. After-sales Phase	Convenient customer support	EOU
	Satisfying customer support	USEF
5. Community Component	Good access to community	EOU
	Good quality and quantity of relationships in community	USEF
	Good quality and quantity of content generated by community	USEF
	Purchasing power	USEF
6. Criteria that concern all phases	Good availability	EOU
	Good user interface	EOU
	Improvement in productivity/time gained	USEF
	Interactivity	USEF
	Good contact possibilities	USEF
	Trustworthy business partners	TRUST
	Trust in Internet as platform and legal situation	TRUST

Table 11 - Criteria for the Extended Web Assessment Method (Schubert & Selz, 2000)

2.2.2 Criteria development using Fortune 1000 companies' web masters

Liu et al. (2000) used literature on information systems and marketing to develop six hypotheses about the important characteristics of an e-commerce web site. These hypotheses are shown below with a brief description of each concept:

H_1: Information quality is positively related to a well-designed web site
The accuracy, currency and reliability of the information provided on the site
H_2: Learning capability is positively related to a well-designed web site
The level of ease with which customers learn how to use the site effectively
H_3: Playfulness is positively related to a well-designed web site
The level of enjoyment of the customer's experience on the site
H_4: System quality is positively related to a well-designed web site
The design quality of the site, including security issues
H_5: System use is positively related to a well-designed web site
The effectiveness of the system with regards to ordering, purchasing, delivery, etc.
H_6: Service quality is positively related to a well-designed web site
The quality level of the service provided on the site, including after-sales service

Liu et al. then wanted to test these hypotheses and generate criteria that can be used to evaluate them on a site. In order to do this, they decided to ask web masters to answer a questionnaire designed to evaluate whether these hypotheses were true. Their methodology in this regard is based on the fact that "web masters may only be considered

surrogates for the ultimate customer, but as a starting point for exploratory research may be more professional, better informed on web design issues, have more experience, and are inexpensively reachable through their published web sites" [Liu et al., 2000:120]. The questionnaire was set up by developing several variables to test each of the hypotheses and asking respondents to indicate the perceived importance of these variables. The responses were then used to test whether the hypotheses were true. From different sources in literature, the following variables were developed for the evaluation of each hypothesis:

Research construct	Measurement variables	
Information quality	Relevancy	Complete description of products/services
	Accuracy	Price information
	Timeliness	Satisfying ethical standards
	Flexibility	Perceived products/services quality
	Customisability	Information to support business objectives
Learning capability	Interactivity	Help function
	Well-defined links	Customised search engine
Playfulness	Enjoyment	Charming
	Excitement	Escapism
	Feeling of participation	
System quality	Security	Precise operation and computation
	Rapid access	Balanced payment method between security and
	Quick error recovery	ease of use
		Coordination
System use	Confidence	On-line tracking of order status
	Control	Privacy
	Ease of use	
Service quality	Quick Responsiveness	Empathy
	Assurance	Follow-up service

Table 12 - Measurement variables for e-commerce web site evaluation hypothesis testing (Liu et al., 2000)

E-mails were sent to 661 reachable web masters of Fortune 1000 companies, asking them to complete the questionnaire (a copy of the questionnaire is shown in Appendix B). They received 119 usable responses from web masters who completed the questionnaire. From the statistical analysis that was performed on the data (see Liu et al. (2000) for details), the researches concluded that all six hypotheses could be strongly supported.

These results were expected, but the real value of the study lies in the variables that were designed to test the hypotheses. These variables can be used as criteria to evaluate e-commerce web sites.

This section gave an overview of two important approaches to e-commerce web site evaluation. The next section digs deeper into the mechanics of e-commerce web sites and focuses on an important part of these web sites, namely the quality of the databases that support the sites.

2.3 Database Quality evaluation

Database quality is a very important aspect of e-commerce that is receiving more and more attention. The quality of Internet companies' databases cannot be separated from their online business processes and form a very big part of a customer's experience with the company. The database is a part of the back-end operations that have to run smoothly for the business to be effective. When these databases contain data that is not of good quality, it can be a very big cause of companies' inability to fulfil promises and provide good customer service. Consider these examples of mistakes in databases that can have a serious negative influence on database quality:

- When customers log in on a site and provide different mailing addresses or e-mail addresses each time, it may happen that duplicate entries are created for the same customer. This expands the already over-large customer database, and makes it difficult to provide effective service to the customer.
- When a book is entered into the product database and a spelling mistake is made, it may be a small error, but it can lead the customer to believe the online retailer does not have the book in stock.

These are only two examples of a growing problem that is threatening the growth of e-commerce. A white paper by Firstlogic (*Data quality in the information age*, 2000:5) recently summed up the current situation in e-commerce databases as follows: "If the users of [an] enterprise system have no confidence in the quality of the data, they will not use it." This is definitely an area that should be included in any evaluation of an e-commerce web site. This section provides an overview of data quality aspects and the cleansing of databases, a technique that lays a valuable foundation for the development of criteria to evaluate e-commerce databases.

2.3.1 Unclean data and its causes

Unclean data is any form of data that is ambiguous, not regularly updated, duplicated, inconsistent or unclear as to its meaning. Flannagan & Safdie (1998) view the following causes as the roots of unclean data:

- **Mistakes**. Human errors include misspellings, typographical errors, out of range values or incorrect data types. An example might be entering a value of 13 in a month field or an alphabetical character in a numeric field.
- **Homonyms**. Especially in the English language, many words exist that have different meanings in different contexts. For example, *St.* can be used in several different ways: for example as part of a last name (St. James) or as an abbreviation for first (1st). The

possible misinterpretation of these homonyms can result in out-of-context data that is unreliable.

- **Lack of standards.** When data entry responsibilities are spread among different employees in the company, inconsistencies will occur if definite standards for data formats do not exist. Different employees may intuitively enter data in different formats. Some may use periods at the end of abbreviations while others may not. Even the choice of fields for entering data can fluctuate from employee to employee.
- **Missing data**. It often happens that data have the proper structure and values, but is not complete—without a company's knowledge. For example, an address like *Avenue 2* may be syntactically correct for a street address, but it is in fact the *name* of a street, and can contain many apartments, houses and buildings. This data goes undetected and results in incomplete data.
- **Phantom data**. Phantom data occurs when invalid data is entered into a field to flag it for some reason. The date 99/99/99, for example, may be used to indicate that a record is no longer valid. This can be very confusing, especially when database authority changes and new employees have no idea what these unwritten conventions are.

Unreliable, unclean data discredit companies and frustrate users endlessly. Data are only valuable if it is accurate, and customers will not do business online if they distrust the data on the site.

2.3.2 Clean data and its benefits

Data quality is measured by the degree of **currency**, **integrity** and **accessibility** of the data. *Currency* refers to the timeliness of the data, in other words how quickly it is updated when changes occur; *integrity* refers to the correctness of the data and its format; and *accessibility* refers to the ease of using and communicating with the database. These three categories fully define the quality of data, and database developers should aim to meet them at all costs. The process of evaluating databases and making sure the data is of high quality is generally referred to as *cleansing the database*. After cleansing, when data then sufficiently meets the criteria within the three data quality categories, it is said to be *clean data*. The following benefits of having a clean database illustrate why companies have to address this problem on an ongoing basis:

- Clean data develop **trust relationships with customers** because they know that they do not have to doubt the quality of the data on the site. This removes the customer's fear of online shopping and ensures that he will come back to the site.

- It **reduces delivery errors** and **promotes excellent customer service and customer interaction** because all the data regarding products and customers are correct.
- It provides companies with **extreme flexibility** and large opportunities to use their data meaningfully. **Complex one-to-one marketing campaigns** can be launched because high degrees of matching customers with products become possible.
- It can **improve profits** as companies are able to model and track customer relationships and product/service preferences correctly.

It is clear that clean data are essential for doing business on the Internet. Low price and high quality of products are no longer enough to ensure competitive advantage—"no data mistakes" is a sure way to get the competitive advantage that seems to be so elusive in online retail. The next section describes the actual data cleansing process.

2.3.3 Creating data quality (data cleansing procedure)

From *Data quality in the information age* (2000) it can be inferred that data quality is obtained by two equally important tasks: the **cleansing of existing data** and **ensuring quality at the points of data entry**. The process normally starts with the cleansing of the existing database, after which the same criteria is implemented as data entry validation procedures to ensure that the database stays clean.

For database cleansing, *Data quality — The foundation of one-to-one customer relationships* (no date) proposes certain checks within the three categories of data quality (see Figure 3 on page 28):

- The **currency check** involves monitoring how often a database is updated, how comprehensive each update is, what happens to unused entities and records, etc.
- The **integrity check** is normally completed in six steps: parsing, correction, standardisation, enhancement, matching and consolidation. Each of these steps will be discussed shortly:

  1. The first thing to do when cleansing a database is to check the structure and content of each field. This activity is called **parsing** and makes sure that all the data appears where it should be—for example the address field should not appear in the name field, etc.
  2. Data **correction** verifies the accuracy of data. It checks whether postal codes exist, eliminates different spellings for the same item, etc.

3. **Standardisation** works in conjunction with correction and parsing to arrange data in a consistent format. Examples include removal of dashes from phone numbers; the use of common abbreviations like DEPT for *Department;* and accurate representation of titles (Mr., Ms., etc.).

4. **Enhancement** adds useful but optional data. This includes data that is not necessary to do core business, but can be used for other purposes such as marketing. Examples are demographic data, geographic data and psychographic data (hobbies, interests, etc.) in the customer database.

5. The first four steps ensure that all the necessary data is present. **Matching** then eliminates duplicates and redundancies. It searches through all existing records and removes all duplicate records or combines two incomplete records that refer to the same customer/item.

6. **Consolidation** is the final step of data cleansing and involves two processes. Firstly, all the data available on a given customer is combined; and secondly, this data is used to identify relationships between different customers.

- The **accessibility check** monitors how easy it is to communicate with the database and how easy it is for customers and administrators to use the database (making updates, searches, etc.).

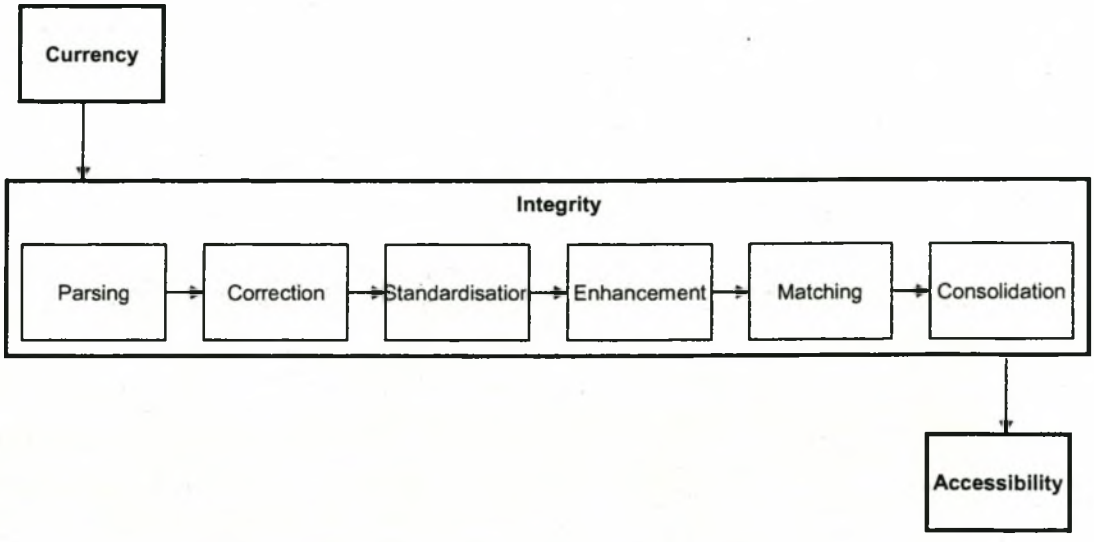


Figure 3 - Data cleansing process

This model and the data cleansing process are essential for the development of criteria to evaluate the databases that support e-commerce web sites. Although clean databases are definitely an important aspect of e-commerce web sites, it was decided not to include it in the evaluation framework. This decision was made because in order to evaluate the quality of a company's database it is necessary to have access to the actual database and its data

entries. For the purpose of web site evaluation by a third party, this is not possible because of privacy considerations and time constraints. Imagine how long it would take to assess the quality of the database of a major online retailer like Amazon.com! Instead, the framework does make provision for the evaluation of certain aspects of databases for which information is more readily accessible.

Database quality is, however, essential for the successful functioning of e-commerce web sites, and Internet companies should be encouraged to keep their databases clean.

2.4 Summary and next steps (evaluation data presentation)

This chapter provided an overview of available literature on the evaluation of general web sites and e-commerce web sites, as well as some important aspects of database cleansing. After reviewing the literature, it was important to decide how it was intended to gather evaluation data and in what form this data would be presented. This decision was essential at this point in the research because it is necessary to know the format in which data will be presented before a decision can be made on the correct technique to analyse the data.

In order to make this decision, it was necessary to know how the web site evaluation criteria were going to be developed. It was decided to develop a framework consisting of **three different levels of detail**. The complete framework and its development process are described in Chapter 4. The basic outline of the process given here is sufficient to make the appropriate decisions relevant to this section. Firstly, building on the literature reviewed, a few main **criteria categories** were identified to serve as a framework for e-commerce web site evaluation. Following this, a few **criteria groups** were developed within each category. For each of these criteria groups **individual evaluation criteria** were developed, with each criterion relevant to a specific criteria group (see Figure 4 on page 30 for a graphical representation of this concept).

The evaluation process works as follows. Using the framework, evaluation data for a given web site is gathered by going through the list of evaluation criteria and awarding a value to each criterion. This value is an indication of how well the web site adheres to the specific criterion. All the individual values for the criteria in a specific criteria group is then added and amounts to a unique score for that group. The criteria group scores within each category is then added, which generates a scalar of criteria category scores indicating how well a web site performs in each of the categories

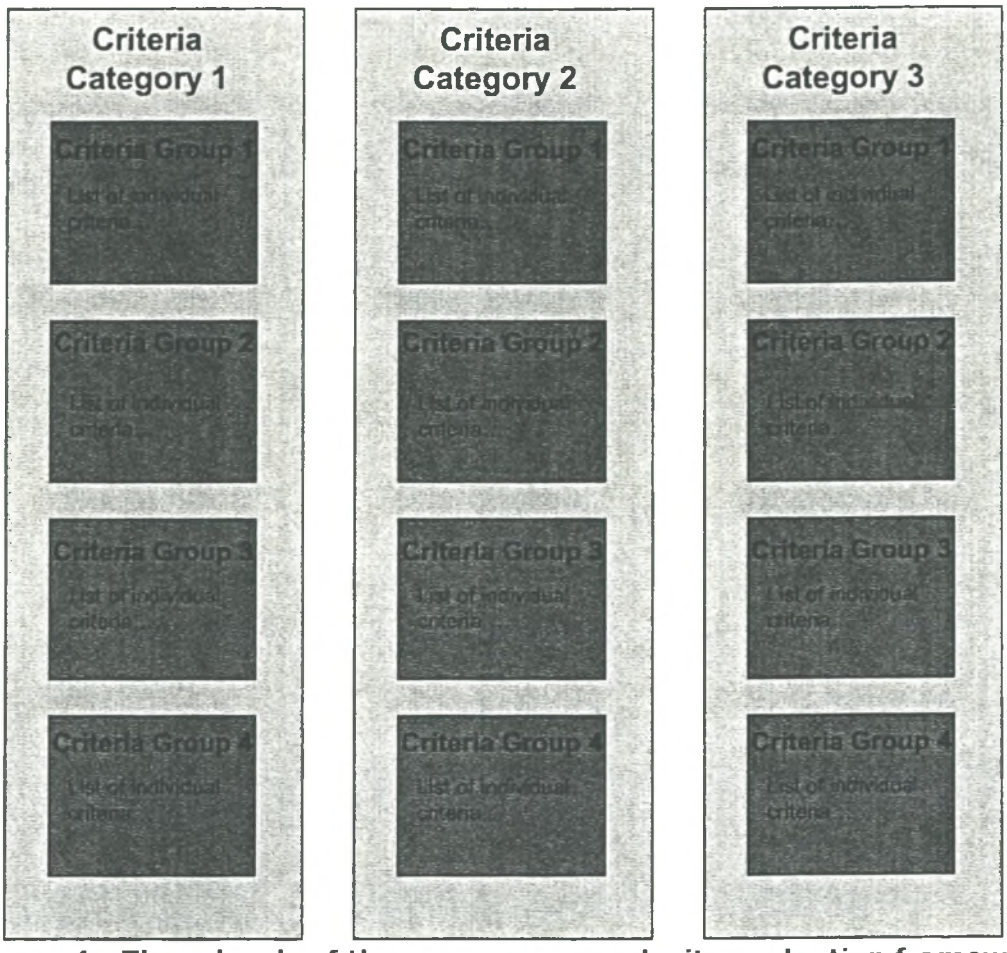


Figure 4 - Three levels of the e-commerce web site evaluation framework

As soon as more than one site is evaluated with this procedure, the data can be presented in tabular form. A matrix of values is the result, with the column variable consisting of criteria categories and the row vector consisting of different web sites. In this thesis, the rows in a specific matrix represent web sites within the same industry. Each different matrix consists of web sites in a different industry. The values in the cells of the matrix represent the score of each site for a specific criteria category, as calculated from the evaluation criteria. Matrices of this form are called *two-way contingency matrices*. An example of such a matrix is shown below:

	Criteria Category 1	Criteria Category 2	Criteria Category 3	Criteria Category 4	Criteria Category 5
Web Site A	67	78	63	78	23
Web Site B	87	74	48	93	67
Web Site C	78	64	49	56	83
Web Site D	37	57	85	73	65

Table 13 - Example of a two-way contingency matrix for web site evaluation

The next step in the thesis methodology is to find a statistical technique to analyse data that are presented in this form. As mentioned in the introduction, Prof. Pitt suggested Multidimensional Scaling as a group of statistical mapping techniques that can be used for this purpose. Because of Prof. Pitt's suggestion and his belief that this group of techniques

is an excellent tool for the graphical representation of data, the researcher focused only on Multidimensional Scaling in this thesis, and other mapping techniques were not investigated. The next section explores Multidimensional Scaling in detail and eventually focuses on *Correspondence Analysis* as the specific technique chosen to evaluate data in this thesis.

3. Multidimensional Scaling and Correspondence Analysis

This chapter discusses the group of statistical techniques called Multidimensional Scaling (MDS). As mentioned in the introduction, MDS is part of a group of statistical analysis techniques called *statistical mapping*. The chapter starts with a definition of MDS concepts as well as a brief history of its development. It then explores some of the most popular MDS techniques. The chapter continues with a detailed account of Correspondence Analysis and its mechanics, as well as the reasons why this specific technique was chosen for the thesis. The chapter concludes with a simple example illustrating the basic principles of Correspondence Analysis.

3.1 Definition and history

The purpose of MDS techniques is to condense large amounts of data to as few dimensions as possible, while still describing it well. Formally, MDS can be defined as **"[a set of mathematical techniques] concerned primarily with the spatial representation of relationships among behavioural data"** [Green et al., 1989:5]. To paraphrase this definition, MDS techniques seek to find an understandable visual representation of the similarities, dissimilarities or ordered rankings between different data points. All that is needed to perform MDS is a matrix of numbers that expresses all the similarities, dissimilarities or rankings within a group of objects—like the matrix discussed at the end of the previous chapter. The term *multidimensional* refers to the fact that different attributes or properties (dimensions) are used to compare objects among each other. These dimensions are mapped onto a single scaled dimension (hence the term *scaling*) so that it can be represented visually and therefore understood more easily.

The obvious first concern about these techniques is how accurate the representation is. This issue will be explored in detail when Correspondence Analysis is discussed[6].

Although authors disagree on how far the mathematical concepts behind MDS can be traced back in history, most sources agree that the roots of MDS as we know it today resulted from work done by Richardson (1938) and fundamental theorems proved by Young and Householder (1938). Authors like Klingberg (1941) continued the research, and in 1958 Torgerson (1958) came forward with the first well-known metric MDS approach. The terms

[6] See Step 6 in Section 3.3.2

metric and *non-metric* MDS, among others, will be discussed in the next section. Coombs (1950) and Bennet & Hays (1960) focused on the non-metric problem, but it is only in 1962 that Roger Shepard (1962) made a breakthrough in this field with the first conceptual paper and computer program using non-metric MDS methods. Since that time, progress in the field of MDS methods has been rapid. Important contributions included Carroll & Chang's (1970) INDSCAL program for fitting the weighted MDS model and Takane, Young & deLeeuw's (1977) ALSCAL program that combined metric, non-metric and weighted MDS into one algorithm for the first time. MDS techniques have come a long way since its inception in 1938; and as will be seen later in this document, it now finds new and exciting applications with the rise of the Internet.

3.2 Overview of MDS techniques

This section aims to give the reader an overview of the many different types of MDS techniques that have evolved over the years. It is not intended to give a thorough description of each method, as this document is primarily concerned with Correspondence Analysis. A holistic view of MDS scaling is important, however, and is therefore included here.

- With **fully metric scaling** methods "interpoint distances can be converted to a matrix of 'scalar products', which can be factored to give stimulus coordinates directly" [Morris et al., 1999]. For this method to be used, data must be in the form of ratio-scaled distances. This means that data is not only ordinal (ranked by order), but also scaled so that it quantitatively preserves the original similarity (or dissimilarity) between the objects. The principal task of fully metric methods is to take a square data matrix of interpoint distances (these distances can represent similarities or dissimilarities between objects) and find the configuration of points that best describe the data in a low-dimensional plane. Factor analysis is most often used as the computational method to find the best configuration of points.
- **Fully non-metric scaling** methods use data that only consist of ordinal relations and then they "find a solution space of minimum dimensionality that maintains the rank order of each point on each dimension in turn" [Green et al., 1989:15]. This range of methods does not produce a configuration of points that can be mapped, it only provides the rank order for each dimension identified. The disadvantage of these methods is that while the input data is only required to be non-metric, the output data is also non-metric, which makes it difficult to interpret.
- **Non-metric scaling** methods (as opposed to *fully* non-metric scaling) also use input data that is of ordinal nature, but the output information is metric and can therefore be

represented on a visual map. The objective of non-metric scaling is as follows: "given a rank order of proximity data, find a configuration whose rank order of ratio-scaled distances, in a specified dimensionality, best reproduces the original rank order of the input data" [Green et al., 1989:40]. It is therefore a very powerful set of techniques, as it only needs ranked similarities (or dissimilarities), which it then represents on a low-dimensional map.

- **Individual differences MDS** includes a variety of methods, the most important of which are *replicated MDS* and *weighted MDS*, to study the nature of individual differences between objects. *Weighted MDS* is important in the context of this thesis as it entails the weighting of each object's individual dimensions as an indication of how important the dimension is relative to the others. The underlying theory is used in Correspondence Analysis.
- There are a number of techniques that are less common, but also part of MDS. The most important of these are **Cluster Analysis** and **Correspondence Analysis**. *Cluster Analysis* is used in conjunction with other MDS techniques to further explain relationships between objects. *Correspondence Analysis*, the main focus of this chapter, displays patterns of association between the rows and columns of a data matrix.
- There have been quite a few **recent developments in the field of MDS**. These advances can be grouped into four main areas, namely *three-way unfolding models*, *stochastic MDS models*, *non-symmetric matrix models* and *hybrid models*. The interested reader is referred to a detailed overview by Carrol and Green (1997).

The next section describes the MDS method used in this thesis to represent web site evaluation data on a low-dimensional scale, namely Correspondence Analysis.

3.3 Correspondence Analysis (CA)

Correspondence Analysis (CA) originated with a French analyst called Jean-Paul Benzécri in the early 1960s. The original French term *analyse des correspondances* literally means *analysis of associations*, and the technique was used to analyse associations between rows and columns in a matrix. For the English term, the French word *correspondance* was wrongly translated as *correspondence*, which changed the original meaning. The correct translation as it was intended in its original context would be *associations analysis*, and it is important to keep this in mind throughout the discussion of the technique.

An important question is why Correspondence Analysis was chosen to work with the web site evaluation framework and criteria that will be developed in this thesis. This question will now be addressed before the technique is discussed in detail.

3.3.1 Definition and motivation for use in this thesis

Correspondence Analysis can be defined as "a technique for displaying the rows and columns of a data matrix (primarily a two-way contingency table) as points in dual low-dimensional vector spaces" [Greenacre, 1984:54]. The power of the technique is that it provides a way to examine relations not only among row variables and column variables individually but also between row *and* column variables together. This implies that, if used in this thesis, it will be possible to compare the similarities that different web sites have with each other or the similarities that different criteria groups have with each other. But more importantly, it will also be possible to compare individual web sites on the basis of the different criteria groups. This is one of the main objectives of the thesis.

This technique has several advantages, the most important of which, according to Hoffman (1986), are as follows:

- CA can reveal relationships between variables that would not be possible with a few pairwise comparisons of the variables.
- CA also shows *how* variables are related, not merely that a relationship exists.
- The fact that rows and columns are displayed on the same graphical space makes it possible to get a large amount of relationship information from a fairly simple graphical representation.
- The only data requirement CA has is a rectangular matrix with non-negative values, which makes the technique highly flexible in that respect.

As explained at the end of chapter 2, evaluation data will be displayed in a two-way contingency table with rows consisting of different web sites and columns displaying the different criteria categories on which the sites are evaluated. CA will be able to examine the data and plot the complex relationships between web sites among themselves, criteria among themselves, and sites and criteria among each other, on a low-dimensional scale (usually 2 dimensions).

From the above discussion it is clear that the reason why CA was chosen as the technique to examine data in this thesis is because of its powerful ability to analyse the complex relationships that are needed, and represent the results graphically. The rest of this section will describe how CA accomplishes this, and also provide a simple example of the technique.

3.3.2 Correspondence Analysis step-by-step

The following is a step-by-step explanation of how to perform CA on an existing data matrix. The generation of the matrices for this thesis will be discussed at a later stage. Although the basic process of CA is fixed, authors differ slightly on some minor points of the technique. The methodology presented here is based primarily on the work of Weller & Romney (1990). Please note that the notation used for the different matrices and computations will be used throughout this document.

STEP 1: CHECK IF THE APPARENT RELATIONSHIP IS STATISTICALLY SIGNIFICANT

With a contingency matrix (consisting of positive values only) set up, the first task is to test for data independence. Data independence will indicate significant differences between the row values of a column and the column values of a row. If not, there is no need to perform CA on the data, as there would be no significant relationships to describe. The traditional Pearson Chi-Square statistic with a predetermined significance level is most commonly used to test for independence. The following procedure is followed:

- **State H_0.** In this case, the null hypothesis is that the relationships between the row and column variables in the contingency matrix are *not* statistically significant. The aim of the test is thus to reject this hypothesis.
- **Compute expected frequencies.** For each cell in the table, the expected frequency must be calculated. The expected relative frequency for a given column is the sum of that column's values divided by *n*, the sum of all entries in the table. Likewise, the expected relative frequency for a given row is that row total divided by *n*. The expected relative frequency for a cell is then the expected column value times the expected row value. The expected cell frequency is simply the expected relative cell frequency multiplied by *n*. This formula reduces to the expected frequency for a given cell equalling its column total times its row total, divided by *n*:

$$\text{expected relative frequency for cell ij} = \frac{\text{column total}_j}{n} \times \frac{\text{row total}_i}{n}$$
$$= \frac{(\text{column total}_j) \times (\text{row total}_i)}{n^2}$$

$$\text{expected frequency for cell ij} = \text{expected relative cell frequency} \times n$$
$$= \frac{(\text{column total}_j) \times (\text{row total}_i)}{n^2} \times n \qquad (3.1)$$
$$= \frac{\text{column total}_j \times \text{row total}_i}{n}$$

- **Apply general Chi-square formula**. The Chi-square value for the table is computed as follows:

$$\chi^2 = \sum_{cells} \frac{(\text{observed frequency - expected frequency})^2}{\text{expected frequency}} \qquad (3.2)$$

- **Calculate the degrees of freedom and critical value**. The chi-square value is not interpretable directly but must be compared to a critical value found in a table of the Chi-square distribution. This table shows critical values for alternative significance levels (α=.001, .01, .05, etc.) and the corresponding degrees of freedom (df). For a two-way contingency table, the following formula apply to compute degrees of freedom:

$$\text{df} = (r-1)\times(c-1), \qquad r = \text{number of rows}, c = \text{number of columns} \qquad (3.3)$$

  The critical chi-square value is then read from the table for a specific significance level and the calculated degrees of freedom.

- **Compare critical value with calculated statistic**. The calculated chi-square value must be greater than the critical value to reject the null hypothesis that the row variable is unrelated to the column variable at the level of significance selected. If this is the case, there is evidence to believe that a significant relationship exists between the row and column variables, and the process can continue to step 2 below. Otherwise the process is terminated.

STEP 2: NORMALISE THE DATA

Suppose a data matrix **F** exists with entries f_{ij}, where f_{ij} refers to an entry in row i and column j. A normalised data matrix **H** is then calculated by dividing each entry by the square root of the product of the corresponding row and column totals:

$$h_{ij} = \frac{f_{ij}}{\sqrt{f_{i.} \times f_{.j}}} \qquad (3.4)$$

where $f_{i.}$ and $f_{.j}$ respectively denotes the row totals and column totals of matrix F.

STEP 3: FIND THE SINGULAR VALUE DECOMPOSITION (SVD) OF THE NORMALISED MATRIX

This step finds the basic structure of the normalised matrix **H** using Singular Value Decomposition (SVD). The SVD of **H** produces a diagonal matrix **S** and matrices **U** and **V** so that $\mathbf{H} = \mathbf{U}\text{x}\mathbf{S}\text{x}\mathbf{V}^{\mathbf{T}}$. SVD separates the rows from the columns in the matrix and puts them into separate matrices (**U** and **V** in this case) so that it can be compared among and between each other. The geometric correspondence between the two sets of points in

these matrices make it possible to merge them into one joint display in 2 dimensions. Most computer software that deals with matrices (e.g. Matlab) has built-in functions to compute SVDs. For a discussion of the complex algebra involved in SVD, the reader is referred to Greenacre (1984:340-351).

STEP 4: RESCALE ROW AND COLUMN MATRICES OF SVD TO OBTAIN ULTIMATE SCORES

The row matrix **U** and column matrix **V** are used to plot the data. However, they must first be rescaled and weighted. The rescaled row and column matrices **X** and **Y** containing the optimal scores of the matrices are computed using the following formulae:

$$X_i = U_i \times \sqrt{f_{..}/f_{i.}}$$
$$Y_j = V_j \times \sqrt{f_{..}/f_{.j}} \qquad (3.5)$$

where $f_{..}$ denotes the total sum of all the entries in matrix F.

The values of the first columns in these two matrices are always 1, and serve as a check on the calculations. This column is commonly referred to as the *trivial dimension* and is ignored when the data is analysed further.

STEP 5: WEIGH THE SCORES AS A FUNCTION OF THE SIZE OF THE SINGULAR VALUES

In order to compare scores within and between rows and columns, they must also be weighed as a function of the size of the singular values. This is done so that the scales of the row vector **X** and the column vector **Y** are the same, because then they can be represented on the same graphical scale. The most common method used for weighing is to **multiply each unweighted value in the matrix with the square root of its corresponding singular value**. The resulting matrices are called **A** and **B**, and they are used directly to plot the original data.

STEP 6: COMPUTE THE INERTIA OF EACH DIMENSION (AXIS)

It is important to remember that this is multidimensional data, which is scaled to only two dimensions in order to represent it graphically. It is inevitable that some of the information portrayed by the original data will be lost. It is therefore important to know how accurate the representation is, and how much each of the two dimensions contributes to accuracy. This is accomplished by computing the so-called *contributions to inertia* of each dimension. A detailed discussion of *inertia* can be found in Greenacre (1984:66-70). For the purpose

of this document it is sufficient to give the formula as it is used to compute the contribution to inertia of each dimension (excluding the trivial dimension). For each dimension:

$$\text{Contribution to inertia} = \frac{(\text{singular value of that dimension})^2}{\sum_{all}(\text{singular values of all dimensions})^2} \quad (3.6)$$

These singular values are found in the matrix **S** computed in step 3.

STEP 7: PLOT DATA ON A 2-DIMENSIONAL PLANE

The last step is to use the data in matrix **X** and **Y** to plot each variable's coordinates on a 2-dimensional plane. The calculation procedure will become clear in the next section, where an example of a simple CA will be shown. Although the data in the example is not relevant to this thesis, it was chosen because of its simplicity in order to make clear the basic procedure for performing CA.

3.3.3 An example of CA

This example uses data found in Cox & Cox (1994:127), which deals with malignant melanoma, a dangerous type of skin cancer. The data matrix shown below represents the site of the tumour and its histological type for a sample of 400 patients.

	Site of tumour			
Histological type	**Head, neck (h)**	**Trunk (t)**	**Extremities (e)**	**ROW TOTALS**
Hutchinson's melanotic freckle (H)	22	2	10	**34**
Superficial spreading melanoma (S)	16	54	115	**185**
Nodular (N)	19	33	73	**125**
Interminate (I)	11	17	28	**56**
COLUMN TOTALS	**68**	**106**	**226**	**400**

Table 14 - Malignant melanoma data for CA example

This data was entered and analysed with Matlab and Microsoft Excel. Detailed calculations and Matlab program code are shown in Appendix C. The following initial two-way contingency matrix was obtained:

```
F =
    22     2    10
    16    54   115
    19    33    73
    11    17    28
```

STEP 1: CHECK IF THE APPARENT RELATIONSHIP IS STATISTICALLY SIGNIFICANT

For this matrix, the calculated chi-square value is $\chi^2_{calc} = 65.813$. The critical chi-square value for (4-1) x (3-1) = 6 degrees of freedom and a chosen significance level of α=0.001 is $\chi^2_{crit}(6,0.001) = 22.457$. As 65.813 > 22.457, the null hypothesis is rejected, and there is no evidence that a significant relationship does not exist between the row variables and column variables. We may thus proceed to the next step.

STEP 2: NORMALISE THE DATA

Equation (3.4) was applied to matrix $\mathbf{F}$, and the following normalised matrix followed:

```
H =
    0.4575    0.0333    0.1141
    0.1427    0.3856    0.5624
    0.2061    0.2867    0.4343
    0.1783    0.2206    0.2489
```

STEP 3: FIND THE SINGULAR VALUE DECOMPOSITION (SVD) OF THE NORMALISED MATRIX

Matlab performs SVD easily with the *svd* command. The following matrices were the result, so that $\mathbf{H} = \mathbf{U}x\mathbf{S}x\mathbf{V}^T$:

```
U =
    0.2915   -0.9233   -0.1722
    0.6801    0.3742   -0.2434
    0.5590    0.0646   -0.2336
    0.3742   -0.0571    0.9255

S =
    1.0000         0         0
         0    0.4029         0
         0         0    0.0468

V =
    0.4123   -0.9083    0.0709
    0.5148    0.2964    0.8044
    0.7517    0.2952   -0.5898
```

STEP 4: RESCALE ROW AND COLUMN VECTORS OF SVD TO OBTAIN ULTIMATE SCORES

After rescaling, the following matrices were obtained through equation (3.5):

```
X =
    1.0000   -3.1670   -0.5908
    1.0000    0.5502   -0.3578
    1.0000    0.1155   -0.4178
    1.0000   -0.1527    2.4735
```

```
Y =
    1.0000   -2.2029    0.1718
    1.0000    0.5758    1.5627
    1.0000    0.3927   -0.7846
```

Note that the first column of each matrix is made up of 1's, which is called the *trivial dimension*, as mentioned earlier.

STEP 5: WEIGHT THE SCORES AS A FUNCTION OF THE SIZE OF THE SINGULAR VALUES

The matrices X and Y were weighted as discussed previously and so the final matrices were obtained:

```
                x-values  y-values
A =
    1.0000   -2.0103   -0.1278    H
    1.0000    0.3493   -0.0774    S
    1.0000    0.0733   -0.0904    N
    1.0000   -0.0969    0.5352    I

B =
    1.0000   -1.3983    0.0372    h
    1.0000    0.3655    0.3381    t
    1.0000    0.2493   -0.1698    e
```

These matrices are used to plot the data. Column 1, the trivial dimension, is left out at this stage and is not used further. Column 2 denotes the first dimension of the plot (referred to as **Dim 1**), and Column 3 is the second dimension of the plot (referred to as **Dim 2**). Each of these dimensions is used to describe the data in the original contingency matrix, and its interpretation is explained in the next section. Each row of each matrix represents an entry in one of the original table's (Table 14) dimensions. The values in Column 2 and Column 3 of each row are used as the X-values and Y-values respectively for each point on the plot. In this case Matrix A plots the *Histological type* data, and Matrix B plots the *Site of tumour* data.

This particular matrix has only 3 columns of which only the second and third is used (i.e. two dimensions, **Dim 1** and **Dim 2**). We therefore expect that no data will be lost in two graphical dimensions, and that the total contribution to inertia of the two dimensions will be 100%. More columns would imply that more dimensions would be needed to describe the original data in its entirety (in other words, **Dim 3**, **Dim 4**, etc.). It becomes clear here that (c-1) dimensions are needed to ensure that the data is represented perfectly (because the first column is the trivial dimension not used). A matrix with 5 columns will therefore need 4 dimensions to ensure that no data is lost. CA aims to describe as much of the data as possible in as few dimensions as possible, so that 2 dimensions will describe enough of the data to make it sufficiently accurate.

STEP 6: COMPUTE THE INERTIA OF EACH DIMENSION (AXIS)

Using equation (3.6), the following calculations can be made. The trivial dimension is left out, and the singular values for each dimension are as follows (from matrix S):

Dim 1: 0.4029
Dim 2: 0.0468

Then,

$$\text{Dim 1 contribution to inertia} = \frac{0.4029^2}{0.4029^2 + 0.0468^2} = 0.987 = 98.7\%$$

$$\text{Dim 2 contribution to inertia} = \frac{0.0468^2}{0.4029^2 + 0.0468^2} = 0.0133 = 1.3\%$$

As **Dim 1** describes 98.7% of the data on its own, it is almost sufficient to describe all the data in its entirety, with **Dim 2** making a small contribution to inertia.

STEP 7: PLOT DATA ON A 2-DIMENSIONAL PLANE

The final step is to plot the data. This plot was done using Microsoft Excel, and is shown in Figure 5.

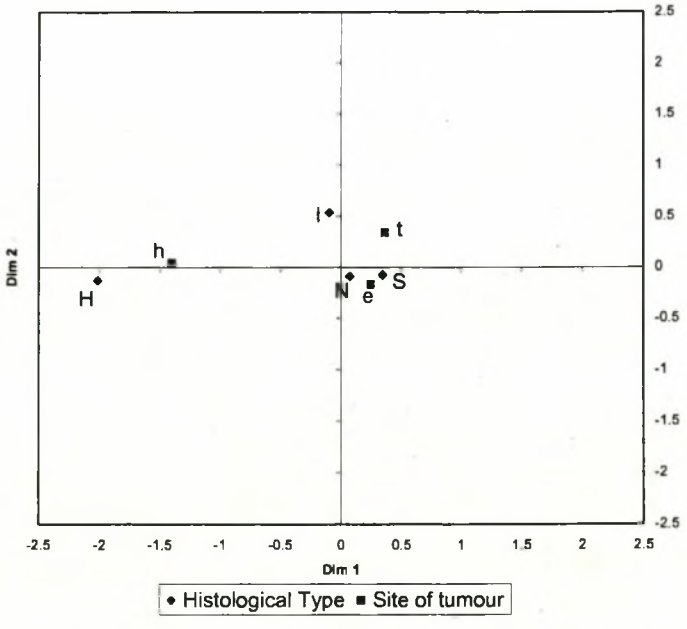


Figure 5 - Correspondence Analysis Graph of cancer data

3.3.4 Interpretation of CA output

Undoubtedly the most important part of Correspondence Analysis is the interpretation of the output data. As mentioned earlier, the power of CA lies in its easy interpretation because of the graphic nature of its output. When interpreting specific points on a plot, there are two main aspects to consider, namely the **position of points relative to each axis** and the **proximity of the points to each other**. Both of these aspects will now be discussed using the graph from the previous section (See Figure 5) as an example.

Position of points relative to each axis

It is firstly important to understand which points are contrasted with each other by each axis. The X-axis (Dimension 1) and the Y-axis (Dimension 2) independently contrasts the profiles of different points with each other[7]. The functions of the axes differ for each unique plot, but it is easy to identify. The axis for Dimension 1 mainly contrasts points that have the most negative X-values with the points that have the largest positive X-values, because these are the points that contribute the most to inertia. These points on the extremes of the axis have profiles that are the most different from each other[8]. The points situated between the extreme groups, especially those near the origin, contribute very little to inertia, and usually indicate rows with an even distribution of values in the Contingency Matrix. The process is the same for points on Dimension 2, but in this case points that have the most negative Y-values are contrasted with those that have the largest positive Y-values.

It is also important to note that both the row profiles and the column profiles of the contingency matrix are considered when the functions of the axes are identified. This point can be illustrated in the example where it is clear that Dimension 1 firstly contrasts *Hutchinson's melanotic freckle* (H) with the other three histological types (S, N and I) for the row profiles, and secondly it contrasts the *Head and neck* (h) with the *Trunk* and *Extremities* (t, e) for the column profiles. This conclusion is also clear from the contingency matrix data where it is evident that the row profiles that are contrasted with each other have very different profiles. The same is true for the column profiles. In this example,

---

[7] Both axes are later considered in unison when the proximities between all points are considered. As a start, however, the function of each axis is identified independently.

[8] When the profiles of two groups are different it means that the values in the contingency matrix are significantly different for each group of points. Individual points in the same group have similar values in the matrix, relative to the other values in the specific point's row.

Dimension 2 contributes very little to the total inertia of the plot, as it is clear that the points are mainly scattered on the first dimension.

When the column and row profiles are considered together, two main groups can be identified. It seems that *Hutchinson's melanotic freckle* (H) and *head and neck* (h) often occur together, while the rest of the histological types often occur on the *trunk* (t) and *extremeties* (e). This conclusion will be confirmed and interpreted in the next paragraph.

Proximity of the points to each other

The second aspect of CA graph interpretation involves the proximities of the points to each other. The proximities are an indication of how similar column profiles and row profiles are individually among themselves, as well as how similar the column and row profiles are between each other. A short distance between points implies similar profiles and a great distance means a significant dissimilarity. Analysis of the proximities should be done in three phases, namely *analysis of column profiles*, *analysis of row profiles*, and *analysis of column and row profiles together*. A discussion of each of these phases, as applied to the example, follows:

- The **analysis of column profiles** is a measure of the similarities of the column profiles (site of the tumour) with regards to the row profiles (histological cancer type). As *h* is far away from the other two sites, this implies that the head and neck (h) has a much different profile as that of the trunk (t) and extremities (e). On inspection of the data it can be seen that this is in fact true, as the head and neck (h) is a common site for all four the tumours, while the other two are more susceptible to one or two specific tumours.
- The **analysis of row profiles** is a measure of the similarities of the row profiles (histological type of cancer) with regards to the column profiles (site of the tumour). It shows that Hutchinson's melanotic freckle (H) stands on its own while the other three cancer types are more or less grouped together. It therefore seems that *H* mainly occurs in different sites than the other three do.
- The **analysis of column and row profiles together** link these two analyses and is a measure of the similarities between the column profiles and row profiles, in this case the *site of the tumour* and the *histological type*. Therefore, the proximity of *H* to *h* leads to the conclusion that Hutchinson's melanotic freckle (H) occur mainly in the head and neck (h), and not so much in the other sites. It is also clear from the other proximities that Nodular (N), Superficial spreading melanoma (S) and Interminate (I) cancer mainly occur at the extremities (e) and near the trunk (t). These three cancer types are very far away from *h*, which implies that they do not occur on the head and neck (h) often in proportion to their total amount of occurrences on the body. From the

perspective of the site of the tumour, it is clear that the extremities (e) are particularly susceptible for Nodular (N) and Superficial spreading melanoma (S) cancer.

Although these conclusions can be made intuitively from the data, this is only an illustrative example to show how the theory works in practice. Larger matrices as the ones used in this thesis are not so easy to interpret, and for that purpose CA is an ideal tool to simplify the interpretation.

3.4 Summary

This chapter has provided an overview of Multidimensional Scaling and has also shown how Correspondence Analysis is a suitable tool to evaluate the complex data matrices that result from the evaluation of e-commerce web sites. It described the basic theoretical principles of the technique as well as a simple illustrative example to clarify the calculations.

This literature review aimed to give the reader an understanding of all the tools that are needed to fulfil the objectives of the thesis set in chapter 1. The theoretical principles established therefore conclude part 1 of this thesis. The second part firstly shows the reader how the e-commerce evaluation framework and evaluation criteria were developed. It then continues to show how Correspondence Analysis can be used to analyse the data and represent the results graphically. A method for the comprehensive evaluation of e-commerce web sites is subsequently developed. In closing it describes how this method was implemented in Microsoft Visual Basic.

Part 2

Development of a computerised e-commerce web site evaluation method

4. Development of e-commerce web site evaluation method

This chapter describes the development of an evaluation method for e-commerce web sites. As mentioned in section 2.4, it was decided to design criteria for the evaluation of individual web sites in three different levels of detail. The **first level** consists of **five criteria categories** to serve as a broad framework for evaluation, covering all the relevant aspects of e-commerce web sites. The **second level** involves **four criteria groups** within each of the criteria categories, which provide a little more detail and definition to the categories. On the **third and most detailed level**, **five individual criteria** were developed within each of the criteria groups.

The chapter starts with a discussion of the development of the criteria categories and the web site evaluation framework that resulted from it. It also reports on the development of the different criteria groups within each of the categories. The next section continues with a thorough explanation of the development of the individual criteria within each of the groups. In closing the chapter integrates the framework and criteria with the use of Correspondence Analysis and describes the three steps of the web site evaluation method that were developed.

4.1 E-commerce web site evaluation framework

Before an evaluation framework could be developed, it was important to find sound business principles to build it upon. The most important criterion for the development of this framework was **comprehensiveness**. The purpose of the framework and its criteria categories is to find a way to evaluate e-commerce web sites in such a comprehensive and thorough manner that no important aspect of the site is left out of the evaluation. To fulfil this purpose it was decided to use the basic **customer buying cycle** as a framework for e-commerce web site evaluation. This was done because of two reasons:

- **Comprehensive evaluation of all aspects of e-commerce web sites**. By following every action a customer will perform on an e-commerce web site when a purchase is made, the buying cycle will ensure that all the relevant parts of the web site are evaluated.
- **Total customer focus**. The customer buying cycle puts the focus on the processes customers follow when they pay for products/services. They are the users of the web sites, and therefore any evaluation framework must focus specifically on their experience of the sites.

There are several approaches to customer buying processes in marketing literature[9]. A basic thread runs through these approaches, and for the purpose of this thesis the customer buying cycle was defined as consisting of four distinct phases. These phases are given below with an appropriate example to explain each phase:

1	**Need recognition**	A potential customer realises a specific need, for example, he would like to read a book on a certain subject, say aviation in World War II.
2	**Gather information**	The customer goes to a bookstore and finds all the resources on aviation during the Second World War.
3	**Evaluate information**	The customer reads the summaries of the books on the subject, and decides which of the resources he is going to buy.
4	**Make purhase**	The customer goes to the checkout counter and buys the book he selected.

Table 15 - Four phases of the customer buying cycle

To use this buying cycle as a framework for the evaluation of e-commerce web sites, it is necessary to find links between each of the buying phases and a specific aspect of a web site that is most important during that phase.

Figure 6 shows how these links were made and five criteria categories were identified as a result. An explanation follows after the figure.

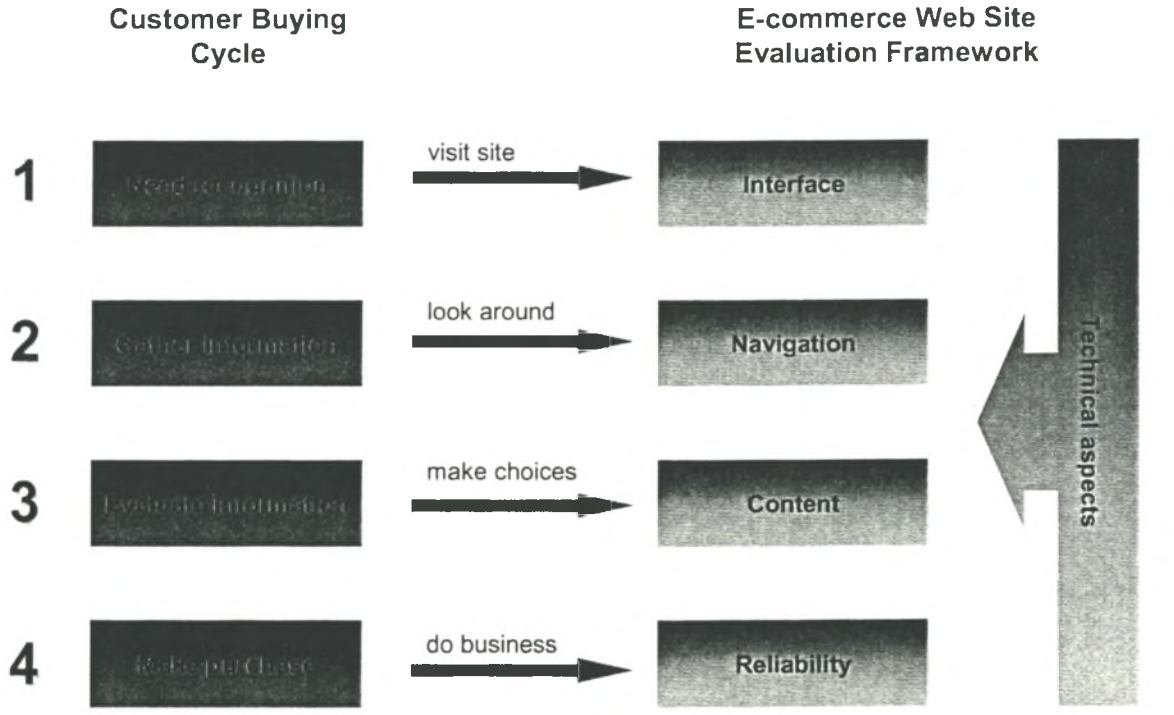


Figure 6 - Development of e-commerce web site evaluation framework

[9] Examples are the *new customer buying phases* by Robinson et al. (1967) and the *complex consumer buying process* by Schoell and Guiltinan (1992).

During the **need recognition** phase, potential customers will visit a web site that they believe will be able to fulfil their specific needs. The **interface** is very important at this stage, because this determines the user's first impression of the site. The interface refers to the visual aspects of the site—how attractive it is visually, which backgrounds, fonts and colours are used, etc.

During the **information-gathering** phase, customers explore the site and try to find the products/services they are interested in. During this phase, **navigation** is the most important aspect of the site. Navigation refers to the process of moving around between the different pages of the web site. Aspects that are important in this phase include the straightforward, logic structure of the site, and menus that are easy to understand.

During the third phase of the buying cycle (**evaluate information**), customers make choices on which products/services they want to purchase. The **content** of the site is the most important aspect of the site in this phase. Content refers to the actual information on the site. This phase covers aspects like the quality of the information and the amount of information on the site.

In the last phase of the buying cycle (**make purchase**), customers purchase the products/services they have chosen. The **reliability** of the site is extremely important at this stage. For the purpose of this thesis, reliability is defined as consisting of two aspects:

- The degree to which a customer is able to use the order process on the site easily and effectively.
- The degree to which the company is able to fulfil its promises to customers every time a purchase is made.

Finally, several **technical aspects** of a web site are essential for its proper functioning. As Figure 6 shows, these aspects are important during all the phases of the buying cycle. It is the backbone of the site on which the other phases are built.

With the evaluation framework and its five criteria categories fixed, the next step in the process was to find criteria groups within each category. These groups were designed by combining literature on web site evaluation with the researcher's own experience with e-commerce web sites. A total of 4 groups per criteria category were identified, as shown in Figure 7.

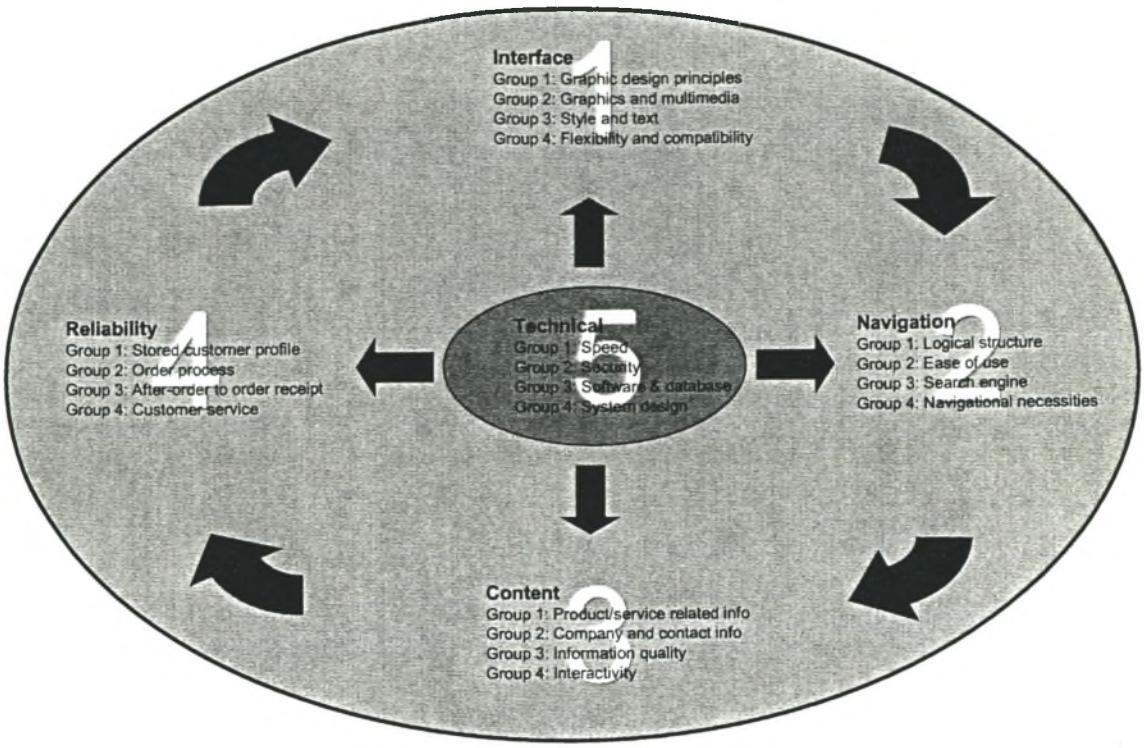


Figure 7 - E-commerce web site evaluation framework and criteria groups

Each of these groups is explained below in Table 16.

Phase	Criteria Group	Description
Interface	Graphic design principles	Evaluates the effective use of colour, text, backgrounds, and other general graphic design principles.
	Graphics and multimedia	Evaluates the effectiveness of the graphics and multimedia used on the site.
	Style and text	Evaluates whether or not the text is concise, relevant and in good style.
	Flexibility and compatibility	Evaluates the degree to which the interface is designed to handle exceptions, for example text-only versions of pages.
Navigation	Logical structure	Evaluates the organisation and menu system of the site.
	Ease of use	Evaluates the ease of navigation in finding the pages that the user is looking for.
	Search engine	Evaluates the search engine's ability to find the correct pages easily and provide with clear descriptions of the search results.
	Navigational necessities	Evaluates other important aspects of navigation like the absence of broken links and under construction pages.
Content	Product/service related info	Evaluates whether or not the products/services are described precisely and thoroughly.
	Company and contact info	Evaluates whether or not it is easy to find information on the company, its employees and its principles.
	Information quality	Evaluates the currency and relevancy of the content on the site.
	Interactivity	Measures how much input the user has on the content displayed on the site.
Reliability	Stored customer profile	Evaluates the registering process and how the company uses the stored customer profile.
	Order process	Evaluates the effectiveness and ease of use of the online order process.
	After-order to order receipt	Evaluates the company's actions from order placement until the order is delivered.
	Customer service	Evaluates how the company communicates and helps its online customers.
Technical	Speed	Evaluates different aspects of the loading speed of the site.
	Security	Evaluates security systems and the ways used by the company to protect customers' privacy on the site.
	Software & Database	Evaluates flexibility in terms of different software used. Also looks at the data software and data communication systems used on the site.
	System design	Evaluates the correct functioning of the site and how well it integrates with internal and external systems.

Table 16 - Description of e-commerce web site evaluation criteria groups

The next step in the development of the evaluation method was to find specific criteria within each criteria group. The next section discusses the criteria that were developed.

4.2 E-commerce web site evaluation criteria

With the 20 criteria groups described above as a guideline, 5 criteria were developed within each of the groups. This means that a total of 100 criteria were eventually developed to evaluate e-commerce web sites. These criteria were again developed through thorough research in current literature, as well as personal experience of the researcher with regards to web sites and what a successful web site constitutes of. A graphical representation of the 100 criteria is shown in Appendix D.

A note on the **subjectivity of the criteria** is in order here. An attempt was made to develop each criterion in such a way that it will require the least amount of subjectivity from the evaluator. Web sites are still used by *people* though—which naturally makes it a subjective experience for users. *People* still have to evaluate the sites, which implicitly brings a certain amount of subjectivity to the criteria. To bring the amount of subjectivity down to a minimum, the use of a **group of evaluators** is suggested. The details of this recommendation are discussed in the next section.

Each of the 100 criteria is used in some way to evaluate the effectiveness of an e-commerce web site. The evaluators will explore the site and assess to what extent it follows the guidelines set by each criterion. This process will be discussed in the next section. A detailed explanation of the criteria is provided on the following pages to show exactly what evaluators must look for on a web site.

1 Interface	
Group 1: Graphic design principles	
Home page concise and clear	The home page should not be so cluttered with too much text and graphics that it becomes difficult to have an immediate understanding of the site, its purpose and how to navigate it.
Effective use of white space	There should be enough open space between text and graphics to give the site a clean feel.
Effective and consistent use of colour	The colours used on the site should adhere to basic graphic design principles and be consistent throughout the site. For example, pink and green backgrounds will not be used because this combination will surely be bad for the user's eyes.
Effective and consistent use of backgrounds	Backgrounds should be simple and not graphic-intensive (a photograph used as a background is generally not a good idea because it drastically influences the download speed). The backgrounds should be consistent throughout the site, except when it is explicitly clear why different backgrounds are used for different pages.
Effective graphics/typeface/colour combinations	Text should be easily readable on the chosen background. Basic common sense should be enough to evaluate this criterion – for example blue text on a pink background is obviously not a good choice. Although considered boring by many, black text on a white (or near-white) background is still the most readable option.
Group 2: Graphics and multimedia	
Site visually attractive	Although fairly subjective, this criterion is easy to evaluate. A 'first impression' approach should be followed. A first, quick glance at the site should be enough to decide whether the site can be considered visually attractive or not.
Graphics and multimedia make a contribution to the understanding and navigation of site	Multimedia like flash animations and java scripts should not only look good, it should have a definite function to make the site more understandable and usable.
Icons easy to understand	Icons (like the graphics used to indicate the user's *shopping basket* or *account info*) should be clearly defined and easy to recognise.
Not excessively used	Too many pictures and multimedia on sites make them slow and difficult to use.
Size of media no negative impact on loading times	Check the size of the graphics and multimedia on the site and use this to determine whether it makes the site slower to load. If necessary, save pictures or backgrounds to the local hard drive to check its size.
Group 3: Style and text	
Style of pages consistent	The general feel of the site should not be different on every page. This only confuses users. The home page should have a clear theme and this theme should be the same throughout the site.
Typefaces consistent and easy to read	Text should be easy to read, and as far as possible the same typeface should be used throughout the site.
Good spelling and grammar	Spelling and grammar mistakes on a site do not leave a good impression.
Text concise and relevant	Users do not want to read unnecessary text on their computer screens. The text should only say what is necessary for the user to understand the site better. When large blocks of text are inevitable, white space should be used effectively to make it easier to read.
Purpose of site made clear on home page	The home page should state clearly what the site is supposed to do for the user. For example, an online bookstore should at least say that it sells books on the home page, and provide the user with an overview of what he/she can expect on the site. Online banking sites should immediately provide the user with information on the services available on the site.
Group 4: Flexibility and compatibility	
Pages sized to fit in browser window	Pages should fit into the screen so that horizontal scrolling is not necessary, and vertical scrolling is kept at a minimum.
Printable versions of certain pages available	For graphic-intensive pages the user should be able to click on a 'printable version' link where the document is sized and structured for printing
Text-only version available	Users should be able to request a text-only version of the site (or parts of it) so that graphics need not be downloaded.
Foreign language support available	Sites providing foreign language support (for example French or Japanese editions of the site) have a considerable advantage over sites that do not provide this service.
Accommodations made for disabled users	Large-text versions or sites providing audio transcripts of certain pages can be very useful and a huge competitive advantage.

Table 17 - 100 Evaluation Criteria discussed

2 Navigation	
Group 1: Logical structure	
Intelligible, straightforward organising scheme	It must be possible to assess the structure and layout of the site at first glance.
Content logically structured in different sections and levels	The navigational system should make logical distinctions between the different pages on the site, and the pages should be structured in different levels of detail.
Menus understandable and straightforward	The menu system used for navigation should not confuse the user. It should be clear what the purpose of each link is, and the system should be consistent throughout the site.
Site map/table of contents available	The user must be able to get an overview of all the pages/sections on the site.
Consistent navigation throughout site	Menus should be situated at the same place on all pages. The same menu should appear on all pages, except for large sites with many different sections. In this case each section's menu should be clearly identified and easy to use.
Group 2: Ease of use	
Easy to find site	It must be easy to find the URL of the site when a search is done on popular search engines like [www.google.com] or [www.yahoo.com].
Easy to explore specific idea or subject	It must be easy to use the navigational system to explore a specific need systematically. For example, when a user wants to know what the most popular books on tap dancing are, the navigational system should guide him to explore this subject.
Easy to return to main page	Each sub-page must have a link to the home page. This link should be easy to find, and always on the same place.
Easy to find specific information	When very specific information is needed, for example the pricing structure for the delivery of overseas packages, it should be easy to know exactly which path to follow in the navigational system to get this information.
Easy to access complete product/service range	It should be easy to find complete information on all the products/services offered on the site, including pricing information.
Group 3: Search engine and help function	
Easy to use search engine	The search engine must be easy to find and it should be easy to enter search criteria. It is also useful if users can either press 'Enter' or click on the 'Search' button to start the search.
Search engine accurate	The pages to which the search engine refers the user should be mostly useful and on the topic.
Good description of search engine findings	It shouldn't be necessary to click on a specific link to find out where it leads. The description of the link should be sufficient for the user to know whether the specific search result is relevant or not.
No search engine errors	No error messages should be displayed when a search is performed. This includes 'server busy' messages or programming errors on the server side. Also, when no results for a specific query can be found, the user should be clearly informed of this fact.
Help function easy to use	It should be easy to request help on a specific aspect of the site. Specific e-mail addresses or feedback forms should be readily available for this purpose.
Group 4: Navigational necessities	
No broken links	There should be no errors when a specific link is clicked.
No 'under construction' pages	No pages should display 'under construction' messages. Pages that are still under construction should not appear on the navigation menu.
Links clearly discernable, well labelled and defined	The specific names given to links should be easy to understand, and the user should know exactly where the links would take him/her. Links should also be clearly indicated as links, either by underlining it or by some other form of consistent identification.
Clear label of current position on site	Each page should have a clear indication of its exact position in the structure of the site. The user should know at which level he finds himself and how to get back to the previous page/level.
Effective use of frames, non-frames version available	When frames are used, it should not confuse the user by making the navigation of the site difficult. As a general guideline, in order to avoid confusion no more than two frames per window should be used. A non-frames version of the site should also be available.

Table 17 - 100 Evaluation Criteria discussed (continued)

3 Content	
Group 1: Product/service related content	
Extensive product/service information available	Each product/service offered on the site should be described and explained extensively.
Price information and cost benefits communicated	It should not be difficult to find prices of the products/services and any cost benefits that may exist must be communicated clearly.
Adequate breadth of product range	There should be enough different products/services on the site to justify its existence.
Adequate amount of advertising of own products	There should be interesting banner ads and advertising of special offers on the site.
Adequate amount of advertising by other companies	If applicable and possible for the site, there should be an adequate amount of banner ads or any other form of advertising on the site by other companies.
Group 2: Company and contact information	
Full company information available	A history of the company, its core values and goals and other relevant information on the general functioning of the company should be readily available.
Terms and conditions easily accessed	It should be very easy to find the terms and conditions concerning sales on the site.
E-mail addresses of employees available	E-mail addresses of individual employees of the company (if applicable) or general e-mail addresses of key personnel should be easy to find.
Telephone and fax numbers available	Not only e-mail addresses but also telephone and fax numbers should be available for relevant departments in the company.
Mailing address and physical address available	The mailing address and physical address of the company and its offices should be available. This provides a means for correspondence by mail if the user prefers it, and it also enables customers to visit the company if they want to.
Group 3: Information quality	
Content current and updated	The site should give the user an indication of how regularly the data is reviewed and updated.
Content relevant to purpose of site	Content should always be relevant to the products/services the site aims to provide for the user. For example, it is debatable whether it is relevant for an online bookstore to provide not only the prices of books in the currency of the home country, but also the estimated price in currencies of other countries.
Content concise and non-repetitive	Descriptions of products/services should be brief and concise, but complete. The same information should not be repeated on many different pages.
High perceived quality of product/service	The content should aim to give the user a realistic idea of the quality of its products/services. This type of content might include for example reviews by other customers, the results of scientific experiments performed on certain products, etc.
Quality of advertisements high	The advertisements on the site should be relevant and the quality of its design should be of a high standard.
Group 4: Interactivity	
User able to customise content to fit needs	If applicable, the user must be able to choose which content is displayed on the site.
Large amount of personalisation possible	If applicable, the user must be able to personalise the site so that he is more than just another user — he is welcomed by name and gets personalised attention based on previous encounters with the site. A good example of the first two criteria in this group is the web site for Crosswalk Ministries [http://my.crosswalk.com]. This Christian portal provides a large amount of customisation and personalisation opportunities.
Easy access to online community	It must be easy to register (become a member) and log in to the online community of users of the site. A good example of an online community can be found at [www.webshots.com], where users are encouraged to share their photos, comment on and even buy each other's photography.
High perceived value of online community	The content displayed and discussed in the community should be relevant and valuable to the user's understanding and experience of the site.
High degree of interactivity in community	Members should be able to communicate not only with the site but also with each other, for example through the use of forums.

Table 17 - 100 Evaluation Criteria discussed (continued)

4 Reliability	
Group 1: Stored customer profile	
Easy to register on site	It should be easy to register as a member on the site.
High perceived benefits from registering	It should be clear what benefits the user can expect from registering. This includes special offers, e-mail updates on specials and easy, fast ordering of products/services.
Easy to log in to site	It should be easy to find the login page for the site and enter the user information and password.
Adjustable customer profile stored	The user must be able to customise his personal information and preferences.
Guided ordering using customer profile available	The customer profile should be used to provide recommendations and guide the user to order the products/services that he will really be interested in. The benchmark site in this field is [www.amazon.com], which provides extensive recommendations to members and make it easy for them to purchase these recommendations through initiatives like "1-Click Ordering".
Group 2: Order process	
Transparent, interactive and easy order process	The order process should not be confusing to the user. It must be clear at which stage he will be asked for his credit card number/other payment details. It must be clear when the purchase is made and how he can cancel the order.
Easy selection of generic services	The choice of generic services (e.g. courier service, payment system) should be integrated with the order process and easy to choose.
Alternative methods of ordering/payment available	Alternatives to the traditional credit card payment system should be available (e.g. electronic cash). The user should also have the option of collecting products directly from the company if he does not want to pay for a courier service.
Good stock availability	Users should be advised when a product is not in stock, and these instances should be kept at a minimum.
Acknowledgement of order sent to customer	An e-mail acknowledging the order and providing the customer with an order number should be sent directly after the order is placed.
Group 3: After-order to order receipt	
On-line order tracking available	The customer must be able to track the progress of his order online.
Effective payment settlement	Payment should be settled effectively. Bank statements should be clear on what the amount was used for.
Confirmation of order dispatch sent to customer	When the order is dispatched to the customer, confirmation e-mail should be sent to the customer immediately.
On-time delivery	Delivery should occur within the time frame communicated on the site.
Full order delivery	Split orders should be kept at a minimum. If absolutely necessary, it must be done at no extra cost to the customer. The company's policy regarding this aspect should be made clear on the site.
Group 4: Customer service	
Feedback forms available	The customer should be able to evaluate the service on the site with pre-designed feedback forms.
Good after-sales support	After-sales support like warranties, easy return policies or availability for installation queries should be available.
Quick reply on e-mail enquiries	Any enquiries sent by e-mail should be replied on promptly. The generally accepted rule is to allow 24 hours for a reply.
General courtesy of company good	E-mails received from the company should be courteous and helpful.
Good after-sales relationship building	Subsequent e-mails should be sent to customers, encouraging them to make use of the company's services again. This should take place on irregular intervals so that the customer does not get irritated with unnecessary e-mails.

Table 17 - 100 Evaluation Criteria discussed (continued)

5 Technical	
Group 1: Speed	
Fast home page loading speed	The home page should load quickly, especially considering the amount of graphics and multimedia on the site.
Fast sub-page loading speed	The sub-pages should load as quickly as the home page.
Good perceived use of caching	The use of caching is important for return visits. It is possible to check whether or not caching is used by checking whether or not certain images and graphics on the site are stored in the computer's 'Temporary Internet Files' folder (if a Microsoft Windows PC is used).
Good consideration of non-broadband users	When a site is very graphic-intensive (for example when Macromedia Flash animations are used), non-graphic intensive alternatives should be available for non-broadband users who do not have the facilities or the time to use these multimedia-enhanced sites.
Good perceived speed of database	When queries that use the company's database are made on the site, results should be available quickly.
Group 2: Security	
Security systems accredited	The site should provide an indication of whether or not their security systems have been accredited, and by whom. This should preferably be indicated on the home page.
Secure payment systems used	Secure payment systems should be used. This can be evaluated by checking whether or not SSL (Secure Sockets Layer) encryption is used on the site, and how many bit encryption is used. Banking sites are usually the most secure (128-bit), while online retailers normally run on slightly lower encryption.
Privacy of users protected	The privacy statement of the site should be easy to find and should provide details on all the privacy aspects regarding the user's personal information.
Security protocols communicated well	The security systems used on the site should be explained thoroughly.
Security certificates adequate	If accredited, the site should also have a security certificate which outlines all the security issues on the site: digital certificates, secure payment systems, privacy issues, etc. These certificates are usually found by clicking on the logo of the accreditation company provided on the site.
Group 3: Software and database	
Good cross-browser capability	The site should look the same in different versions of both Microsoft Internet Explorer and Netscape Navigator.
Users advised on ideal browser and resolution	The user should be advised on the ideal browser and resolution for which the site was designed.
Database software adequate for size of database	A large customer base calls for extremely complicated database management. The software used on the site should be able to handle all management actions without problems. It is usually possible to deduce what kind of programming language is used to communicate with the database by viewing the source code of appropriate pages (Click on *View* | *Source* in Microsoft Internet Explorer). This programming language should be powerful enough to perform all necessary database functions.
Good data transfer between systems	Data should be transferable without difficulty between different systems on the site. For example after a user bought a book on a web site, it should not be necessary to re-enter his information if he now wants to auction an item on the same site.
No perceived duplication of data in database	The database software should take precautions to ensure that it is not be possible to register the same user more than once on the same site, even if different information is entered.
Group 4: System design	
Precise operation and computation	In general, the site should function fast and effective with no problems in the navigational or technical aspects.
Good resolution compatibility	The site should be designed to display correctly on different resolution settings.
Good integration with systems of users	If applicable, relevant data should be available for download to users' own systems. For example, when banking statements are drawn, it should be possible for users to download the statements into their local Microsoft Money applications.
Good integration of different systems on site	All the systems on the site should function properly and complement each other. Ordering, payment and logistics should be integrated into one complete experience.
Global accessibility to products/service available	If applicable, users from all over the world should be able to buy the products/services offered on the site.

Table 17 - 100 Evaluation Criteria discussed (continued)

From the above discussion it is clear that **there are some criteria that cannot be evaluated simply by looking at the site**. It is necessary actually to order and pay for a product/service on the site before a proper evaluation can be done. All the criteria are very relevant and it is therefore not possible to leave some of them out of the evaluation because of this fact. It is therefore important that the evaluators be people who have used

all the sites in the specific industry extensively and consequently has the knowledge to evaluate all the criteria. If this is not possible, it is suggested that e-mail is sent to the company in which information is requested on the criteria in question. None of the information that is needed can be considered as confidential, and therefore it should not be a problem for companies to provide it to the evaluators. Answers to the following questions will provide adequate information to evaluate the criteria in question:

- Is an 'acknowledgement of order' e-mail sent to the customer immediately after an order is placed?
- Is it possible to track the progress of the order online?
- Is a 'confirmation of shipment' e-mail sent to the customer as soon as the order is dispatched?
- Percentage wise, how many orders are filled within the time limit provided on the site?
- What is the company's policy on split orders?
- What does the company do to ensure good after-sales relationships building?
- What type of database software is used for the customer and product databases?

The actual reply received can also be used to evaluate the criteria concerning *prompt e-mail replies* and *general courtesy of the company*.

If it is not possible to send e-mail, or some companies refuse to disclose some of the information, those specific criteria should be regarded as 'Not Applicable'[10] for all the sites in the industry so that no unfair advantage is given to some sites.

An important issue that should be addressed is **what the course of action should be if some of the web site features that appear in the criteria are not present on the site**. For example what should be done if a specific site does not have a search engine or online community to evaluate? In such cases the evaluators should first of all decide whether or not the specific feature is applicable to the site, in other words whether or not the feature will enhance the performance of the site. Then, if the feature is indeed applicable to the site and it does not exist, a strong negative score should be awarded. Otherwise, the criterion should be regarded as 'Not Applicable'.

The next section describes the process followed to develop the finalised method for the evaluation of e-commerce web sites by linking the criteria given above with the Correspondence Analysis theory given in Chapter 3.

[10] The exact scoring method used for evaluation is discussed in section 4.3

4.3 E-commerce web site evaluation method

After the evaluation criteria were developed, the next step was to develop a method to evaluate e-commerce web sites and analyse the data with Correspondence Analysis to give a graphical representation of the findings. Before evaluation is commenced, it is firstly important to identify a specific industry in which certain web sites are going to be evaluated. This industry should be clearly defined so that no uncertainty can exist whether or not a certain web site belongs to the industry or not. Typical industries include the online books industry, the online banking industry, the online auction industry and the online flowers industry. It is then necessary to identify the specific web sites within the industry that are going to be evaluated with the method. When each web site for evaluation is clearly identified and grouped within a specific industry, the evaluation method can be initiated.

The following evaluation method, which consists of three steps, is proposed for the evaluation of e-commerce web sites:

STEP 1: GATHER E-COMMERCE WEB SITE EVALUATION DATA AND PRESENT THE DATA IN A MANNER SUITABLE FOR FURTHER ANALYSIS

Before starting the evaluation of a specific industry, it is suggested that a group of between 3 and 5 evaluators be assembled to evaluate the sites. Each of the evaluators will first use the evaluation criteria to perform his own evaluation of the site without input from the others. Afterwards, they will come together and discuss their evaluations. A final score will then be awarded to each criterion based on a combination of their discussions and the averages of their individual scores. A consensus decision should be made regarding the score for each site. **Consensus decisions based on the average scores of a group of evaluators will remove the amount of subjectivity that exists within the criteria to a large extent.**

A few characteristics are important for the group of people who evaluate the site:

- The same group should be used for all the sites in a specific industry. Different groups can be used for different industries. As long as there is consistency in the way all the sites in a specific industry is evaluated, the results will be reliable.
- None of the evaluators in the group should be an employee of any of the companies that are evaluated or in any way directly involved with their web sites.
- The evaluators should be people who (1) regularly use the Internet for e-commerce purposes; (2) spend much time surfing the Internet and therefore have a clear understanding of what the important characteristics of a good web site are; and (3) as far as possible, have made use of the web sites that are evaluated.

After the evaluation group is assembled, each of the evaluators should, on their own, assess how well the web site in question performs when compared to the 100 evaluation criteria developed. To facilitate this it was decided to use a **scoring method** where the evaluator has to score the web site according to the level he agrees with the statement made by each criterion. For each criterion a score between -2 (strongly disagree) and +2 (strongly agree) is awarded. It was decided to provide only four scoring options (-2, -1, 1 or 2) so that the evaluator is forced to decide whether he agrees or not—a neutral answer is not possible, as would have been the case if five scoring options were given. A fifth option, *Not Applicable*, is provided for cases in which the criterion is not at all relevant to the site that is being evaluated. In this case a score of 0 is awarded. This score of 0 has no positive or negative effect on the outcome of the evaluation.

After the evaluations are completed, the group should go through the criteria together, discuss their opinions, and award a final score to each criterion. The sum of the final scores of the five criteria in each criteria group gives the 20 criteria group totals. The four group totals within each of the five criteria categories in the evaluation framework are then summed and these five totals provide a scalar of evaluation data for the site (see Figure 7 on page 50 for a graphical representation of the framework).

This process should be repeated for all the web sites that were identified for evaluation in a specific industry. As soon as more than one site has been evaluated, a contingency matrix is created that can be used for Correspondence Analysis. At this stage, there is still one problem with the contingency matrix. The scoring method implicitly makes it possible for negative values to exist in this matrix, while Correspondence Analysis can only work with matrices with all-positive values. For the purpose of CA the data are therefore linearly shifted by adding a constant value to each value in the matrix. This constant is equal to the absolute value of the lowest possible (negative) value that an entry in a matrix can assume. This ensures that all values in the matrix are positive. As the results obtained from CA are relative and not absolute values, this does not compromise the integrity of the results.

It is also important to note that a minimum of three evaluated sites is necessary to perform CA. From the theory it can be seen that only two rows in the matrix will result in a one-dimensional presentation of the data, which is not what is required. As soon as three rows or more appear in the matrix, CA scales the results to two dimensions.

After this step is completed, a two-way contingency matrix of non-negative values exists, which contains the evaluation data of different e-commerce web sites in a specific industry. The matrix is now in the correct form for Correspondence Analysis to be performed on it.

Step 2: Perform Correspondence Analysis on the matrix and represent the results graphically

In this step, the evaluation group will perform Correspondence Analysis on the matrix, and draw the plot of the results. This process was described in section 3.3.2. Normally, a combination of Matlab and Microsoft Excel or some other statistical software is needed to do this. The software described in the next chapter performs this entire step easily.

Step 3: Interpret the results in the graphical representation and make recommendations

This is the most important step in the evaluation method. The group of evaluators should come together and interpret the plot in a manner similar to that described in section 3.3.4. This interpretation process should be done in three distinct phases:

- Use the CA Graph to form basic conclusions on the performance of each of the web sites and the profile groups that can be identified.
- Use these preliminary findings and combine it with the original data in the contingency matrix to refine the conclusions.
- For each web site, use its individual criteria group scores to make recommendations on possible ways to improve the sites, and how to do it.

When the group has finished with the interpretation of the results, they should write a report consisting of the findings and recommendations on the following aspects:

- Report on which of the web sites have **similar profiles** and why this is so.
- For each **web site**, report on its performance with reference to the criteria categories on which it performed best as well as those on which it performed worse. Recommend steps that could be taken to improve performance on the categories on which the site did not perform so well.
- For each **criteria group**, report on which web sites perform best and what the common factors in these sites are. Recommend which aspects of these sites other companies should investigate to help them to improve the performance of their sites on each of the different groups.

This completes the method for comprehensive evaluation of e-commerce web sites. A graphical representation of the method is shown in Figure 8.

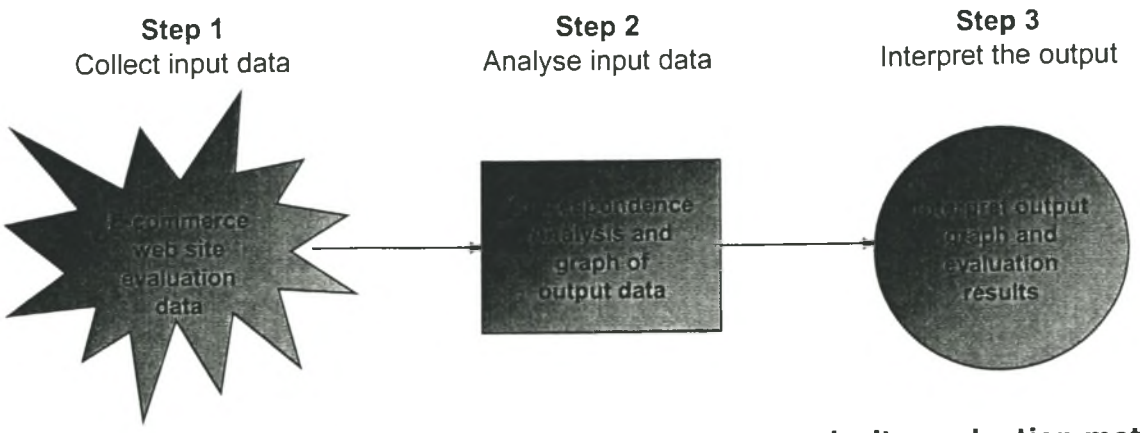


Figure 8 - Graphical representation of e-commerce web site evaluation method

This chapter systematically described how the method was developed. An examination of the effectiveness of the method is performed in Chapter 6. The next chapter will describe how the evaluation method was implemented using Microsoft Visual Basic.

5. Software development

This chapter describes how Microsoft Visual Basic was used to create software to implement the e-commerce web site evaluation method developed in the previous chapter. The chapter begins with a general description of the program, how it was designed and how it works. It then concludes with examples of how e-commerce web sites in three different industries were evaluated using the software. The procedure and findings are discussed in detail.

5.1 Overview of the software

Microsoft Visual Basic was used for this thesis because of its ease of use and its versatility. There are four main challenges when programming an application like this in Visual Basic:

- The first challenge to consider is **where to store the evaluation data that is generated and used by the software**. A database is needed to store the industry names, web site details and each site's individual scores so that it can be used to perform Correspondence Analysis.
- The mechanics of Correspondence Analysis consists of fairly **complicated mathematical computations**, and Visual Basic is not primarily designed to perform the matrix algebra demanded by Correspondence Analysis. A program like Matlab performs these computations easily. The challenge is therefore to find a way to call Matlab procedures from Visual Basic.
- Drawing a basic bar graph or scatter plot in Visual Basic is not difficult, but **drawing the type of graph used in Correspondence Analysis is more complicated**. A scatter plot consisting of two matrices must be drawn on the same graph. This is also fairly easy in Matlab, so again the challenge is to find a way to integrate certain Matlab procedures with Visual Basic.
- **Performing the Chi-Square test for statistical independence** is very difficult in Visual Basic, again because of its mathematical complexity. It is easy to perform the test in software like Microsoft Excel. Except for incorporating Matlab in Visual Basic, another challenge is thus also to use Microsoft Excel from within Visual Basic.

To respond to the challenges outlined above, it was firstly decided to use a *Microsoft Access* database to store the data. With the help of Visual Basic's *Microsoft Access Object Library*, it is possible to use Visual Basic directly to add, delete and retrieve data from a Microsoft Access database. The database was designed to consist of two entities, namely **INDUSTRIES** and **SITES**. The **INDUSTRIES** entity consists of the attributes **Industry ID** and **Industry Name**. All the industry names are stored in this entity. The **SITES** entity consists of the

attributes **Site ID**, **Industry ID** (to identify the industry to which the site belongs), **Site Name**, **Site Address**, 100 attributes to store each of the scores obtained in the evaluation of the 100 criteria, and 5 attributes (**Interface**, **Navigation**, **Content**, **Reliability** and **Technical**) to store the total scores of each of the five phases of the framework. These last 5 attributes are used to populate the contingency matrix and perform Correspondence Analysis. Individual site data is stored in this entity. A 1-to-Many relationship exists between the **INDUSTRIES** and **SITES** entities.

These entities were implemented with Microsoft Access tables. The relationships between the entities are shown below. The **SITES** table is only shown in part because it has a total of 109 attributes as outlined in the paragraph above. The numbers shown in **tblSites** refer to the specific criterion of which a score is stored. For example, *1_1_4* refers to *Criteria Category 1*, *Criteria Group 1* and *Individual Criterion 4*, which evaluates *effective and consistent use of backgrounds*.

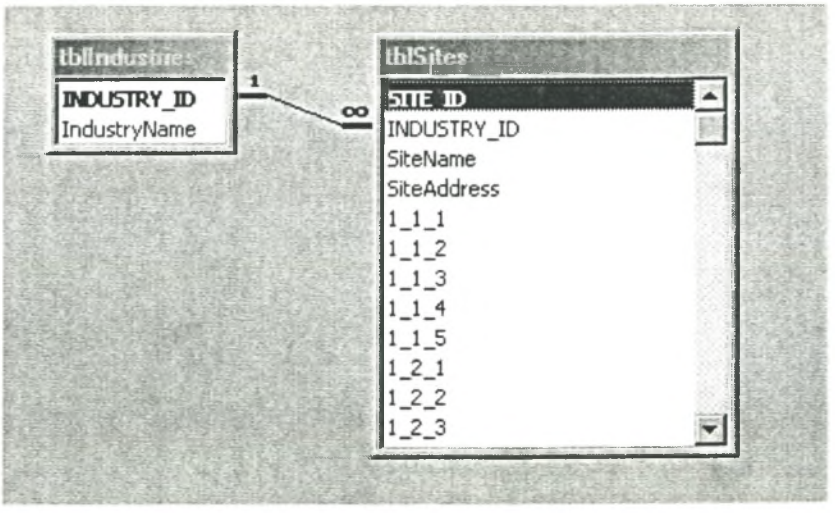


Figure 9 - Relationships between database entities

When the Matlab challenges were investigated, it was decided to use software called *MatrixVB*[11] to solve the problem. *MatrixVB* is a library of matrix math functions for Visual Basic, enabling the user to perform several Matlab functions in Visual Basic. A big advantage of the library is that it does not require Matlab to be installed on the computer. The library includes a linear equations solver, a singular value decomposition function, linear and quadratic optimisers, algorithms to solve and fit polynomials, filtering functions, and much more. Except for immediately providing a means to perform the mathematical side of Correspondence Analysis, the software also makes it possible to draw complicated plots easily, which solves the graphical challenge as well.

[11] The software was developed by *Mathtools, Ltd.*, a leading developer and supplier of technical computing software. *MatrixVB* was downloaded from the company's web site, [www.mathtools.com].

To perform the Chi-Square test, Visual Basic's *Microsoft Excel Object Library* was installed. Microsoft Excel was then utilised from Visual Basic to perform the mathematical computations needed.

With these initial stumbling blocks out of the way, the user interface was designed and program code was written to perform the steps in the evaluation of e-commerce web sites. The rest of this section will provide an overview of the program and its abilities.

When the program is started, the following main menu is displayed:

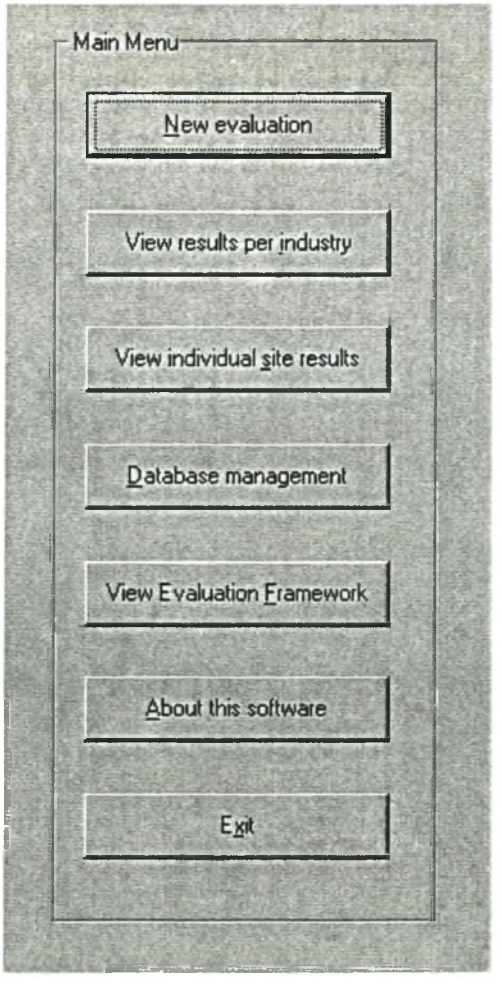


Figure 10 - Main Menu of the e-commerce web site evaluation software

This menu shows what functions the software is able to perform. Each option will now be discussed in detail.

5.1.1 New Evaluation

The **New evaluation** option is used to perform an evaluation of a specific site. The user is first prompted to enter the industry within which the site can be classified, the site's name and its address:

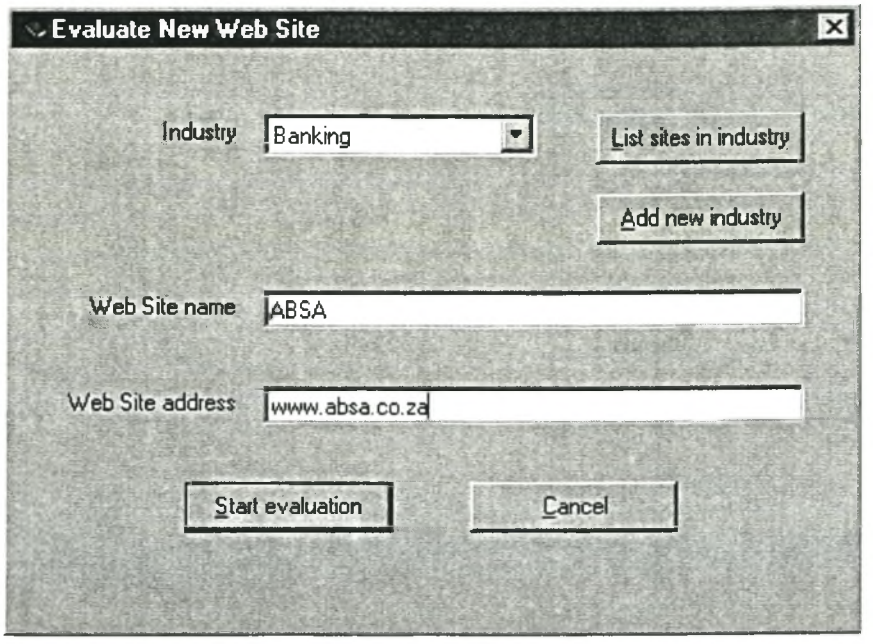


Figure 11 - "Evaluate New Web Site" form

The **List sites in industry** command shows the user which sites have already been evaluated in a specific industry. The **Add new industry** command is used to add a new industry to the database if the sites to be evaluated cannot be grouped in an industry that already exists.

After entering the information, the user clicks on **Start evaluation**. The user is then prompted to either confirm that the information is correct, or go back to change the details. With the information confirmed he is then guided through 20 forms, each representing a criteria group as developed in the e-commerce web site evaluation framework. Each form prompts the user to evaluate the site in question on certain criteria. An example of one of these forms is shown in Figure 12.

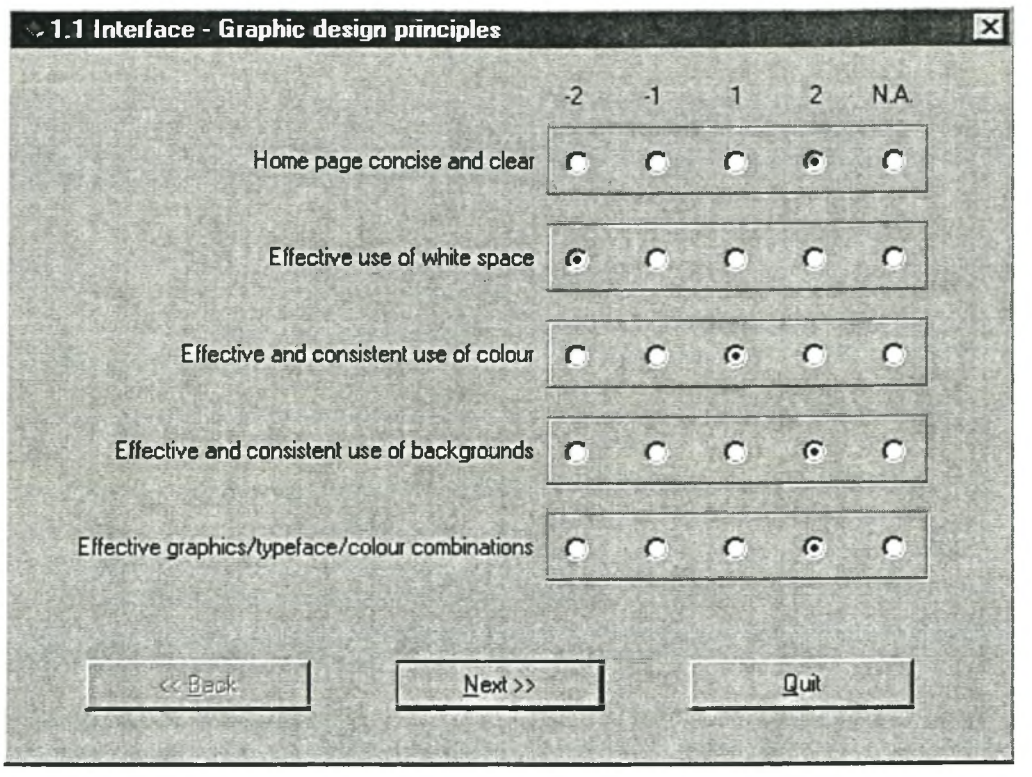


Figure 12 - Example of an evaluation form

It is important to note the **title** of each of these 20 forms, because it indicates at what stage of the evaluation framework the user currently finds himself. In the example above, it is clear that the user is currently evaluating *Criteria Category 1, Criteria Group 1*, which is Graphic Design Principles (see Figure 7 on page 50).

When each form is completed, the user clicks on the **Next** button to advance to the next form. The **Next** button on each form is disabled when the form loads, and is only enabled when all the criteria on a form have been evaluated. This ensures that the user cannot accidentally move on to the next criteria group before all the criteria in the current group have been evaluated. On each of the forms, the user can use the **Back** button to return to the previous form/forms if he needs to make some changes. The **Quit** button is used to terminate the evaluation. No evaluation data is saved in this case, which means that the user will have to start from the beginning if he wants to evaluate the site again.

When the evaluation is completed, the user clicks on the **Finish** button. A message box showing the total scores that the site received for each of the criteria categories are displayed, after which the user is returned to the main menu of the program.

5.1.2 View results per industry

This option is used to view the results of evaluations per industry. The user is prompted to choose the specific industry and also the format in which he wants to view the results[12]:

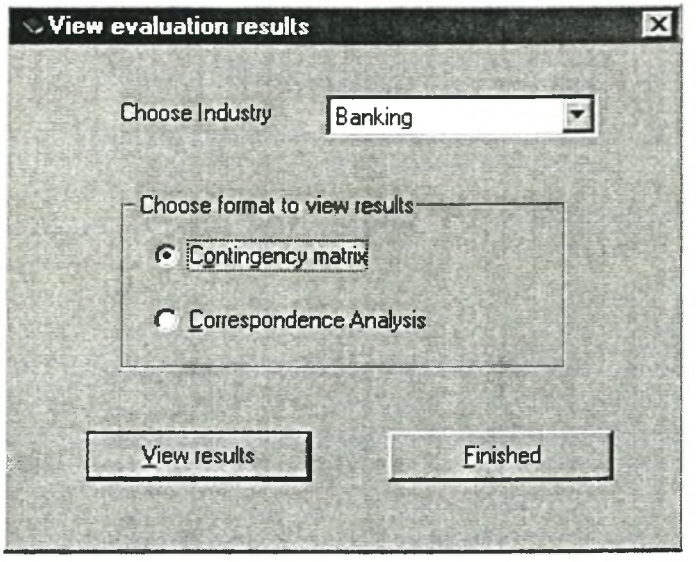
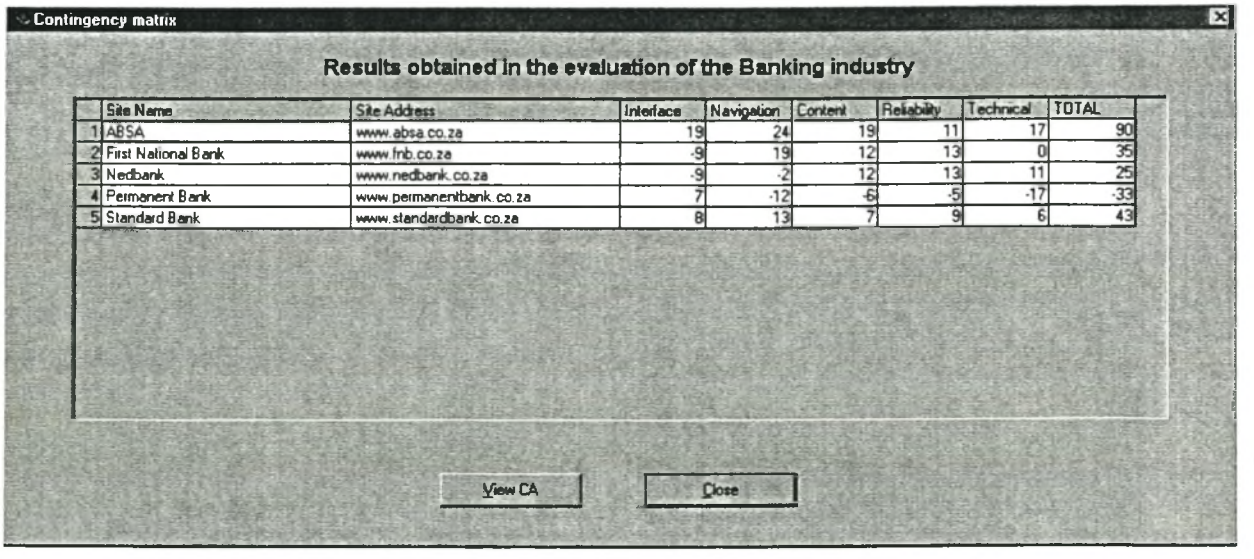


Figure 13 - "View evaluation results" form

The **contingency matrix** option will show a matrix of the final scores obtained for the criteria categories of the evaluation framework. As an example, results from the banking industry are shown below:



	Site Name	Site Address	Interface	Navigation	Content	Reliability	Technical	TOTAL
1	ABSA	www.absa.co.za	19	24	19	11	17	90
2	First National Bank	www.fnb.co.za	-9	19	12	13	0	35
3	Nedbank	www.nedbank.co.za	-9	-2	12	13	11	25
4	Permanent Bank	www.permanentbank.co.za	7	-12	-6	-5	-17	-33
5	Standard Bank	www.standardbank.co.za	8	13	7	9	6	43

Figure 14 - Evaluation results - Contingency Matrix example

[12] Once the initial format is chosen, it is possible to view the results in the other format by clicking a button on the form that is shown at that stage. The Contingency Matrix form has a **View CA** button that shows the Numeric Results form, and the Numeric Results form has a **View Matrix** button that shows the Contingency Matrix form.

When the user chooses the **Correspondence Analysis** option, two consecutive forms are shown. The first form provides an overview of the CA results, consisting of three parts:

- It firstly shows the results of the **Chi-Square test**. When the results are not reliable, CA should not be pursued further. Otherwise, CA provides reliable results that can be very valuable when interpreted.
- Secondly, it shows the **Numerical results of Correspondence Analysis**. These two matrices show the X-values and Y-values as they are plotted on the CA graph.
- The form also shows each dimension's **contribution to inertia**. Refer to section 3.3.2 for a description of this concept.

An example of this form is shown below:

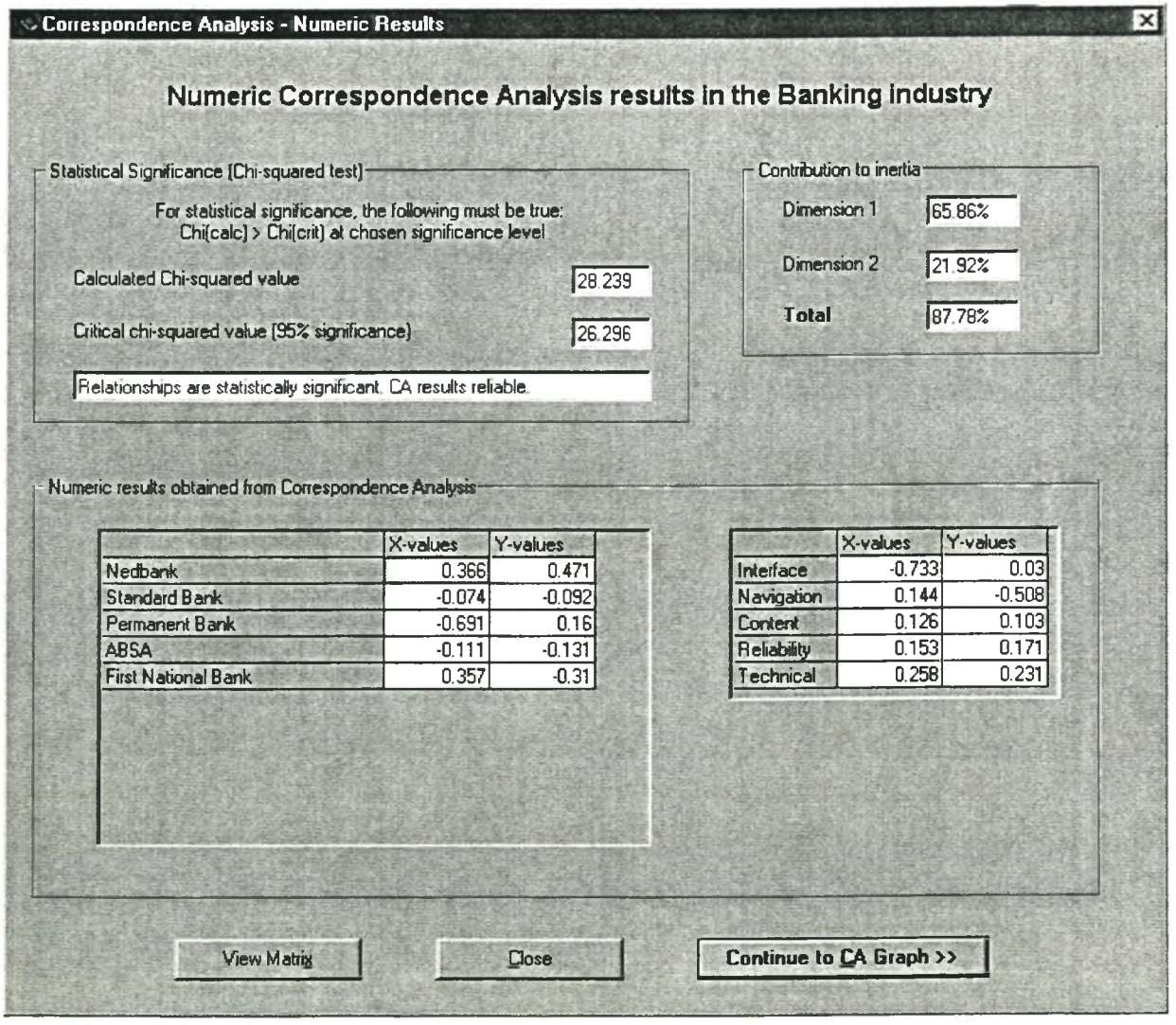


Figure 15 - "Correspondence Analysis – Numeric Results" form

It is important to review the information on this form. If the CA results are reliable, the user can continue to the CA Graph, which shows the actual comparison between different e-commerce web sites within the industry. An example of the graph is shown below:

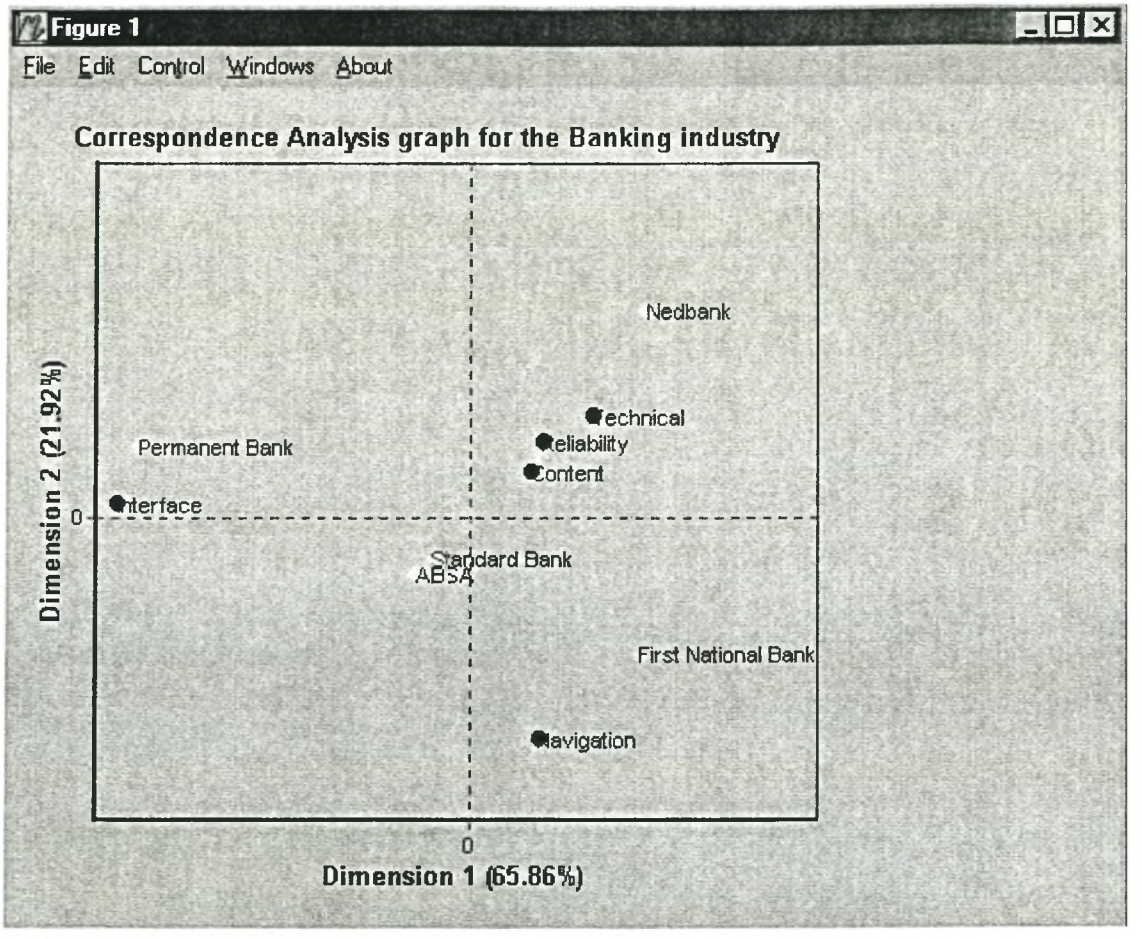


Figure 16 - Example of Correspondence Analysis Graph

The red dots on the graph represent the profiles of the criteria categories, while the yellow diamonds represent the different site profiles. It is possible to zoom in on a specific part of the graph by using the left mouse button to draw a rectangle around the area that needs to be inspected more closely. It is also possible to move the interior of the graph by holding down the right mouse button on the graph and moving the mouse. The menu on this form can be used to print and even customise several aspects of the graph. It is also possible to change the size of the window (or maximise it) to have a clearer view of the graph.

To ensure that the Chi-Squared test results, Correspondence Analysis results and Correspondence Analysis graph were calculated correctly, the results were tested extensively. Each step of the Correspondence Analysis process was implemented separately in Visual Basic program code. To test the correctness of the code, Visual Basic was used to perform the calculations on sample data, and the same data was then used to perform those calculations in Matlab or Microsoft Excel, whichever was appropriate. The

results were compared with each other and any Visual Basic code errors were corrected and tested again until the results were exactly the same. On completion of the code, full Correspondence Analysis were performed on different sets of sample data, and the results obtained in Visual Basic were compared with Matlab results. No errors occurred, and it can be concluded that the technique was implemented successfully in Visual Basic for the input data.

When the user has finished viewing the results of all the industries he is interested in, he clicks on the **Finished** button, which displays the main menu again.

5.1.3 View individual site results

This option is used to display the scores that a specific site received in each of the 20 criteria groups evaluated. The user is prompted to choose the site for which he wants to see the results, and then a form is shown with the results. The next two figures give an example of this process.

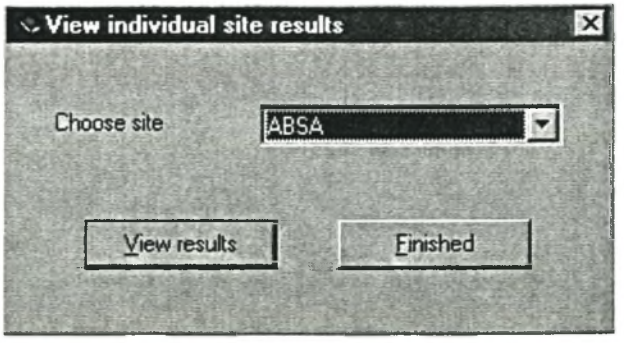


Figure 17 - "View individual site results" form

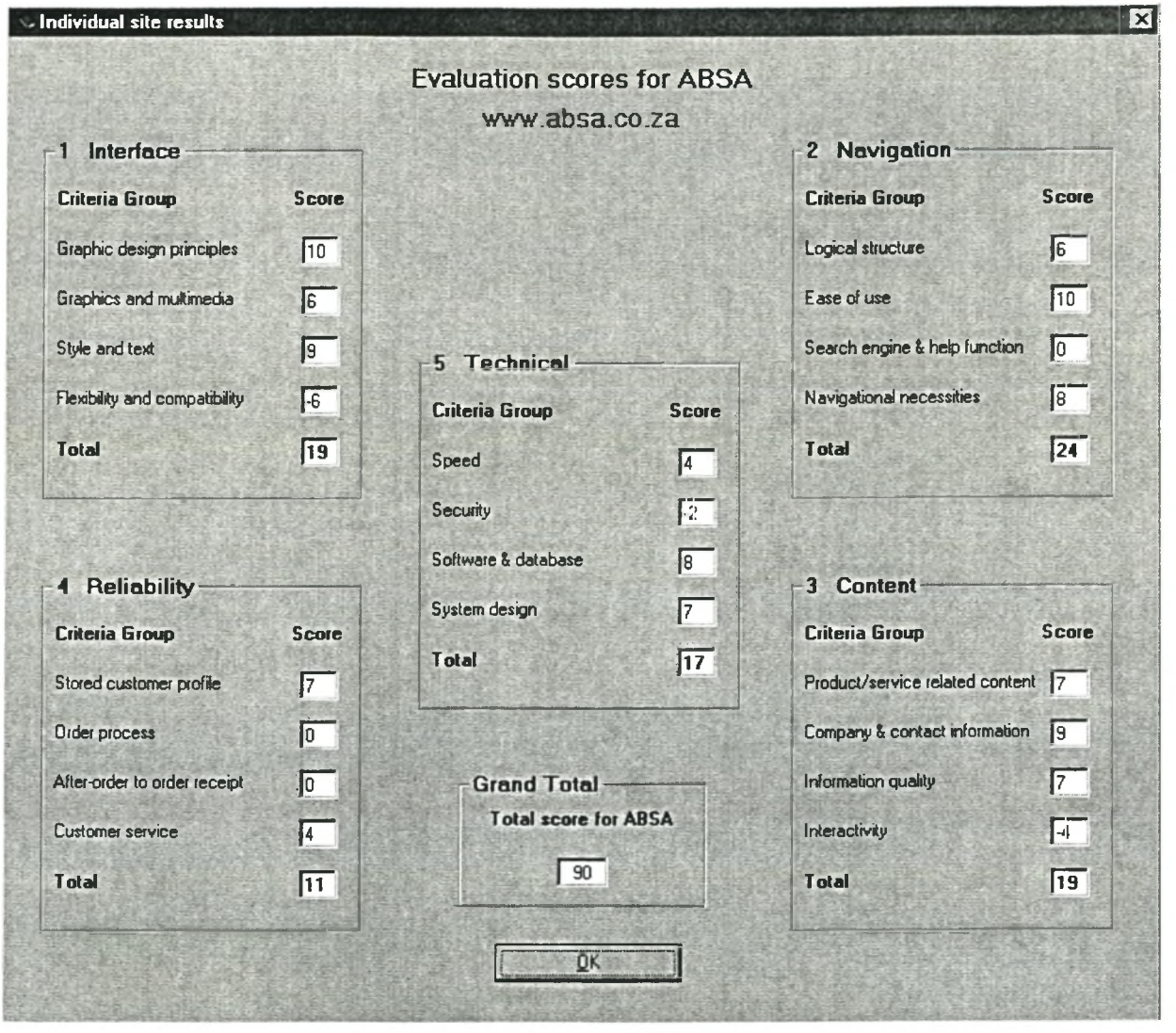


Figure 18 - "Individual site results" form

If the user wants to visit the web site of the company whose results are shown on the form, he can click on the blue hyperlink showing the web site address. The user's Internet browser will then launch and load the site in question.

These results could be used as an indication of where the company should improve their web sites. This form should be used in conjunction with the CA Graph and the individual criteria to draw conclusions on which aspects of the site need improvement. The examples later in this chapter will make this process clear.

As previously, the **Finished** button is used to return to the main menu when the user has completed viewing the individual web site results.

5.1.4 Database management

This option displays a form on which the user can perform three database management actions:

- Delete an industry and all its sites and evaluation data from the database.
- Delete a specific site and its evaluation data from the database.
- Add an industry that does not currently exist in the database.

The user simply uses the appropriate drop-down box or text box to either perform the deletion or add an industry. The command buttons for each of the actions are disabled until a valid choice is made in either of the drop-down boxes or text is typed in the text box. The user is also prompted to confirm his actions when a button is clicked, which further prevents errors from occurring. It is possible to add an industry on the **'Evaluate new web site' form** as well, because this option fits in logically on both forms. Two examples of this screen are shown in the next two figures, namely how to delete a specific site and how to add an industry.

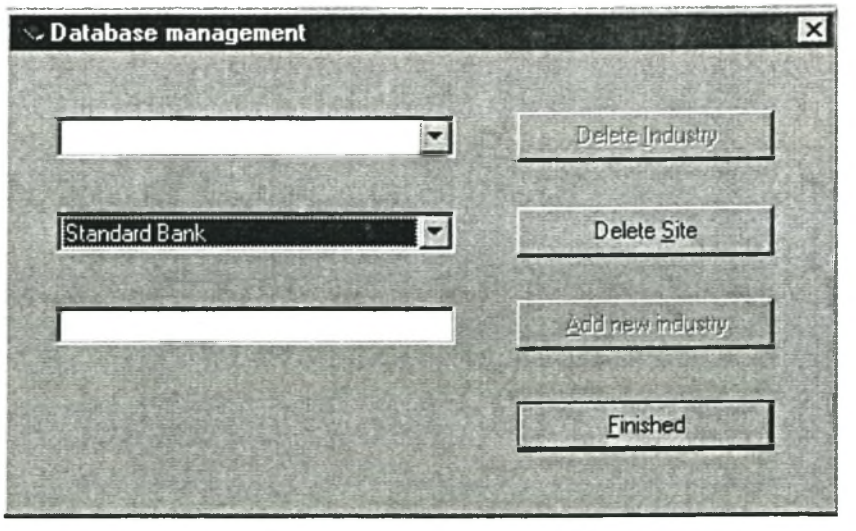


Figure 19 - "Database Management" form: Delete Site

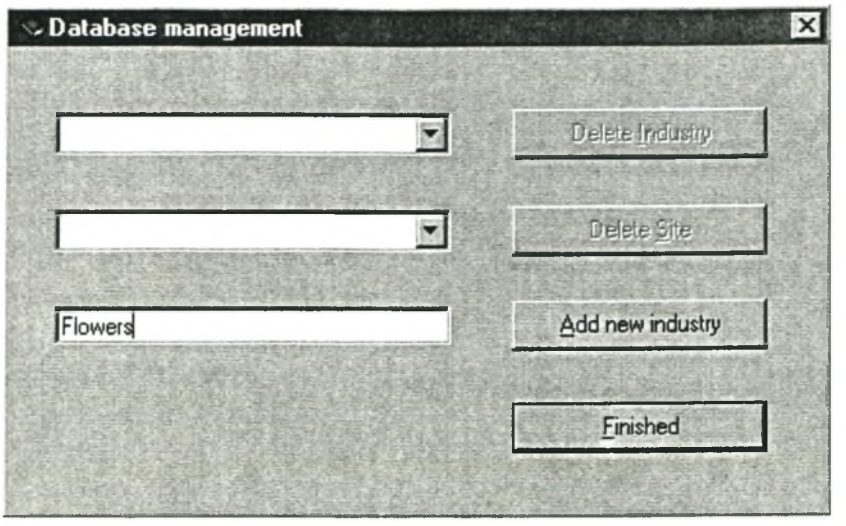


Figure 20 - "Database management" form: Add new industry

5.1.5 View evaluation framework

This option simply displays the e-commerce web site evaluation framework shown in Figure 7 on page 50. It is important for the evaluators to have an understanding of the evaluation framework when they are busy with specific criteria in order to remember how their current position in the framework fits into the holistic view.

5.1.6 About this software

This option displays a form with information about the author of the software:

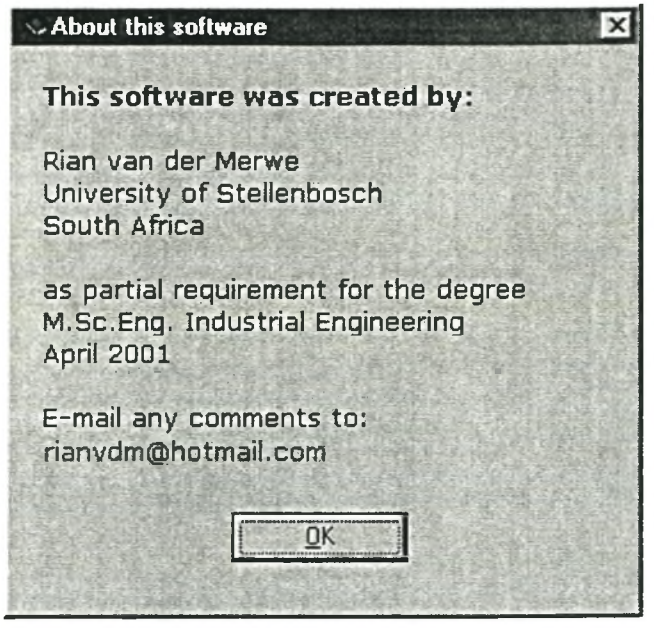


Figure 21 - "About this software" form

5.1.7 Exit

This option is used to quit the program. The user is asked to confirm that he wishes to exit, and then the application is terminated.

This section gave a brief overview of the software and its use. The complete source code of the program can be found on the web site of the Department of Industrial Engineering, University of Stellenbosch ([www.bing.sun.ac.za], follow the links to *graduate students*). Details on obtaining a copy of the software are also available on the site.

To test the abilities of the software, it was used to evaluate sites in three different industries. The next section describes the findings in each of the three chosen industries.

5.2 Illustrative examples in three different industries

In order to find out how well the software and the CA tool work, the e-commerce web site evaluation software was used to evaluate sites in three different industries, namely **online bookstores**, **online banking** and **ammunition reloading equipment**. It was decided to choose two industries (*books* and *online banking*) that the researcher knows well and have spent considerable time using. The third industry (*ammunition reloading* equipment) was

chosen in order to evaluate a speciality industry that does not have such a large market penetration. It is important to note that these examples are merely for illustrative purposes. Although the researcher took every precaution to ensure the accuracy of the evaluations, time constraints prevented him from forming a team of evaluators as suggested in the evaluation method. The evaluations were therefore done without the help of other evaluators.

In order to ensure that all the criteria are evaluated effectively, standard e-mail was sent to the companies whose sites the researcher have not used at least once for online business. This e-mail requested the information needed to complete the evaluation thoroughly, as discussed in section 4.2. The content of the e-mail is shown in Table 18.

To whom it may concern

I am a Masters Student in Industrial Engineering at the University of Stellenbosch in South Africa. I am doing my Masters Thesis on E-commerce Web Site Evaluation. I came across your site, and would like to do an objective and independent evaluation of the site for thesis purposes. In order to do this, I need some information regarding your order process. This is not confidential information. Instead of ordering products directly from your site and finding out that way, I was hoping that you would be able to answer these few simple questions regarding your online order process:

- Is an 'acknowledgement of order' e-mail sent to the customer immediately after an order is placed?
- Is it possible to track the progress of the order online?
- Is a 'confirmation of shipment' e-mail sent to the customer as soon as the order is dispatched?
- Percentage wise, how many orders are filled within the time limit provided on the site?
- What is the company's policy on split orders? When an order needs to be split in two or more shipments, does the company pay for the extra shipping costs, or do you wait for the whole order to be available before the shipment is dispatched to the customer?
- What do you do to ensure good after-sales relationship-building? For example, do you send regular e-mails to customers regarding new offers?
- What type of database software is used for the customer and product databases?

I hope that you will be able to help me with this information. I will use it strictly for academic purposes.

I thank you for your time and your help.

Regards,
Rian

Rian van der Merwe
M.Sc. Industrial Engineering (Electronic Business)
University of Stellenbosch
South Africa

Tel. +27 83 635 1141
E-mail: rianvdm@ing.sun.ac.za

Table 18 - E-mail requesting order information from companies

Responses were received from most companies. For the purpose of this thesis, the sites of the companies that did not respond (*PC Books*, *Lee Precision* and *RCBS*) were explored in detail and the best possible assumptions made on the criteria in question.

After a note on the interpretation of the results, the rest of this section is devoted to a discussion of the evaluations conducted.

5.2.1 A note on the interpretation of the results

As mentioned in section 3.3.4, the most important part of the evaluation is the interpretation of the Correspondence Analysis results. Section 3.3.4 explained how the CA graphs should be analysed considering two aspects: firstly the **position of the points relative to each axis** to determine which points are mainly contrasted by each axis; and secondly the **proximities of the points to each other** to determine the similarities between different row and column profiles. The nature of the data is always the same in this thesis—different web sites are evaluated on the same criteria and therefore, although they receive different scores, the data is always in the same format. This makes it possible to combine the two analysis aspects mentioned by making certain assumptions based on consistency in the graphs. This greatly simplifies the interpretation of the graphs. The contrasts between different point groups highlighted by considering the position of the points on the axes can intuitively be incorporated with the consideration of the proximities of the points. Therefore, considering both aspects as outlined in section 3.3.4, page 43, in the current context of web site evaluation the following general guidelines can be given when the graphs are interpreted:

- Web sites whose points lie on the periphery of the graph and are not in close proximity to one or more criteria category points have generally received fairly low overall evaluation scores.
- Criteria category points that lie on the periphery of the graph and are not in close proximity to one or more web site points have generally not been met well by any of the sites.
- Web sites whose points lie on the periphery of the graph and are in close proximity to one or more criteria category points have generally received a fairly low overall evaluation score, except for the criteria category/categories nearest to it.
- Criteria category points that lie on the periphery of the graph and are in close proximity to one or more web site points have generally been met well only by the sites with points closest to it.
- Web sites whose points lie near the origin of the graph or more or less in a position where the category groups are evenly distributed around it, have generally received a fairly high overall evaluation score, especially for those criteria nearest to it.
- Criteria category points that lie near the origin of the graph with the site points evenly distributed around it have generally been met well by most of the sites.

Because these general evaluation guidelines implicitly consider both aspects of the interpretation, the explicit definition of the points contrasted by each dimension is not necessary, and is therefore not performed here. With the guidelines as basis, the **proximities of points** is analysed in the usual way using the three phases described in section 3.3.4, namely **analysis of column profiles** (criteria categories), **analysis of row profiles** (web sites) and the **analysis of column and row profiles together**.

The rest of this section will discuss the example evaluations performed in each of the three industries identified at the beginning of this section.

5.2.2 The online books industry

Six online bookstores were evaluated for this industry:

- **Amazon.com Books** [www.amazon.com/books] is the bookstore of the major international online retailer Amazon.com.
- **Barnes & Noble** [www.bn.com] started as a chain of bookstores in America, and this site is the online store of the company. They specialise in all kinds of books, including academic and second hand books and rare, out-of-print editions.
- **Exclusive Books** [www.exclusivebooks.com] is the South African counterpart of Barnes & Noble. This site is the online store of the major South African chain of bookstores.
- **Kalahari.net** [www.kalahari.net] can, in many ways, be compared to Amazon.com. It also started as South Africa's first online bookstore, and is in the process of broadening their product range considerably. At the moment, their main focus is still selling books.
- **PC Books, Inc.** [www.pcbooks.co.za] is a South African bookstore that sells computer reference manuals online.
- **VS Online** [www.vsonline.co.za] is the online counterpart of the South African chain of **Van Schaik** bookstores. They specialise in academic textbooks for both school and university students.

Considerable time was spent getting to know these sites, and eventually the evaluation revealed the following contingency matrix:

	Site Name	Site Address	Interface	Navigation	Content	Reliability	Technical	TOTAL
1	Amazon.com Books	www.amazon.com/books	17	35	30	31	28	141
2	Barnes & Noble	www.bn.com	15	29	25	27	19	115
3	Exclusive Books	www.exclusivebooks.com	14	35	18	27	23	117
4	Kalahari.net	www.kalahari.net	-15	11	13	24	12	45
5	PC Books, Inc.	www.pcbooks.co.za	-19	-4	1	16	22	16
6	VS Online	www.vsonline.co.za	6	14	-9	6	11	28

Figure 22 - Contingency Matrix for the Books industry

Correspondence Analysis yielded the results shown in Figure 23.

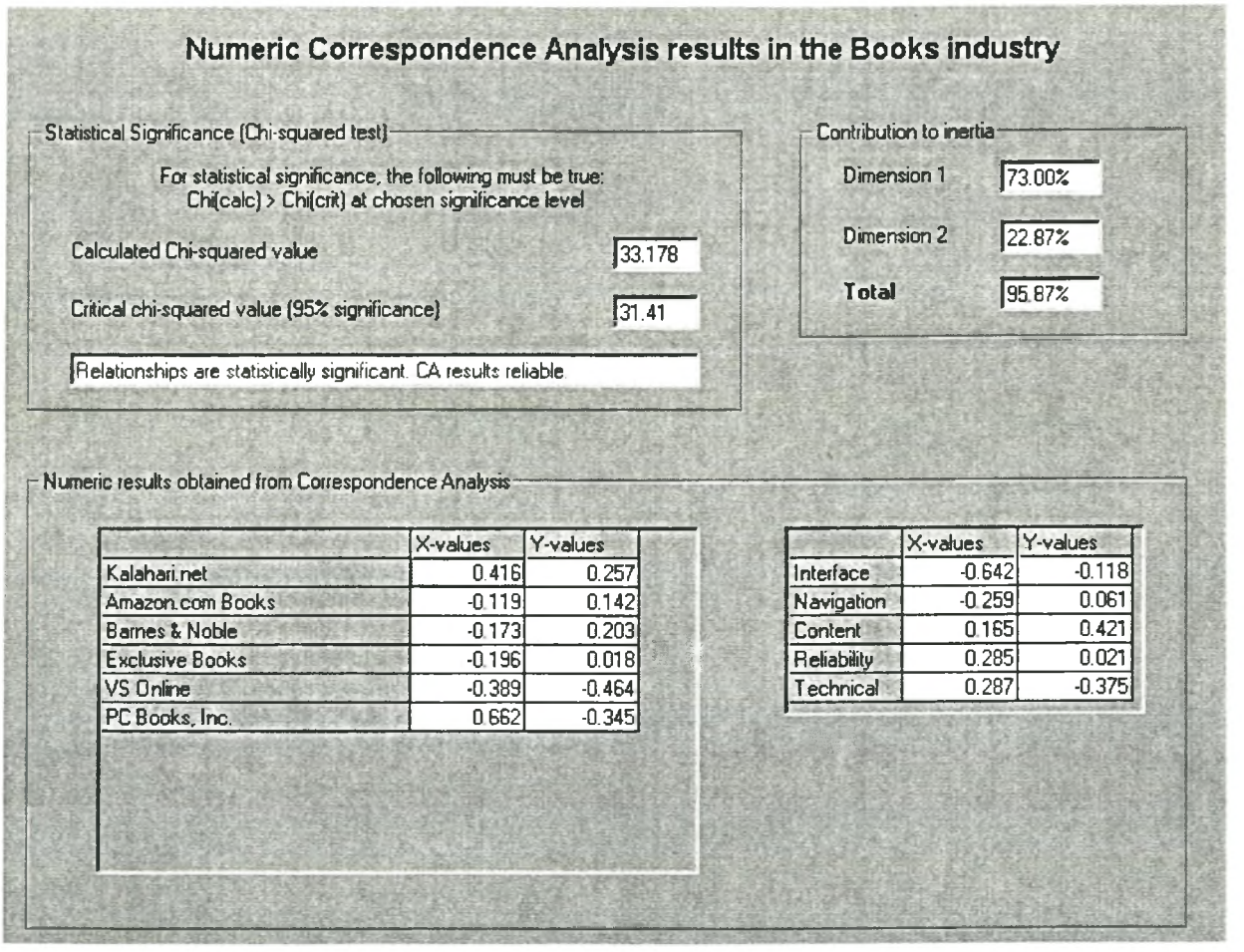


Figure 23 - Correspondence Analysis results in the Books industry

From these results it is clear that the relationships between row and column variables are significant on a 95% significance level. The Correspondence Analysis results are therefore reliable. From the *Contribution to inertia* it is also evident that 95.87% of the data are represented accurately with this analysis. The first dimension contributes 73% to the inertia, and the second dimension 22.87%. The CA graph is shown in Figure 24.

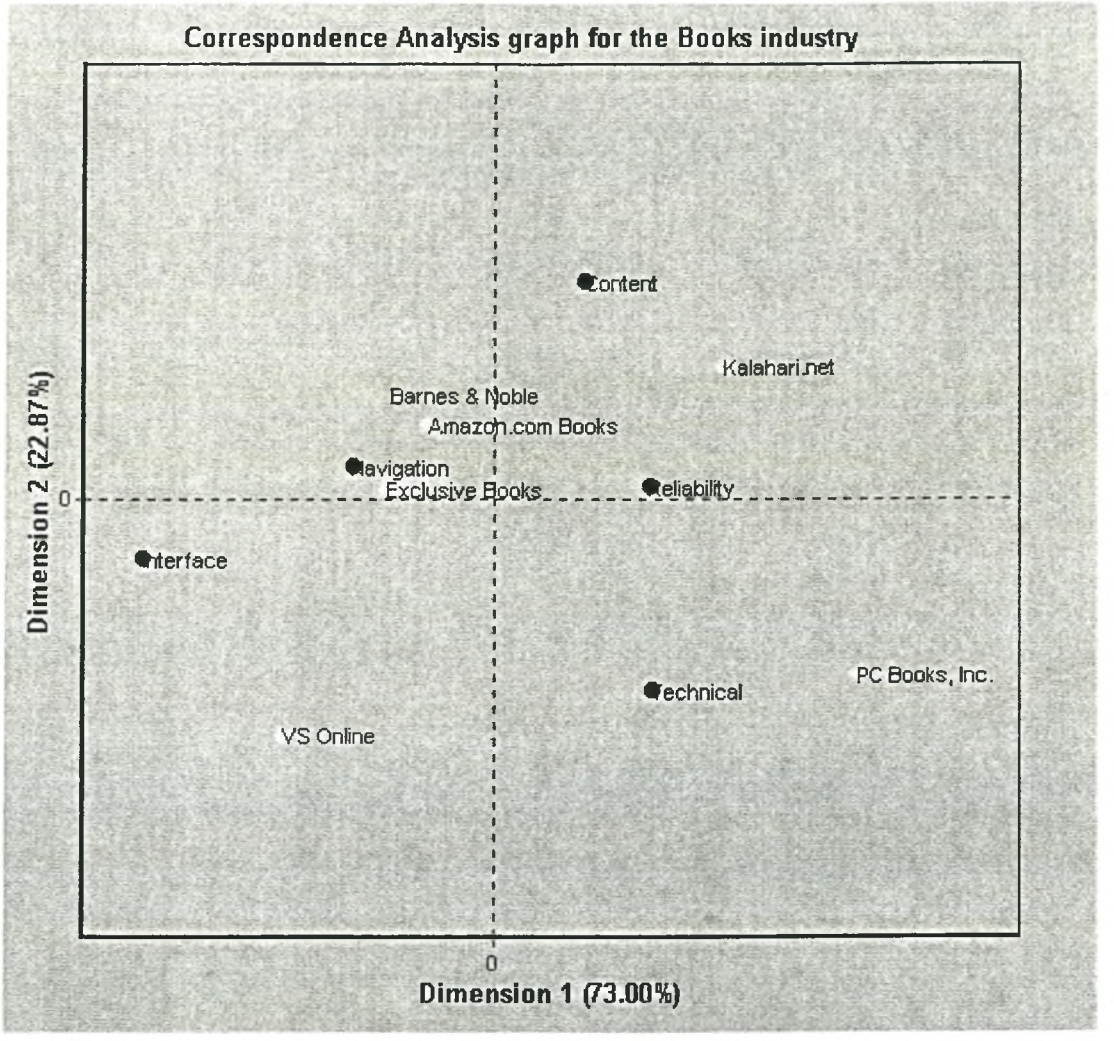


Figure 24 - Correspondence Analysis Graph in the Books Industry

When the graph is analysed, the following conclusions can be drawn.

Analysis of criteria category profiles

Based on the lack of proximity between criteria category profiles, it is evident that the criteria categories form no distinct similar groups in this case. This implies that their profiles differ, in other words two or more specific criteria do not feature considerably more on some sites than on others.

Analysis of web site profiles

Based on proximities, it is evident that the sites of *Amazon.com Books*, *Barnes & Noble* and *Exclusive Books* form a group of similar profiles. It seems therefore that these sites received similar scores for some of the criteria. The other three sites have quite different profiles from this group and from each other.

Analysis of row and column profiles together

Firstly, it is evident that *Amazon.com Books*, *Barnes & Noble* and *Exclusive Books* all lie more or less in the centre of the graph, with the five criteria forming a rough circle around them. This implies that these three sites generally received the highest scores for the criteria. From their close proximity to the *Navigation* criteria category, it can be deduced that these three sites not only received higher scores for *Navigation* than for the other categories, but also that their scores for *Navigation* is higher than the scores that the other three sites received for this category.

Kalahari.net finds itself position almost halfway between the *Content* and *Reliability* criteria categories. Although its total score is not so high, it seems therefore that compared to the other categories, this is where *Kalahari.net* received the highest scores.

PC Books and *VS Online* lie on the periphery of the graph and not in close proximity to any of the categories. It is therefore clear that they have not received particularly high scores for any of the criteria.

Conclusions

Based on the discussion above, the following conclusions can be made:

- Clearly, *Amazon.com*, *Barnes & Noble* and *Exclusive Books* are the industry leaders. Especially in terms of navigation, the other sites can learn a lot from these three.
- *Kalahari.net* needs major improvements, but its overall performance is not so bad. They should start by improving the interface, technical and content aspects of the site.
- *VS Online* and *PC Books* are in a bad state, and need to improve their sites drastically on all terrains.
- When the CA graph and the contingency matrix are analysed together, it is clear that all sites need to improve their interface. *Kalahari.net* and *PC Books* in particular have to make drastic improvements in this area.

To illustrate what further steps can be taken, the situation of *Kalahari.net* is considered in more detail in this example. From the above conclusions, it is clear that *Kalahari.net* should start by improving the interface of their web site. The obvious next question is *how* they should do it. In order to answer this question, the individual results of the site, shown in Figure 25, should be studied.

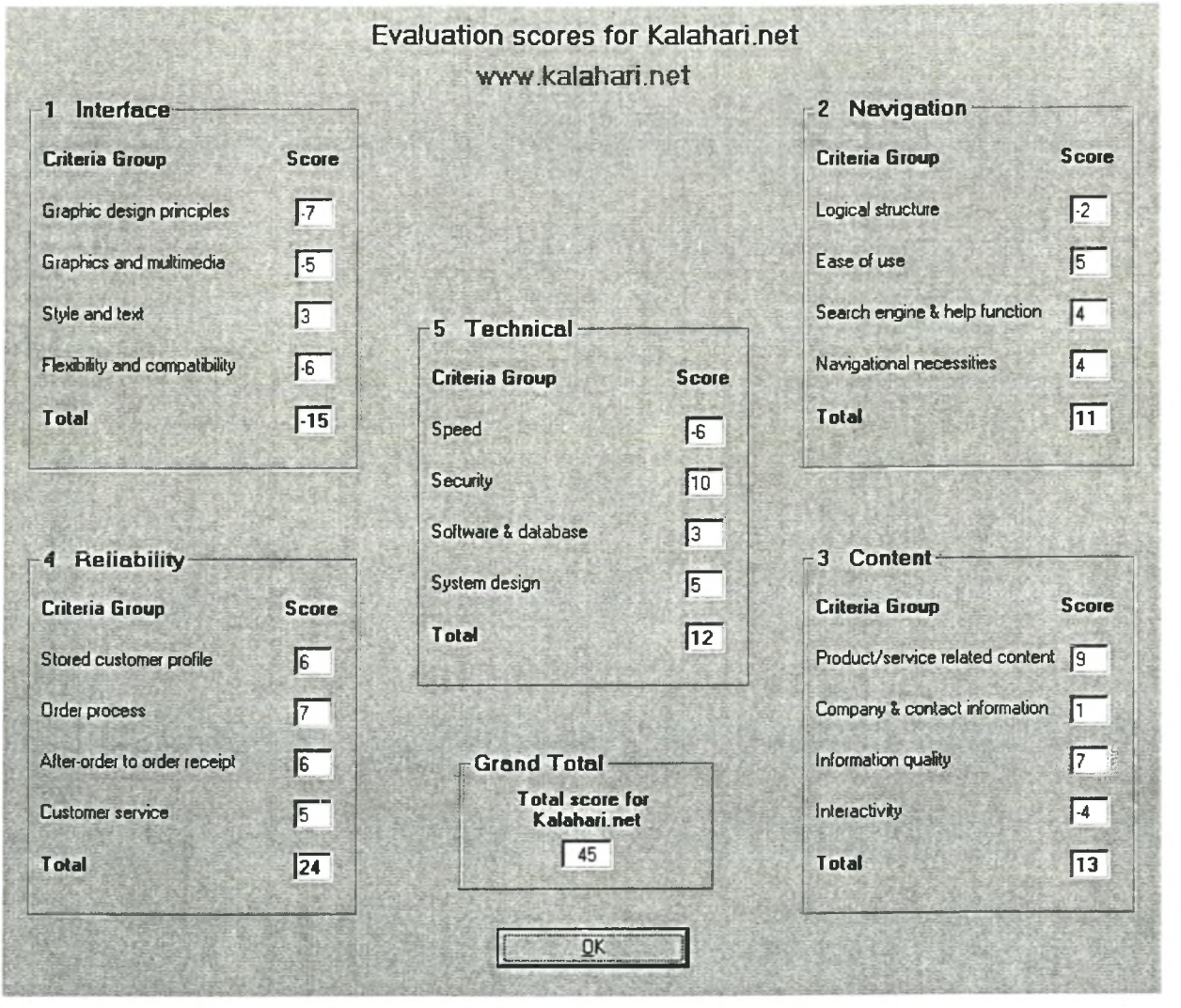


Figure 25 - Evaluation scores for *Kalahari.net*

It is obvious that *Kalahari.net* should start improving their interface by focusing on those criteria groups for which they received the lowest scores. In this case they should take steps to improve the *Graphic design* of the site, the use of *Graphics and Multimedia* and the *Flexibility and compatibility* of the site. When they have done this successfully and move on to improving the next category—their navigational system—the *Logical structure* of the site is the first aspect that should be investigated. At this stage it can be very helpful to study the navigational systems on the sites of *Amazon.com*, *Barnes & Noble* and *Exclusive Books*, as they received the highest scores in this area.

In this manner the scores of each category group that needs improvement can be studied and strategic decisions made on the aspects that should be addressed. Although only *Kalahari.net* is considered here as an example, this process should be repeated for all the sites when recommendations are made on the steps that should be taken to improve them.

5.2.3 The banking industry

Five web sites were evaluated in the banking industry, namely **ABSA** [www.absa.co.za], **First National Bank** [www.fnb.co.za], **Nedbank** [www.nedbank.co.za], **Permanent Bank** [www.permanentbank.co.za] and **Standard Bank** [www.standardbank.co.za]. It was decided to use the sites of five of the most well known banks in South Africa. As these sites provide online banking services and do not sell any products, several criteria in the *Reliability* category were *Not Applicable* for all the sites because no product order process exists. The following contingency matrix was set up by the evaluation software.

	Site Name	Site Address	Interface	Navigation	Content	Reliability	Technical	TOTAL
1	ABSA	www.absa.co.za	19	24	19	11	17	90
2	First National Bank	www.fnb.co.za	-9	19	12	13	0	35
3	Nedbank	www.nedbank.co.za	-9	-2	12	13	11	25
4	Permanent Bank	www.permanentbank.co.za	7	-12	-6	-5	-17	-33
5	Standard Bank	www.standardbank.co.za	8	13	7	9	6	43

Figure 26 - Contingency Matrix for the Banking industry

Performing Correspondence Analysis yielded the numeric results shown in Figure 27.

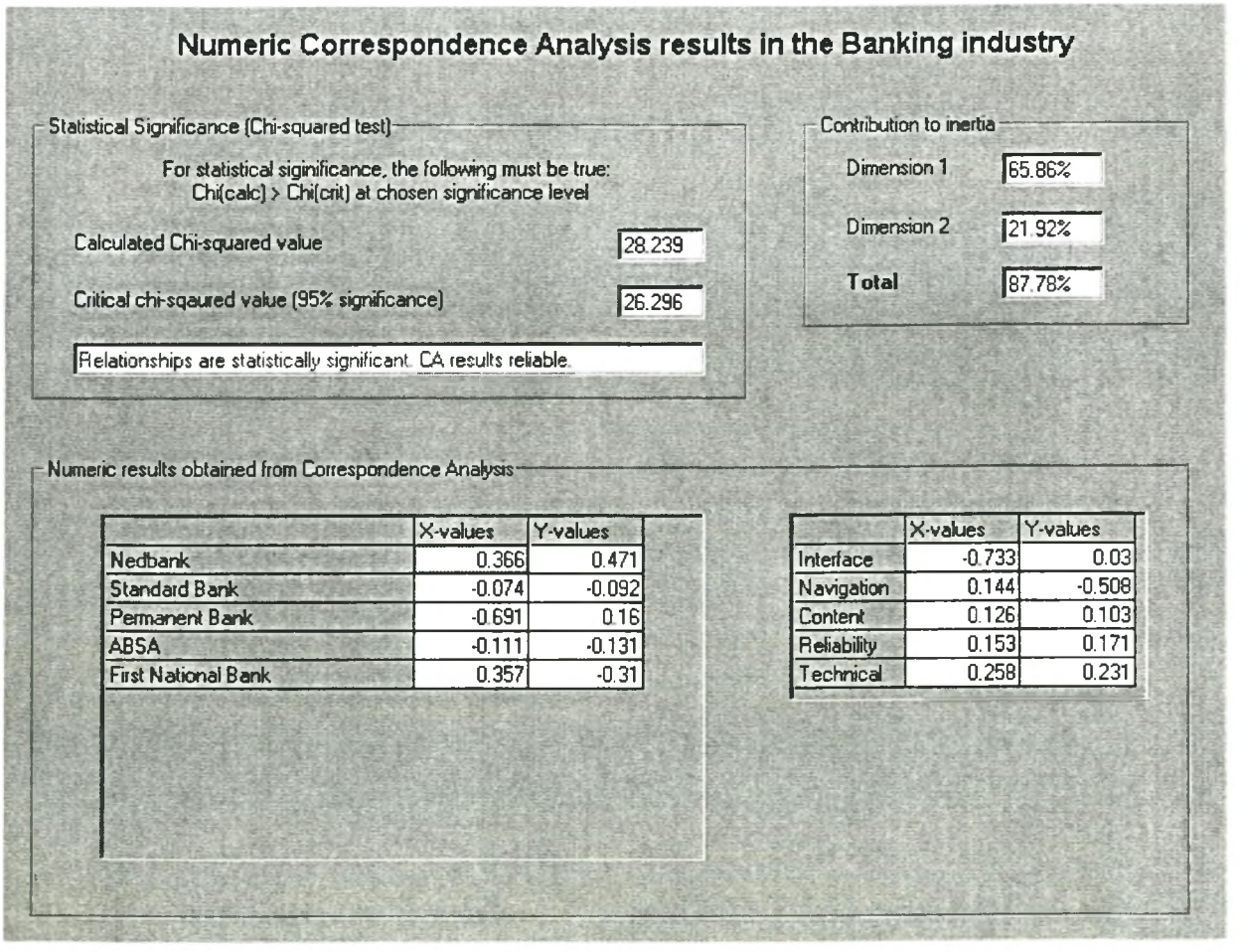


Figure 27 - Correspondence Analysis results for the Banking industry

The results are statistically significant at a 95% confidence level, and the analysis can therefore be continued. As in the previous example, the overall contribution to inertia is also good. The CA graph is shown in Figure 28.

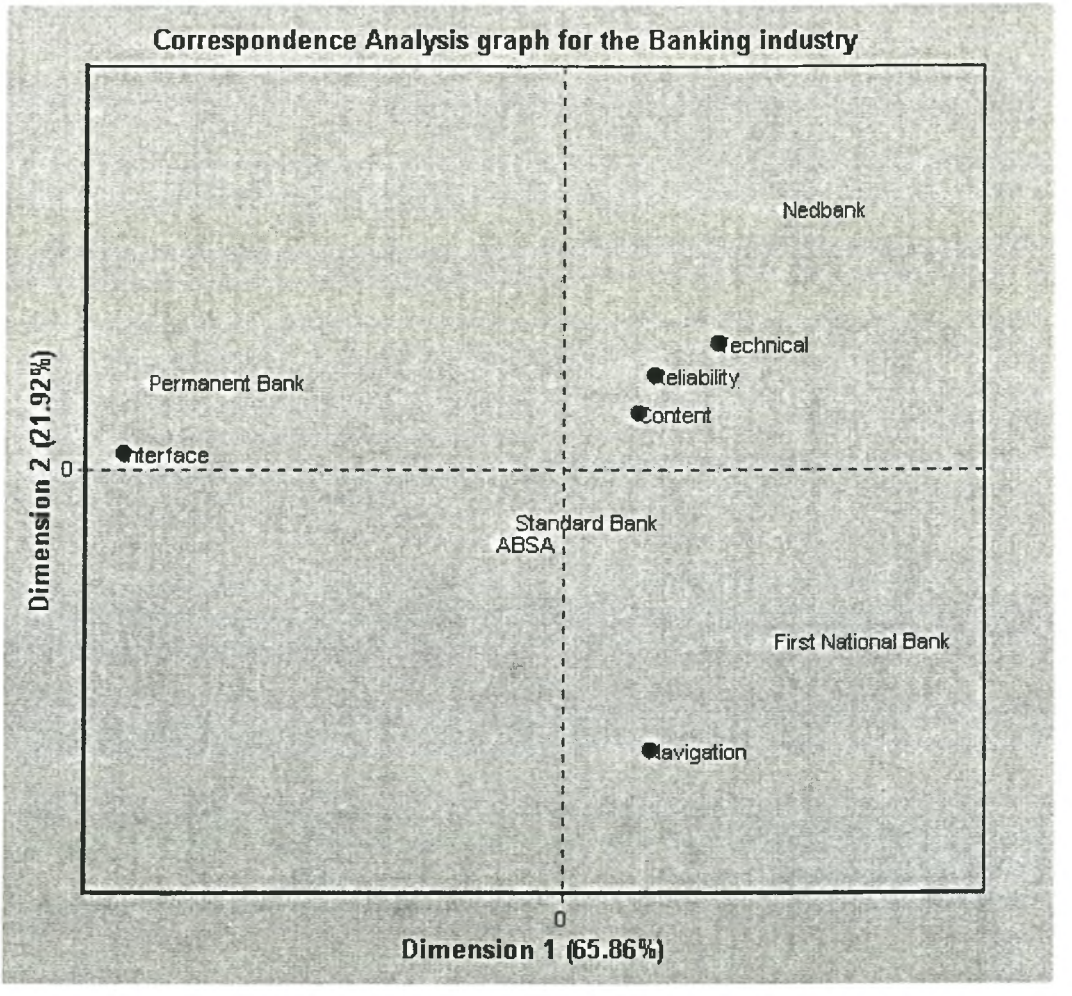


Figure 28 - Correspondence Analysis graph for the Banking industry

The analysis of the graph follows below.

Analysis of criteria category profiles

From the graph it is clear that the criteria categories *Technical*, *Reliability* and *Content* form a group of similar profiles. In these categories, therefore, the ordering of the scores that web sites received are more or less the same compared to their scores on other categories. The *Interface* and *Navigation* categories do not share profile similarities with any of the others.

Analysis of web site profiles

ABSA and *Standard Bank* form a definite group of similar profiles. The other sites are spread out on the graph and form no significant groups.

Analysis of row and column profiles together

Standard Bank and *ABSA* lie more or less in the centre of the graph, with the categories situated all around them. Based on this fact and the sites' close proximity to the origin, it can be deduced that these sites received the best overall scores. As they are not situated particularly close to any of the categories, it can be assumed that they received relatively high scores for all the criteria, not only for some.

Permanent Bank lies far from the origin and near *Interface*, indicating a low overall score with a high score for *Interface* when compared to the other categories.

Although *First National Bank* lies somewhat on the periphery of the graph, it is situated more or less in the centre of the *Technical*, *Reliability*, *Content* and *Navigation* categories, which indicate good scores on these categories. A particularly bad score on *Interface* is evident because of the great distance between *First National Bank* and that point.

Nedbank forms a definite group with the *Technical*, *Reliability* and *Content* categories. Although the point is situated relatively far from the origin, its close proximity to the three category groups indicates a fairly high total score. *Nedbank* is situated far from the *Interface* and *Navigation* categories, which indicates that it received low scores for those categories.

Conclusions

The following conclusions can be made:

- *Nedbank*, *First National Bank* and *Permanent Bank* need to improve their sites on certain aspects because they do not perform as well as the other sites.
- *Nedbank* should start by improving their interface and navigational system.
- *First National Bank* should start by improving their interface and then move on to the other categories situated farthest from it, starting with technical.
- *Permanent Bank* should start by improving their navigational and technical aspects, and then move on to reliability and content.
- *ABSA* and *Standard Bank* seem to be the industry leaders in this case. The other sites should use them as a benchmark for improvements. *Permanent Bank*'s interface and *First National Bank*'s navigation also received high scores relative to their other scores and can be useful to study for improvement ideas.

The individual results for *Nedbank* is shown in Figure 29 to illustrate how they can improve the aspects identified in this section.

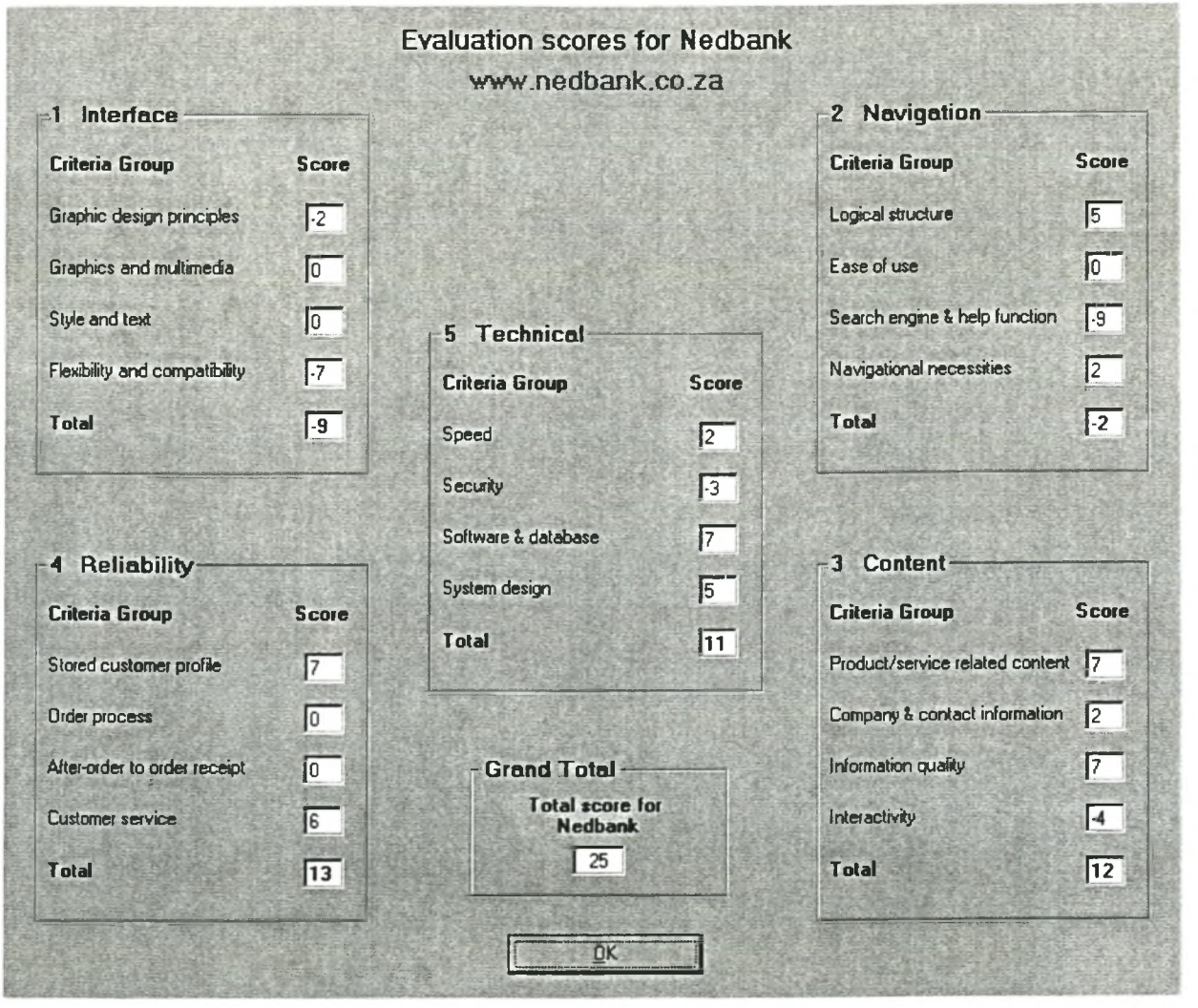


Figure 29 - Evaluation scores for *Nedbank*

When improving their interface, it is clear that *Nedbank* should start by improving the *Graphic design* and *Flexibility* of their site, and then also spend some time on the other two criteria groups. It may be helpful to study *Permanent Bank*'s site—as well as those of the industry leaders (*ABSA* and *Standard Bank*)—as they seemed to have scored well in this area. To improve their navigational system, *Nedbank* should firstly improve their *Search engine and help function*, and then move on to *Ease of use* and the *Navigational necessities* not present on the site.

5.2.4 The ammunition reloading equipment industry

The ammunition reloading equipment industry sells bullets, moulds, dies and other attire used to reload different kinds of ammunition. The following four sites were chosen for evaluation in this industry:

- Joyce Harnady founded **Hornady Manufacturing** [www.hornady.com] in 1949. His philosophy of "Ten bullets through one hole" is still the driving force behind the company's focus on bullets with extreme accuracy.
- **Lee Precision, Inc.** [www.leeprecision.com] is a family-owned business providing reloading equipment since 1958. Their headquarters is situated in Wisconsin, USA.
- William Lyman founded the **Lyman gun sight company** [www.lymanproducts.com] in 1878. He was an inventor of sights, shooting boxes and boating gear. Over the years Lyman steadily built a worldwide reputation for their products, especially for their gun sights and reloading equipment product lines.
- **RCBS** [www.rcbs.com] founded the term *Precisioneered® Reloading*, meaning that they believe their equipment is 'engineered for precision'. They have dealers all over the United States.

All these companies now have web sites with e-commerce capabilities. Evaluation of these sites revealed the following contingency matrix.

	Site Name	Site Address	Interface	Navigation	Content	Reliability	Technical	TOTAL
1	Hornady	www.hornady.com	-18	-5	6	4	-8	-21
2	Lee Precision, Inc.	www.leeprecision.com	-1	5	2	12	0	18
3	Lyman	www.lymanproducts.com	-13	6	-11	2	-15	-31
4	RCBS	www.rcbs.com	18	0	-4	12	-6	20

Figure 30 - Contingency Matrix for the Reloading Equipment industry

The numeric Correspondence Analysis results are shown in Figure 31.

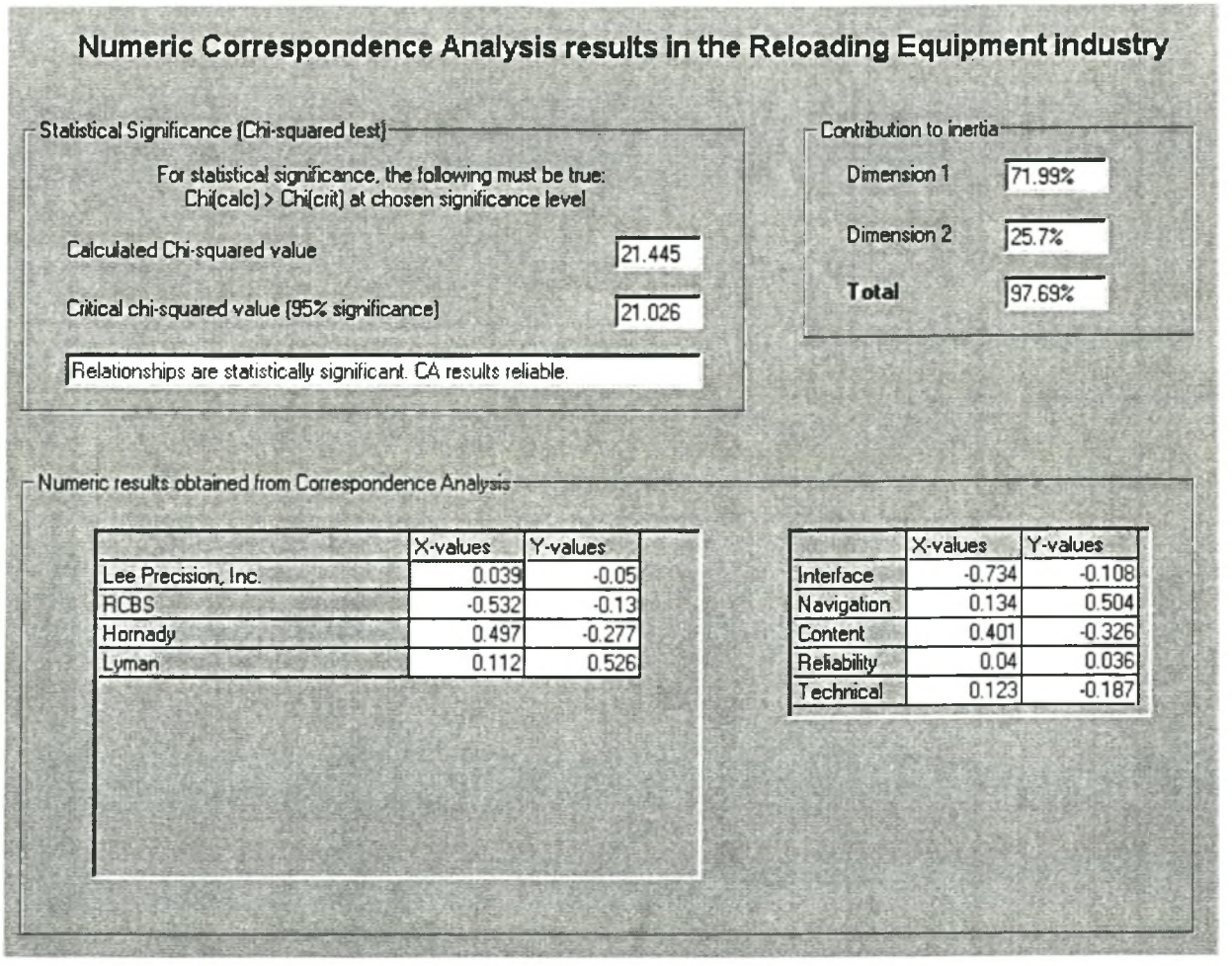


Figure 31 - Correspondence Analysis results for the Reloading Equipment industry

With the results statistically significant, and a very good overall contribution to inertia, the analysis was continued. The CA graph is shown in Figure 32, and the interpretation of the graph follows afterwards.

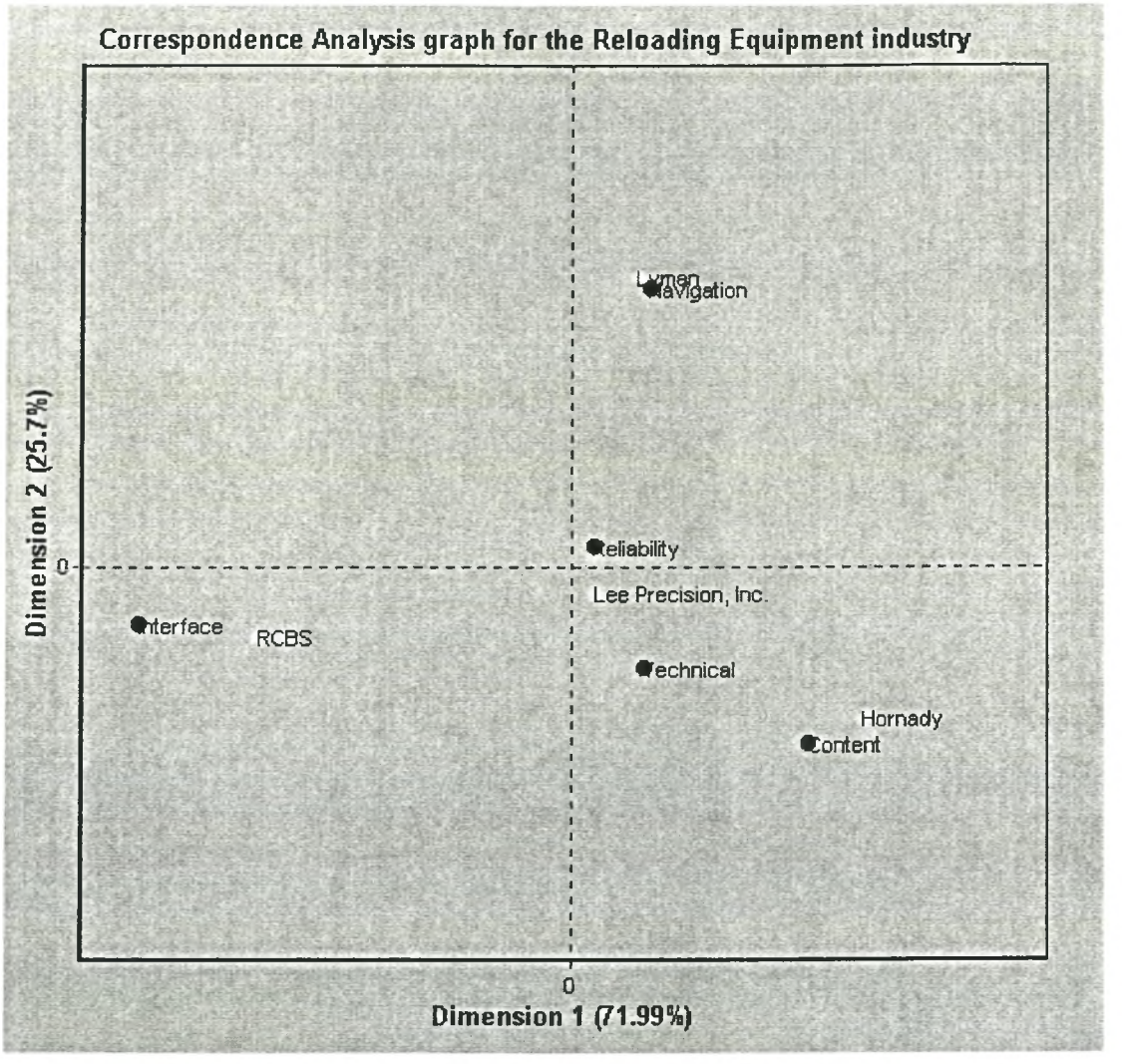


Figure 32 - Correspondence Analysis graph for the Reloading Equipment industry

Analysis of criteria category profiles

The categories are spread out over the graph, which indicates no specific profile groups.

Analysis of web site profiles

The web sites are also spread out over the graph, and it is therefore not possible to form profiles of similar groups. There seems to be much variation in this industry.

Analysis of row and column profiles together

When considering rows and columns together, definite groups can be identified. Based on proximities and its closeness to the origin, it is evident that *Lee Precision* received high scores, particularly on *Reliability* and *Technical*. In the same way *RCBS* forms a group with *Interface*; *Lyman* with *Navigation*; and *Hornady* with *Content*.

When comparing the graph with the contingency matrix, it is interesting to note that *RCBS*, which received the highest score of all the sites, is not situated near the origin. This is because *RCBS* scored particularly well on *Interface*, but fairly bad on the other categories. *Lee Precision*, which also received a good score, has a much more even distribution of scores, and is therefore situated closer to the origin (and also more in the centre of the five categories).

Although *Lyman* received the lowest overall score, it lies very close to *Navigation* on the graph because this is the one category where it *did* score well compared to its other scores.

Conclusions

The following conclusions can be made:

- Considering the fairly low scores most of the sites received, it is clear that there is still much room for improvement for the sites in this industry.
- All the sites need improvement, and they should start with the category groups situated farthest from them. Based on the CA graph, *Lyman* should start with their content, technical and interface aspects; *RCBS* with navigational, content and technical aspects; *Hornady* with their navigational system and interface; and *Lee Precision* with their interface, navigational system and content.

There seems to be some discrepancies between the contingency matrix and the CA Graph in this industry. This problem will be addressed in the next chapter.

5.3 Summary

In this chapter e-commerce web sites in three different industries were evaluated with the e-commerce web site evaluation software, and the results were discussed. The final chapter will examine this software and the Correspondence Analysis technique critically and provide suggestions to solve some of the problems identified. It will also give a general overview of what was accomplished in this research, the quality of the results and the need for further research in certain areas.

6. Conclusions and recommendations

This chapter discusses the main findings of the thesis by drawing together the results from the previous chapters. The chapter starts with three sections that **examine the degree to which each of the three main objectives set in this thesis were realised**. This is done by critically examining **firstly** the e-commerce web site evaluation framework and criteria that were developed; **secondly** the methodology developed to perform e-commerce web site evaluation and the appropriateness and success of Correspondence Analysis as the Multidimensional Scaling technique used in that context; and **thirdly** the software that was developed to implement the framework and statistical analysis. These sections will show how the results obtained in the thesis relate to the literature and theory in this field of study. It also discusses any anomalies and surprising results that deviate from what was expected. Explanations and solutions to these anomalies are also discussed. The chapter concludes with a section that provides a **summary of the recommendations made earlier in the chapter** as well as **some final remarks** in closing.

6.1 Critical examination of the e-commerce web site evaluation framework and criteria

The first objective of the thesis was to **develop a framework and criteria for the comprehensive evaluation of e-commerce web sites**. During the literature study it became apparent that there is a very large amount of literature on general web site evaluation, but that the evaluation criteria advocated in these sources are either obsolete in the context of the Internet as it is operated today, or not comprehensive, or both. Literature on e-commerce web site evaluation proved to be more reliable and revealed a fair amount of research on which the concepts were built. Still, the criteria developed by these authors were mostly very subjective and not properly defined to eliminate ambiguities. Also, a proper framework built on solid academic principles was lacking. The framework and criteria developed in this thesis for e-commerce web site evaluation aimed to address these main needs in the research by doing the following:

- Founding the basic framework on solid business principles and thorough research.
- Making sure the framework provides comprehensive coverage of all aspects of an e-commerce web site.
- Developing criteria within the framework that are as objective as possible.

It is the researcher's opinion that these main aims were fulfilled to a large extent with the framework and criteria that were developed. This conclusion is based on the following remarks:

- The criteria categories developed in the e-commerce web site evaluation framework (see Figure 6 on page 48) is based on the customer buying cycle, which is a solid business principle to ensure comprehensive coverage of all customer processes during a buying cycle.
- The four criteria groups developed within each of the five criteria categories (see Figure 7 on page 50) are based on a combination of firstly all the approaches followed in literature sources reviewed, and secondly the researcher's personal experience as a regular user of e-commerce web sites. This ensures a holistic view of e-commerce web site evaluation and results in a comprehensive evaluation framework that considers all the important aspects of the site.
- The five individual criteria developed within each criteria group (see Section 4.2 on page 51) aimed to capture the essence of each criteria group in the most objective way possible. These criteria were developed by reviewing the multitude of criteria advocated in literature, choosing the ones most relevant to each criteria group. In those groups where the criteria in literature were not sufficient, the researcher added relevant criteria based on personal experience.

The framework and criteria are a good integration between literature and creative thinking, resulting in an e-commerce web site evaluation framework that facilitates the comprehensive evaluation of e-commerce web sites. There are, however, a few reservations about the individual criteria:

- Although the objectivity of criteria was a major driving force behind the development process, the criteria still require a certain measure of human subjectivity, as mentioned in the document. Although the evaluation framework and criteria groups seem to be well developed, it may be necessary to modify or even replace some of the criteria to ensure greater objectivity in the process.
- As discussed in the document, some of the criteria cannot be evaluated without intricate knowledge about the order process of the companies being evaluated. In the thesis, this problem was overcome by sending e-mail to the companies requesting the information needed. These criteria are essential to the evaluation; therefore they can not be left out. It may, however, be possible to modify the individual criteria to make it possible to evaluate these aspects without any information from the companies.

The framework and criteria made it possible to gather evaluation data in a contingency matrix form (see Section 2.4 on page 29). This approach opened the door to a multitude of possible ways to analyse the data.

The researcher therefore concludes that one of the most important contributions of this research is the development of the evaluation framework and 100 criteria to evaluate web sites comprehensively. This framework can be used as a checklist by all e-commerce companies wishing to design or improve their web sites. The framework is a good mixture of literature and real-world situations that contributes much to the need for effective e-commerce web site evaluation. Further research may be helpful to modify some of the individual criteria to make them more objective and easier to evaluate.

6.2 Critical examination of the success of the evaluation method and Correspondence Analysis

The second objective was to **use the framework and sound statistical reasoning to develop a method that can be used to evaluate e-commerce web sites quantitively, using a Multidimensional Scaling technique to plot the evaluation results of different sites within the same industry on a two-dimensional scale**. The e-commerce web site evaluation method uses its three steps to systematically evaluate e-commerce web sites and interpret the results effectively (see Section 4.3 on page 58). The method facilitates this firstly by gathering evaluation data using the e-commerce web site evaluation framework and criteria; secondly by then performing Correspondence Analysis on the results and representing the results graphically; and finally by interpreting these results in a systematic manner. This was the essence of the second objective of the thesis, which was therefore met in full. A few thoughts on the success of the use of Correspondence Analysis in this regard are necessary.

Correspondence Analysis is in its essence a statistical technique that represents complex comparison data on a two-dimensional graphical scale. From this point of view it seems like the perfect technique for this thesis, as it addresses three important aspects of the thesis objectives—namely to have a (1) graphical representation of data based on (2) sound statistical techniques so that (3) useful conclusions about the differences between and similarities of different web sites can be made easily. Statistically, the technique functioned well and generally provided an accurate representation of the data. The theory was implemented correctly and the results were interpreted as dictated by common literature. In the context of this thesis, however, a few anomalies occurred that would have to be investigated further:

- Interpretation of Correspondence Analysis results seemed to become less accurate as the number of sites that were evaluated decreased. For example, when the Correspondence Analysis graph of the Reloading Equipment industry was compared to the Contingency Matrix for that industry, the graph was less clear on some aspects than

the data seemed to indicate (see Section 5.2.4 on page 86). An examination of the χ^2-test for Reloading Equipment reveals that χ^2_{calc} and χ^2_{crit} are very close to each other, which implies that the results are only just significant. This may be due to the small number of sites that were evaluated. It is therefore recommended that the graph is used to make initial conclusions, and that the Contingency Matrix is then used to validate and elaborate on these conclusions, especially when less than 5 sites are evaluated for a specific industry.

- One drawback of the technique is that the results deal with relative positions. Suppose, for example, that the point of a specific web site lies very close to a certain criteria category on the plot. This *does* mean that the site scored well on that category, compared to its other scores. But it does *not* necessarily mean that it received the highest score for that category of all the sites that were evaluated. Other sites with high overall scores and a more even distribution of scores may very well have scored higher on that category. Again, it is recommended that the Correspondence Analysis graph be used in conjunction with the Contingency Matrix, so that this problem can be avoided.
- On examination of evaluation scores it becomes clear that absolute scores differ greatly for different industries. The purpose of the web site evaluation method is to evaluate web sites *per industry*, and the reader should therefore keep in mind that different industries cannot be compared with each other. A good example of this is found in the large difference between the results obtained in the *Books* industry and the *Weapon reloading equipment* industry in section 5.2. Some of the web sites in the *Books* industry received total evaluation scores well into the 100s, while the highest total score for a web site in the *Ammunition reloading equipment* industry was just 20. This indicates that the sites in the *Ammunition reloading equipment* industry are generally of a much lower quality than the web sites in the *Books* industry. The CA graph has no way of communicating this piece of information to the user, and is not intended to do so. It can, however, be useful for organizations in low-scoring industries to examine and benchmark relevant aspects of web sites in high-scoring industries.

It is recommended that further study be conducted in the wider field of statistical mapping. In this thesis, focus was placed on Multidimensional Scaling and Correspondence Analysis as a specific technique within this field. A better group of statistical mapping techniques may exist to perform the analysis of evaluation data. Apart from that, research should also be conducted specifically in the field of Correspondence Analysis output interpretation, as there is still some confusion in that aspect, e.g. how to provide an interpretation that most accurately reflects the output data.

In spite of some drawbacks, it is the researcher's conclusion that, if used correctly, Correspondence Analysis can be a very effective technique to analyse web site evaluation data. Its graphic nature makes it possible to form an overview of the data very quickly, and make some very useful conclusions. It would have been much more difficult to form these conclusions without the help of this graphical representation, which makes it an essential first step to identify important tendencies. When the Correspondence Analysis graph is used with the contingency matrix—and later with each individual site's results as well—it becomes a very powerful tool that can help managers to find definite and concrete ways to improve the performance of their web sites.

If users understand the danger of using the Correspondence Analysis in isolation, it can be concluded that the objective to develop a comprehensive method for the evaluation of e-commerce web sites was met successfully.

6.3 Critical examination of the evaluation software

The third objective of the thesis was to **implement the e-commerce web site evaluation method by developing computer software that enables users to evaluate web sites and plot the results**. Microsoft Visual Basic was used to develop an application with an easy user interface to facilitate the automated evaluation of web sites as proposed by the evaluation method. The objective was met successfully, as users now have the ability to evaluate sites and then view (1) the numerical Correspondence Analysis results per industry, (2) a graph of the overall results per industry and (3) the individual site results quickly. It is possible for users to interpret these results and use it firstly to make general statements about the current performance of companies' e-commerce web sites, and secondly to make valuable suggestions on ways to improve the performance and functionality of these companies' web sites.

No technical errors could be found in the software, and the statistical calculations were done without error with the data supplied. A few drawbacks regarding the general functioning of the program do exist, however:

- Each of the 20 forms used to evaluate criteria groups has a **Next** button and a **Back** button to navigate through the forms. The **Back** button is used to go back to the previous form/forms if the evaluator wants to change some of the previous data. The drawback in the software is that the previous form's data is not remembered when the **Back** button is clicked. In other words, if a user clicks on the **Back** button, he has to re-evaluate the whole criteria group. This also implies that if, for example, he is almost finished with the evaluation, and wants to go back to a previous form much earlier in

the evaluation, he will have to re-evaluate all the forms following the one he actually wanted to change.

- Each evaluation form also has a **Quit** button to terminate the evaluation before it has been completed. When the user clicks this button, all data on that particular web site is lost, and the site is deleted from the database.
- Because of the previous drawback, it is also not possible for a user to stop an evaluation temporarily and continue it later. Once the evaluation is started, it has to be completed; otherwise the data will not be updated in the database.
- A Microsoft Access database is used in this thesis. For Visual Basic to communicate with this database, it is necessary to provide a 'path' so that Visual Basic will know in which directory to look for the database. It is possible to write code so that the software asks the user for the path of the database. However, many users will be confused by such a request, because they may not even know where the database is situated. To solve this problem, the path was hard-coded in Visual Basic to be the same as the directory in which the program files are copied during installation, namely C:\Program Files\Evaluation. No input is thus required from users, but they have to bear in mind that they should install the software in this directory; otherwise the program will not function correctly.
- It is possible for the user to see the framework of the evaluation if he clicks on 'View evaluation framework'. The drawback is that he cannot see exactly where in this framework he finds himself during the evaluation process. Although each evaluation form indicates its position in the framework by using a numbering system, it could be very useful to find some graphic way of indicating to the user where in the evaluation framework he currently finds himself—throughout the evaluation process.

It is possible to eliminate these problems with extra program code. However, this software is for illustrative purposes only, and not for commercial use. The researcher therefore decided not to address these secondary problems that are not essential to the fulfilment of the thesis objectives.

It is also important to note that it takes a considerable amount of time to complete an evaluation. The researcher found that it took an average of two to three working days per industry to evaluate the web sites thoroughly. The evaluators have to spend time to get to know the site (if they have not used it considerably before), and a dedicated effort is needed to complete each of the 20 evaluation screens. It is, however, not a trivial undertaking to evaluate an e-commerce web site, and it is therefore necessary that enough time be spent to perform an accurate and useful evaluation.

There is still some concern about the subjectivity of the criteria and subsequently the reliability of the evaluation data. An area of future research could be to minimize the

problem of subjectivity by expanding the software so that every assessor in the evaluation group can input his data for a particular site into the same database. The software could then report on discrepancies at the lowest level of data entry. Based on this report, the evaluators will then be able to reach agreement and ensure reliable consensus ratings at the lowest level. The contingency matrix can then be generated automatically by the software based on the combined consensus score of all evaluators.

It is possible to write algorithms to automate the interpretation of the graphical output. This is not recommended, however. The interpretation of output is very industry-specific and demands consideration of a combination of many factors relating to the industry and the web sites that are evaluated. Automating the interpretation would definitely speed up the process, but it would also provide clinical answers that may miss some of the important interpretations and even misinterpret some data. Evaluators skilled in Internet behaviour and the industry that is evaluated should conduct the interpretation of the output.

The objective to write software that implements the e-commerce web site evaluation software was met successfully. It is recommended that further projects be established to address the problems that still exist, and therefore to provide added functionality to the software. It is also recommended that a Help file be developed for the software.

6.4 Final statements

Many recommendations were made in the previous three sections. Table 19 provides a summary of these recommendations.

Objective	Recommendations
Framework for e-commerce web site evaluation	▸ Conduct further research in the field of e-commerce web site evaluation and individual criteria to quantify the evaluation. ▸ Based on the research, modify or even replace some of the criteria in the framework to make them (1) more objective and (2) easier to evaluate without the need for internal company information.
E-commerce web site evaluation method	▸ Conduct further research in the wider field of statistical mapping to find out whether or not other alternatives to Multidimensional Scaling exist. ▸ Conduct further research in the field of Correspondence Analysis to get a better understanding of the interpretation of Correspondence Analysis graphs. ▸ Use this research to refine the process of interpretation of the Correspondence Analysis graphs (in conjunction with the contingency matrix data and individual site results data).
Software implementation of the e-commerce web site evaluation method	▸ Write additional code to address some general problems in the software, as outlined in section 6.3.

Table 19 - Summary of recommendations

In the introduction, it was stated that the purpose of this thesis was **to address important needs in the field of web site evaluation as applied to e-commerce in particular**. It is the belief of the researcher that this purpose was accomplished, and that a big step was taken to meet the current needs in this field of study. Although many factors have to be considered for an Internet company to be successful, the web site is one

of the essential aspects that have to be in place effectively for e-commerce business to succeed. Internet companies should be committed to excellence in all areas of business, and this is especially true for their web sites. This belief is summed up perfectly by the author Buck Rodgers who once said, "There are countless ways of attaining greatness, but any road to reaching one's maximum potential must be built on [...] a commitment to excellence and a rejection of mediocrity." This thesis made a contribution towards the achievement of business excellence by providing a systematic approach to improve the performance of e-commerce web sites.

References

Abels, E.G., M.D. White & K. Jahn (1999). A user-based design process for web sites. *OCLC Systems & Services* **15**(1):35-44.

Alexander, J. & M. Tate (1996). *Evaluation criteria*. From [http://hoku.lcc.hawaii.edu/kaz/100/evalCriteria.htm].

Barnes, S.J. and B. Hunt (editors) (2000). *E-Commerce and V-Business : An international money making machine*. Butterworth-Heinemann Publishers, London, England.

Bauer, C. & A. Scharl (2000). Quantitive evaluation of web site content and structure. *Internet Research: Electronic Network Applications and Policy* **10**(1):31-43.

Bennett, J.F. & W.L. Hayes (1960). Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika* **25**: 27-43.

Carrol, J. & P.E. Green (1997). Psychometric methods in marketing : Part II, multidimensional scaling. *Journal of Marketing Research* **34**(3): 193-204.

Carroll, J. & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of the "eckhard-young" composition. *Psychometrika* **35**: 283-319.

Coombs, C.H. (1950). Psychological scaling without a unit of measurement. *Psychological Review* **57**: 148-158.

Cox, T.F. & Cox, M.A.A. (1994). *Multidimensional Scaling*. Chapman & Hall, London, UK.

Cunliffe, Daniel (2000). Developing usable web sites – a review and model. *Internet Research: Electronic Network Applications and Policy* **10**(4): 295-307.

D'Angelo, J. & S.K. Little (1998). Successful web pages: what are they and do they exist? *Information Technology and Libraries* **17**(2):71-81.

Data quality — The foundation of one-to-one customer relationships (no date). White paper by Firstlogic. Downloaded from [http://www.dmreview.com].

Data quality in the information age — the key to a strategic enterprise resource (2000). White paper by Firstlogic. Downloaded from [http://www.dmreview.com].

Davis, F.D. Jr. (1985). *A technology acceptance model for empirically testing new end-user information systems: theory and results*. Doctoral Thesis, Sloan School of Management, Massachusetts Institute of Technology.

Day, A. (1997). A model for monitoring web site effectiveness. *Internet Research: Electronic Networking Applications and Policy* **7**(2): 109-115.

Evaluating websites (no date). From [http://servercc.oakton.edu/~wittman/find/eval.htm], accessed 24/10/2000.

Figallo, C. (1998). *Hosting web communities: building relationships, increasing customer loyalty, and maintaining a competitive edge*. Wiley Computer Publishing, New York.

Flannagan, T. & Safdie, E. (1998). *A practical guide to achieving enterprise data quality*. The Applied Technology Group, Massachusetts. Downloaded from [http://www.techguide.com].

Green, P.E., Carmone, F.J., Smith, S.M. (1989). *Multidimensional scaling : Concepts and applications*. Allyn & Bacon, Needham Heights, Massachusetts.

Greenacre, M. J. (1984). *Theory and applications of Correspondence Analysis*. Academic Press, London.

Hawkins, Steve, D. C. Yen & D.C. Chou (2000). Awareness and challenges of Internet security. *Information Management & Computer Security* **8**(3): 131-143.

Hoffman, D.L. & Franke, G.R. (1986). Correspondence Analysis: Graphical representation of categorical data in marketing research. *Journal of Marketing Research* **23**, August 1986.

Klingberg, F.L. (1941). Studies in measurement of the relations among sovereign states. *Psychometrika* **6** (December): 335-352.

Labuschagne, L. & J.H.P. Eloff (2000). Electronic commerce: the information-security challenge. *Information Management & Computer Security* **8**(3): 154-157.

Liu, C., K.P Arnett & C. Litecky (2000). Design quality of websites for electronic commerce: Fortune 1000 webmasters' evaluations. *Electronic Markets* **10**(2): 120-129.

Misic, M.M. and K.L. Johnson (1999). Benchmarking: a tool for web site evaluation and improvement. *Internet Research: Electronic Network Applications and Policy* **9**(5): 383-392.

Morris, Michael H., P. Berthon & L.F. Pitt (1999). Assessing the structure of industrial buying centers with multivariate tools. *Industrial Marketing Management* **28**: 263-276.

Nielsen, J (1998). *Failure of corporate web sites*. Alertbox for 18 October, available at [www.useit.com/alertbox/981018].

Olsina, L., D. Godoy, G.J. Lafuente & G. Rossi (1999). *Specifying quality characteristics and attributes for web sites*. Paper presented at the First ICSE Workshop on Web Engineering (WebE-99), Los Angeles, USA. Downloaded from [http://citeseer.nj.nec.com/265461.html].

Richardson, M.W. (1938). Multidimensional Psychophysics. *Psychological Bulletin* **35**: 659-660.

Robinson, P.J., C.W. Faris and Y. Wind (1967). *Industrial buying and creative marketing*. Allyn and Bacon, Boston.

Schoel, W.F. and J.P. Guiltinan (1992). *Marketing,* 5th edition. Allyn and Bacon, Needham Heights.

Schubert, P & D. Selz (1999). *Web Assessment – Measuring the effectiveness of electronic commerce sites going beyond traditional marketing paradigms*. Proceedings of the 32nd Hawaii International Conference on System Sciences.

Schubert, P & D. Selz (2000). *Extended Web Assessment Method (EWAM) – Evaluation of e-commerce applications from the customer's viewpoint*. Unpublished article submitted to ECIS 2001 conference, received by e-mail on 21/11/2000 from Petra Schubert (E-mail: Petra.Schubert@unisig.ch).

Shepard, R.N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function, part one. *Psychometrika* **27**: 287-315.

Spiliopoulou, M. (2000). Web usage mining for web site evaluation: Making a site better fit its users. *Communications of the ACM* **43**(8): 127-134.

Takane, Y., F.W. Young, J. DeLeeuw (1977). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika* **42**: 7-67.

Torgerson, W.S. (1958). *Theory and methods of scaling*. John Wiley & Sons Inc., New York.

Weller, S.C. & Romney, A.K. (1990). *Metric Scaling: Correspondence Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences, Series nr. 07-075: Sage Publicatins Ltd, Newbury Park, CA.

Young, G. & A.S. Householder (1938). Discussions of a set of points in terms of their mutual distances. *Psychometrika* **3** (March): 19-22.

Bibliography

Abels, E.G., M.D. White & K. Jahn (1999). A user-based design process for web sites. *OCLC Systems & Services* **15**(1):35-44.

Alexander, J. & M. Tate (1996). *Evaluation criteria*. From [http://hoku.lcc.hawaii.edu/kaz/100/evalCriteria.htm].

Alves, S. (1997). *World Wide Web site evaluation for information professionals*. From [http://www.bf.rmit.edu.au/Dimals/rguides/website.htm]

Barnes, S.J. and B. Hunt (editors) (2000). *E-Commerce and V-Business : An international money making machine*. Butterworth-Heinemann Publishers, London, England.

Bauer, C. & A. Scharl (2000). Quantitive evaluation of web site content and structure. *Internet Research: Electronic Network Applications and Policy* **10**(1):31-43.

Bennett, J.F. & W.L. Hayes (1960). Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika* **25**: 27-43.

Berthon, P., L.F. Pitt & M. Ewing (2000). *Evaluating web sites using multidimensional scaling*. Working Paper Series, February 2000, downloaded from [http://www.cbs.curtin.edu.au/mkt/research/0001.htm].

Bishop, R. (1997). *New packaging, old commodicty: evaluation criteria for web sites*. From [http://www.slis.ualberta.ca/598/robin/pap598.htm].

Borg, I. & P. Groenen (1997). *Modern multidimensional scaling: Theory and applications*. Springer-Verlag, New York.

Carrol, J. & P.E. Green (1997). Psychometric methods in marketing : Part II, multidimensional scaling. *Journal of Marketing Research* **34**(3): 193-204.

Carroll, J. & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of the "eckhard-young" composition. *Psychometrika* **35**: 283-319.

Christoper, M., A. Payne & D. Ballantyne (1991). *Relationship Marketing*. Butterworth Heinemann, Oxford.

Clausen, H. (1999). User-oriented evaluation of library and information centre web sites. *New Library World* **100**(1146):5-10.

Coombs, C.H. (1950). Psychological scaling without a unit of measurement. *Psychological Review* **57**: 148-158.

Cox, T.F. & Cox, M.A.A. (1994). *Multidimensional Scaling*. Chapman & Hall, London, UK.

Cunliffe, Daniel (2000). Developing usable web sites – a review and model. *Internet Research: Electronic Network Applications and Policy* **10**(4): 295-307.

D'Angelo, J. & S.K. Little (1998). Successful web pages: what are they and do they exist? *Information Technology and Libraries* **17**(2):71-81.

Data quality — The foundation of one-to-one customer relationships (no date). White paper by Firstlogic. Downloaded from [http://www.dmreview.com].

Data quality in the information age — the key to a strategic enterprise resource (2000). White paper by Firstlogic. Downloaded from [http://www.dmreview.com].

Davis, F.D. Jr. (1985). *A technology acceptance model for empirically testing new end-user information systems: theory and results*. Doctoral Thesis, Sloan School of Management, Massachusetts Institute of Technology.

Day, A. (1997). A model for monitoring web site effectiveness. *Internet Research: Electronic Networking Applications and Policy* **7**(2): 109-115.

Druse, J. (accessed 24/10/2000). *Web site evaluation criteria*. From [http://www.washburn.edu/mabee/reviews.html].

English, Larry P. (1999). *The intelligent enterprise—beyond data cleansing, a white paper for Vality*. Information Impact International, Inc. Downloaded from [http://www.dmreview.com].

Evaluating websites (no date). From [http://servercc.oakton.edu/~wittman/find/eval.htm], accessed 24/10/2000.

EWAM Evaluation Form: Assessing the attractiveness of web sites (no date). From [http://e-business.fhbb.ch/survey/ewam.nsf/assessment], accessed 21/11/2000.

Figallo, C. (1998). *Hosting web communities: building relationships, increasing customer loyalty, and maintaining a competitive edge*. Wiley Computer Publishing, New York.

Flannagan, T. & Safdie, E. (1998). *A practical guide to achieving enterprise data quality*. The Applied Technology Group, Massachusetts. Downloaded from [http://www.techguide.com].

Grassian, E. (1997). *Thinking critically about World Wide Web resources*. From [http://www.library.ucla.edu/libraries/college/help/critical/index.htm].

Green, P.E., Carmone, F.J., Smith, S.M. (1989). *Multidimensional scaling : Concepts and applications*. Allyn & Bacon, Needham Heights, Massachusetts.

Hawkins, Steve, D. C. Yen & D.C. Chou (2000). Awareness and challenges of Internet security. *Information Management & Computer Security* **8**(3): 131-143.

Hoffman, D.L. & Franke, G.R. (1986). Correspondence Analysis: Graphical representation of categorical data in marketing research. *Journal of Marketing Research* **23**, August 1986.

Klingberg, F.L. (1941). Studies in measurement of the relations among sovereign states. *Psychometrika* **6** (December): 335-352.

Labuschagne, L. & J.H.P. Eloff (2000). Electronic commerce: the information-security challenge. *Information Management & Computer Security* **8**(3): 154-157.

Lebart, L., A. Morineau & K.M. Warwick (1984). *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices.* John Wiley & Sons, New York.

Liu, C., K.P Arnett & C. Litecky (2000). Design quality of websites for electronic commerce: Fortune 1000 webmasters' evaluations. *Electronic Markets* **10**(2): 120-129.

Mandernack, S. (2000). *Web site evaluation chart.* From [http://www.lib.purdue.edu/StudentInstruction/web_site_evaluation_chart.html].

Misic, M.M. and K.L. Johnson (1999). Benchmarking: a tool for web site evaluation and improvement. *Internet Research: Electronic Network Applications and Policy* **9**(5): 383-392.

Morris, Michael H., P. Berthon & L.F. Pitt (1999). Assessing the structure of industrial buying centers with multivariate tools. *Industrial Marketing Management* **28**: 263-276.

Nielsen, J (1998). *Failure of corporate web sites.* Alertbox for 18 October, available at [www.useit.com/alertbox/981018].

Olsina, L., D. Godoy, G.J. Lafuente & G. Rossi (1999). *Specifying quality characteristics and attributes for web sites.* Paper presented at the First ICSE Workshop on Web Engineering (WebE-99), Los Angeles, USA. Downloaded from [http://citeseer.nj.nec.com/265461.html].

Parsons, J.J. and D. Oja (2000). *New Perspectives on Computer Concepts* (4th edition). Course Technology, Cambridge.

Reinken Information Solutions (1997). *Web site evaluation criteria.* From [http://www.reinken.com/web_eval.html]

Richardson, M.W. (1938). Multidimensional Psychophysics. *Psychological Bulletin* **35**: 659-660.

Robinson, P.J., C.W. Faris and Y. Wind (1967). *Industrial buying and creative marketing.* Allyn and Bacon, Boston.

Rogers, A., C. Bakken & S. Armstrong (2000). *Web page evaluation criteria*. From [http://www.ux1.eiu.edu/~cfmgb/web.htm]

Schiffman, S.S., M.L. Reynolds & F.W. Young (1981). *Introduction to multidimensional scaling: Theory, methods and applications*. Academic Press, New York.

Schoel, W.F. and J.P. Guiltinan (1992). *Marketing,* 5th edition. Allyn and Bacon, Needham Heights.

Schubert, P & D. Selz (1999). *Web Assessment – Measuring the effectiveness of electronic commerce sites going beyond traditional marketing paradigms*. Proceedings of the 32nd Hawaii International Conference on System Sciences.

Schubert, P & D. Selz (2000). *Extended Web Assessment Method (EWAM) – Evaluation of e-commerce applications from the customer's viewpoint*. Unpublished article submitted to ECIS 2001 conference, received by e-mail on 21/11/2000 from Petra Schubert (E-mail: Petra.Schubert@unisig.ch).

Selz, D. & P. Schubert (1998). *Web Assessment – A model for the evaluation and the assessment of successful electronic commerce applications*. Proceedings of the 31st Hawaii International Conference on System Sciences.

Shepard, R.N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function, part one. *Psychometrika* **27**: 287-315.

Sowards, S.W. (1997). "Save the time of the surfer": evaluating web sites for users. *Library Hi Tech* **15**(3-4):155-158.

Spiliopoulou, M. (2000). Web usage mining for web site evaluation: Making a site better fit its users. *Communications of the ACM* **43**(8): 127-134.

Takane, Y., F.W. Young, J. DeLeeuw (1977). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika* **42**: 7-67.

Tillman, H.N. (2000). *Evaluating quality on the net*. From [http://www.hopetillman.com/findqual.html].

Torgerson, W.S. (1958). *Theory and methods of scaling*. John Wiley & Sons Inc., New York.

Watson, R.T., P. Berthon, L.F. Pitt & G.M. Zinkhan (2000). *Electronic Commerce*. The Dreyden Press, Orlando.

Weller, S.C. & Romney, A.K. (1990). *Metric Scaling: Correspondence Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences, Series nr. 07-075: Sage Publicatins Ltd, Newbury Park, CA.

Young, G. & A.S. Householder (1938). Discussions of a set of points in terms of their mutual distances. *Psychometrika* **3** (March): 19-22.

Appendix A – Full list of QEM quality characteristics developed by Olsina et al. (1999)

1. Usability

1.1 Global Site Understandability
- 1.1.1 Global Organization Scheme
  - *1.1.1.1 Site Map*
  - *1.1.1.2 Table of Content*
  - *1.1.1.3 Alphabetical Index*
- 1.1.2 Quality of Labeling System
- 1.1.3 Student-oriented Guided Tour
- 1.1.4 Image Map (Campus/Buildings)

1.2 On-line Feedback and Help Features
- 1.2.1 Quality of Help Features
  - *1.2.1.1 Student-oriented Explanatory Help*
  - *1.2.1.2 Search Help*
- 1.2.2 Web-site Last Update Indicator
  - *1.2.2.1 Global*
  - *1.2.2.2 Scoped (per sub-site or page)*
- 1.2.3 Addresses Directory
  - *1.2.3.1 E-mail Directory*
  - *1.2.3.2 Phone-Fax Directory*
  - *1.2.3.3 Post mail Directory*
- 1.2.4 FAQ Feature
- 1.2.5 On-line Feedback
  - *1.2.5.1 Questionnaire Feature*
  - *1.2.5.2 Guest Book*
  - *1.2.5.3 Comments*

1.3 Interface and Aesthetic Features
- 1.3.1 Cohesiveness by Grouping Main Control Objects
- 1.3.2 Presentation Permanence and Stability of Main Controls
  - *1.3.2.1 Direct Controls Permanence*
  - *1.3.2.2 Indirect Controls Permanence*
  - *1.3.2.3 Stability*
- 1.3.3 Style Issues
  - *1.3.3.1 Link Color Style Uniformity*
  - *1.3.3.2 Global Style Uniformity*
  - *1.3.3.3 Global Style Guide*
- 1.3.4 Aesthetic Preference

1.4 Miscellaneous Features
- 1.4.1 Foreign Language Support
- 1.4.2 What's New Feature
- 1.4.3 Screen Resolution Indicator

2. Functionality

2.1 Searching and Retrieving Issues
- 2.1.1 Web-site Search Mechanisms
  - *2.1.1.1 Scoped Search*
    - *2.1.1.1.1 People Search*
    - *2.1.1.1.2 Course Search*
    - *2.1.1.1.3 Academic Unit Search*
  - *2.1.1.2 Global Search*
- 2.1.2 Retrieve Mechanisms
  - *2.1.2.1 Level of Retrieving Customization*
  - *2.1.2.2 Level of Retrieving Feedback*

2.2 Navigation and Browsing Issues
- 2.2.1 Navigability
  - *2.2.1.1 Orientation*
    - *2.2.1.1.1 Indicator of Path*
    - *2.2.1.1.2 Label of Current Position*
  - *2.2.1.2 Average of Links per Page*
- 2.2.2 Navigational Control Objects
  - *2.2.2.1 Presentation Permanence and Stability of Contextual (sub-site) Controls*
    - *2.2.2.1.1 Contextual Controls Permanence*
    - *2.2.2.1.2 Contextual Controls Stability*
  - *2.2.2.2 Level of Scrolling*
    - *2.2.2.2.1 Vertical Scrolling*
    - *2.2.2.2.2 Horizontal Scrolling*

- 2.2.3 Navigational Prediction
  - *2.2.3.1 Link Title (link with explanatory help)*
  - *2.2.3.2 Quality of Link Phrase*
- **2.3 Student-oriented Domain-related Features**
  - 2.3.1 Content Relevancy
    - *2.3.1.1 Academic Unit Information*
      - *2.3.1.1.1 Academic Unit Index*
      - *2.3.1.1.2 Academic Unit Sub-sites*
    - *2.3.1.2 Enrollment Information*
      - *2.3.1.2.1 Entry Requirement Information*
      - *2.3.1.2.2 Form Fill/Download*
    - *2.3.1.3 Degree Information*
      - *2.3.1.3.1 Degree Index*
      - *2.3.1.3.2 Degree Description*
      - *2.3.1.3.3 Degree Plan/Course Offering*
      - *2.3.1.3.4 Course Description*
        - *2.3.1.3.4.1 Comments*
        - *2.3.1.3.4.2 Syllabus*
        - *2.3.1.3.4.3 Scheduling*
    - *2.3.1.4 Student Services Information*
      - *2.3.1.4.1 Services Index*
      - *2.3.1.4.2 Healthcare Information*
      - *2.3.1.4.3 Scholarship Information*
      - *2.3.1.4.4 Housing Information*
      - *2.3.1.4.5 Cultural/Sport Information*
    - *2.3.1.5 Academic Infrastructure Information*
      - *2.3.1.5.1 Library Information*
      - *2.3.1.5.2 Laboratory Information*
      - *2.3.1.5.3 Research Results Information*
  - 2.3.2 On-line Services
    - *2.3.2.1 Grade/Fees on-line Information*
    - *2.3.2.2 Web Service*
    - *2.3.2.3 FTP Service*
    - *2.3.2.4 News Group Service*

3. Site Reliability

- **3.1 Nondeficiency**
  - 3.1.1 Link Errors
    - *3.1.1.1 Dangling Links*
    - *3.1.1.2 Invalid Links*
    - *3.1.1.3 Unimplemented Links*
  - 3.1.2 Miscellaneous Errors or Drawbacks
    - *3.1.2.1 Deficiencies or absent features due to different browsers*
    - *3.1.2.2 Deficiencies or unexpected results (e.g. non-trapped search errors, frame problems)*
    - *3.1.2.3 Dead-end Web Nodes*
    - *3.1.2.4 Destination Nodes (unexpectedly) under construction*

4. Efficiency

- **4.1 Performance**
  - 4.1.1 Static Page Size
- **4.2 Accessibility**
  - 4.2.1 Information Accessibility
    - *4.2.1.1 Support for text-only version*
    - *4.2.1.2 Readability by deactivating Browser Image feature*
      - *4.2.1.2.1 Image Title*
      - *4.2.1.2.2 Global Readability*
  - 4.2.2 Window Accessibility
    - *4.2.2.1 Number of panes regarding frames*
    - *4.2.2.2 Non-frame Version*

Appendix B – Questionnaire to Fortune 1000 companies' web masters

Please share your thoughts about the critical design factors of a website. This survey is completely voluntary and your responses will be confidential. If you would like a copy of the results, please request by e-mail to cc2@cobilan.msstate.edu.

Please rate the following in Section A, B, C, D, E, and F by checking the number that corresponds most closely to your perception of its critical (importance) to the Web site design.

1 = completely unimportant
2 = moderately unimportant
3 = somewhat unimportant
4 = neither important/unimportant
5 = somewhat important
6 = moderately important
7 = completely important

Section A: In order to attract customers, please consider indicators to enhance Information Quality of a website.

(1) flexible information presentation (appropriate amounts of sound, graphic, text, image, animation uses considering users' hardware constraints)	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(2) customized information presentation (provide different user interfaces for different types of customers)	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(3) information that is relevant to the customer (such as company, product, service information)	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(4) provide a feature to compare products/services with those of competitors	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(5) provide a feature to differentiate products/services with those of competitors	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(6) provide accurate information	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(7) provide price information	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(8) provide complete products/services descriptions	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(9) provide timely information	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(10) enhance perceived quality of products/services	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(11) meet ethical standards (such as no misleading or deceptive info.)	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(12) Provide information to support business objectives	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			

Section B: In order to attract customers, please consider indicators to enhance Learning Component of a website.

(1) well organized hyperlinks	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(2) a help function	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(3) customized search functions	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(4) interactive feedback between customers and business	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(5) interactive communications among the customers	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			

Section C: In order to attract customers, please consider indicators to enhance Playfulness of a website.

(1) user enjoyment during the Web site visit	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(2) user participation/interaction during the website visit	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(3) user excitement during the website visit	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(4) charming features to attract customers	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(5) capturing the user's attention during the website visit	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			

Overall, items in above sections represent the critical success factors for the design of a website to ATTRACT customers

completely disagree 1 2 3 4 5 6 7 completely agree

Section D: In order to gain customers trust, please consider indicators to enhance System Quality of a website.

(1) provide security	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(2) design to include information on all functional areas involved in the business process	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(3) balanced security and ease of use payment methods (such as credit card, cash, electronic check, etc.)	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(4) high speed of accessing the web site	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(5) ease of correcting server's errors such as run time error, inability to connect, etc.	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(6) insure correct transactions	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			

Section E: In order to gain customers trust, please consider indicators to enhance System Use of a website.

(1) customers control a transaction process	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(2) gaining customers confidence during the transaction	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(3) providing ease of use for the transaction	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(4) allowing customers to track-order status	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(5) keeping the customers information confidential	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			

Overall, items in sections D and E represent factors for a well designed website, to encourage customers to TRUST on-line transactions.

(1) the site is reliable	completely disagree	1	2	3	4	5	6	7	completely agree
(2) the site is dependable	completely disagree	1	2	3	4	5	6	7	completely agree
(3) the site is trustworthy	completely disagree	1	2	3	4	5	6	7	completely agree

Section F: In order to obtain customer satisfaction, please consider indicators to enhance Service Quality of a website.

(1) provide quick responsiveness to customers	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(2) provide assurance to solve customers' problems	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(3) empathy to customers' problems	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			
(4) provide follow-up services to customers	1	2	3	4	5	6	7
Does your website incorporate this design factor?			Yes	No			

Overall, items in sections F represent well-designed factors for a website to obtain CUSTOMER SATISFACTION.

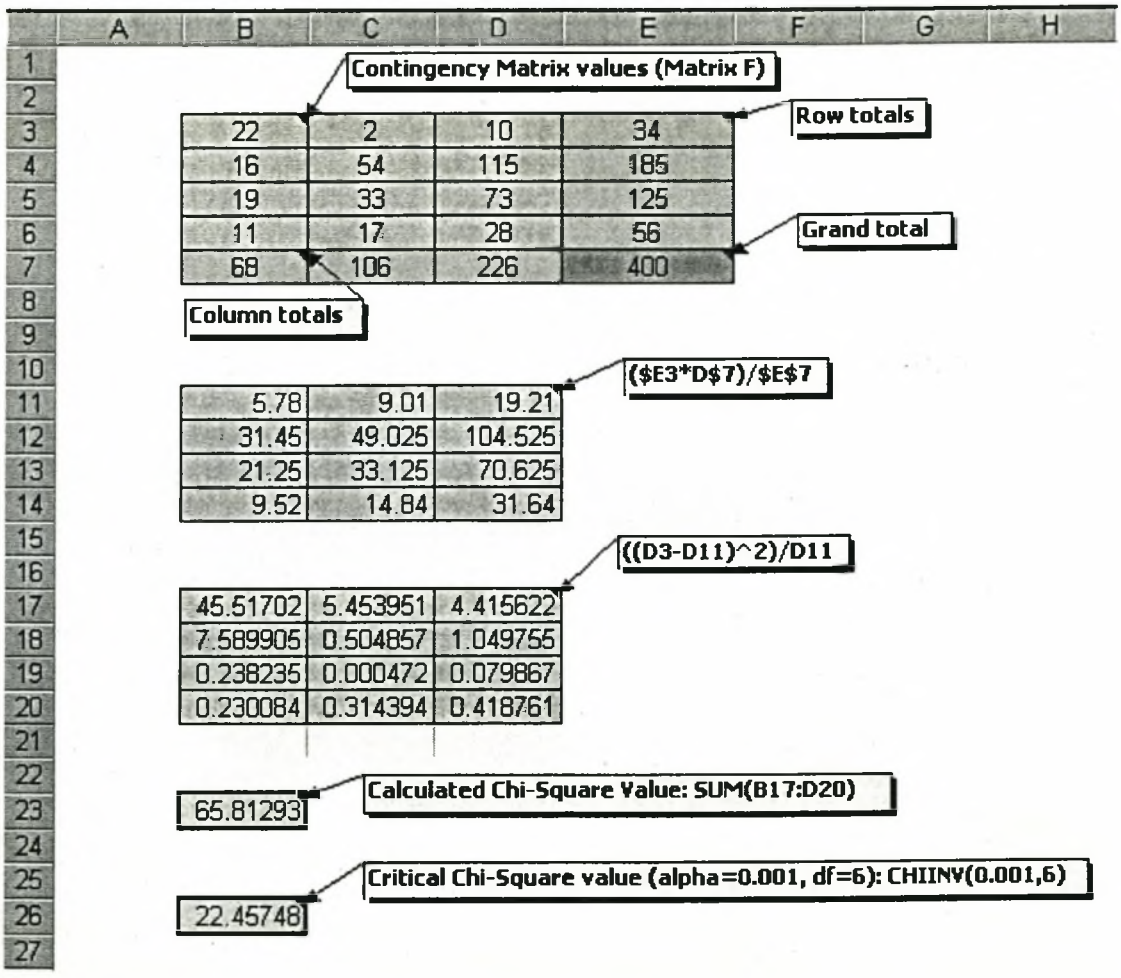
(1) the site meets customer demands								
completely disagree	1	2	3	4	5	6	7	completely agree
(2) customer will be pleased for the service								
completely disagree	1	2	3	4	5	6	7	completely agree

Overall, items in above sections provide an adequate representation of well-designed factors for an electronic market on a website (for example, to attract customers on-line, make them trustfully purchase, and obtain customer satisfaction)

completely disagree	1	2	3	4	5	6	7	completely agree

Appendix C – Calculations for Correspondence Analysis example in section 3.3.3

Chi-Square calculations in Excel:



Matlab code to perform Correspondence Analysis on a Contingency Matrix F:

```
'Put contingency matrix in F';

'Normalise F --> Matrix H';
'------------------------';
clear H
[r,k] = size(F);

for i = 1:r;
   for j = 1:k;
      rowsum(i) = sum(F(i,:));
      colsum(j) = sum(F(:,j));
      H(i,j) = F(i,j)/sqrt(rowsum(i)*colsum(j));
   end
end


'Get the SVD of H --> Matrices U, S and V';
'----------------------------------------';
clear U
clear S
clear V
clear s

[U,S,V] = svd(H,0);
s = svd(H);


'Rescale SVD --> Matrices X and Y';
'--------------------------------';
clear X
clear Y

[r,k]=size(F);

total=0;
for i = 1:r
   f(i) = sum(F(i,:));

   for j = 1:k;
      g(j) = sum(F(:,j));
      total=total + F(i,j);
   end

end

[rn,kn] = size(U);

for i = 1:rn;
   for j = 1:kn;
      X(i,j) = U(i,j)*sqrt(total/f(i));
   end
end

[rm,km] = size(V);
```

```
for i = 1:rm;
   for j = 1:km;
      Y(i,j) = V(i,j)*sqrt(total/g(i));
   end
end


'Weight X and Y --> Matrices A and B';
'-----------------------------------';
clear A
clear B

r = size(X,1);
k = size(s,1);
rn = size(Y,1);
km = size(s,1);

for i = 1:r;
   for j = 1:k;
      A(i,j) = sqrt(s(j,1))*X(i,j);
   end
end

for i = 1:rn;
   for j = 1:km;
      B(i,j) = sqrt(s(j,1))*Y(i,j);
   end
end
```

Appendix D – Graphical representation of 100 Evaluation Criteria

E-commerce Web Site Evaluation Framework and Criteria

1 Interface	
Group 1: Graphic design principles	
1	Home page concise and clear
2	Effective use of white space
3	Effective and consistent use of colour
4	Effective and consistent use of backgrounds
5	Effective graphics/typeface/colour combinations
Group 2: Graphics and multimedia	
1	Site visually attractive
2	Makes contribution to understanding and navigation of site
3	Icons easy to understand
4	Not excessively used
5	Size of media no negative impact on loading times
Group 3: Style and text	
1	Style of pages consistent
2	Typefaces consistent and easy to read
3	Good spelling and grammar
4	Text concise and relevant
5	Purpose of site made clear on home page
Group 4: Flexibility and compatibility	
1	Pages sized to fit in browser window
2	Printable versions of certain pages available
3	Text-only version available
4	Foreign language support available
5	Accommodations made for disabled users

2 Navigation	
Group 1: Logical structure	
1	Intelligible, straightforward organising scheme
2	Content logically structured in different sections and levels
3	Menus understandable and straightforward
4	Site map/table of contents available
5	Consistent navigation throughout site
Group 2: Ease of use	
1	Easy to find site
2	Easy to explore specific idea or subject
3	Easy to return to main page
4	Easy to find specific information
5	Easy to access complete product/service range
Group 3: Search engine and help function	
1	Easy to use search engine
2	Search engine accurate
3	Good description of search engine findings
4	No search engine errors
5	Help function easy to use
Group 4: Navigational necessities	
1	No broken links
2	No 'under construction' pages
3	Links clearly discernable, well labelled and defined
4	Clear label of current position on site
5	Effective use of frames, non-frames version available

5 Technical	
Group 1: Speed	
1	Fast home page loading speed
2	Fast sub-page loading speed
3	Good perceived use of caching
4	Good consideration of non-broadband users
5	Good perceived speed of database
Group 2: Security	
1	Security systems accredited
2	Secure payment systems used
3	Privacy of users protected
4	Security protocols communicated well
5	Security certificates adequate
Group 3: Software and database	
1	Good cross-browser capability
2	Users advised on ideal browser and resolution
3	Database software adequate for size of database
4	Good data transfer between systems
5	No perceived duplication of data in database
Group 4: System design	
1	Precise operation and computation
2	Good resolution compatibility
3	Good integration with systems of users
4	Good integration of different systems on site
5	Global accessibility to products/service available

4 Reliability	
Group 1: Stored customer profile	
1	Easy to register on site
2	High perceived benefits from registering
3	Easy to log in to site
4	Adjustable customer profile stored
5	Guided ordering using customer profile available
Group 2: Order process	
1	Transparent, interactive and easy order process
2	Easy selection of generic services
3	Alternative methods of ordering/payment available
4	Good stock availability
5	Acknowledgement of order sent to customer
Group 3: After-order to order receipt	
1	On-line order tracking available
2	Effective payment settlement
3	Confirmation of order dispatch sent to customer
4	On-time delivery
5	Full order delivery
Group 4: Customer service	
1	Feedback forms available
2	Good after-sales support
3	Quick reply on e-mail enquiries
4	General courtesy of company good
5	Good after-sales relationship building

3 Content	
Group 1: Product/service related content	
1	Extensive product/service information available
2	Price information and cost benefits communicated
3	Adequate breadth of product range
4	Adequate amount of advertising of own products
5	Adequate amount of advertising by other companies
Group 2: Company and contact information	
1	Full company information available
2	Terms and conditions easily accessed
3	E-mail addresses of employees available
4	Telephone and fax numbers available
5	Mailing address and physical address available
Group 3: Information quality	
1	Content current and updated
2	Content relevant to purpose of site
3	Content concise and non-repetitive
4	High perceived quality of product/service
5	Quality of advertisements high
Group 4: Interactivity	
1	User able to customise content to fit needs
2	Large amount of personalisation possible
3	Easy access to online community
4	High perceived value of online community
5	High degree of interactivity in community