

From interactions to transactions : designing the trust experience for business-to-consumer electronic commerce

Citation for published version (APA):

Egger, F. N. (2003). *From interactions to transactions : designing the trust experience for business-to-consumer electronic commerce*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Engineering and Innovation Sciences]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR572640>

DOI:

[10.6100/IR572640](https://doi.org/10.6100/IR572640)

Document status and date:

Published: 01/01/2003

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

FROM INTERACTIONS TO TRANSACTIONS:

**Designing the Trust Experience for
Business-to-Consumer Electronic Commerce**

Florian N. Egger

The work described in this thesis has been carried out under the auspices of the J.F. Schouten School for User-System Interaction Research.

© 2003 Florian N. Egger – Eindhoven – The Netherlands

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Egger, Florian N.

From interactions to transactions: designing the trust experience for business-to-consumer electronic commerce / by Florian N. Egger. – Eindhoven: Technische Universiteit Eindhoven, 2003. – Proefschrift. –

ISBN 90-386-1778-X

NUR 778

Keywords: Electronic commerce / Trust / Usability / Human-computer interaction

Trefwoorden: Electronic commerce / Vertrouwen / Gebruikersvriendelijkheid / Mens-machine interactie

Cover: Jan-Willem Luiten, JWL Producties, Eindhoven

Printing: Universiteitsdrukkerij Technische Universiteit Eindhoven

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the author.

FROM INTERACTIONS TO TRANSACTIONS:

Designing the Trust Experience for Business-to-Consumer Electronic Commerce

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr. R.A. van Santen, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op donderdag 11 december 2003 om 16.00 uur.

door

Florian Nicolas Egger

geboren te Lausanne, Zwitserland

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. D.G. Bouwhuis
en
prof.dr. J.B. Long

Copromotor:

dr. P. Markopoulos

ACKNOWLEDGEMENTS

Many people have contributed, be it directly or indirectly, to the success of this research. First of all, I would like to express my gratitude to John Long, without whom I would probably never have crossed the Channel to join the Eindhoven University of Technology!

This project was partly funded by SOBU, the Cooperation Unit Brabant Universities in the strategic research programme *Enabling Electronic Commerce*. Thanks to my SOBU colleagues at the University of Tilburg for their multidisciplinary insights.

Special acknowledgements are due to Boyd de Groot for his early collaboration and interest, to Susan Farrell for asking me to review the NN/g trust report and to Roland Schegg for his collaboration on the hotel study.

I would also like to thank the students I worked with: Agnieszka Matysiak, Bhiru Shelat, Lei Huang, Natalia Kirillova and Raghavi Koti. I certainly learnt as much as they did and I wish them the best of luck in their future careers.

Finally, special thanks go to my family and friends for supporting, encouraging and, most importantly, entertaining me!

Florian N. Egger

CONTENTS

Tables & Figures.....	xi
Tables.....	xi
Figures.....	xii
CHAPTER 1: Introduction	1
1.1 Introduction.....	2
1.2 From EDI to B2C.....	2
1.2.1 EDI-Based Business-to-Business E-Commerce	2
1.2.2 Internet-based Business-to-Business E-Commerce	2
1.2.3 Internet-Based Business-to-Consumer E-Commerce	3
1.3 The Trust Challenge.....	3
1.3.1 Security.....	4
1.3.2 Privacy.....	4
1.3.3 Unfamiliar Services	6
1.3.4 Lack of Direct Interaction.....	6
1.3.5 Credibility of Information	7
1.3.6 Conclusions	7
1.4 Approach & Objectives	7
1.4.1 Human-Computer Interaction.....	7
1.4.2 Objectives.....	8
1.5 Methodological Framework.....	8
1.5.1 Introduction	8
1.5.2 Discipline Model of HCI.....	9
1.5.3 HCI General Design Problem.....	10
1.6 Research Approach.....	12
CHAPTER 2: Trust.....	13
2.1 Introduction.....	14
2.2 Semantics of Trust	14
2.2.1 Morality of Trust	14
2.2.2 Defining Trust	15
2.3 Risk.....	15
2.3.1 Risk Perception.....	16
2.3.2 Heuristics & Biases	16
2.4 Trust in Personal Relationships	17

2.4.1 Predictability	17
2.4.2 Dependability	18
2.4.3 Faith.....	18
2.4.4 Reputation	19
2.4.5 Cooperation	20
2.4.6 Familiarity	20
2.5 Trust in Business Relationships	21
2.5.1 Calculative Process.....	21
2.5.2 Prediction Process	21
2.5.3 Capability Process	21
2.5.4 Intentionality Process	21
2.5.5 Transference Process	22
2.6 HCI & Trust in E-Commerce.....	22
2.6.1 Trust in Machines.....	22
2.6.2 Trust in E-Commerce: Academic Research.....	23
2.6.3 Trust in E-Commerce: Industry Reports.....	27
2.7 Conclusions.....	28
CHAPTER 3: MoTEC: A Model of Trust in E-Commerce	29
3.1 Introduction.....	30
3.2 Initial Model Development.....	30
3.3 Revised Model.....	34
3.4 Pre-interactional Filters.....	36
3.4.1 User Psychology.....	36
3.4.2 Pre-Purchase Knowledge.....	37
3.5 Interface Properties.....	39
3.5.1 Branding.....	40
3.5.2 Usability	40
3.6 Informational Content.....	41
3.6.1 Competence	41
3.6.1.1 Company.....	41
3.6.1.2 Products & Services.....	42
3.6.2 Risk	42
3.6.2.1 Security	43
3.6.2.2 Privacy	44
3.7. Relationship Management	44
3.7.1 Pre-Purchase.....	44
3.7.2 Post-Purchase	45
3.8 Conclusions.....	45
CHAPTER 4: Trust Toolbox.....	47
4.1 Introduction.....	48
4.2 GuideTEC: Trust Design Guidelines	48
4.2.1 Process-Oriented Trust Design Guidelines	49
4.2.1.1 Pre-interactional Filters.....	49
4.2.1.2 Interface Properties	50
4.2.1.3 Informational Content.....	50
4.2.1.4 Relationship Management	51
4.2.2 Product-Oriented Trust Design Guidelines.....	51
4.2.2.1 Interface Properties	51
4.2.2.2 Informational Content.....	52
4.2.2.3 Relationship Management	54
4.2.3 Conclusions	55
4.3 CheckTEC: A Checklist for Expert Evaluations.....	55
4.4 QuoTEC: A Questionnaire for Trust in E-Commerce	58

4.4.1 Objectives.....	58
4.4.2 Questionnaire Development	58
4.5 Conclusions.....	59
CHAPTER 5: QuoTEC Applications	61
5.1 Introduction.....	62
5.2 Trust in Online Services	62
5.2.1 Experimental Design	62
5.2.2 Websites	63
5.2.3 Scenarios	63
5.2.4 Procedure.....	63
5.2.5 Participants.....	64
5.3 Trust in Online Retail	65
5.3.1 Experimental Design	65
5.3.2 Websites	65
5.3.3 Scenarios	66
5.3.4 Procedure.....	66
5.3.5 Participants.....	67
5.4 Analysing the Main Constituents of Trust	67
5.4.1 Combining Data	67
5.4.2 Reliability	67
5.4.3 Factor Analysis.....	68
5.4.3.1 Hotel Websites.....	68
5.4.3.2 Retail Websites	69
5.4.3.3 Combined Data	71
5.4.4 Final QuoTEC Questionnaire	74
5.4.5 Trust Performance Visualisation	75
5.4.6 Regression Analysis	77
5.4.6.1 Hotel Websites.....	78
5.4.6.2 Retail Websites	78
5.4.6.3 Combined Data	79
5.5 Conclusions.....	81
CHAPTER 6: Toolbox Validation	83
6.1 Introduction.....	84
6.2 Hypotheses & Approach	84
6.2.1 Hypothesis 1	84
6.2.2 Hypothesis 2	85
6.2.3 Hypothesis 3	85
6.2.4 Approach	86
6.3 Trust Problems Predicted by Expert Evaluators	87
6.3.1 Experimental Design	87
6.3.2 Participants	87
6.3.3 Websites	88
6.4 Study 1: Unguided Expert Evaluations	91
6.4.1 Procedure.....	91
6.4.2 Data Analysis	91
6.4.3 Results of Study 1 for the Flower Website	92
6.4.5 Results of Study 1 for the Perfume Website	93
6.4.6 Study 1: Conclusions across Websites.....	95
6.5 Study 2: Checklist-Guided Expert Evaluations.....	95
6.5.1 Procedure.....	95
6.5.2 Data Analysis	96
6.5.3 Results of Study 2 for the Flower Website	97
6.5.4 Results of Study 2 for the Perfume Website	99
6.6 Comparing the Results from Studies 1 and 2.....	101
6.6.1 Number of Problems Found.....	101
6.6.2 Problem Distribution	102

6.6.3 Time & Satisfaction.....	103
6.6.4 Study 1: Methods Feedback	105
6.6.5 Study 2: CheckTEC Feedback.....	105
6.7 Studies 3 & 4: Trust Problems Reported by Users	106
6.7.1 Experimental Design	106
6.7.2 Participants	107
6.8 Study 3: User Tests.....	107
6.8.1 Procedure.....	107
6.8.2 Data Analysis for Study 3a.....	108
6.8.3 Results of Study 3a for the Flower Website	108
6.8.4 Results of Study 3a for the Perfume Website	109
6.9 Comparing Unguided, Checklist & User Tests Results	110
6.9.1 Comparisons for the Flower Website	110
6.9.2 Comparisons for the Perfume Website	112
6.9.3 Comparisons across Studies	112
6.10 Testing Hypothesis 3: Questionnaire Studies	114
6.11 Study 3b: Questionnaire after User Tests	114
6.11.1 Results of Study 3b for the Flower Website	114
6.11.2 Results of Study 3b for the Perfume Website.....	115
6.12 Study 4: Questionnaire without Facilitator	116
6.12.1 Procedure.....	116
6.12.2 Data Analysis	116
6.12.3 Results of Study 4 for the Flower Website	117
6.12.4 Results of Study 4 for the Perfume Website.....	118
6.13 Comparisons of Questionnaire Results	119
6.14 Conclusions.....	120
CHAPTER 7: Discussion	121
7.1 Introduction.....	122
7.2 Recapitulation	122
7.3 Contribution	123
7.3.1 The MoTEC Model	124
7.3.2 The Trust Toolbox.....	124
7.3.2.1 GuideTEC Guidelines.....	124
7.3.2.2 CheckTEC Checklist.....	125
7.3.2.3 QuoTEC Questionnaire.....	126
7.4 Implications for HCI Research	126
7.5 Implications for HCI Practice	127
7.6 Limitations.....	128
7.6.1 Scope	128
7.6.2 Internal Validity	128
7.6.3 External Validity	129
7.6.4 Construct Validity	130
7.6.5 Ecological Validity.....	130
7.7 Future Research	130
7.9 Ethical Considerations.....	132
7.9 Conclusions	133
BIBLIOGRAPHY.....	135
APPENDIX.....	143
APPENDIX 1: Abstracts of Papers Produced in this Research.....	143
APPENDIX 2: Background of the evaluators in Studies 1 and 2.....	146

Legend.....	146
APPENDIX 3A: Raw data of the user tests for the Flower Website	147
APPENDIX 3B: Raw data of the user tests for the Perfume Website	149
APPENDIX 4: Comparative Results for the Unguided, Checklist and User Tests Conditions for the Flower and the Perfume Websites	151
Legend.....	151
Units.....	151
SUMMARY	153
SAMENVATTING.....	155
CURRICULUM VITÆ.....	157

Tables & Figures

Tables

Table 1 – Rationale for initial model component selection (MSc project).....	33
Table 2 – Pre-interactional Filters: Components & sub-components.....	36
Table 3 – Interface Properties: Components & sub-components	40
Table 4 – Informational Content: Components & sub-components.....	41
Table 5 – Interface Properties: Components & sub-components	44
Table 6 – Checklist for expert evaluations	56
Table 7 – QuoTEC questionnaire (analytical).....	59
Table 8 – Screenshots of the six hotel websites by predicted trust category.....	64
Table 9 – Screenshots of the four retail websites by predicted trust category.....	66
Table 10 – Main attributes of the questionnaire studies.....	67
Table 11 – Initial Cronbach Alphas for the QuoTEC components	68
Table 12 – Hotel Websites: factor loadings.....	70
Table 13 – Retail Websites: factor loadings.....	70
Table 14 – Combined Websites: factor loadings (initial).....	70
Table 15 – Combined Websites: factor loadings (3 rd iteration).....	71
Table 16 – Final Cronbach Alphas for the QuoTEC components.....	74
Table 17 – Final QuoTEC questionnaire.....	75
Table 18 – Multiple linear regression for the hotel data.....	78
Table 19 – Multiple linear regression for the retail data	79
Table 20 – Multiple linear regression analysis on the combined data.....	80
Table 21 – Study 1-Flowers: Number of problems found per participant.....	92
Table 22 – Study 1-Flowers: Most frequently reported problems with their severity ratings.	93
Table 23 – Study 1-Perfumes: Number of problems found per participant	94
Table 24 – Study 1-Perfumes: Most frequently reported problems with their severity ratings	94
Table 25 – Study 2-Flowers: Number of problems found per participant.....	97
Table 26 – Study 2-Flowers: Concluding comments by checklist users	98
Table 27 – Study 2-Flowers: Number of problems found per participant.....	99
Table 28 – Number of problems found by expert evaluations in Studies 1 and 2.....	101
Table 29 – Proportion (%) of problems per MoTEC component.....	102
Table 30 – Number of & overlap between predicted and observed problems	113
Table 31 – Correlation coefficients for unguided, checklist and user tests	113

Figures

Figure 1 – Discipline Model of HCI: General.....	9
Figure 2 – Discipline Model of HCI: Present Research	10
Figure 3 – Thesis Structure.....	12
Figure 4 – The four dimensions of MoTEC	35
Figure 5 – QuoTEC items in the 2-factor space	73
Figure 6 – Trust performance visualisation as a function of efficient access to information and perceived risk.....	76
Figure 7 – Toolbox validation approach.....	86
Figure 8 – Homepage of the flower website.....	89
Figure 9 – Homepage of the perfume website.....	90
Figure 10 – Screenshot of the online version of CheckTEC	96
Figure 11 – Study 2-Flowers: Components and trust performances	99
Figure 12 – Study 2-Perfumes: Components and trust performances	100
Figure 13 – Reported subjective performance in unguided and checklist studies.....	104
Figure 14 – Number of problems found by the different methods (Flowers).....	110
Figure 15 – Number of problems found by the different methods (Perfumes)	112
Figure 16 – Study 3b-Flowers: Components and trust performances	115
Figure 17 – Study 3b-Perfumes: Components and trust performances	115
Figure 18 – Study 4-Flowers: Components and trust performances	117
Figure 19 – Component performances based on questionnaires only (Perfumes)	118

CHAPTER 1

Introduction

Electronic Data Interchange (EDI) is a protocol that enables businesses to exchange information and transact via proprietary networks. This early form of Business-to-Business (B2B) electronic commerce (e-commerce) was quickly supplanted by Internet-based e-commerce, as it is much cheaper and more flexible. With increasingly more private users on the Internet, Business-to-Consumer (B2C) e-commerce flourished in the late 90s, giving them access to products and services from all over the world. However, adoption and usage of e-commerce websites were found to be particularly affected by trust concerns. Lack of trust is mostly due to security and privacy concerns, unfamiliar online services, lack of direct interaction with products and people, as well as the poor credibility of online information. In order to address the problem of trust in B2C e-commerce, we have adopted an approach that stems from the discipline of Human-Computer Interaction (HCI) that focuses on the total user experience rather than on operational efficiency alone. The objectives of this research are: (1) To build up substantive knowledge about what makes customers trust e-commerce websites and (2) To build up and validate methodological knowledge to help practitioners design and evaluate trust-shaping factors in e-commerce websites.

1.1 Introduction

Electronic commerce (e-commerce) as we know it today has evolved through various stages of technological development. To understand human-computer interactions in an e-commerce situation, we must first look at the context in which business-to-consumer (B2C) e-commerce emerged. E-commerce has developed from the tradition known as electronic data interchange (EDI) which takes place between known business partners, using expensive, proprietary networks. With the advent of the Internet, business-to-business (B2B) e-commerce flourished, thanks to a much cheaper way of connecting companies through a single and open network. As the number of private Internet users increased, so did the number of B2C e-commerce websites.

1.2 From EDI to B2C

1.2.1 EDI-Based Business-to-Business E-Commerce

Historically, companies eager to re-engineer their business processes have integrated their systems with those of their suppliers and distributors, using proprietary networks for electronic data interchange. Since these interactive systems are used to exchange not only informational, but also commercial data, they represent the first form of business-to-business electronic commerce. It is noteworthy that all parties engaged in a commercial relationship in EDI are bound to know and trust each other, since they have been individually connected using a private value-added network (VAN).

One implication of such systems is that users consider EDI-based computer applications as just another tool by means of which they can carry out a task. That is, although the human-computer interactions can involve transactions, users do not have to worry about the legitimacy of their business partner. Besides, should users wish to switch, say, from one supplier to another, they would not be able to unless a new, costly, EDI network connection is created. From a Human-Computer Interaction (HCI) perspective, designing commercial systems of that nature do not pose any special problems, as traditional user-centred design methods can successfully be applied to ensure a high task quality.

1.2.2 Internet-based Business-to-Business E-Commerce

With the advent of the World-Wide Web, many companies wishing to engage in B2B e-commerce have switched from their proprietary EDI networks to the Internet as the mediator between business partners. The advantages of Internet-based B2B commerce are, for instance, the low initial cost of the IT infrastructure and the sheer number of potential business partners. However, since the Internet is publicly accessible, data can be more easily intercepted, which seriously undermines the security of online transactions, as well as the privacy and confidentiality of the commercial exchange. Moreover, the legitimacy and the trustworthiness of online vendors cannot be guaranteed as adequately as on a private network, because there is no control as to who will enter the system and how parties will authenticate themselves. That is why so-called trusted third parties (TTPs) play an increasingly important role by guaranteeing a

vendor's authenticity (e.g. VeriSign), its commitment to customer privacy (e.g., TRUSTe), or the security of online transactions (e.g. American Express). Since users will often have the choice between a large number of different business partners and since the cost of switching from one vendor to another is negligible, it is imperative that online vendors stand out by addressing not only users' functional business needs, but also their concerns in terms of security, confidentiality and trustworthiness.

It must be noted that the user experience with such systems is likely to vary according to the amount of freedom and responsibility they allow in selecting business partners. Indeed, if employees are told by management to do business with one particular vendor, they will tend to perceive the interaction with the vendor as a mere extension of their other computer-based activities, just like in EDI. On the other hand, if business users are left free to select which party to trade with, it will be in their and the company's interest to identify a party that can genuinely be trusted. Thus, traditional HCI methods still apply for the design of the operational effectiveness and usability aspects of these systems, but they might fail to deliver when it comes to designing trust-inducing features susceptible to turn users into customers, as such considerations have been outside their scope.

1.2.3 Internet-Based Business-to-Consumer E-Commerce

The Internet has notoriously democratised direct networked access to vendors, putting them only a few mouse-clicks away from consumers. Not only does the context of the human-computer interaction shift from work to home, the private user's mindset vis-à-vis the system and its functionality is also likely to be considerably different from that of the business user's. Indeed, it seems reasonable to assume that a user engaging in Business-to-Consumer (B2C) e-commerce does not perceive the system as a work tool (cf. B2B), but rather as a means to order some goods or services for personal use.

For private users to adopt e-commerce, it is imperative that the benefits of using the new commercial medium (e.g. convenience, decreased transaction costs) significantly outweigh potential risks. Indeed, the private user's freedom to select appropriate vendors tends to be correlated with greater concerns regarding financial risk, privacy and trust (cf. below). This can be accounted for by the fact that private users are more directly involved in the commercial exchange, since they are using their own equipment, giving sensitive information about themselves as individuals, and spending their own money. Again, existing HCI analysis and design methods are expected to be well-suited for the usability aspects of the e-commerce interface, but they are likely to be ill-adapted to address issues characteristic of the transactional dimension of commercial relationships.

1.3 The Trust Challenge

Because of the important personal risk involved in B2C transactions, consumer trust in online vendors has emerged as an important barrier to transacting online (GVU, 1998). This section gives an overview of the main psychological barriers to the adoption of e-commerce.

1.3.1 Security

Trust is often equated with security concerns in online transactions. A secure transmission implies that the two parties in a transaction have been properly authenticated and that the information exchanged via the network remains unaltered. However, there are three main ways in which confidential information can be obtained (Camp, 2000):

1. Information can be copied during transmission: By eavesdropping, i.e. monitoring a communication, it is possible to get access to sensitive information like passwords (password sniffing). Replay attacks can then be carried out where this information is duplicated. For example, this would enable merchants to accept payment twice by duplicating the purchase authorisation process. This problem has been addressed by increasingly sophisticated encryption technologies. However, even encrypted transmissions can be analysed and the encryption algorithms sometimes broken (cryptanalysis).

2. Information can be accessed during storage: Although security experts urge online businesses not to store confidential information of their customers in systems directly linked to the Internet, such cases are still relatively frequent. Hackers can take advantage of this carelessness by sending a large number of popular words to a system and see if they match any passwords. This example shows that even the most advanced security solutions can fail if they are not properly implemented.

3. Information can be obtained from an authorised party: Whom can you trust? Apparently, 95% of all security incidents are caused by insider attacks (Bernstein et al., 1996). This means that secure systems that have been properly set up are still at risk from people who have legitimate access to the system.

While objective security risks exist, they do not necessarily correlate with the risks perceived by consumers. For instance, the media have been focusing on cases of credit card fraud on the course of payments over the Internet, thus giving a distorted image of the frequency of fraud. Traditional use of credit cards is by no means more secure, as it is extremely easy to commit fraud by copying someone's credit card details in an offline transaction. This shows that consumers may minimise risk in one case, while maximising it in another. Perceived risk and the biases associated with it will be discussed in more details in the next chapter.

1.3.2 Privacy

Privacy concerns usually follow security concerns in surveys (CommerceNet, 1997; GVU, 1998). It is true that, within a site, log files can provide extremely useful information to marketers. Starting from the referring web page, a user's complete browsing activity within a site can be recorded and analysed, i.e. which pages were visited, in which order and for how long. Data mining software helps making sense of such data by identifying patterns. Typically, browsing patterns are therefore not analysed for individual users but are presented in the form of aggregate profiles.

Another rich source of information is users' IP addresses. Reverse DNS lookup automatically uses this series of numbers to determine what network the user comes from.

Provided the network is not that of an Internet service provider (ISP), it is possible to gather information, such as the name and size of the users' organisation and geographical location. The following, imaginary, example shows how meaningful information can be derived from a string of numbers. A user visits a web site and leaves behind her IP address of the form "123.456.78.90". DNS lookup recognises this address as "smith.cs.abc.ac.uk", which, in turn, can be translated into the rather informative "*User "Smith" in the Computer Science department of the UK-based ABC University*".

Websites can also store certain pieces of information about a user and his/her preferences directly on that user's computer. Such files called "cookies" typically record users' browser type and version, operating system, language preferences and any other data the user might have supplied to the site. Since one site cannot read the cookies from another site, cookies-related risks are fairly minimal. However, privacy can be at risk when a person's information is collected across sites sharing a common database.

Such a threat to privacy happened when the online advertiser DoubleClick merged with Abacus, which tracks consumer purchases (Macavinta, 2000). They claimed that, by combining their databases, they would be able to target their banner ads more effectively. However, one implication is that this merger made it possible to combine non-personally identifiable with personally identifiable information. Of course, every time consumers shop online, they have to provide their personal details for delivery purposes. A less obvious example is when an affiliated company offers a sweepstake on its site, thereby asking for a name, physical address and e-mail. Consumers are not aware of when such combinations of data happen. The DoubleClick case is a good example of two a priori trusted entities that become untrustworthy when being combined.

To address consumers' privacy concerns, both traditional and Internet-only companies have introduced services providing audits of privacy policies, as well as privacy seals guaranteeing that a particular company operates according to its posted policy. Examples of Internet-only seals are TRUSTe or WebTrust, while offline associations include the BBBOnline Privacy Program or the Which? Web Trader seal in the UK (in existence from 1999 to early 2003). It is noteworthy that these privacy seals do not offer any legal protection, as they only assess the extent to which online businesses conform to their promises.

The real effectiveness of privacy seals is questionable. Given the number of different seals on the Internet, many of them are not recognised, let alone trusted. The meaning of such seals and their legal implications have also been found to be quite unclear in consumers' minds (Egger, 1998; Egger & De Groot, 2000). For example, an early study reported in Egger (1998) indicated that a minor privacy seal aimed at US consumers placed on a UK site was hardly noticed by the British participants and, when pointed out to them, was not at all perceived to be relevant to them, nor trustworthy. Similarly, a study by Egger and De Groot (2000) revealed that Dutch consumers would place significantly more trust in a seal of their local Consumers' Association than in a US Internet-only seal like TRUSTe. Therefore, the challenge for e-businesses is to provide a trusted service globally, while addressing consumers' con-

cerns locally.

1.3.3 Unfamiliar Services

Internet-based businesses have created hosts of new services and business models. Novelty implies unfamiliarity, and unfamiliarity is more likely to breed mistrust than trust. This can be explained by the lack of previous experience with such services and the difficulty of understanding radically new business models.

Borderline cases in this category are online auctions such as eBay. While auctions are by no means a novelty per se, the Internet has remarkably democratised them by bringing them to the masses. Related to that is a new business model based on traditional auctions: consumer-to-business (or inverted) auctions, such as Priceline.com. Such auctions are new in that a consumer can set a maximum price for a service, be it a flight or a hotel room, while businesses compete to offer the best deal.

Marketing also introduced a new paradigm called *permission marketing* (Godin, 1999). Permission marketing lets consumers specify a detailed profile of their interests, so that they only receive advertisements and special offers that are highly relevant to them. Thus, consumers permit advertisers to send them commercial emails. The benefit to consumers, apart from more relevant offers, is that they get rewarded whenever they react to an offer. Rewards can take the form of points, discounts or cash. Although Mypoints.com has been around in the USA for some time, EuroClix was one of the first such programs in the Netherlands. In a study investigating consumers' attitudes towards the EuroClix web site, it turned out that people had indeed difficulties understanding the whole concept and how they could benefit from them. Most of them tended to be negatively biased towards the idea of direct marketing and were afraid of junk mail and other non-desirable side effects (Egger & De Groot, 2000). Therefore, preconceptions about the familiar and misconceptions about the unfamiliar can be clear impediments to the development of consumer trust.

1.3.4 Lack of Direct Interaction

The fact that the commercial exchange is mediated via a computer screen can also constitute a barrier to the development of trust. First, there is the lack of direct interaction with people. That is, both salespeople and fellow shoppers can give cues about a business's trustworthiness in face-to-face interaction. On the Internet, salespeople are often replaced by a collection of "frequently asked questions" (FAQ) sections and search engines, or are only available through electronic media. One way to give web sites a more human face is to create agents or avatars to assist customers (e.g. Lumkin, 2003). The reassurance provided by fellow shoppers can only be found in online forums, which is much less direct and which introduces the additional problem of trust in online advice (Briggs et al., 2002).

The other kind of direct interaction missing online is obviously the interaction with the products themselves. The lack of experiential interactions proves a particularly tough challenge for non-standard products such as groceries, crafts or textiles. Although sophisticated haptic feedback devices might, one day, enable a pretty accurate tactile rendering of surfaces and, thus, communicate factors associated with the intrinsic quality of a non-standard product, it is unlikely that this will happen in the fore-

seeable future.

1.3.5 Credibility of Information

Since everybody can register a domain name and set up a website, it is sometimes hard to tell the difference between websites of legitimate companies and those of crooks. This has important implications in terms of liability, accountability and legal recourse once a transaction has taken place.

Sometimes, published information can purposely be wrong or misleading. A common example is the case of alleged objective product reviews that are being sponsored by the manufacturer or fake testimonials. This problem is made worse since information can be altered simply and quickly, leaving no trace of the original text.

In the opposite case, although information can be updated and published instantly, not all websites are always up-to-date. This can lead to outdated information crucial to the transaction, such as price, description or availability. Fogg's (2003) extensive work on the credibility of websites will be reviewed in the next chapter.

1.3.6 Conclusions

Consumer's trust concerns appear to be related to a number of factors, including security, privacy, unfamiliarity, distance in time and space and unreliable information. As prospective customers of a website need to have enough trust before placing an order, it is essential that online merchants address trust concerns by means of the information they provide on their website. Instead of focusing on trust in terms of technological solutions to increase security, we will adopt a global, user-centred approach that focuses on user's initial perception of an online merchant's trustworthiness. This approach is described in greater detail in the next section.

1.4 Approach & Objectives

1.4.1 Human-Computer Interaction

Historically, Human-Computer Interaction (HCI) has its roots in applied psychology and ergonomics. HCI has been defined as "the discipline concerned with the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" (ACM SIGCHI, 1992). In the context of new interactive media, HCI gives rise to several sub-disciplines, such as: usability engineering (Nielsen, 2002), interaction design (Preece et al, 2002) or information architecture (Rosenfield & Morville, 2002). These different approaches all stress users' functional needs as the starting point for designing systems that allow efficient access to information and rapid task completion.

In recent years, HCI in the context of new media has often been referenced to as "user experience" to stress that it does not exclusively focus on operational effectiveness and efficiency, but also on softer aspects of the interaction. This approach therefore borrows a lot from Marketing, as they both aim at getting an insight into user's attitudes and behaviours at several stages of the interaction (Garrett, 2002). As such, user experience strategy should be seen as a user-centric approach, the goal of which is to develop products and services that meet both functional and non-functional require-

ments, such as pleasure (Jordan, 2000), persuasion (Fogg, 2003) or fun (Blythe, Overbeeke & Monk, 2003). In this research, trust is also conceptualised as a non-functional requirement, which justifies the more global approach of user experience. Other studies, such as that by the Nielsen Norman Group (2000), also discuss trust as being a crucial part of the user experience. Of course, in the context of e-commerce, users can also be seen as customers, hence our usage of the phrases “user experience” and “customer experience” interchangeably.

The main benefit of this approach for this research is the ability to gather systematic user feedback about what specific factors, whether they are directly related to the site’s design or not, do impact adoption and usage, and, therefore, implicitly, user trust in that website. On the other hand, the main limitation of this approach is that we will only deal with subjective reports about perceived trustworthiness and that we will therefore not directly deal with objective risk and trustworthiness.

1.4.2 Objectives

The objectives of this research are twofold:

(1) To build up substantive knowledge about what makes people trust e-commerce websites.

The first objective of this research is to identify off- and online factors likely to affect consumers’ trust in an e-commerce merchant. After these factors have been identified, they will be analysed and classified into appropriate categories to form a model of trust in e-commerce. It is important, at this stage, to distinguish *trust in e-commerce* from *trust in an e-commerce website*. The first phrase refers to a person’s general trust in online transactions, although the second one refers to trust in a specific website. In this research, we shall focus on the issue of what makes people trust one website more than another, thereby assuming that our target users have the necessary general trust to transact online

(2) To build up and validate methodological knowledge to design and evaluate trust-shaping factors in e-commerce websites.

The second objective is to use the model of trust from the first phase to develop tools to help HCI practitioners evaluate a site’s perceived trustworthiness, as well as design a site in which trust-shaping factors are maximised. Concretely, this will take the form of trust design guidelines (prescription), as well as two tools for diagnosis: a checklist for expert evaluations and a questionnaire to measure consumer trust in a specific website.

1.5 Methodological Framework

1.5.1 Introduction

This section introduces the methodological framework with reference to which this research has been carried out. The first section starts by demonstrating why a theoretical model for the discipline of Human-Computer Interaction is necessary. This will be followed by the presentation of Long and Dowell’s (1989) conceptions of the HCI

discipline and its general design problem. The method used to address the problem of trust in e-commerce will then be exposed and rationalised.

1.5.2 Discipline Model of HCI

The paradigm of HCI adopted here is an *engineering* conception, as put forward by Long and Dowell (1989) and Dowell and Long (1989). These authors present HCI as a super-ordinate discipline integrating Human Factors (HF) and software engineering (SE) “as (ideally) constituted of (HF and SE) engineering principles, and its practices (HF and SE practices) as (ideally) specifying then implementing designs”. The particular scope of HCI as an engineering discipline is designing “users interacting with computers to perform effective work”, whereas SE focuses on the design of computers interacting with humans.

In order for Human-Computer Interaction research to be effective in terms of its relevance to the specific design problem, it is necessary to specify its interdependence with other epistemological stages of the HCI discipline. Long and Dowell (1989) propose a formal framework specifying the relations between the three main entities constituting the HCI discipline: *Research*, *Knowledge* and the *General Design Problem*. The interrelations between these entities are illustrated by their view that the HCI discipline should consist of:

“The use of knowledge to support practices, seeking solutions to a general problem having a particular scope” (Long & Dowell, 1989)

Figure 1 illustrates this view of the HCI discipline by showing the logic and the hierarchy of the different entities. HCI research produces HCI knowledge through research practices. Knowledge is thus acquired by research, which, in turn, seeks to validate it. The link between HCI knowledge and the general design problem with its particular scope is made by design practices. These practices use HCI knowledge to diagnose the general problem and/or to prescribe solutions.

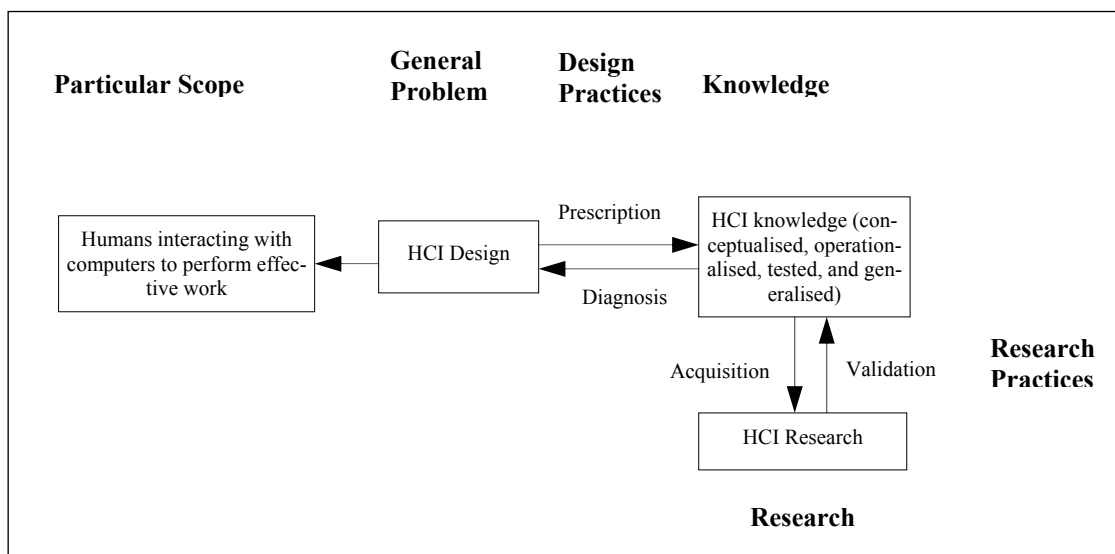


Figure 1 – Discipline Model of HCI: General

We will use the discipline model as a means to expose the methodological approach chosen to tackle the research problem at hand. Figure 2 illustrates the relations between the different processes that will be encountered in this research.

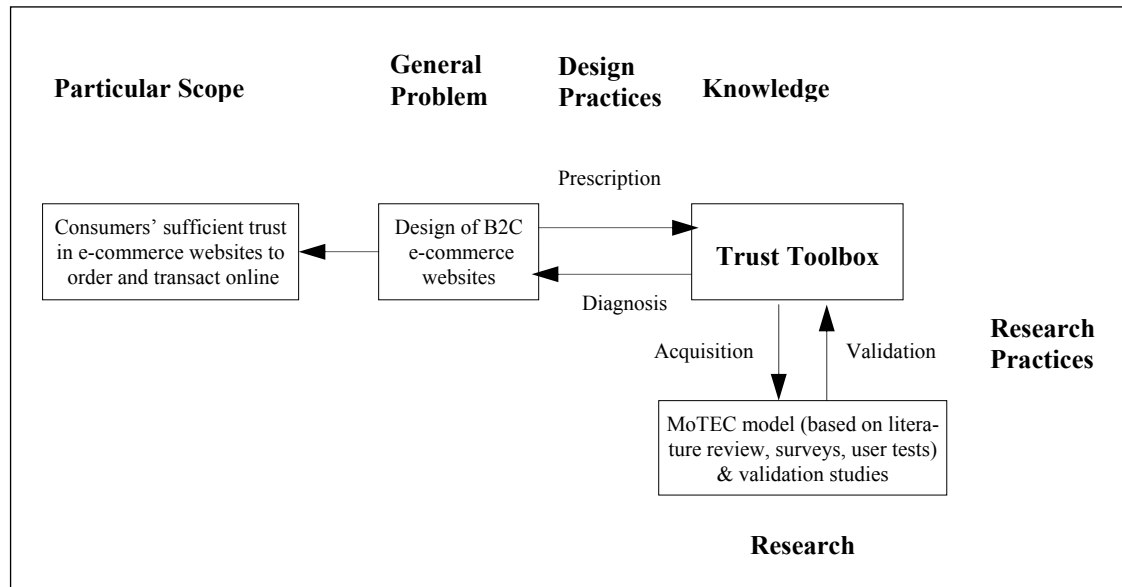


Figure 2 – Discipline Model of HCI: Present Research

The specific scope can thus be redefined as “consumers’ sufficient trust in e-commerce websites to order and transact online”. The general problem is the HCI design of B2C e-commerce systems. The design and the evaluation of such systems are supported by knowledge that can either be scientific (and experimental) or non-scientific (and experiential). Given the softness of the specific scope (viz., trust), it will be argued later that the most appropriate type of knowledge is a suite of validated tools to be used by practitioners. These elements will be acquired through research practices involving the creation of a model of trust in e-commerce, based on surveys, questionnaire studies and user tests. To test the validity of the tools, this knowledge will be applied to concrete design and evaluation cases – an activity which the discipline model calls “design practices”. Long and Dowell (1989) also indicate that only research findings that have been validated can constitute general knowledge, i.e. tools that are coherent, complete and fit-for-purpose.

1.5.3 HCI General Design Problem

Dowell and Long (1989) also offer a more complete definition of the general design problem. The model they propose reflects the characteristic of the engineering conception of HCI, namely the specification of *work*. Central to their conception is the notion of domain of application, “where work originates, is performed, and has its consequences”. The domain contains a certain number of objects characterised by attributes which can have different values. In this context, work should be understood as a change occurring in an object or an entity from one state into another state. In other words, work affects objects of a domain by changing the value of their attributes.

Work is performed by a work system constituted of humans, with their structures and behaviours, interacting with computers, with their structures and behaviours. When interacting with each other, each subsystem incurs costs that can affect both its structures and behaviours. The effectiveness of this transformation of objects by the work system is expressed by its performance; more specifically, in terms of a desired performance, either absolute or relative to an actual performance. The application of the general design problem model forces one to consider the problem at hand at a higher level and retain only its quintessential ontological substance. According to this model, the work system in this study is thus constituted of consumers interacting with networked computers. The specific scope is to increase consumers' trust in online vendors. It is therefore essential to conceptualise the very notion of trust in terms of the HCI general design problem. Where does trust fit in?

The multifaceted nature of trust makes it only possible, at this stage, to speculate about its actual place in the domain model presented above. An analysis will now be attempted to shed some light on the nature of trust. It is suggested that the best way to tackle this problem is to consider the user and computer sub-systems in turn, which will be followed by an analysis of their interaction. Besides, one should note that the following investigation will be made with reference to a threefold distinction explained below:

“The user is conceptualized as having cognitive, conative and affective aspects. The cognitive aspects of the user are those of knowing, reasoning and remembering, etc; the conative aspects are those of acting, trying and persevering, etc; and the affective aspects are those of being patient, caring, and assured, etc. Both mental and overt human behaviours are conceptualized as having these three aspects” (Dowell & Long, 1989).

User Structures

Trust formation may be influenced by cognitive user structures, such as information or knowledge which consumers have about a particular vendor. Conative user structures might similarly play a part, since they refer to the *impetus* users have to implement certain behaviours, as well as to actual actions; one could imagine that some consumers are generally risk-averse or mistrusting, which would explain their reluctance to try something new. Finally, one could likewise envisage that affective user structures such as personality or temperament (e.g. introversion-extraversion, curiosity) have a direct influence on one's proclivity to trust another party.

User Behaviours

Cognitive user behaviours can refer, in this case, to the assessment of another party's trustworthiness; this might include general knowing and reasoning, as well as a mental risk assessment. A large amount of information processing, coupled with complex mental computations might constitute unmanageable costs, which would explain the safer behaviour of not trusting the other party. Conative aspects affecting trust could be a non-willingness to try and see if the other party is trustworthy. Affective user behaviours refer to mistrust caused by the interaction with the e-commerce, e.g. confusion, apprehension, etc.

Computer Structures

The lack of trust exhibited by consumers could be explained by their general dislike of computer-mediated commerce compared to traditional or other types of home-based shopping. An alternative explanation could be that accessibility, availability and other technological restrictions are to blame for the poor level of interactivity between buyers and sellers, which can have detrimental implications for the formation of trust.

Computer Behaviours

Consumers' trust will undoubtedly be influenced by their interaction with computer behaviours. Computer behaviours refer mainly to the interface and interaction design of the e-commerce system. A major assumption in this study is that a user-centred design approach can increase the experienced trustworthiness of e-commerce systems.

These different conceptualisations of trust illustrate its fuzzy nature and indicate that it bears specific relationships with the structures and behaviours of both consumers and computers. This research will indicate to what extent the different constituents of the work system can be manipulated to increase consumers' trust in e-commerce.

1.6 Research Approach

Chapter 2 will present a survey of literature to analytically deduce potential factors for trust formation. This review will focus on the role of trust in interpersonal relationships, in business relationships and in people's interactions with technology. A formal theory is proposed in Chapter 3, which integrates these findings in a model of trust in e-commerce (MoTEC). In Chapter 4, the knowledge contained in the model is operationalised into design knowledge in the form of a Trust Toolbox for design and evaluation, containing design guidelines (GuideTEC), a checklist for expert evaluations (CheckTEC), as well as a questionnaire (QuoTEC). Chapter 5 will show how the QuoTEC questionnaire has been validated by a factor analysis conducted on 320 sets of data. Chapter 6 will present an extensive validation study of the different tools derived from the model and show the concrete benefit they provide to HCI practitioners. Chapter 7 will contain a discussion, reflecting on the research's results and implications both for HCI research and practice. The structure of the thesis, resulting from this approach, is illustrated in Figure 3.

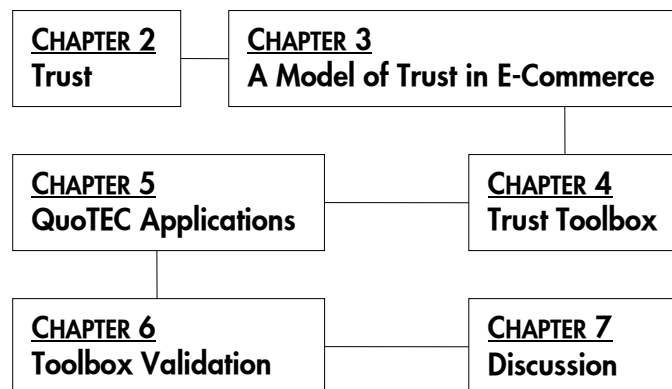


Figure 3 – Thesis Structure

CHAPTER 2

Trust

Before focusing on trust issues characteristic of the B2C e-commerce situation, this chapter reviews literature from Psychology, Marketing and Human-Computer Interaction. Since without risk, there would be no need to trust, the next section describes how people perceive risk, what heuristics they use in assessing risk and what biases they can be subjected to. Several models of interpersonal trust are introduced. Rempel et al.'s (1985) model of trust in romantic relationships focuses on Predictability, Dependability and Faith. In our analysis, these components were complemented by Reputation, Cooperation and Familiarity. Doney et al.'s (1997) more Marketing-oriented model of trust introduces five distinct processes whereby buyers assess the trustworthiness of sellers: Calculative, Prediction, Capability, Intentionality and Transference. After reviewing studies of trust in automation, the last sections present academic and commercial studies of trust in e-commerce. Although these studies all mention different factors that may influence consumer trust in an e-commerce website, none has attempted to relate these factors to existing models of trust. That is why the model of trust we will propose in Chapter 3 will explicitly refer to the concepts introduced in this chapter. This would ensure a top-down development that will be tested bottom-up in subsequent chapters.

2.1 Introduction

We all have an inherent need to interact with other individuals and institutions. These interactions can take the form of short- to long-term relationships in which exchanges take place. However, given the other party's independence, we can never fully understand its actions, let alone control it. It is the social need for mutually beneficial relationships, coupled with the unpredictability of other parties, which creates the need for trust. Luhmann (1988) describes trust as a mental strategy that reduces the complexity of our environment and that allows us to take decisions even though their outcomes may potentially be harmful. Ganster (1988) refers to trust as an illusion of control over our environment.

Trust has been studied extensively in a number of disciplines. For instance, personality psychology focuses on trust as an individual characteristic, while social psychology focuses on the dynamics of trust between individuals. Economics and marketing look at the trust in the context of commercial exchanges and transactions. Despite the multidimensional character of trust, the different conceptions all share common elements.

Essentially, the trustor-trustee relationship is characterised by dependency, under conditions of uncertainty and risk (Luhmann, 1988; Curral & Judge, 1995). Some have also distinguished between trust as the willingness to accept risk and trusting behaviour as the actual assumption of risk (Meyer, Davies & Shoorman, 1995). The decision to trust or not can be affected both by cognitive and emotional elements (McAllister, 1995). The cognitive element refers to a rational assessment of risk, the other party's reliability and competence, and is therefore more task-oriented. On the other hand, the emotional element refers to attraction, in the short term, and loyalty, in the long term. Its orientation is therefore more inter-personal. This view implies that trust can refer to several objects – in this case, the task (the transaction) and the trustee (the merchant). Although trust develops over time by ongoing interactions, the focus here is on initial trust, implying limited interactions and no earlier transaction with that merchant.

2.2 Semantics of Trust

2.2.1 Morality of Trust

Baier (1992) distinguishes between trust and reliance. She argues that reliance has more to do with technical competence that is hoped to benefit the trustor. For example, if a person says "I trust the plumber will fix this quickly", what this person really means is that she can rely on the plumber's skills and experience. However, trusting others does not only mean that they are dependable, but also that they have goodwill. Our decision to trust is based on our belief that we can predict other people's behaviour based on our perception of their motives, intentions and past behaviour. Baier (1992) redefines trust as "accepted vulnerability to another's possible but not expected ill-will" (p. 15). However, according to Glaser (1994), this distinction breaks down

when talking about trust in strangers. Most of the trust that keeps together social and economic agents in society is trust in strangers (Fukuyama, 1992). In the business world, a basic form of trust in competitors is also necessary. The argument goes that someone who has never met me and who knows nothing about me cannot be said to have goodwill towards me. Are these cases of reliance as Baier (1992) would argue? I would be betrayed if a stranger gave me wrong directions on purpose. To address this problem, Glaser (1994) distinguishes between partial goodwill and impartial goodwill. Goodwill towards a person qua person is partial goodwill (friendships, love). On the other hand, impartial goodwill refers to basic respect towards people we do not know as individuals. In the context of e-commerce, consumers trust online businesses' impartial goodwill in providing quality services.

2.2.2 Defining Trust

It is useful at this point to make explicit what is meant by trust. The uncertainty characteristic of trust relationships can be found in Deutsch's (1960) earlier account that the need for trust arises under the following contextual parameters:

- There is an unambiguous course of action in the future;
- The outcome depends on the behaviour of another party;
- The strength of the harmful event is greater than the beneficial event.

In other words, trust should be seen as "a generalized expectancy that the word, promise, oral or written statement of another individual or group can be relied on" (Rotter, 1980). Koller (1988) expands on the expectancy view of trust, as he defines it as:

"A person's expectation that an interaction partner is able and willing to behave promotively towards the person, even when the interaction partner is free to choose among alternative behaviours that could lead to negative consequences for the person. The degree of trust can be said to be higher the stronger the individual holds this expectation" (Koller, 1988).

Arion et al. (1984) introduce the concept of the Faith-Trust-Confidence continuum. They argue that all three notions are types of beliefs. The difference between them is the amount of concrete evidence available to the decision-maker. Thus, for situations where there is no available evidence at all, people's belief can be seen as an act of faith. In the case of incomplete evidence, people can use the available information to rationalise their belief – an act of trust. Lastly, if there is a great deal of data or evidence to back up a decision, that belief would be confidence.

Trust, therefore, can be seen as a mental mechanism that helps reduce complexity and uncertainty in order to foster the development or the maintenance of relationships even under risky conditions (Luhmann, 1988).

2.3 Risk

As Luhmann pointed out (1988), there is a close connection between trust and risk. If there is no risk, there is no need to trust. Indeed, the absence of risk implies confidence, i.e. certainty in positive outcomes. On the other hand, risk, implying unpredict-

able future events, requires trust to overcome uncertainty and enable constructive interpersonal relations. In other words, trust should be seen as confidence in the face of risk (Lewicki, McAllister & Bies, 1997). This section discusses two research avenues related to risk: risk perception, as well as heuristics and biases.

2.3.1 Risk Perception

Decision theory attempts to formalise alternative choices and their consequences, so as to help people take an optimal decision in the face of risk and uncertainty. Ranking of alternatives can be produced, based on criteria set by agents in view of their preferences and objectives. Thus, the probability of alternatives and criteria set by an individual can be used to calculate a function showing the expected utility for each decision. The utility of a risk is computed as the loss (or gain) multiplied by the probability of that event. The theory states that the most rational decision is the one offering the highest expected utility (Coombs, Dawes & Tversky, 1970).

However, people do not always take the most rational decisions. For instance, it has been found that people are much more afraid of losing than they are happy of winning, as subjective utilities are not linear functions of value and probability. Indeed, the negative side of the utility curve is much steeper than the positive side, which shows the importance of trust as a mechanism that mitigates risk and losses.

2.3.2 Heuristics & Biases

Many studies have shown that people use various heuristics to cope with decisions under uncertainty. In situations where people have to decide how frequently a certain event occurs, they often rely on *availability*, i.e. they try to remember situations that are representative of this special event. This means that their judgement is largely based on the retrievability of such instances. That heuristic is a very useful one because availability will be a representative of the size of a certain class: large classes will have a better availability than small classes. However, the availability not only depends on frequency, but also on how easily instances of that class can be retrieved (Kahneman & Tversky, 1973).

In general, people underestimate small probabilities and overestimate large probabilities. People tend to be overconfident in general knowledge items that are moderately to very difficult, such as the answer to the question “how many kings or queens are there in Europe?” (Fischhoff, Slovic & Lichtenstein, 1977). Experts are, in general, much better calibrated. Thus, there can be large discrepancies between what lay people and experts perceive as a risk.

Another bias that can affect judgement under certainty is the *base rate fallacy* (Tversky & Kahneman, 1980). If it is known a priori that an object has a specific probability of belonging to a certain class, this information should be used in decision making. Unfortunately, when people are given a problem containing irrelevant information, they do not use the a priori probabilities in their estimation, but base their estimation on the more salient, yet statistically irrelevant, information. For instance, if users were told that ten percent of all Internet sites were owned by crooks, and they were asked to rate a certain site, it may well be that they would only use the information from that

site and not their knowledge about the facts given to them . People's judgements of the frequency of an event are also affected by *anchor points* (Kahneman et al., 1982). When people start from an initial value, they tend to stay close to this value. If they are told that a certain event is rare, they will arrive at another estimate than if they are told that this particular event is rather common. With complex events, this is even more complicated. People tend to overestimate conjunctive events and underestimate disjunctive events (*ibid.*). Together with the effects already mentioned, this can have serious consequences for the outcome. Highly unlikely outcomes can be judged as likely events.

A dominant factor in bias effects has been labelled *representativeness* (Tversky & Kahneman, 1974). In an e-commerce situation, some people may believe that a reliable business probably has a well-designed website. From this, people infer that if a website is well designed, then that company must be reliable. People are very vulnerable to this type of judgement error, and this type of error tends to override factors that should influence the judgement of probability.

A factor related to this is that people tend to generalise from a too small sample-size and they do not appreciate the variability of small samples. If they are successful in a transaction for a few times, their trust in that type of transaction can grow much faster than can be defended objectively.

Since decisions always have rational and emotional components, it is important to understand the fallacies consumers can be subjected to when evaluating a company's trustworthiness on the basis of its website.

2.4 Trust in Personal Relationships

Before examining trust issues in business relationships, it is important to understand the basics of interpersonal trust. That is why we will start by looking at Rempel et al.'s (1985) model of trust dynamics in romantic relationships. These authors addressed the issue of trust formation by proposing a model consisting of three components that reflect increasing levels of attributional abstraction. The components of the model are *Predictability*, *Dependability*; and lastly, *Faith*. Their model is discussed below, together with references to other models.

2.4.1 Predictability

Predictability should be understood as one's general expectations about another party's future behaviour, based on an inference from observed past behaviours. It thus refers to the other party's consistency of behaviour. This consistency can, in turn, be influenced by the social environment in which the relationship takes place. Accordingly, the stability of the psycho-social environment should also be taken into account when assessing another party's predictability.

Rempel et al. (1985) argue that the judgement of another party's predictability is facilitated if one party possesses specific information about the other party, such as its reinforcements and restraints. This approach to predictability entails therefore that the first party's beliefs about the other party's behavioural consistency is related to and

influenced by the amount of time both parties were involved in a relationship. It would seem, consequently, that the formation of trust is dynamic and evolutionary in that it is unlikely to happen in the early moments of a relationship. The effect of first encounters on subsequent trust formation is discussed below, in the Faith component.

2.4.2 Dependability

Rempel et al.'s (1985) model of trust states that, with time, there is a shift from assessing another party's specific behaviours (cf. Predictability) to evaluating the *qualities* and *characteristics* attributed to that party; indeed, as they put it, "trust is placed in a person, not his or her specific actions". This also refers to what Hwang and Burgers (1997) mean when they define trust as "confidence in another's goodwill". Dependability thus refers to the partner's moral integrity, encompassing factors such as benevolence, reliability, honesty and concern with providing expected rewards. The evaluation of such personal qualities will largely be influenced by experiences involving risk and personal vulnerability, as such experiences are genuine test cases for trust.

Therefore, an emphasis on experiences that involve personal risk is essential to understand the growth of feelings of security and trust; in other words, as Deutsch (1973) expressed it: "Trust involves a willingness to put oneself at risk".

The Dependability component is clearly related to Predictability, although it is more concerned with a "sub-class of behaviours that involves personal vulnerability and conflicts of interests" (Rempel et al., 1985). Like Predictability, Dependability also assumes a relationship long enough to allow for a detailed analysis of the other party's trustworthiness. When faced with novel situations within an existing relationship, or when starting a new relationship altogether, one must consider the third component: *Faith*.

2.4.3 Faith

In the context of trust, *faith* refers to a pre- or pseudo-trust which is *not* based on past experience. When there is no evidence from previous interactions whereby one party can assess another party's trustworthiness, i.e. its predictability and dependability, it can only hold *beliefs* about it. Indeed:

"Beliefs are held in the presence of equally plausible alternatives, and pertinent but inconclusive evidence is acknowledged as insufficient to either confirm or refute them" (Rempel et al., 1985).

This is why the Faith component is crucial in explaining the formation of trust in novel situations. First impressions will give rise to beliefs, which in turn can become convictions. According to Rempel et al. (1985), people expect that future events will prove their convictions to be correct. Faith is seen as an emotional security that allows individuals to go beyond shared experiences and hope they will not be harmed by entering into a new relationship. Given that the Faith component refers to the absence of previous interactions on which to base one's assessment of trustworthiness, it is surprising that Rempel et al.'s (1985) model makes no mention of another party's *reputation*. It is suggested that reputation is also a significant factor when it comes to judge

whether one should trust another party. The influence of reputation is discussed next.

2.4.4 Reputation

In order to be trusted, individuals, as well as economic agents, strive to establish a favourable reputation (Good, 1988). Indeed, the reputation of a little known party can supplement *and* influence the first impression it makes on other parties. Although reputation does not necessarily imply direct interaction, it can nevertheless be used as data on which to base one's judgement of trustworthiness (i.e. predictability and dependability). In addition:

“Not only will the perceivers of a reputation usually have access to information which the reputation holder does not control, but also the manner in which [this] information is interpreted is not straightforward” (Good, 1988).

What Good (1988) means is that information concerning a party's reputation will often be ambiguous and, thus, will be interpreted differently by different people. This implies that not everyone will attach the same significance to a particular piece of information. To understand what process underlies this phenomenon, it is necessary to consider how new information is handled and integrated with existing views.

Good (1988) explains the effect of reputation on subsequent trust formation by the cognitive *confirmation bias* discussed in the decision-making literature (Wason, 1968; Kahneman & Tversky, 1973; Kahneman, Slovic & Tversky, 1982). This cognitive bias refers to people's tendency to confirm their theories or *Weltanschauungen* rather than seek to falsify them (cf. Popper, 1968). It is therefore a bias towards the preservation of one's beliefs. Robinson (1996) also refers to this bias as cognitive consistency, due to selective attention. The reason why we generally seek to confirm our views is that being able to reduce a wealth of sometimes contradictory information to a manageable size constitutes a clear advantage, even though we may disregard some important pieces of evidence. In other words:

“Even events which might offer disconfirmation of one's views, and which are simply made available by chance, can yield to this confirmation bias by being interpreted in a way which hides or denies their potential as counter-example to a view” (Good, 1988).

Kahneman and Tversky (1973) account for the persistence of this bias by claiming that, when decisions are made under uncertainty and/or ambiguity, information is even more likely to be interpreted in line with one's preconceptions. Good (1988) further argues that this reflects the fact that computations are not cost-free, which explains why some beliefs can sometimes dominate an individual's behaviour even though they are totally inappropriate. This phenomenon is also referred to as the *set effect* or *Einstellung* (Luchins, 1942).

The implications of the confirmation bias for the formation of trust is thus evident. Whatever comes first, be it the first impression made by an unknown party or the knowledge of its reputation, will have an irrationally strong effect on a person's judgement of that party's trustworthiness. In addition, Good (1988) claims that the *Einstellung* effect is likely to influence trust formation only in one direction, given the

asymmetric nature of trust. That is, if one is faced with a clear breach of trust by a person one used to trust, then it is very likely that trust in that person will be lost. However, if a person one does *not* trust happens to behave particularly well on one occasion, then it is very unlikely that this is sufficient to trust that person in the future. As Rempel et al. (1985) put it, “it is a curious paradox that, whereas trust is slow and difficult to build up, it appears notoriously easy to break down”. Good (1988) describes another factor important for trust formation, viz., Cooperation. This seems all the more sensible, as it relates to Rempel et al.’s (1985) Dependability component.

2.4.5 Cooperation

Dependability and Cooperation both refer to a set of beneficial behaviours between two parties. In that respect, they could be interchangeable. However, Cooperation, as conceptualised by Good (1988), also implies the important influence of direct communication on trust formation. For example:

“In conditions where the long-term interests of the participants are stressed, where only small initial or additional rewards are at stake, where there is no potential for threat and great potential for communication in that the ambiguity of the situation is reduced, and where the participants are in free and easy contact, then cooperation and, one might suggest, a certain level of trust can develop” (Good, 1988).

The importance of “free and easy contact” for the development of trust can be related to the effect of *depersonalisation*, as reported in Milgram’s (1963) seminal experiment. Milgram (1963) reported that subjects did not mind giving what they thought were electric shocks to people they could not see. However, when these people were brought into the same room and thereby made visible to the subjects, the former were less willing to give electric shocks. This study shows that physical proximity and a high level of interaction constitute a psycho-social framework in which the behaviours of other parties can be observed and evaluated. Proximity thus facilitates the formation of trust, as it provides experiential data on which to base one’s judgement of trustworthiness.

2.4.6 Familiarity

Another factor that affects a person’s assessment of another party’s trustworthiness is *familiarity*. Luhmann (1988) distinguishes between *familiarity* and *trust* by claiming that the first concept refers to an unavoidable fact of life, whereas the notion of trust refers to a “solution for specific problems of risk”. However, these two concepts are thought to be related, to the extent that trust requires an element of familiarity in order to develop. Even seemingly unfamiliar situations have certain features that are familiar, which may form the basis for an embryonic trust. This implies that strategies like abstraction or the use of metaphors can make a situation trustworthy, by creating associations with familiar items. When transposed to relationships, it is clear that previous experience with similar people, parties or situations can affect the way a person judges new people, parties or situations.

2.5 Trust in Business Relationships

Trust is an important element within organisations and in business transactions, as it facilitates risk taking (Luhmann, 1988). Trust has been found to be a remarkably efficient lubricant to economic exchange that reduces complex realities far more quickly and economically than prediction, authority, or bargaining (Powell, 1990). Also, the costs of acquiring new customers is significantly higher than retaining existing customers, which explains that relationships based on trust constitute a strong competitive advantage.

Doney and Cannon (1997) identify five distinct processes whereby commercial parties evaluate another party's trustworthiness. The descriptions of these processes will be compared to Lewicky et al.'s (1997) 3-stage model of trust.

2.5.1 Calculative Process

Is the partner worth trusting? This is a cost/benefit analysis in which the costs and rewards associated with either cheating or staying in the relationship are assessed. Factors thought to influence this process are the investments already done, the partner's size and reputation, willingness to customise, information sharing, and the length of the relationship. This is similar to Lewicky et al.'s (1997) first stage which they call *calculus-based trust*, based on assuring consistent behaviour. Individuals will do what they say because they fear the consequences of non-compliance.

2.5.2 Prediction Process

Is the partner likely to stay trustworthy? Here one party tries to forecast the other party's behaviour. This is not only influenced by duration and number of interactions, but also by the sharing of experiences in order to learn more about the other party. This process corresponds to Lewicky et al.'s (1997) second stage of trust called *knowledge-based trust*. This trust is based on the other's predictability, on experience, so that the other's behaviour can be anticipated. This type of trust relies on information rather than deterrence.

2.5.3 Capability Process

Does the other have the means to stay trustworthy? This component focuses on the other party's credibility. A party estimates the other party's resources to fulfil its promises. Doney and Canon (1997) argue that a salesperson's expertise and power are significant factors influencing this process.

2.5.4 Intentionality Process

Why can the other party be trusted? In this process, the other party's motives are assessed. In order to make the assessment, one should look at the situation from the other party's perspective (Dasgupta, 1998). The establishment of trust is facilitated when two parties share common values. This is similar to Lewicky et al.'s (1997) third stage called *identification-based trust*, which is based on similarity with the other party's desires and intentions.

2.5.5 Transference Process

Do third parties trust the other party? One will trust a party more easily when it is trusted by others one trusts. It can be compared to asking references from another party. This process can be of particular importance when one has no prior experience with the other party.

2.6 HCI & Trust in E-Commerce

While trust in romantic and business relationships is exclusively interpersonal in nature, the domain of electronic commerce introduces another variable: trust in machines. In other words, this extension introduces a shift from assessing motives and intentions to assessing reliability. In e-commerce, as will be shown below, both human intentions and system reliability can affect consumer trust. The first sub-section will present an overview of the work that has been done on trust in automation. This will be followed by a review of academic research on HCI and trust issues in B2C e-commerce. The last sub-section will present industry reports on the same issues.

2.6.1 Trust in Machines

When people start interacting with a system, they learn how this system works. In this learning phase, the outcome becomes more predictable and the user starts to develop some confidence that the system will perform as predicted. Through this confidence, the user will build trust in the system, learning more about its behaviour, its reliability, and the risks involved in using it. The user develops an attribution of dependability, taken as evidence that the system can be relied on (Rempel et al., 1985). This requires extensive interaction between the user and the system so that the user can develop a reference model.

Most of the definitions of trust focus very much on predictability, reliability, and risk of a situation, not so much on the question of whether the user can actively manipulate the system and intervene in the process or not. Only very few studies cover this aspect, mostly in the field of ergonomics and human-machine interaction (Muir, 1987; Lee & Moray, 1992; Muir & Moray, 1996). There are good reasons to include the risk of failure, which can be a malfunctioning machine, a human error or intentional failure, together with the possibility of control. If it is possible to intervene in a system when the initial outcomes of the process are unfavourable, the perceived total risk will be different from when this intervention is impossible. It is not even needed that the possibility of control indeed exists: the illusion of control is enough.

There is very little literature on the relationship between control and trust in consumer to business situations. Some more literature can be found in the field of process control. Muir and Moray (1996) have shown that trust in automation was mainly based on the perception of the control unit's competence. The overall performance had hardly any effect on trust in the automation. It was also shown that there was a high correlation between the operators' trust in the system and the use of automatic control. This trust also had a significant effect on their monitoring behaviour.

Arion et al. (1994) discuss the issue of trust in machines with reference to a frame-

work called the User-Tool-Task triangle. The access point of an IT system is its user interface, as this is where information, such as feedback about internal processes, is being communicated. It follows that the user interface is the point where trust is generated. They argue that trust should be considered at two different levels: the individual level, as trusting consists in a cognitive mechanism to reduce complexity and uncertainty, as well as the social level, as the decision to trust can be affected by relational factors between agents using the system. They define trusting as "the mental action, based upon inconclusive evidence, of expecting an agent to behave promotively towards the goals the trusting person has" (p. 358). Trust is seen as a dynamic process, initially based on faith due to the lack of evidence, that seeks to reach a certain level of confidence, i.e., where there is conclusive evidence in favour of trusting behaviour. When systems are very complex and where users can never fully understand how they work, trust can help bridge this knowledge gap and thereby help users focus on attaining their goals. Thus, trust is seen as a cognitive structure that facilitates the action regulation process.

They further identify three categories of trust. The first one is trust based on external resources, e.g., a social effect such as the way other people react to the system or people claiming that the system can be trusted. Secondly, one's decision to trust a system can also be based on the objective knowledge of successes, in terms of the accuracy of the computer's recommendations under specific conditions. This would clearly show that the system is reliable and, consequently, dependable. The third type of trust is conceptualised as a combination of the first two types, together with the use of logical inferences, such as extrapolation or interpolation. In this case, trust is derived from an internal process that uses external facts. If the outcome of trusting behaviour is positive, the person may, on a subsequent occasion, skip the cognitively demanding reasoning process altogether and directly arrive at the conclusion that trust will be beneficial, transforming the third type of trust into the second type, one that occurs more automatically. The authors equate this transformation with a move from explicit to implicit knowledge.

In their consideration of computer-supported cooperative work (CSCW) systems, Arion et al. (1994) argue that trust refers both to the human-computer interaction and to the computer-mediated human-human interaction. Specifically, they claim that the development of trust between two parties should be supported as much as possible by the system. The problem with CSCW systems is that communication does not necessarily happen at the same time, nor at the same place, something that Giddens (1990) calls "disembedding mechanisms". Indeed, "there would be no need to trust anyone whose activities were continuously visible and whose thought processes were transparent, or to trust any system whose workings were wholly known and understood" (Giddens, 1990).

2.6.2 Trust in E-Commerce: Academic Research

Taking a very low-level approach to HCI factors in e-commerce, Lohse and Spiller (1998) report a study, the aim of which was to predict store traffic and dollar sales as a function of interface design features (e.g., the number of products and links in the store, search modes, image sizes etc.). These design features were grouped according to four marketing attributes identified by Lindquist (1975), viz., *Merchandise, Ser-*

vice, Promotion and Convenience. Using a stepwise regression, the authors finally retained a total of 13 predictor variables. It is noteworthy, however, that the causality of design factors could not be determined, as this regression model only revealed correlations between variables. In substance, their results suggest that traffic and sales can be positively affected by improving browsing and navigational features of commercial websites. Detailed product descriptions and representations were also found to have a great effect on sales. It might be argued that the approach adopted by Lohse and Spiller (1998) lacks scientific rigour, insofar as the selection of design factors was arbitrarily made by the authors and distributed into an *ad hoc* form of an existing Marketing classification. In addition, no distinction was made in terms of the qualitative properties of the interface features, such as information vs. graphics. The findings reported in that study stem from a purely empirical approach without any theoretical backup, which has serious implications for their validity and generalisability.

A different approach was put forward by Kim (1997), who introduces a more conceptual, higher-level approach to HCI and B2C e-commerce. He starts off by distinguishing between the *user* interface, more concerned with ease of learning and ease of use and the *customer* interface which, in addition, "should provide a pleasant shopping environment" (p. 12). Indeed, it is crucial that e-commerce interfaces should attract consumers, thereby converting them into potential customers. Indeed, If consumers do not feel attracted by an online shop's interface, they will simply switch to other online vendors. Kim (1997) then puts forward a research framework that outlines the HCI variables to take into account when designing customer interfaces. He identifies four dimensions of customer interface design:

1. Content design refers to the type and scope of information provided about products and services. This information, Kim argues, should be appropriate for users to construct an appropriate mind map to assess whether a product is worth purchasing.
2. Structure design refers to the way knowledge of the domain is organised in the electronic shop, so that it is in accordance with the customers' mental model of the domain. This is especially important in the case of product categorisation.
3. Navigation design refers to the site's architecture and to the design aspects that minimise user costs when navigating the site, such as, e.g., user support in the form of search engines.
4. Graphic design refers to the graphical representation of the site's architecture, navigational aids, use of logos, colours, layout, etc. It is thus assumed that different graphic elements can have crucial effects on the feelings of customers.

Kim and Moon (1998) conducted a study to investigate precisely which graphic design elements were most likely to communicate trustworthiness in cyber-banking interfaces. That is, they focused exclusively on the impact of visual design features on the feeling of trustworthiness, at the expense of the system's informational content. Their results indicate that a cyber-banking interface induces more trust if it contains a clipart image that is "3D, dynamic and covers half of the total screen size" and if the colours used have got a "cool tone", if the main colours are "pastel and of low brightness". This, they argue, would lend support to the hypothesis that manipulating visual

properties of a user interface can affect its experienced trustworthiness. However, Kim and Moon (1998) admit to some methodological flaws, as, for example, the passive presentation of stimuli and the homogeneity of the subjects' experience with the Internet and socio-cultural background. In addition, no comprehensive analysis of the interactions between the different design factors was carried out. This implies that, depending on which design factors are combined, the trustworthiness of the interface can either increase or decrease. Their focus on graphics alone is all the more surprising, as Kim (1997) acknowledged the importance of content design in the classification he had proposed one year earlier. The argument remains that a more holistic approach to the design of trustworthiness into interfaces would yield more valid results.

To address this problem, Tan and Thoen (1999) propose a generic model of trust for e-commerce. They introduce the notion that, for consumers to engage in a commercial relationship, their level of trust must exceed a personal threshold. This threshold is determined by personality (e.g., risk seeking vs. risk-averse) and by the potential profit or utility gained from entering the transaction. Central to their conception of trust is the idea of information asymmetry since it allows for opportunistic behaviour. In electronic forms of commerce, the level of hidden information is even more difficult to determine due to the unobservability of the other party. Their model distinguishes between two kinds of trust, Party trust and Control trust. Party trust refers to a subjective feeling of trust about the vendor. If that level of trust is not above one's threshold, then Party trust needs to be complemented by Control trust, which refers to more objective and independent control mechanisms. According to Tan and Thoen (1999), Transaction trust can only be achieved if Party trust and Control trust are positive. Interestingly for HCI designers, the authors claim that their descriptive model, which has not been validated empirically, can be applied to design. Although they do not say how prescriptive design knowledge can be derived from their model, which would guarantee a certain level of trust-related performance, the fact remains that their model is a useful starting point for further HCI developments.

Jarvenpaa et al. (1999) investigated the effect of national culture on consumer trust in e-commerce. The study was originally carried out in Australia, later replicated in Israel and Finland. The focus was on initial trust rather than on trust developing over time. Their assumption was that people from individualistic countries have a greater pre-disposition to trust. An individualist country, as defined by Hofstede (1980), is one where relationships between individuals are loose, mostly taking care of themselves and their close relatives. The argument goes that relying on other people is important in a competitive environment. Therefore, it was predicted that these people might trust impersonal e-commerce sites more easily than people from collectivist countries. A collectivist country is one where individuals, from their birth onwards, are part of a group in which individuals give protection in exchange for loyalty. For this reason, collectivists are more likely to trust a close network of people and are therefore more risk-adverse outside their trusted group (Hofstede, 1980). However, the authors report no cultural antecedents regarding the antecedents of trust. Limitations include a sampling bias and the fact that this study did not take into account factors such as the site's aesthetics, language or usability.

Jarvenpaa et al. (2000) report a study where consumers recognised differences in size and reputation among e-businesses, which influenced their judgement of trustworthi-

ness and their perception of risk. Trust was positively affected by lower perceived risk, larger company size and increased reputation. Size is taken as an indication that other businesses trust the seller and conduct business with it successfully and that the seller will deliver on its promises. Also, large size suggests that the seller has the necessary resources to provide customer and technical service, which will also help reduce the complexity of the commercial exchange and thus increase trust. It also implies that the seller can assume risk if something goes wrong and that it could offer compensation to its customers. Large companies have invested more in their business and therefore have more to lose than smaller companies (cf. Doney & Cannon, 1997).

Reputation can be defined as the extent to which consumers believe a company is honest and concerned about its customers (Rempel et al., 1985; Doney & Cannon, 1997). As such, it is an asset that requires a long-term investment of resources, effort and attention to customer relationships. Again, such companies are seen as unlikely to jeopardise their reputation by acting in an opportunistic way to secure short-term benefits. Jarvenpaa et al. (1999) conclude that a company's size is more important when buying a high-risk product like an airline ticket than a low-risk product like a book. One should bear in mind that their sample consisted of MBA students in their 20s who were frequent users of the Internet. Thus, this must be borne in mind when generalising these findings to other consumer segments.

Fogg and Tseng (1999) approached the subject more analytically by discussing the interrelations between computer credibility, expertise and trustworthiness. They put forward a more holistic approach as they argued that users' evaluation of computer trustworthiness and credibility is a function of both system design features and psychological factors ascribed to the entity behind a system. Fogg et al. (2000) presented the results from an online survey about what design features are likely to increase the credibility of websites. The key findings were that sites should convey a real-world presence, by showing a physical address or displaying the photographs of the staff. Also, professional design, ease of use and frequent updates all contribute to a site's credibility. On the other hand, small typographical errors, technical problems, such as a site's reliability and uptime (availability), as well as long download time can reduce the credibility of a website. It should be noted that this study used a self-selected sample from only two countries (US and Finland) and the measurement of attitudes rather than actual behaviour.

The same group from Stanford states that credibility and trust are similar, but not identical constructs. Their research framework stipulates that perceived trustworthiness and perceived expertise result in perceived credibility. They distinguish four dimension of credibility:

1. Presumed Credibility, based on general assumptions we hold about what certain sites should look like and contain in terms of information;
2. Reputed Credibility, based on a reference from a third party;
3. Surface Credibility, based on what can be found on simple inspection;
4. Experienced Credibility, based on past experience with the site or company.

A large quantitative survey by Fogg et al. (2001) using questionnaires identified five main factors that increase the credibility of a website: real-world feel, ease of use, expertise, trustworthiness and tailoring. On the other hand, commercial implications and amateurism were found to significantly decrease the perceived credibility of a website.

It is important to note that this selection presented the main findings that were used as *early input* to this research project. More recent findings will be referenced where appropriate in later chapter.

2.6.3 Trust in E-Commerce: Industry Reports

In addition to the academic studies presented above, it is important to consider commercially-produced reports on the same topic, as they tend to be more reactive and concrete with respect to consumer behaviour and website design.

The most influential industry report on trust in e-commerce was produced by Cheskin Research and Studio Archetype/Sapient (1999). Although trust develops over time, the authors acknowledge the fact that trustworthiness can be communicated at the very outset of the interaction. Indeed, with time and increasing Internet experience, consumers build up a sense of what a truly professional and trustworthy site should look like. Through a questionnaire study, site reviews and interviews with experts, the authors identified six major factors that help communicate trustworthiness, namely: brand, navigation, fulfilment, presentation, up-to-date technology and security, as well as privacy seals. They argue that combining effective navigation with a strong brand is the best way of communicating trustworthiness.

Branding is to the product or company what reputation is to a person. It would, however, be inaccurate to restrict branding to a company's visual identity, such as its logo and colour scheme. Rather, one should think of branding as affecting all touch points between a customer and a company, be they on- or offline. It is therefore important to project an image that is consistent across different media. Creating brand awareness means making its company and services known to consumers in a way that differentiates them from competitors. For instance, TV ads or online banners are typically used for indirect messaging. Knowing about a brand is certainly not as powerful as experiencing it. Direct experience is employed to create an emotional association with a brand, e.g., by creating a unique, memorable shopping experience.

For new companies with no existing brand name, the report suggests that strong navigation, effective fulfilment and, thus, satisfaction can be effective in communicating trustworthiness. US respondents also found that web-based seals of approval also contribute to communicating trustworthiness. Use and mention of up-to-date technology, in particular as regards encryption was also observed to have a positive effect on consumer trust. It turned out that the most trusted websites were well-known classic brands. This can be explained by the reputation of, as well as personal experience with, these brands offline. The least-trusted sites were obscure, web-only businesses. It is noteworthy that web-only privacy and security seals were perceived as trustworthy only by the people who knew them. An obvious pre-requisite is that consumers should trust the third party and its seals in the first place. Strangely enough, familiar

brands of credit card companies, such as VISA or MasterCard, were less of an indication of trustworthiness than web-only trusted third parties like VeriSign or TRUSTe. Given that the participants in this study were all American, it is unlikely that the same applies in other parts of the world.

The international validity of these results was investigated in a report by Cheskin Research (2000) entitled "Trust in the Wired Americas". It turned out that the VISA brand was most trusted in Latin America while TRUSTe was most trusted in the US. The authors conclude that cultural differences require different strategies to minimise risk and increase trust. Interestingly, US consumers and Brazilians were found to be more cynical as to the ability of governments and hackers to get hold of personal information than Spanish-speaking Latin Americans.

In 2000, the Nielsen Norman Group (NN/g) also published a report on trust as part of their E-Commerce User Experience Series. The definition of trust they used was "the person's willingness to invest time, money, and personal data in an e-commerce site in return for goods and services that meet certain expectations". This is also the definition we will have in the context of the present research. In their user trials, they tested 64 users (US and European) on 20 e-commerce sites. The NN/g study confirmed that people want to have very detailed information about the company and the products they offer, if possible, with objective reviews. Privacy and return policies written clearly were greatly appreciated by the participants. In addition, website design was also found to be important, as was content that was out of date, spelling mistakes, long download times and unclear error messages. When ordering, participants either wanted a secure connection or alternative means of ordering. It also appeared that users found it important to be able to have easy access to company representatives, either through e-mail or chat. One should not that the author was asked to review the NN/g Trust report before publication and that some cross-pollination took place. The extent of the overlap between the NN/g guidelines and the guidelines derived from this research will be discussed in Chapter 4.

2.7 Conclusions

In conclusion, this review of trust-related literature and models showed how initial trust is acquired in first interaction with a party and what factors can affect trust development over time. Regarding antecedents of trust, it is interesting to note that they may be objective facts, associations with familiar situations or mere speculations. In addition, trust can have several objects, such as the other party and the medium used to communicate with that party. That distinction is all the more in an e-commerce situation where system-related factors can be as important as attributes of the online merchant. This chapter presented several different approaches that have been attempted to shed some light on the human factors to affect consumers' trust in e-commerce. It appears that studies that put forward more abstract models and frameworks generally fail to discuss methods whereby substantive knowledge can be applied to design. On the other hand, studies, such as the NN/g report, that are concerned with the identification of concrete design factors sometimes lack a firm theoretical basis. That is why an alternative approach would be to develop both substantive and methodological knowledge, as will be argued in the next two chapters.

CHAPTER 3

MoTEC: A Model of Trust in E-Commerce

This chapter describes the development of the Model of Trust in E-Commerce (MoTEC). Initially presented in Egger (1998), the model was first developed on the basis of the literature presented in Chapter 2, as well as on results from surveys. It consisted of eight components: Transference, Reputation, Attitude, Familiarity, Risk, Cooperation, Benevolence and Transparency. A questionnaire study and user tests helped to refine the analytical model on the basis of empirical results. The initial model went through several iterations (Egger, 2000; Egger & De Groot, 2000; Egger, 2002) to better reflect changing consumer concerns, new findings and common concepts used in related disciplines, such as web design and marketing. The final MoTEC component contains four dimensions, divided into different components and sub-components. The first dimension, Pre-interactive Filters, is made up of User Psychology and Pre-purchase Knowledge. The second, Interface Properties, refers to Branding and Usability. Thirdly, Informational Content contains information about Competence (Company, Products & Services) and Risk (Security & Privacy). The last dimension, Relationship Management, refers to Pre-purchase and Post-purchase interactions with the online vendor and to the development of trust over time.

3.1 Introduction

Chapter 1 gave an overview of the factors generally associated with trust concerns in B2C electronic commerce. Chapter 2 introduced several models of interpersonal trust, both in romantic and business relationships, along with findings on online trust. Given the need to categorise the different trust-shaping factors that have been reported, this chapter presents a classification of the factors in the form of a Model of Trust for E-Commerce (MoTEC).

The first section will describe the different phases that led to the present version of the MoTEC model. The structure of the model will then be presented, along with a thorough description of its components and sub-components.

3.2 Initial Model Development

This section summarises how the MoTEC model was first developed in the MSc project reported in Egger (1998). The very first phase in the development of the model consisted of a review of trust literature and general surveys about consumers' adoption of electronic forms of shopping. The following (early) surveys were of particular interest:

- Equifax/Harris Consumer Privacy Survey (1996)
- CommerceNet Survey (1997)
- Boston Consulting Group (1998)
- 8th Georgia Tech Visualisation and Usability WWW Survey (1998)

A first integration of trust-related literature and concrete e-commerce findings resulted in an analytical version of the MoTEC model. That is, a model entirely developed on the basis of existing literature and survey results. The early model was constituted of the following eight components:

Transference

The first component identified as having a significant influence was Transference, which had been taken directly from the model of Doney and Cannon (1997). It means that if an information source is trusted, this trust can be transferred to another party. Its relevance to e-commerce can be explained by the fact that the Internet is a network of computers where e-businesses, e-magazines, newsgroups, etc. are all linked together. Therefore, a positive attitude to or positive prior experience with one of these sources of information is prone to make it trustworthy. It follows that an e-commerce website mentioned, presented or advertised on a trusted site is more likely to be trusted. In addition, Moorman et al. (1993) observed that information provided by a trusted party is used more and that it, therefore, provides greater value to the recipient. Of course, the notion of transference also encompasses factors such as traditional media (TV, newspapers, etc.), as well as the opinions of people one trusts, e.g. friends or family who recommend certain online businesses. The importance of Transference resides also in the fact that, since it is likely to happen at the start of the computer-mediated interaction with a vendor, an initial confirmation bias is very likely to affect

subsequent assessments of this site's trustworthiness. In other words:

“Prior trust moderates the relationship between psychological breach and subsequent trust.” (Robinson, 1996)

Reputation

The influence of reputation on trust is evident, as these two scenarios demonstrate. The first scenario refers to consumers who know a specific online vendor, either because they have visited the site before or because they know the off-line stores of the same company. In this instance, it is predicted that previous experience with a vendor will give rise to a belief about its trustworthiness which will be difficult to change due to people's confirmation bias. In the case where one has had no previous experience with a specific vendor, one would assume that consumers would want to see information about the company and its development. This information would thus contribute to the assessment of the vendor's predictability of behaviour and trustworthiness. This is explained by the fact that:

“One builds probabilistic beliefs about the [other] party based on rational reasons, such as the past behaviour of or experience with that other party.” (Robinson, 1996)

Attitude

One remembers from Section 2.5 that Doney and Cannon (1997) stressed the importance of salespeople in interfacing with clients and thereby building trust. Like a salesman representing his company, it is argued that the homepage of an e-commerce site represents the online merchant. Thus, an e-commerce website should be able to attract consumers' attention and induce positive emotions in them (cf. Kim & Moon, 1998 in Section 2.6.2). It has already been shown that a good first impression is most valuable because of the confirmation bias that is associated with it. The term *Attitude*, referring to a system's likeability, has been chosen, because it is a common evaluation criteria in usability testing (cf. Dix et al., 1993 and Preece, 1994).

Familiarity

Familiarity, as conceptualised by Good (1988) in the context of trust, was also included in the model, as relating new situations to familiar ones contributes positively to the formation of trust, by a process of transference. Familiarity is particularly important in the context of electronic commerce, as this would entail that commercial websites where shopping procedures are familiar to consumers are likely to be more easily trusted. Familiarity in this sense should be understood as familiarity in terms of information presentation, procedures and usability.

Risk

Given consumers' apprehensions regarding online payments, it seemed imperative to include a component reflecting financial risk. In additions, since the psychological literature stresses that “trust requires a willingness to place oneself in a position of risk” (Rempel et al, 1985), it is argued that the way transactional risk is dealt with by an online vendor constitutes a crucial factor for testing its trustworthiness. This corresponds therefore to an assessment of the vendor's cooperation and benevolence under risky conditions. By the same token, the Risk component should also be related to Doney and Cannon's (1997) calculative process, whereby the costs and the benefits of

a commercial relationship with the vendor are evaluated.

Cooperation

The notion of cooperation was discussed in Section 2.4.5, where it was associated with Rempel et al.'s (1985) Dependability component. Indeed, the Cooperation component of the MoTEC model also refers to a benevolent and, as its name implies, *cooperative* relationship with online vendors. However, it is distinct from both the Risk component above and the Benevolence component below, in that the Cooperation component should also be taken as referring to the level of communication and interactivity between consumers and online vendors (cf. Good, 1988). One should indeed remember from Milgram's (1963) experiment that trust develops more easily when there is a high level of interactivity between the parties involved. In the context of e-commerce, the amount of cooperation is mainly reflected in the quality of the company's customer service and support (i.e. guarantees and after-sales service). Interactivity refers to the means by which consumers can communicate with the company, be it online or offline, and thereby render their relationship more personal.

Benevolence

The Benevolence component of the MoTEC model should not be taken exclusively as meaning dependability or cooperation. For example, in personal relationships:

“Benevolence is defined as the extent to which an individual is genuinely interested in a partner's welfare and motivated to seek maximum joint gain.” (Rempel et al, 1985)

It is suggested that trust in the electronic commerce context is similarly affected by the motivation “to seek maximum joint gain”. It is evident that the vendor's motivation is to make money, while the consumers' motivations are to maximise the value of a transaction while minimising costs, such as wasted time and risks. Thus, the maximum joint gain in a commercial context should be interpreted as a decrease of user costs for which consumers would be willing to spend more money. The Benevolence of the MoTEC model therefore refers to services proposed by online vendors which facilitate the shopping experience by providing intrinsically beneficial and time-saving features. This investment on the part of the e-merchant is thus likely to be interpreted as evidence of beneficial behaviour, which would increase the trustworthiness of the online merchant.

Transparency

The last factor judged to be an important indicator of another party's trustworthiness was transparency. This refers primarily to one aspect of Rempel et al.'s (1985) Dependability component, namely to the other party's honesty. In romantic or commercial relationships, the assessment of honesty is facilitated if personal or confidential information is given by the trustee. The overlap with other factors such as intentionality and predictability is also evident. Indeed, the more one knows about the other party and its intentions, the easier it will be to predict its future behaviour and thereby assess its trustworthiness (cf. Doney & Cannon, 1997). In e-commerce, the Transparency component refers to a sort of *glasnost* on the part of the online vendor; that is, to its openness with respect to its business policies. The importance of this component has also been stressed by Hoffman (1998) when she claimed that:

“Businesses don’t say what they are doing, and consumers have no idea what is happening to their information. The Web is an environment where there is absolutely no trust.” (In Claymon, 1998)

A summary of the eight components constituting the first, analytical MoTEC model is given in Table 1, along with the rationale behind their selection. This allows us to explicitly trace back these analytical components to trust-related literature.

Table 1 – Rationale for initial model component selection (MSc project)

Components	Rationale for Selection
Transference	Effect of transference if little or no previous interaction with other party; confirmation bias (Good, 1988; Robinson, 1996; Doney & Cannon, 1997)
Reputation	Effect of reputation if no previous interaction with other party; confirmation bias; basis for predictability of future behaviour and capability (Rempel et al, 1985; Good, 1988; Doney & Cannon, 1997)
Attitude	Effect of first impression, confirmation bias; effect of system likeability on acceptability (Good, 1988; Dix et al, 1993)
Familiarity	Facilitating effect of familiarity on judgment of trustworthiness; conformity with existing schemata (Luhmann, 1988; Dix et al, 1993; Preece, 1994)
Risk	Dependability under risky conditions; sharing of confidential information; calculative process (cost/benefit) (Rempel et al, 1985; Doney & Cannon, 1997)
Cooperation	Importance of proximity, communication and interaction; dependability; benevolence (Milgram, 1963; Rempel et al, 1985; Good, 1988; Doney & Cannon, 1997)
Benevolence	Importance of benevolent behaviour in risk-related situations; dependability; intentionality (Rempel et al, 1985; Doney & Cannon, 1997)
Transparency	Honesty; dependability; sharing of confidential information; intentionality (Rempel et al, 1985; Doney & Cannon, 1997)

In order to get user feedback on the relevance of these components to consumers’ trust in an online merchant, a questionnaire study was conducted with 14 participants, as is reported in Egger (1998). Forty statements of the form “I trust websites that/if/when...” were developed to investigate people’s *general* attitude towards e-commerce and not one vendor in particular. Participants had to rate the statements on a scale from 1=strongly agree to 9=strongly disagree. The results indicate that all components but one had average scores above 5. The one that did not was the component labelled “Benevolence” which referred to added-value functionalities of a website that would make shopping more convenient. That is why this component was not retained in the final version of the model presented below.

User tests were conducted subsequently, in which 8 other participants had to visit 3 online grocery websites and comment on their perceived trustworthiness. One of the most salient results was that the effect of Transference was hard to distinguish from Reputation. That is why it was proposed to integrate these two aspects under a com-

mon heading. Attitude, and especially the first impression a website makes on users, was found to be crucial. In one case where an American way of classifying food items was unfamiliar to British users, we had evidence that Familiarity should not be restricted to usability in terms of operational efficiency but that it should also extend to cultural differences. Many users found it very important that a website should have a link to a privacy or security policy; however, most of them would not actually access and assess information related to Risk and Transparency. Regarding Cooperation, participants reported that prominent ways of contacting online merchants would help trust formation. However, due to time constraints, the responsiveness and the quality of the buyer-seller communication could not be tested.

The results from the literature-based model development, the questionnaire study and the user tests showed that, in general, the concepts articulated in the initial development of the model were genuine reflections of consumer concerns when evaluating the trustworthiness of an online merchant. The results also showed that Benevolence-related factors were not seen as being directly related to trustworthiness and that other components could be integrated. That is why this initial model was subjected to a complete revision, based on new literature and further user tests, as is described in the next section.

3.3 Revised Model

The present doctoral research was conducted between 1999 and 2003 and produced the papers presented in Appendix 1. Its starting point was the initial MoTEC model that underwent further development cycles to accommodate additional trust-shaping factors and therefore increase its descriptive power. These additional factors were based on new literature, as well as on new user tests. In addition, the model components were renamed to be more in line with commonly encountered terms in website design and marketing. This was to increase the model's understandability and applicability. One should add that some descriptions of the new components are based on user tests that will be presented in later chapters. In other words, the following sections present the results of a series of development-application-evaluation cycles more than of a linear development.

The MoTEC model aims to explain the factors that can affect one person's judgement of an e-commerce site's trustworthiness. It is important to distinguish between initial trust (related to perceived trustworthiness) and trust acquired over time (related to experienced trustworthiness). Indeed, the latter presupposes repeated interactions and a constant monitoring of the other party's trustworthiness, while the former is only based on relatively superficial cues. Our stress in this research will be on initial trust as it is intimately linked with the design of a good customer experience with a company's website. Indeed, following a Human-Computer Interactions approach as explained in Chapter 1, we are particularly interested in what website design elements – be they graphical, informational or navigational – affect the consumer trust experience.

Section 3.4 to 3.7 present the Model of Trust for Electronic Commerce (MoTEC), initially developed by Egger (1998) and refined in Egger (2000), Egger and De Groot

(2000) and Egger (2002). The model attempts to regroup an important number of factors that have been observed to affect consumers' judgement of an online vendor's trustworthiness. Not only does the model list these factors, it also classifies them into different components or interaction phases. This model applies to the selling of products and services in a business-to-consumer situation. Most of it would also apply to business-to-business or consumer-to-consumer exchanges. Nevertheless, this would require slight adjustments to the model presented below.

Given our focus on initial trust, the model has been structured around the different phases a visitor goes through when exploring an e-commerce website for the first time. It is constituted of the following four dimensions, which all contain several components as will be shown below (cf. Figure 4):

- Pre-interactional Filters
- Interface Properties
- Informational Content
- Relationship Management

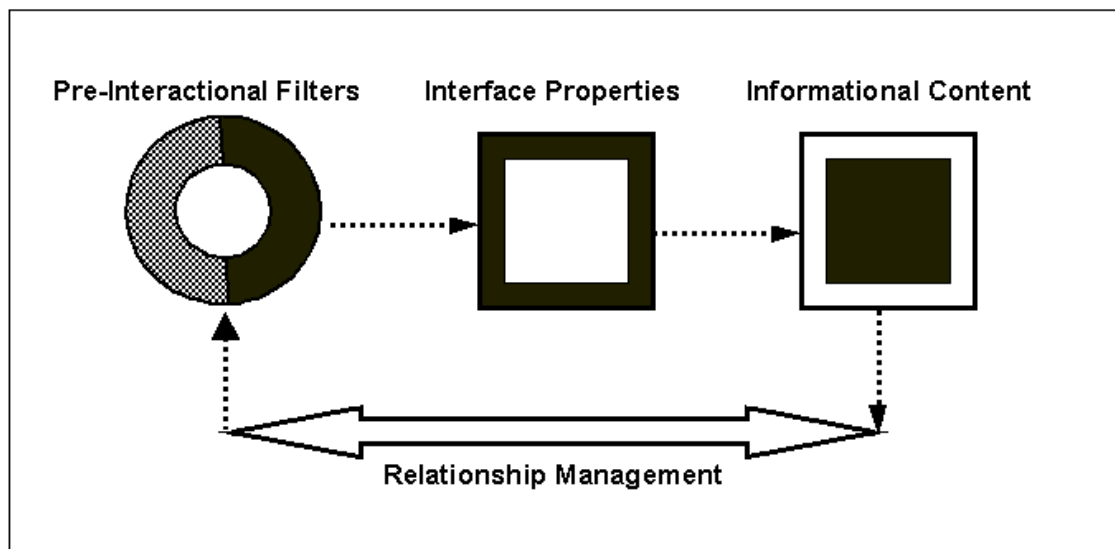


Figure 4 – The four dimensions of MoTEC

Although, eventually, the decision to trust is binary, one's level of trust fluctuates during the interaction with the website, as increasingly more information about the other party is processed (cf. Fogg & Tseng, 1999). The model is based on the metaphor that people's predisposition to trust and pre-knowledge determine an initial trust value even before a merchant website is accessed. As one explores a new site for the first time, the first impression made by a system, in terms of graphic design and usability, will lead to a re-assessment of that trust value. As one examines cognitively more demanding factors, such as the company's competence or the risk of a transaction, one's trust value is bound to change once again. The fourth main dimension, Relationship Management, refers to the handling of inquiries or orders over time. Whether communication happens before or after ordering, the responsiveness and the quality of the help may also affect one's level of trust.

This model shows that trust can be affected by both emotion and cognition, in both implicit and explicit ways. One should add that the evaluation of Interface Properties does not necessarily precede the evaluation of Informational Content; however, the former, *qua* means, is often required for the latter. It remains that the model holds that it is likely that users first pick up superficial cues and that they would then focus their assessment of trustworthiness on the quality of the content. It is noteworthy that this view has been given support both by McKnight et al (2000) and Briggs et al., (2003). The first authors propose a trust model consisting of two stages: *exploration* and *commitment*. These stages correspond to our distinction between superficial and deeper trust assessments. Briggs et al. (2003) also support this view by distinguishing between a “heuristic or impressionistic judgement” and an evaluation that is “cognitively intensive and dynamic”.

3.4 Pre-interactional Filters

The first model dimension consists of Pre-interactional Filters (PIFs), that is, factors that can affect people's perceptions even before a particular e-commerce system is accessed for the first time. As shown in Table 2, there are two main types of PIFs: user psychology and foreknowledge.

Table 2 – Pre-interactional Filters: Components & sub-components

Dimension 1: Pre-interactional Filters	
User Psychology	Pre-purchase Knowledge
General propensity to trust	Reputation of the industry
Trust in IT and the Internet	Reputation of the company
General attitude towards e-commerce	Transference (offline and online)

3.4.1 User Psychology

When trying to identify factors susceptible to have an effect on a consumer's judgement of a vendor's trustworthiness, there are three types of psychological predispositions that need to be highlighted.

General Propensity to Trust. Research has shown that there are large individual differences in terms of readiness to trust another party, be it a person, a group or a business (cf. Rempel et al., 1985). This fact can be accounted for by a variety of philosophical and moral attitudes about the goodness of others, as well as by different personal experiences. Cultural factors are also likely to play a role as it has been found, by example, that that Americans and Japanese trust more readily than Chinese and French (Fukuyama, 1995). Jarvenpaa et al. (1999) have also investigated the possible effect of culture on people's trust in e-commerce websites but did not find a significant effect of culture in the sample they tested.

Trust in IT and the Internet. People's amount of experience with information technology (IT) has a direct effect on how confident they feel using this technology. For example, previous experience helps them discriminate between task-critical and less

important error messages. Of course, prior experience with the Internet is particularly relevant, given the potential reliability problems a person might have experienced with connections to the network. Additional factors that may affect a person's attitude towards the Internet are the context of use (home and/or work), the type of Internet access (modem, cable, etc.) and cost. Generally, a person's expertise in the underlying technology also affects the extent to which the medium is perceived to be reliable and, thus, technologically trustworthy.

General Attitude towards E-Commerce. Although a third of the population forms the "early adopters of almost anything" (Keen, 2000), two-thirds will need good arguments and the benefit of other people's experiences to feel confident enough to embrace e-commerce as a new commercial medium. It appeared in user tests reported in Egger (1998) that people are extremely influenced by the media. As negative shopping experiences are more likely to be reported than positive ones, the way online security is portrayed may be rather biased. Interestingly, novice users with no or little knowledge about encryption might not be worried by the lack of it when transacting over the net. As to expert users, they can typically be split into two categories: those who are aware of the importance of secure connections and trust them; and those who are sceptical about the actual level of security provided by these so-called "secure" connections. Besides, the problem may not lie solely in the connection per se, but also in the way data are stored on a merchant's system. Diversity in predispositions to trust also entails that trust-inducing design features, however well implemented, will never be enough to convert a generally mistrusting individual into a trusting customer.

3.4.2 Pre-Purchase Knowledge

A second important factor is people's foreknowledge and expectations with respect to a certain domain, industry or company. Such pre-purchase knowledge can strongly influence their attitude towards one particular e-business before its website is actually accessed (Good, 1988; Jarvenpaa et al., 1999). As Robinson (1996) puts it: "One builds probabilistic beliefs about the [other] party based on rational reasons, such as the past behaviour or of experience with that other party" (p. 578).

Reputation of the industry. For instance, some people may have a rather negative perception of direct marketing companies (cf. Egger & De Groot, 2000) or second-hand car dealers. This entails that these people will approach online systems of such companies with more mistrust than other systems. When it comes to international business, a similar line of reasoning can be held for associations one might have with particular regions. For instance, people might have more trust in Swiss banks than in banks from a politically and economically unstable country.

Reputation of the company. This refers to one's offline experience with a specific company and/or associated businesses. This experience can be both indirect and direct. Indirect experience is mostly related to brand awareness, i.e., knowing that a particular company exists, what it offers and how it positions itself in the market. Direct experience, on the other hand, implies active interaction with a particular business, be it walking into a store, buying from a store or talking to a customer representative on the phone. In that respect, both brand awareness and experience are likely to play a crucial role as trust in a brand offline is very likely to transfer to its online extension.

Transference. In addition to one's own experience with a company, as mentioned before, one can also rely on the experience or advice of sources one trusts. The phenomenon of trust transference means that one is very likely to trust another, previously unknown or little known, party if a party one trusts recommends it (Doney & Cannon, 1997).

Offline: The author's research has indicated that people rely a lot on the advice of friends, as well as on traditional media. It appears that risk-averse individuals wait until early-adopters have tried out a new service and reported on their experience. As far as the media are concerned, their focus on security problems tends to convey a slightly biased image of online business. For instance, an influential Dutch TV program reported, in the late 1990s, that the online system of a reputable bank had been hacked. According to the bank, this information was inaccurate but it nevertheless had a lasting effect on its customers' adoption of the online banking facilities.

Online: There are numerous sources of information on the Internet that report on the reliability and quality of service of online businesses. These include reviews on influential websites or newsletters, as well as postings in newsgroups, bulletin boards and mailing lists. An increasingly popular destination for exploring customers are websites where people can consult reports on products and companies, as well as ask specific questions to self-proclaimed experts in a variety of fields. An example of such a site is Epinions, which provides this service for free. In order to help assess the credibility of a review or the expertise of the reviewer, Epinions has implemented an internal reputation manager that lets people rate the reviews (Nielsen, 1999). The idea is to make sure that reviewers' advice can be trusted, so as to facilitate the transference process.

Another influential factor in the online world is the presence (or absence) and the ranking of a company's website in directories and search engines. As search engines widely differ as to how they include and rank websites, the average consumer tends to have little knowledge of what a site's ranking actually means. Generally, sites listed first can be perceived to be the leaders in their field, which suggests that they must be reliable and trustworthy. Obviously, expert users know that this is a misconception, as increasingly more search engines propose a service whereby businesses can pay to have their site listed at the top of the list. In addition, a good meta-tag strategy can fool search engines' spiders in believing that a particular web page is highly relevant. To address this problem, search engines like Google also take into account the number of backwards links, i.e. the number of other web pages linking to the target website. The assumption is that a high quality website is more likely to be referred to, and thereby linked to, than a poor site. However, popularity or high content does not necessarily mean trustworthiness. Some directories, such as Yahoo! or Dmoz, hand pick the sites they list and are, therefore, highly selective. Being listed in such a directory is undoubtedly an advantage as it implies an endorsement by an expert; what it does not imply, however, is that the company is truly trustworthy.

3.5 Interface Properties

Interface properties refer to more superficial aspects of an e-commerce website. Superficial should be taken to mean those aspects of an interface that are mostly cosmetic in nature and thus relatively easily changeable. The importance of interface properties lies in the general fact that, in general, “emotional responses precede intellectual ones” (Goleman, 1996). That entails that, superficial though they may seem, interface design features are likely to have a non-negligible effect on a user’s subsequent decision to trust and to buy from an online vendor.

The first component has been termed Branding and refers to a site’s visual design. Design is crucial as it can make a strong first impression when accessing a site for the first time. As Lindgaard (1999) noted, “an immediate negative impression may well determine our subsequent perception of the site’s quality and usability, whereas we may inherently judge a site making a good first impression to be ‘better’” (p. 2). As we have seen in Chapter 2, literature from Psychology also stresses the important role of a party’s first impression, as someone’s confirmation bias would entail that all user actions will unconsciously seek to confirm the first impression rather than falsify it (Kahneman & Tversky, 1973; Good, 1988).

In traditional commerce, Doney and Cannon (1997) distinguished between trust in the salesperson and trust in the company. They identified the salesperson’s expertise, likeability and similarity to the customer as determinant factors to engender trust. Transposed to online commerce, one can hypothesise that the appeal of the interface, the quality of the information provided, as well as the customer-centredness of the system are also likely to have a positive impact on customers’ feeling of trust (cf. Kim & Moon, 1998). Fogg et al. (2002) report that, in their large study about how people evaluate the credibility of websites, almost 50% of all comments made by participants referred to graphic design. They therefore argue that, in the context of online credibility (and trust), findings indicate that *looking* good is often interpreted as *being* good - and credible.

The second important component in the Interface Properties is Usability. According to Nielsen (1993), usability should be understood in terms of a system’s learnability, efficiency, memorability, error prevention and user satisfaction. Conceptually, the Usability component is the necessary link between Branding and Informational Content (the third dimension). Indeed, visual design is presented to the user passively, while the user actively needs to navigate the website in order to access relevant information. Usability is all the more important in the context of online shopping as it is known to be an important condition for the acceptance and adoption of new technologies. The Technology Acceptance Model (TAM), as defined by Davis (1989), holds that usefulness and ease of use are both strong predictors of adoption. The TAM model has also been explicitly applied to trust and e-commerce in studies by Gefen and Straub (2000) or Pavlou (2001), amongst others.

As shown in Table 3, two main aspects of interface properties can thus be distinguished: Branding and Usability.

Table 3 – Interface Properties: Components & sub-components

Dimension 2: Interface Properties	
Branding	Usability
Appeal	Organisation
Professionalism	Navigation
	Relevance
	Reliability

3.5.1 Branding

Appeal. This refers to the former *Attitude* component and the first impression one gets when accessing a site for the first time, notably from the home page. Appeal has largely to do with the site's graphic design and layout. Other elements of branding include prominent information the name of the business, what it does, and what distinguishes it from its competitors. In other words, the online business should be clearly identifiable, be it by a logo or a slogan. A clear statement of what the site is about should also be present, along with the company's key selling points. This also refers to what Dix et al. (1993) call system likeability and acceptability.

Professionalism. Customer-centredness, as well as attention to detail, can help the site convey a professional image. A company's investment in setting up a professional-looking site can be perceived as a sign of a financially viable business with a reputation to defend. Therefore, the company may seem less likely to act opportunistically as it would have more to lose than to gain.

3.5.2 Usability

Organisation. This refers to the extent to which the site's commercial offerings and resources are made explicit by organising its content in a manner relevant to the end user (cf. Kim's 1997 "structure design"). Familiarity in terms of domain knowledge, classification schemes and terminology also fall into this category. In the case of unfamiliar websites, the amount of guidance available, be it in the form of FAQs or a guided tour, can also help first-time visitors familiarise themselves with the system.

Navigation. This has to do with the site's ease-of-use, in particular, the ease of finding relevant information (minimal click stream). An additional factor is the amount of information users can access without having to register, i.e. engage in a transaction of personal information. Generally, ease-of-use could be perceived as a sign that the company understands, cares for and respects its customers. Related to that is the design of dialogues, i.e. system responses to user input. This encompasses confirmation of actions, feedback, constructive error messages, etc.

Relevance. The degree to which consumers feel that the website is relevant to them also has an influence on their willingness to explore the site further. In addition to localisation issues, such as language, date, time, currencies and other measurement units, customisation and personalisation might also contribute to the one-to-one experience.

Reliability. System reliability can be affected by a number of factors, e.g., the performance of the website's servers, the user's ISP or local network, as well as hardware and software problems. The unreliability perceived by a user can often not be attributed to any specific factor. Novice users have even more problems diagnosing causes of unreliability, often blaming themselves if their system crashes. An important factor in this category is a page's download time. User expectations were 8s in a 1999 study and down to 4s over any kind of connection in 2000 (Nelson, 2000).

3.6 Informational Content

As shown in Table 4, the Information Content dimension is constituted of two main components, Competence and Risk, which are further broken down into sub-components.

Table 4 – Informational Content: Components & sub-components

Dimension 3: Informational Content			
Competence		Risk	
Company	Products & Services	Security	Privacy
Identity	Description	Policy	Policy
Values	Objectivity	Encryption	Registration
Contact	Costs	Payment method	Data access
Achievements		Third parties	Subscriptions
Partnerships		Samples	
		Contractual terms	
		Consumer redress mechanisms	

3.6.1 Competence

The competence component directly refers to Doney and Cannon's (1997) capability process, whereby one party assesses the other party's ability to fulfil its promises. In an e-commerce environment, this can be achieved by assessing and comparing information the company's profile and the services it offers.

3.6.1.1 Company

Identity. In interpersonal relationships, knowing who the other party is and how it has behaved in the past proves to be a very good basis to decide whether it is trustworthy (cf. Reputation). The lack of direct interaction in e-commerce can be addressed by providing complete information about the history of the company, its legal status and the people behind it.

Values. Given the moral dimension of trust, companies that manage to communicate their philosophy and values in a credible way can help bridge the identity gap. One way of doing that would be to provide a list of charities or events the company supports.

Contact. Since the buyer-seller relationship is mediated by the online interface, people appreciate easy ways to contact the company, be it online (e-mail, chat, etc.) or

offline (phone, fax, mail, etc.). This can contribute to a sense of closeness and re-embedding, according to Giddens's (1990) definition.

Achievements. A powerful indication of a company's competence is its success. Success can be conveyed by listing important clients, presenting recent projects in a portfolio or simply by providing a link to an annual report. This is also linked to the company's expertise in its area.

Partnerships. Trust transference being such a powerful force in face of the unknown, little known e-commerce sites that have created strategic alliances with high-profile and trusted companies are likely to benefit from those partnerships in terms of added credibility (e.g. Egger & De Groot, 2000).

3.6.1.2 Products & Services

Description. Whether people have a precise goal in mind or simply browse a site, detailed descriptions of the products and services offered helps them make informed decisions about their purchases. Features that reduce user costs, such as comparisons with competitive products, may also be seen as a sign of honesty and competence, although respondents in the questionnaire study reported in Section 3.2 minimised the link between benevolence and trustworthiness. In addition, the provision of related content, if relevant, can also be interpreted as the company truly understanding its customers' needs.

Objectivity. The credibility of the information has also been observed to be very important as unreasonable or misleading claims can say a lot about a company's ethical standards. That is why offsite reviews by trusted entities can also complement content provided by the company or manufacturers and make this information appear to be more objective. Briggs et al. (2002) also report that one of the three factors that came out of their factor analysis in the context of online advice was termed *source credibility*. This provides support for the inclusion of this factor in the MoTEC model.

Costs. As the price can be a very compelling reason to buy from one site rather than another, it is something that people want to have displayed prominently. In addition, transparency with respect to additional costs (e.g. shipping) early in the buying process prevent people from being negatively surprised at a later stage. Besides, if prices are unusually low or high, people expect some additional information about why this is the case. Only then can prospects make an informed decision as to how reasonable the proposed prices are.

As consumers are generally reluctant to buy unfamiliar products or services from an unknown company, one way to address this problem is by offering free samples or free subscriptions. That way, people can experience first hand whether a product or a service conforms to their expectations. A great number of Internet services function on such a basis, offering free membership for basic functions to get acquainted with the system, as well as full membership for a fee.

3.6.2 Risk

As argued in Chapter 2, risk and trust are closely related. This model component can also be related to Doney and Cannon's (1997) *calculative process*, whereby the costs

and the benefits of a commercial relationship are evaluated. In the context of e-commerce, this component can be split into the contractual terms of the exchange, the security of financial transactions and the privacy of confidential information.

3.6.2.1 Security

Policy. Consumers have been observed to react positively to a company's explicit security policies. Such policies typically list the measures taken by a company to ensure that data is transferred, processed and stored at the highest security standards. Although consumers appreciate the sense of security such policies provide, it has been observed that they rarely read them in great detail.

Encryption. Consumers' concerns about security have made them alert about recognising secure connections, characterised by a URL starting with "https://" and the image of a padlock in the browser window. When consumers notice that a particular page they expect to be encrypted is not, they are often not willing to proceed further. Thus, clear textual and graphical feedback about the different levels of security throughout an exchange can reassure prospective consumers.

Payment methods. Although credit card payments have become the main means of transaction on the web, many consumers still feel uncomfortable about giving out their details online (cf. Abrazhevich, 2002). That is why some consumers appreciate alternative methods of payment, be they online, offline or hybrid. Besides, consumers who choose to pay by credit card generally feel reassured if e-businesses stress the fact that consumers are only liable for a relatively small amount in case of fraud.

Third parties. Some companies use third parties such as escrow services to take part in the transaction process. Typically, escrow services ensure that a product is only shipped once payment has been made, which mitigates risk between the transacting parties. Regarding the payment process, precise feedback about what party is involved at what stage of the process can clearly reassure consumers who are sometimes required to access the third party site for payment, before being brought back to the vendor's site.

Contractual terms. Consumers judge the competence and the professionalism of a company by the presumed quality and validity of any warranties, return policies or customer service. Unexpected and unreasonable obligations on the part of the customer once a product or a service is purchased are also likely to affect the customer's judgement of the vendor's trustworthiness.

Consumer redress mechanisms. The presence or absence of a dispute-resolution policy, as well as financial compensation should fraud be committed as a result of an online purchase, also indicate to what extent the company is committed to providing a secure and trusted business environment (cf. Schellekens & Van Der Wees, 2002).

Samples. Risk can be reduced when customers can be given a free sample of an unfamiliar product, a sneak preview or a free trial of a new online service. Since prospects do not have to spend any money at such a stage, as the financial burden lies with the company, consumers can experience the value and reliability of a service at a minimal risk to themselves (cf. above).

3.6.2.2 Privacy

Policy. As with security policies, people like seeing that a website has a privacy policy, although most of them hardly ever read it. That explains why consumers often do not really know what it contains and what protection such a policy gives them. Typically, a privacy policy makes explicit the use and dissemination of any personal information collected on the site.

Registration. People usually want to see as much of a website as possible, before having to register. The type of personal information required in registration forms has been observed to be quite determinant, as information judged to be irrelevant to the transaction at hand is usually not given or is even made up. Thus, the way companies justify the importance of a particular piece of information is important if they want to retain the prospect. Also, the way the registration process is designed, e.g., on one page vs. on different screens, can also affect consumers' willingness to start the process.

Data access. Once a personal profile has been created on a website, the ease with which details can be accessed and modified also contributes to a feeling of control. The same also applies to the retrieval of a forgotten password needed for a particular site.

Subscriptions. Companies consider e-mail addresses as very valuable assets for marketing campaigns. However, experienced Internet users have come to recognise companies that put the consumers' interest before their own. An example of this is the use of opt-in rather than opt-out modes of subscription (cf. Godin, 1999).

3.7. Relationship Management

Relationship Management reflects the facilitating effect of relevant and personalised vendor-buyer interactions on trust development (pre-purchase) and maintenance (post-purchase), as shown in Table 5.

Table 5 – Interface Properties: Components & sub-components

Dimension 4: Relationship Management	
Pre-purchase Interactions	Post-purchase Interactions
Means of contact	Order processing
Responsiveness	Fulfilment
Quality of help	After-sales
Personal touch	

3.7.1 Pre-Purchase

Means of contact. Different means of contact, both online (chat, email) and offline (phone, fax, mail) can be perceived as an indication that the company attaches a great deal of importance to customer care and service. This helps convey a real-world feel, as was observed by Fogg et al. (2001). It also provides a reasonable alternative to

face-to-face interaction.

Responsiveness. Once contact has been initiated by the prospective customer, it is important that the company's response be prompt and informative. Indeed, if a customer never gets an answer to his transaction-critical question, he or she is not very likely to buy from that vendor. Immediate feedback that the customer's enquiry is queued for processing can help communicate responsiveness.

Quality of help. In addition to the promptness of the reply, its quality in terms of relevance and completeness is also of paramount importance.

Personal touch. Customers value a personal touch, e.g., that they are addressed by their name and that the message is written by an identifiable individual, as opposed to general appellations such as "Helpdesk" or "Customer Service".

3.7.2 Post-Purchase

Order processing. Once an order has been placed, consumers generally value a confirmation of their order. The tracking facility proposed by some vendors to follow progress on order processing also constitutes feedback and helps consumers feeling in control of the situation. In case unexpected problems should arise, the way they are handled by the company is also taken as an indication of competence and customer care.

Fulfilment. This refers to the customer experience related to the delivery of the products, viz., the physical delivery to the customer's house, its condition, presentation, packaging, as well as the correctness and completeness of the order. In addition, the amount actually charged by the company should be identical to the amount authorised through the online payment system.

After-sales. Should something be wrong with the order or should help be required, the ease with which customer service can be notified is crucial to the maintenance of trust. Again, the way returns are handled and alternatives proposed all put consumers' trust in a vendor to the test. Fogg et al. (2002) also stress the importance of customer service as a trust cue in e-commerce.

3.8 Conclusions

In conclusion, the section has presented the initial and the revised MoTEC models. The strength of the revised model is that it contains very concrete factors likely to affect trust, while at the same time, being grounded in theory. As most of the Pre-interactive Filters are acquired before the interaction with the website and Relationship Management takes place over time, we shall focus on factors that can easily be observed and manipulated, namely those contained in the Interface Properties and Informational Content dimensions. The next chapter will describe how concrete tools for design and evaluation can be deduced from the MoTEC model.

CHAPTER 4

Trust Toolbox

The MoTEC model presented in Chapter 3 constitutes substantive knowledge about what factors are likely to affect a customer's judgement of an online vendor. In its present form, the model cannot directly be applied to diagnose a site's trust performance or prescribe design strategies to maximise a site's perceived trustworthiness. That is why this chapter shows how methodological knowledge can be derived from the model, in the form of a Trust Toolbox. Three different tools have been developed. The first, GuideTEC, refers to a set of design principles and guidelines. The second, CheckTEC, is a 54-item checklist to be used in expert evaluations by HCI practitioners. Lastly, QuoTEC is a 23-item questionnaire to be used to get trust-related feedback directly from the target customers. As such, QuoTEC can be used either following user tests or on its own, as will be discussed in Chapter 6. These three Trust Toolbox instruments all share a common structure, based on the MoTEC model. While this chapter only discusses the different tools, Chapters 5 and 6 will report a series of applications and comparative validation studies.

4.1 Introduction

Trust-related research in HCI has mostly been concerned with producing substantive knowledge about the determinants and the objects of trust in e-commerce. Although some findings can directly be recruited by HCI practitioners and applied to concrete design cases, there is still a need for independently validated methodological knowledge for trust design and evaluation.

The MoTEC model constitutes substantive knowledge about the factors that can affect a consumer's trust in a specific e-commerce website. That is, the model makes explicit what factors, both on- and offline, are likely to have an impact on consumers' decision to trust an online merchant. MoTEC not only lists these factors, it also makes explicit reference to the psychology underlying their effect on trust. When it comes to a specific e-commerce website, the model can be employed to *design* it in such a way that trust-shaping factors are optimised, to *explain* user interactions and to *evaluate* the site's experienced trustworthiness (thereby *predicting* users' perception of the site's trustworthiness).

The most direct applicability of the model is in explaining consumer interactions with e-commerce websites. Thus, observations and user comments can be analysed using the model as a framework. This helps distinguish between pure usability problems, lack of perceived usefulness (low value) and lack of trust. In the latter case, the model helps evaluators account for user behaviours on the basis of psychological accounts of trust. As far as design is concerned, the next section will show how concrete trust guidelines (*GuideTEC*) can be derived from the model. Section 4.3 will present a checklist (*CheckTEC*) for expert evaluations. Another tool for evaluation is described in Section 4.4 in the form of a questionnaire (*QuoTEC*) to get trust-related feedback about a website directly from the target customers.

4.2 GuideTEC: Trust Design Guidelines

The MoTEC model can also be employed to maximise trust-shaping properties. Design principles and guidelines have been derived from the model and were first presented in Egger (2001). Although guidelines can be useful for design, they also come with certain costs to the guideline user, such as the selection of an appropriate guideline and its translation to an applied setting (Arnfeld & Rosbottom, 1998). One can have either general guidelines that are difficult to apply to a particular situation or very concrete guidelines that may be too context-specific. Here, we have chosen general guidelines that require domain expertise to use them in applied settings. In general, one can distinguish between the three following types of guidelines (Usability.gov, 2003):

1. Guidelines based on experiments: hypothesis testing;
2. Guidelines based on observational evaluations or performance-based usability tests;
3. Guidelines based on observations or expert opinions

The guidelines presented in this section belong between the second and the third cate-

gory. Although these guidelines will not directly be validated in this research, they will be indirectly validated as the checklist presented in the next section contains the same general information. The actual contribution of the checklist will be investigated in Chapter 6.

It is also important to note that only one part of the guidelines concerns traditional HCI aspects in terms of the design of the user interface. The rest refers to a more general trust design strategy that transcends the design of the interface. The guidelines presented below will collectively be referred to as *GuideTEC*, viz. Guidelines for Trust in E-Commerce. The transformation from descriptive to prescriptive knowledge is non-trivial. In this case, the MoTEC model presented two kinds of descriptions:

1. “X has an effect on/plays a role in how trustworthy sites are perceived to be”
2. “Y is known to increase the perceived trustworthiness of e-commerce sites”

The first example refers to factors such as individual differences, which are likely to play a role in the assessment of trustworthiness. Of course, designers cannot change users, but they can adapt designs to suit different user groups (e.g. novices vs. experts). Another example for this first kind of description would be that the quality of a site’s design (e.g., amateurish vs. professional) affects trust. In that case, designers can and should develop websites in a professional manner. However, “professional” remains subjective and is not defined any further. In summary, the first kind of description is transformed into prescriptive knowledge in such a way that designers are told to pay attention to and research such factors. That is, it produces a guideline about an aspect that should be considered, without specifying in detail *how* it should be considered and *what* constitutes a result that would comply with the guideline.

The second example refers to a much clearer case where one knows that Factor Y increases perceived trustworthiness. That is, we know *exactly what* impacts trust and *how*. For example, we know that prominent links to security and privacy policies affect trust positively. From that description, we derive the prescriptive “Provide prominent links to the security and privacy policies”. That is, provided the description is correct in the first place, we would produce a concrete guideline that would invariably guarantee a superior trust performance if applied to a B2C e-commerce website.

We will distinguish between two types of trust design guidelines: process-oriented and product-oriented.

4.2.1 Process-Oriented Trust Design Guidelines

Process-oriented guidelines refer to the procedures designers should follow to ensure that the end product would be perceived as being trustworthy by the target population.

4.2.1.1 Pre-interactional Filters

To fully address customers’ Pre-interactional Filters, it is essential that designers conduct thorough background research. First, the different market segments need to be precisely identified. Secondly, people’s attitudes towards the industry the website belongs to must also be analysed to identify influential preconceptions and misrepresentations. Lastly, the company’s brand equity also needs to be studied to capture con-

sumer's perception of the brand, experiences and expectations. This is also when a clear online branding and user experience strategy must be formulated.

Know the customers:

- Identify the customer segments targeted by the company's marketing strategy.
- Establish a profile for each group: pay particular attention to age, gender, cultural and socioeconomic background, as well as to likely personality traits (e.g., Early adopters? Risk averse?).
- Determine their levels of proficiency with IT, the Internet and e-commerce.

Examine attitudes towards the industry:

- Analyse consumers' familiarity with and perception of the industry.
- Determine the objective risks characteristic of the industry.
- Determine the perceived risks associated with the industry.
- Identify ways others have addressed those risks and concerns, both off- and online.

Analyse the company's brand equity:

- Determine the company's brand position with respect to its competitors.
- Determine consumers' perception of the brand: reputation, quality of experiences, expectations.
- Identify any associations and values connected to the brand.
- Define a clear online user experience strategy.

4.2.1.2 Interface Properties

As far as interface properties are concerned, offline marketing campaigns should be integrated into the design of the online interface. It is important that the different channels all convey the same brand identity to take advantage of people's familiarity with and expectations about the company and its products. When it comes to functional design, the use of user-centred, iterative design methods is strongly encouraged in order to closely monitor people's experiences conveyed by the website.

Take advantage of a familiar brand experience:

- Ensure that the different channels all convey the same image.
- Integrate offline marketing campaigns into the design of the website.

Create an interactive brand experience:

- Take advantage of people's familiarity with related on- and offline companies.
- Complement online with offline branding channels to facilitate the transfer of trust to the website.

Convey a professional image:

- Invest considerable resources for brand positioning, UX strategy and implementation.

4.2.1.3 Informational Content

Competitive analyses should be carried out to position the company's product and

services in a way that combines familiarity of the offering with added value that justifies risk-taking. Also, user testing should be performed to ensure that consumers understand the offering and its value.

Recruit familiar procedures:

- Conduct competitive analyses to understand customers' expectations.

Create value:

- Create added value that may motivate and justify risk taking.

4.2.1.4 Relationship Management

Pre-purchase relationship management also calls for customer inquiries to be handled efficiently, i.e. promptly, precisely and in a personalised way. Post-purchase trust can be greatly aided by providing feedback about the order (cf. next section) and by selecting reliable third parties for the logistics. An effective after-sales service should also be viewed as a key factor for enhancing trust.

Handle customer inquiries efficiently:

- Provide feedback that inquiries are queued for processing.
- Reply to e-mail inquiries within 24 hours.
- Provide complete and personalised responses.

Choose trustworthy commercial partners:

- Choose reliable partners for the website: e.g. hardware, software, hosting, etc.
- Choose reliable partners for the logistics: e.g. for warehousing, packaging, delivery, etc.

Provide an effective after-sales service:

- Make it easy to return products and get refunds.
- Bear in mind that it is cheaper to retain satisfied customers than to acquire new ones!

4.2.2 Product-Oriented Trust Design Guidelines

Product-oriented guidelines specify required attributes of the website, both in terms of interface properties and informational content.

4.2.2.1 Interface Properties

As far as branding is concerned, care must be taken to transpose offline brand attributes of a trusted brand to the online system. This not only includes logotypes, corporate colours, fonts and style guides, but also communication style. This helps transfer familiarity of the online company to the online extension. In case a company does not have an offline presence, one can recruit brand attributes from familiar and related trusted companies, so as to constitute an initial trust capital by similitude. Competence-related trust can be facilitated by a professional appearance of the website, both in term of graphic design and writing style (Fogg et al., 2001).

Take advantage of a familiar brand experience (traditional companies):

- Transpose trusted offline brand attributes to the website (colour scheme, style

- guide, etc.).
- Meet or exceed people's expectations about the look-and-feel and functionality of the website.
- Take advantage of the medium's interactivity for efficient experience branding.
- Allude to the company's investment in its operations and the size of its customer base (if it is large).
- Pay attention to details, be they graphic, textual or navigational.
- Have a domain name consistent with the brand or company name.

It is evident that good usability is a pre-requisite for people to be able to find information on which to base their trust decision. In that respect, web usability guidelines as proposed by Spool (1999) and Nielsen (2000) should be applied whenever possible. Trust-specific user interface guidelines refer mostly to the provision of prominent feedback throughout the shopping process. Allowing for customisation, for instance, by allowing users to select their own content and display preferences (e.g. measurement units), also helps communicate the company's commitment to customer-centricity, while making customers feel in control of the interaction.

Provide easy access:

- Design for cross-platform and cross-browser compatibility.
- Avoid the need for plug-ins and downloads on the homepage.
- Only use plug-ins if they add value to the presentation of information.

Be customer-centric:

- Structure the site in accordance with customers' domain model and expectations.
- Present information in a way relevant to the customer: e.g. thoroughly test localised systems.
- Minimise click stream for greater efficiency and satisfaction.
- Learn and anticipate customers' preferences: e.g. personalisation over time.

Let the customer be in control:

- Support the browsing behaviours of both novice and expert users.
- Inform customers about the procedures required to transact: e.g. overview of steps.
- Provide clear feedback to user actions: allow for easy error management.
- Allow for customisation: e.g. content, language or measurement units.

4.2.2.2 Informational Content

Guidelines for the creation of informational content can directly be inferred from the model, which is why only the most salient ones are presented here. The company behind the site being a main object of trust, it is essential that sufficient background information be provided. Next to complete contact details, this can also cover history, legal status or management profiles. The development of trust can also be facilitated by providing evidence that other parties have trusted the company before and have benefited from it, for instance, by providing a list of clients or projects in the case of the service industry. The company's values and philosophy can also be illustrated by concrete examples, such as sponsoring or charity work. The effect of photographs on

the development of trust in e-commerce websites is an interesting one. Steinbrück et al. (2002) report that the inclusion of photographs in an e-banking site led to greater perceived trustworthiness. On the other hand, Riegelsberger et al. (2003) report that the inclusion of photos increased the perceived trustworthiness of vendors that had a bad reputation, while it decreased that of vendors who had a good reputation. Our advice is to test photographs for content, labelling and positioning with users, so as to avoid possible decreases in perceived trustworthiness.

Regarding content about products and services, the main principle is to create value, while minimising risk. That translates into creating a value-added offering, as increased usefulness leads to increased adoption, which pre-supposes trust (Gefen & Straub, 2000). The credibility of information should be ensured by backing up claims by independent sources, possibly linked to external websites. Any sponsored content should clearly be labelled as such. Complete pricing information should be displayed prominently and early in the process, so as to help the value assessment process.

The risk of the transaction should be minimised by clearly addressing privacy and security concerns upfront. This can be done by providing policies and having them audited by trusted third parties relevant to the target audience. Personal data should be able to be accessed and modified easily. Textual feedback about the use of encryption should be given throughout the ordering process.

Create value:

- Meet or exceed customers' expectations about the quality of descriptions: e.g. multimedia features.
- Support the decision-making process: e.g. by offering comparisons or alternatives.

Be credible:

- Back up objective content with data and references: e.g. external links.
- Provide credentials and affiliations of reviewers.
- Acknowledge any content that is sponsored by or affiliated to another party.

Be transparent:

- Display all costs prominently and early in the process.
- Provide explanations as to unusually high or low prices.
- Be clear as to any implicit costs: e.g. cost of ownership.

Present the company:

- Provide complete contact details: e.g. physical address, phone and fax numbers, etc.
- Provide information about the company's legal status, associations and partnerships.
- Show that there are real people behind the company: e.g. provide key names and photographs.

Describe the company's achievements:

- Provide company background: e.g. history and development.
- Provide a portfolio of high-profile customers.

- Provide investors information: e.g. stock price and annual reports.

Communicate the company's values:

- Stress moral values in the company's philosophy.
- Mention sponsoring and charity activities the company is involved in.

Address security concerns up-front:

- List the measures taken to ensure that data is transferred, processed and stored securely.
- Provide prominent links to the security policy.
- Mention what hardware and software solutions are used: provide external links to providers.
- Complement browser feedback with text to inform users that they are on a secure page.
- Provide several payment options.

Provide reassurance in case of fraud:

- Be clear about consumers' liability: e.g. policies of credit card companies.
- Provide consumer redress mechanisms and financial compensation.

Provide a privacy policy:

- Communicate the company's commitment to the privacy of its customers.
- Provide prominent links to the privacy policy.
- Be audited by and display the seals of an independent trusted third party.

Let customers be in control of their data:

- Delay the need for registration as long as possible.
- Give customers a complete overview of the information required in registration forms.
- Justify the inclusion of seemingly irrelevant details.
- Provide easy means to access and modify data.

Be transparent about the fine print:

- Strive to produce legal documents that are easily understandable.
- Check that terms and conditions are compatible with the legislations of the target countries.

4.2.2.3 Relationship Management

To facilitate customer inquiries despite the lack of face-to-face contact, it is important to provide a number of different means of contact. Besides offline details, such as postal address, phone and fax numbers, online media, such as e-mail or instant messaging, should also be provided.

Provide different means of contact:

- Provide traditional means of contact: e.g. postal address, phone and fax numbers.
- Provide online means of contact: e.g. e-mail addresses or instant messaging.

Automated feedback as to the reception and handling of e-mail enquiries can also help

communicate the reliability of company's technological infrastructure and business processes. Existent customers should be able to easily modify orders, return products or contact customer service.

Provide feedback about the order:

- Send a confirmation message immediately after a customer has placed an order.
- Allow customers to track orders in real time.
- Make it easy for customers to modify and cancels orders.

4.2.3 Conclusions

It is noteworthy that the GuideTEC guidelines presented in this research partly overlap with the findings made by Cheskin Research & Studio Archetype (1999) and the Nielsen Norman Group (2000). The novelty of the present approach lies in the scope of the knowledge it produces and the way descriptive knowledge about online trust is translated into methodological knowledge for trust design. The wide scope of the model is reflected in the fact that the GuideTEC guidelines refer to factors that affect trust before, during and after the online interaction with an e-commerce website. As the NN/g guidelines are intended for usability and user experience designers, they mostly deal with product-oriented guidelines. That is, they do not mention Pre-Interactional Filters and how they can be addressed in the design process. In general, apart from a guideline about fulfilment, their report does not include process-oriented guidelines at all. Another difference is that the GuideTEC guidelines are classified into semantically-related components and sub-components, which might help making sense of all the different trust-shaping factors. No higher-level classification into a trust model or framework was attempted in the NN/g report.

Some limitations of our approach must also be noted, as keeping the design guidelines too high-level comes at a price. Indeed, it might sacrifice the direct applicability of the principles by HCI practitioners. Although each principle was illustrated by concrete guidelines, it was felt that one really needs high-level design principles to gain a general understanding of the trust issues in business-to-consumer (B2C) e-commerce. These general design principles can then be adapted to different industries and business models. Indeed, systems for healthcare services, online banking or online gambling all come with their own set of trust issues (cf. Shelat & Egger, 2002). In other words, these services are bound to attach different weights to the different model components.

4.3 CheckTEC: A Checklist for Expert Evaluations

Traditionally, expert evaluations of an interface consist in inspecting its usability, using, for example, the heuristic evaluation method developed by Nielsen and Molich (1990). Expert evaluation methods are important, as they are more practical than checking for conformity with all existing guidelines (e.g. Brown, 1988; Nielsen, 2000) and more systematic than merely relying on the evaluator's common sense and experience. Expert evaluations can also be more cost- and time-efficient, as they do not involve testing users.

This section describes how the knowledge contained in the MoTEC model can be used by HCI practitioners to evaluate a website's trust performance. Based on the model and the guidelines of the previous section, a checklist was developed to be used in expert evaluations. The checklist has been called *CheckTEC*, which stands for Checklist for Trust in E-Commerce. Its use is intended for summative evaluations of existing, fully-functional B2C e-commerce websites. The main difference between the guidelines and the checklist resides in the phrasing of items and in the fact that only the most salient properties of a component have been retained. The checklist contains 54 items classified per dimension, component and sub-component (cf. Table 6 below).

Table 6 – Checklist for expert evaluations

#	Checklist Items per MoTEC Dimension, Component & Sub-Component
1.	Pre-interactional filters
1.1	This industry which this company belongs to is reputable
1.2	This company is known from the offline world or from advertisements
2.	Interface Properties
2.1	Branding
2.1.1	The purpose of the website is clear from the start
2.1.2	The graphic design of the website is professional
2.1.3	The colour scheme and graphical elements are appropriate for this kind of website
2.1.4	The homepage incites users to explore the site further
2.1.5	The site pays attention to details, be they graphic, textual or navigational
2.1.6	Good use of grammar and spelling can be found throughout the site
2.1.7	The tone used in the texts is appropriate
2.1.8	The site is up to date
2.2	Usability
2.2.1	The pages display correctly in the most popular browsers
2.2.2	Legibility is high thanks to appropriate font sizes and contrast
2.2.3	The website is structured logically
2.2.4	Navigation across different sections of the site is consistent
2.2.5	Finding relevant information is made easy
2.2.6	The site contains no broken hyperlinks
2.2.7	It is easy to select items to purchase
2.2.8	It is easy to access the shopping basket and view its contents
2.2.9	It is easy to edit items in the shopping basket
2.2.10	The checkout and ordering process is intuitive
2.2.11	Appropriate feedback is given about the different steps in the transaction process
3.	Informational Content
3.1	Company
3.1.1	The site provides complete offline contact details: e.g. physical address, phone and fax numbers, etc.
3.1.2	The site contains detailed information about the company's background
3.1.3	The site shows that there are real people behind the company: e.g. it contains key names, photographs and/or short biographies.
3.1.4	The site contains information about the company's legal status, e.g. registration with a Chamber of Commerce
3.1.5	The company mentions partners involved in manufacturing, complementing or shipping products to show that it is reliable
3.1.6	The site contains precise information about when ordered items will be delivered
3.1.7	The site contains meaningful figures about the size of its customer base

3.2	Products & Services
3.2.1	Product descriptions are detailed and complete
3.2.2	Pictures (or other multimedia supports) effectively complement the textual descriptions
3.2.3	Product descriptions are objective
3.2.4	Sponsored content and advertisements are clearly labelled as such
3.2.5	All costs are displayed prominently and early in the transaction process
3.2.6	The site explains why some prices are unusually high or low
3.2.7	The site contains precise information about when orders will be delivered
3.2.8	The products/services sold on this website are familiar or from reputed brands
3.3	Security
3.3.1	The site features a prominent link to the security policy
3.3.2	The security policy clearly describes the measures taken to ensure that data is transferred, processed and stored securely.
3.3.3	The ordering process takes place on secure pages
3.3.4	When on a secure page, browser feedback is complemented with informative text and/or icons on the page itself
3.3.5	The site proposes alternative payment methods (i.e. not only credit cards)
3.3.6	The site features a detailed return policy
3.3.7	The site contains information about consumer redress mechanisms or financial compensation in case of fraud
3.3.8	The site contains a seal from a trusted third party that guarantees the company's commitment to security
3.4	Privacy
3.4.1	The site features a prominent link to its privacy policy
3.4.2	The privacy policy clearly states what personal information is collected, how it will be used within the company and whether it will be sold to other companies
3.4.3	The site features a seal from a trusted third party that audits the company's privacy practices
3.4.4	The need for registration is delayed as long as possible
3.4.5	The site offers a clear overview of the information required in the registration or ordering form
3.4.6	Only personal information that is absolutely necessary is asked for in the registration or ordering form
4.	Relationship Management
4.1	The site provides easily accessible online contact possibilities, e.g. e-mail addresses, live chat option, etc.
4.2	The site has a dedicated customer service area with frequently asked questions (FAQs) or a help section
4.3	The site makes it possible to manage and track orders online

All items are positively-phrased statements about a website. The role of the evaluators is to check the compliance of the e-commerce website at hand with these statements. This can be done in two ways, a formal and an informal one. The formal way of checking compliance involves the use of rating scales. Several evaluators can thus give a numeric value (e.g., from 1 = very low compliance to 5 = very high compliance) to each item and compute average values per component or dimension. This use of the checklist will be documented in Chapter 6. The informal way of using the checklist involves using (a subset of) items as mere guidance. Although informal use can be very efficient, it does not allow sharing and integrating evaluation results in a structured and quantifiable way.

It should be noted that the content of the checklist is, logically, closely related to the content of the GuideTEC guidelines. Therefore, the validation of the checklist in Chapter 6 will indirectly validate the guidelines.

4.4 QuoTEC: A Questionnaire for Trust in E-Commerce

Traditionally, user tests consist of set tasks which participants have to carry out on a system. The main stress has traditionally been on people's efficiency, effectiveness and satisfaction, as defined by Nielsen (1993). When it comes to trust, the model specifies that a pre-test interview can be very useful to explore participants' Pre-Interactional Filters, which will influence their perception of the site during the test. When the test begins, the model makes clear that the first impression made by the website needs to be analysed in great detail. It is therefore advised to have the participant spend some time commenting the homepage before starting an actual, representative information-seeking and product-selection task. In our experience, the think-aloud method has been found to prove very effective in eliciting spontaneous reactions to presentation, content and process (Egger & De Groot, 2000).

Qualitative data can be complemented by quantitative data in the form of a multidimensional trust questionnaire. A questionnaire study may (but does not have to) follow a user test. Therefore, it can be a very time-efficient way to test numerous respondents. Such a questionnaire does not provide a single trust score for each participant, but a score for each major component in each dimension. This helps visualise the exact causes of poor trust performance.

4.4.1 Objectives

The objective of this study was to develop a resource-effective way of diagnosing a system's trust performance, based on user feedback and to be used for formative evaluation. We chose to develop a Questionnaire for Trust in E-Commerce (*QuoTEC*) as questionnaires are known to be very effective for gathering large amounts of data from a large sample at a relatively low cost (Kirakowski, 1994).

The questionnaire we envisaged would measure the perceived trustworthiness of certain website attributes, rather than how much a person would trust the website. Indeed, knowing how much a person trusts a website is little informative for formative evaluation. On the other hand, knowing the trust performance of individual website attributes helps redesign the website in an effective way. We intend the QuoTEC questionnaire to be used on a fully-functional prototype just before launching a website (formative evaluation) or on an existing system (summative evaluation).

4.4.2 Questionnaire Development

Since we were interested in how people experience a website before transacting, we only focused on the Interface Properties and Informational Content dimensions. We selected three pairs of components: Branding and Usability referring to Interface Properties, Company and Products & Services referring to information about Competence, and Security and Privacy referring to information about Risk. Each model component was operationalised by means of several statements reflecting the essence of that component. That is, each component was turned into a scale. The resulting, analytically-derived, questionnaire is presented in Table 7.

Of course, at this point we do not know how well these 23 statements reflect the MoTEC components, nor how they relate to consumers' true trust concerns. That is

why the next chapter will present a thorough application study of the questionnaire, as well as a statistical validation. Chapter 6 will also compare results from questionnaire studies with results from user tests.

Table 7 – QuoTEC questionnaire (analytical)

Questionnaire Items per MoTEC Dimension, Component & Sub-Component	
Interface Properties	
Branding	
1	The site makes a good first impression
2	It seems the website has invested a lot of resources in the design of its website
3	I find the site's visual design clear and professional
4	The site seems up to date
5	The site meets my expectations
Usability	
6	Access to information is quick and efficient
7	This site is organised in a logical way
8	It is easy to find the information one is looking for
9	Navigation is predictable and consistent
10	The website is reliable
Informational Content	
Company	
11	This company seems to be legitimate and respectable
12	This company seems to have a lot of respect for its customers
13	The company seems to be run professionally
Products & Services	
14	The information about products/services is complete
15	Prices are transparent
16	The information this site provides is objective and credible
17	It seems easy to request help or additional information about products/services
Security	
18	The website provides sufficient information about its terms and conditions
19	The website has adequately addressed consumers' concerns about security
20	I would feel safe giving my payment details on this site
Privacy	
21	I would feel confident giving this website my personal details when transacting
22	The website has adequately addressed consumers' concerns about privacy
23	I trust this website will take good care of my personal details

4.5 Conclusions

The strength of the MoTEC model resides in the breadth of its coverage, i.e. including trust-shaping factors before, during and after the online interaction, and its explicit grounding in theory. While its high level of abstractness allows it to be applied to a wide variety of electronically-mediated commercial transactions, very concrete tools can nevertheless be derived from it. However, the GuideTEC guidelines, the Check-TEC checklist for expert reviews, as well as the QuoTEC questionnaire for user trials

still need to be tested for validity and reliability. The QuoTEC questionnaire has been applied in two studies described in the next chapter. A major validation study testing the real contribution of the Trust Toolbox will be presented in Chapter 6.

CHAPTER 5

QuoTEC Applications

In order to test the analytically-derived QuoTEC questionnaire, this tool was applied to concrete website evaluation cases. This chapter reports two studies aimed at uncovering the main constituents of trust by means of the questionnaire. The first study focused on trust in online services by having 20 participants evaluate 6 hotel websites using the questionnaire. The second study focused on trust in online retail by having 50 participants evaluate 2 online bookstores and 2 computer stores using the same questionnaire. The results from the two studies were then combined, resulting in 320 sets of data. A series of factor analyses was conducted, which uncovered 2 main underlying factors. The factors were clearly identifiable as “efficient access to competence-related information” and “perceived risk”. The validated QuoTEC questionnaire has been shortened to 15 items, while retaining a high overall reliability coefficient of 0.9392. The 2-factor structure was also shown to allow efficient visualisations of the trust performances of individual websites. Lastly, several linear regressions are presented, that point to potential differences in assessing the trustworthiness of service and retail websites.

5.1 Introduction

The Trust Toolbox introduced in the previous chapter contained a tool that HCI practitioners can use to assess the effects of certain website attributes on consumers' trust. The objective of this chapter is to apply the analytically-derived QuoTEC questionnaire in two empirical studies. This will test whether items generated to reflect a specific model component predominantly load on the same factor and investigate the individual components' power at predicting consumer trust. Statistics will also be applied to reduce the number of questionnaire items, while minimising effects on explanatory power and reliability. Besides, these results will also be used to reflect on the appropriateness of the MoTEC model's content and structure.

Data collection was made in two separate, independent, studies. The first one focused on trust in online services and looked at hotel websites. The second one focused on trust in online retail and looked at computer and book websites. The methodology used to collect data in either study is presented first. Data analysis of the combined data will be presented next. Lastly, the revised QuoTEC questionnaire will be presented.

5.2 Trust in Online Services

The first study looked at hotel websites as an example of non-standard services. As increasingly more people book travel arrangements online, the importance of a professionally-designed and comprehensive website has become paramount. Hotel websites are especially interesting as many criteria are assessed in the hotel selection process. In addition, staying in a hotel that did not correspond to one's expectations can have profoundly annoying consequences. This study was designed and conducted in collaboration with Dr Roland Schegg, Scientific Collaborator at the Ecole Hôtelière de Lausanne, a leading hotel management school based in Switzerland.

Since the questionnaire is intended as a tool for trust diagnosis and not as a general survey, we administered it in the same way it would be in a real evaluation case. That is, unlike the Lee et al. (2000) and Fogg et al. (2001) studies, we did not ask people questions about websites in general. We administered the questionnaire in relation to specific websites that participants were asked to visit and evaluate.

5.2.1 Experimental Design

The design for this study was a within-subject design, where 6 different websites were tested. The 6 sites were selected as follows: 3 sites were from hotels located in one particular resort in Switzerland, while 3 were from hotels located in an equally popular resort in the Netherlands. All websites had to have similar guest facilities and be of the same price class. The only difference that was sought after was significant differences in terms of website design and content. Thus, each location included one (expected) low-trust, medium-trust and high-trust website, all selected by two experts on the basis of the MoTEC model. The QuoTEC questionnaire was used to measure the dependent variables, i.e. the performance of six model components, as well as a pure

trust measure. Since the aim of the study was primarily to gather data by means of the questionnaire, no explicit hypothesis was formulated. The experimental conditions were counterbalanced in 3 ways: (1) The location of the hotels, (2) The order of the websites within a location and (3) The order of questions in the questionnaire. This was to minimise learning effects due to filling out the questionnaire between tasks.

5.2.2 Websites

Since the tests were conducted both in Switzerland and the Netherlands, the tests included 3 websites from hotels in Interlaken (CH) and 3 from hotels in Scheveningen (NL). The rationale for selecting these locations is their popularity as holiday destinations for the same kind of travellers. Table 8 shows which hotel websites were included in this study.

The high predicted trust websites (H1 and H2) were characterised by their sober graphic design, their ease of use and the completeness of the information they contained. The medium trust websites (H2 and H5) stood out by a less professional graphic design, a confusing site structure and the incompleteness of booking-related information. Lastly, H3 and H6 were predicted to be low trust websites, because of the amateur graphic design and the bad organisation of their content resulting in difficult access to information.

5.2.3 Scenarios

One scenario was created and adapted to the two countries. This scenario stated that a close friend invited the participant to a special party on a specific week-end in one of the selected cities. Furthermore, the scenario asserted that, since that location was worth visiting a bit longer, the participant should plan his or her stay around the specified week-end. The last part informed them that their friend had included the URLs of three hotels in that city without knowing how appropriate these hotels would be. The participants' task was therefore to find out which of the three hotels appeared to be the most trustworthy to book online and stay at. This scenario forced the participants to thoroughly explore each of the websites, critically evaluating the hotel's value proposition and comparing it against its alternatives.

5.2.4 Procedure

The participants were first asked a few general background questions such as their expertise using the Internet, their experience buying products and services online and their general trust in online services. They were then given the scenario described above and introduced to the think aloud procedure. While they were carrying out the tasks set by the scenario, the tester took notes of their browsing behaviour and comments.

The QuoTEC questionnaire consisted of 23 items which participants were asked to rate on a 7-point Likert scale. To minimise order effects, two versions of the trust questionnaire were created. At the end of both versions, an additional item was included to provide a pure measure of trust, by asking respondents to rate the statement: "I would trust this website". This was to have a direct and unambiguous measure of the dependent variable under scrutiny.

Table 8 – Screenshots of the six hotel websites by predicted trust category

Predicted Trust	Interlaken (CH)	Scheveningen (NL)
High	<p>H1</p>	<p>H4</p>
Medium	<p>H2</p>	<p>H5</p>
Low	<p>H3</p>	<p>H6</p>

5.2.5 Participants

Twenty subjects, six of whom were females, took part in this study. Twelve were based in Switzerland and 8 in the Netherlands. The tests were conducted in English, French and German. Participants were recruited locally at both locations and consisted of members of the general public that were representative guests for these hotels. Pre-test interviews determined that 7 participants were newcomers to the Internet, 4 were advanced users and 9 were expert users. Regarding general trust in e-commerce, 3 participants reported low trust in e-commerce, 7, medium trust and 10, high trust.

5.3 Trust in Online Retail

The second study dealt with online retail sites that sold standard products. Two sites were online bookstores and two others sold computer equipment, two very popular product classes to be bought online. That is why it seemed very important that the questionnaire should also be applied to those types of e-commerce websites. This study was designed and conducted in collaboration with Lei Huang, then an MSc student in HCI at University College London (UK).

5.3.1 Experimental Design

As in the hotel website study, this was a within-subject design, in which participants had to evaluate four different websites: 2 online bookstores and 2 online computer stores. The QuoTEC questionnaire was used to measure the dependent variables, i.e. the performance of six model components, as well as a pure trust measure.

The test sessions were counterbalanced in three ways: (1) The order of the industries, (2) The order of the websites within an industry, and (3) The order of the questionnaire items.



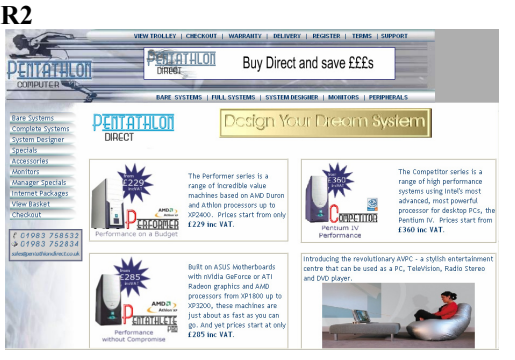
5.3.2 Websites

Our strategy was to select more than one industry to ensure that the questionnaire could be generalised to a number of business-to-consumer retail sites. We selected online book and computer retail websites, two popular industries, to increase the ecological validity of the results. The idea was to have two sites per industry, where one site was predicted to inspire low trust (LT) and the other one, high trust (HT). Another criterion was that the sites should be relevant to a UK audience and would therefore also ship to the UK. The sites used in this study are described in Table 9.

R1 and R3 were predicted to be high trust websites. The most noteworthy characteristic of these two sites was that they had brand names known to the participants. Both had a clear visual design and offered clear guidance throughout the ordering process. They also contained comprehensive risk-related information.

R2 and R4 had a much more amateurish graphic design and confusing navigational structure. In particular, R2 had an awkwardly implemented shopping cart functionality and did not say anything about its privacy policy. R4, on the other hand, did not even have photos of book covers, nor reviews, and also had an unfamiliar way of adding products to one's shopping basket.

Table 9 – Screenshots of the four retail websites by predicted trust category

	Predicted Trust	Computer Stores	Bookstores
High	R1		
	R2		R4
Low			

5.3.3 Scenarios

The objective of this experimental set-up was to let participants evaluate each website in a way that was most similar to a real situation. That is why, for each industry, we created one scenario that would require participants to look for products with given attributes, as well as evaluate the trustworthiness of the vendor. The reason we included the product search task is that it would not make sense for people to worry about the trustworthiness aspect of a website if the site did not contain a relevant product and would therefore not be considered as a candidate in the first place. To make the trustworthiness evaluation process as natural as possible, we did not ask people explicitly to look for, say, the privacy policy or encryption on a payment page. Rather, we urged them to look for any information that might help them assess the trustworthiness of the vendor.

5.3.4 Procedure

At the beginning of each test session, participants were asked to fill in a background questionnaire asking for their age group, gender, education level, Internet usage, as well as their general trust in electronic commerce. These questions all refer to the model's Pre-Interactional Filters dimension.

The same two versions of the questionnaire were used in this study. The only difference was that HTML versions of these questionnaires were created to let participants fill them in online. The online versions made use of a form validation rule that precluded respondents from submitting their answers if a question had been omitted. Results were automatically sent by e-mail and later transferred to an SPSS spreadsheet.

5.3.5 Participants

Fifty UK-based participants, 27 of which were female, took part in this study. Forty-one belonged to the 18-30 age group and the same number accessed the Internet daily. The participants, about three quarters of which were university students, received a financial reward for their participation. The tests were conducted in English. The background questionnaires revealed that the values for their general trust in e-commerce roughly followed a normal distribution. This suggests that the test sample represents a good cross-section of the general population with respect to attitude towards e-commerce. Since we tested 50 people on 4 websites, 200 sets of data were collected in this study.

5.4 Analysing the Main Constituents of Trust

5.4.1 Combining Data

The first step consisted in looking at the data for each study, independently. This was to see whether there were any differences between the data collected in the two studies. The main analysis, however, was conducted on the combined data sets from the service and retail studies. Table 10 summarises the main attributes of the questionnaire studies.

Table 10 – Main attributes of the questionnaire studies

Attributes	Service Industry	Retail Industry
Industries	Hotel	Computer & Books
Test Locations	Switzerland & Netherlands	United Kingdom
Number of Websites	6	4
Participants	20	50
Completed Questionnaires	120	200

Thus, 320 questionnaires of 23 items were collected. This amounts to 7360 items that were combined in a single SPSS data sheet.

5.4.2 Reliability

To test the reliability of the questionnaire, we looked at the Cronbach's Alpha coefficients given by SPSS's Scale Reliability test. As a rule of thumb, Alpha coefficients above 0.70 are generally regarded as acceptable for psychometric measurements. As shown in Table 11, the overall reliability coefficient for the 23-item questionnaire (on

the combined data) is extremely high, at 0.9580.

Table 11 – Initial Cronbach Alphas for the QuoTEC components

Model Component	Initial Cronbach Alpha
Presentation	0.8713
Navigation	0.8983
Company	0.8448
Products & Services	0.7667
Security	0.7119
Privacy	0.7226
Overall	0.9580

5.4.3 Factor Analysis

The statistical procedure selected to analyse this data was factor analysis. In general, factor analysis is used to uncover the latent structure of a set of variables. It can therefore validate a scale by showing that its constituent items load on the same factor and identify items that cross-load on more than one factor. Its use can also help reduce the number of items in a scale by identifying clusters of cases in a graphical representation. In this analysis, we opted for exploratory factor analysis as it presupposes no validated theory. Had the MoTEC model been validated before, confirmatory factor analysis would have been more appropriate to determine whether the number of factors and the loadings conform to what is expected on the basis of pre-established theory. The factor analyses below were conducted using principal component analysis with Varimax rotation. Principal components analysis is a method used to form uncorrelated linear combinations of the observed variables. That is, the first component has maximum variance, while successive components explain progressively smaller portions of the variance and are all uncorrelated with each other. The number of factors extracted in each analysis is thus a function of an internal rule of SPSS, based on the proportion of variance explained by the first factors, and the size of the contribution by the remaining ones.

Before looking at the factor analysis on the combined data, we shall first examine the structure and factor loadings for each study separately.

5.4.3.1 Hotel Websites

Table 12 shows the results of a unique factor analysis (using principal component analysis with Varimax rotation) on the complete hotel website data. The resulting model is a 4-factor structure that explains 69.79% of the variance. For each item, the highest factor loading has been highlighted in bold to help find a pattern in the resulting factor structure. The factors can be described as follows:

Factor 1 regroups almost all Usability items and some other items related to efficient access to information. It also contains one item from the Privacy component. This factor can thus be primarily interpreted as **efficient access to information**.

Factor 2 mostly regroups items from the components Company, Products & Services and Privacy. It can be interpreted as the **integrity** of the company in providing quality information about its offering and its privacy practices. Factor 3 contains all the Branding items but one, as well as one Company item about the perceived respect the company has for its customers. The combination of graphic design and the impression made by the company can best be interpreted as **presentation**. Factor 4 has remarkably few items that highly load on it: only two Security and one Privacy item. It can thus be interpreted as **perceived risk**.

What this structure stresses is the relative independence of Interface Properties from Informational Content. In particular, Factors 1 and 3, constituting Interface Properties are kept distinct from the other properties. As to Informational Content, two factors regroup four initial components. The most interesting aspect is Factor 2, “integrity”, as it seems to refer to three components at once. In doing so, it merges two distinct objects of trust, namely, the company and the information it provides.

5.4.3.2 Retail Websites

A factor analysis on the retail websites data revealed a 3-factor structure that accounts for 67.73% of the variance. Again, the three factors will be interpreted in the light of their factor loadings (cf. Table 13).

Factor 1 mostly refers to the way the company presents itself and its offering via the website. It seems to regroup the notions of *integrity* and *presentation* we had in the previous section. Therefore, this factor could be interpreted as **professionalism**. Factor 2 regroups all the Usability items but one, as well as other items referring to the process of retrieving relevant information. It can therefore also be interpreted as **efficient access to information**. Factor 3 mostly refers to items in the Security and Privacy categories. Therefore, it will also be called **perceived risk**, as above.

The most striking difference between this model and the one from the hotel study is obviously the number of uncovered factors. It is also interesting to note that, based on our interpretation of the factor loadings, two factors seem to refer to the same underlying concepts, namely, *efficient access to information* and *perceived risk*. The only distinction that got lost in this study was that between *integrity* and *presentation*. This could be explained by the very different nature of the industries involved. Indeed, in the hotel case, it does make sense to have this distinction as people, if they book a room, will physically go to that hotel and interact with its staff. Hence the need for presentation (“What will it look like?”) and integrity (“How will they treat me?”) as separate concepts. In the case of retail websites, the products are standard and therefore presentation-related risk, minimal. Also, since customers will not physically engage in face-to-face exchanges with their staff, integrity might indeed be less important. Our interpretation of Factor 1 as *professionalism* combines the two concepts in a way that makes more sense in a retail context.

A recent, analytical, trust model by Corritore et al. (2003) distinguishes three main factors: *perception of credibility*, *ease of use* and *risk*. Interestingly, these three factors can almost perfectly be mapped to the three factors identified in this study.

Table 12 – Hotel Websites: factor loadings

	Factors			
	1	2	3	4
Brand1	0.260	0.145	0.819	0.181
Brand2	0.025	0.455	0.635	0.325
Brand3	0.422	0.394	0.590	0.228
Brand4	0.253	0.169	0.810	0.182
Brand5	0.208	0.618	0.389	0.016
Usab1	0.706	0.206	0.493	0.118
Usab2	0.699	0.227	0.439	0.108
Usab3	0.801	0.245	0.348	0.199
Usab4	0.703	0.267	0.411	0.108
Usab5	0.469	0.567	0.346	0.131
Comp1	0.223	0.784	0.197	0.098
Comp2	0.214	0.544	0.555	0.084
Comp3	0.225	0.554	0.495	0.134
Proser1	0.573	0.393	0.442	0.202
Proser2	0.618	0.306	-0.087	0.303
Proser3	0.571	0.579	0.220	0.003
Proser4	0.334	0.384	0.189	0.364
Sec1	0.170	0.172	0.239	0.673
Sec2	0.140	0.050	0.137	0.897
Sec3	0.623	0.504	0.027	0.132
Priv1	0.482	0.643	0.073	0.140
Priv2	0.345	0.638	0.216	0.215
Priv3	0.101	0.067	0.118	0.902

Table 13 – Retail Websites: factor loadings

	Factors		
	1	2	3
Brand1	0.306	0.654	0.358
Brand2	0.319	0.392	0.634
Brand3	0.604	0.537	0.180
Brand4	0.677	0.219	0.347
Brand5	0.478	0.573	0.219
Usab1	0.359	0.782	0.200
Usab2	0.298	0.706	0.387
Usab3	0.295	0.743	0.339
Usab4	0.564	0.612	0.103
Usab5	0.833	0.141	0.090
Comp1	0.754	0.274	0.350
Comp2	0.429	0.611	0.393
Comp3	0.701	0.326	0.390
Proser1	0.353	0.369	0.573
Proser2	0.644	0.282	0.102
Proser3	0.443	0.678	0.312
Proser4	0.019	0.547	0.393
Sec1	0.264	0.364	0.594
Sec2	0.260	0.198	0.799
Sec3	0.666	0.319	0.398
Priv1	0.558	0.326	0.484
Priv2	-0.033	0.414	0.609
Priv3	0.434	0.124	0.705

Table 14 – Combined Websites: factor loadings (initial)

	Factors		
	1	2	3
Brand1	0.662	0.312	0.275
Brand2	0.519	0.303	0.457
Brand3	0.543	0.571	0.212
Brand4	0.413	0.494	0.314
Brand5	0.537	0.531	0.058
Usab1	0.747	0.372	0.187
Usab2	0.755	0.339	0.215
Usab3	0.775	0.335	0.234
Usab4	0.599	0.558	0.080
Usab5	0.204	0.780	0.170
Comp1	0.274	0.771	0.206
Comp2	0.622	0.439	0.277
Comp3	0.383	0.679	0.249
Proser1	0.573	0.365	0.392
Proser2	0.230	0.573	0.212
Proser3	0.628	0.542	0.135
Proser4	0.539	0.077	0.397
Sec1	0.312	0.201	0.681
Sec2	0.193	0.192	0.836
Sec3	0.302	0.710	0.283
Priv1	0.310	0.660	0.336
Priv2	0.450	0.171	0.460
Priv3	0.086	0.309	0.805

5.4.3.3 Combined Data

The initial factor analysis on the combined data resulted in a 3-factor solution accounting for 63.69% of the total variance (cf. Table 14). It is noteworthy that the combined data gives rise to the same number of underlying factors as the retail study above.

Factor 1, interestingly, shows almost the same pattern of factor loadings as Factor 2 in the retail study. Therefore, our interpretation as **efficient access to information** will remain unchanged. Factor 2, on the other hand, is very similar to Factor 1 in the retail study. Again, we will label it as **professionalism**. Factor 3, in the combined data analysis, results in an even cleaner structure of the **perceived risk** factor in the retail study. Given the striking resemblance between the factor loadings in the retail and the combined analyses, it seems that implications of the data combination are minimal.

Since our primary objective was to reduce the number of items in the QuoTEC questionnaire, we conducted further analyses on the combined data set. For each scale, we identified the item that either did not show the same loading pattern as the rest or that had the lowest loading on the common factor. One item per scale was removed as a consequence. A new factor analysis was conducted on the remaining 17 questionnaire items. This resulted in a 2-factor structure accounting for 61.15% of the total variance. As clusters of cases were identified, 2 further items (Brand2 and Usab2) were removed from the data set.

The third and last factor analysis resulted in a 2-factor structure explaining 61.25% of the total variance. The rotated component matrix below shows the factor loadings of the remaining items. Which items primarily load on which factor has been emphasised through the use of bold typeface. As Table 15 shows, all items belonging to the Risk component load more strongly on Factor 2.

Table 15 – Combined Websites: factor loadings (3rd iteration)

Label	Questionnaire Item	Factors	
		1	2
Brand1	1. The site makes a good first impression	0.661	0.359
Brand4	2. I find the site's visual design clear and professional	0.611	0.382
Brand5	3. The site seems up to date	0.775	0.093
Usab1	4. Access to information is quick and efficient	0.768	0.282
Usab3	5. It is easy to find the information one is looking for	0.769	0.332
Usab4	6. Navigation is predictable and consistent	0.803	0.188
Comp1	7. This company seems to be legitimate and respectable	0.722	0.250
Comp3	8. The company seems to be run professionally	0.721	0.320
Proser1	9. The information about products/services is complete	0.644	0.444
Proser2	10. Prices are transparent	0.545	0.269
Proser3	11. The info. this site provides is objective and credible	0.817	0.215
Sec1	12. The site provides sufficient info. about its terms and conditions	0.307	0.700
Sec2	13. The site has adequately addressed cons. concerns about security	0.191	0.880
Priv2	14. The site has adequately addressed cons. concerns about privacy	0.399	0.498
Priv3	15. I trust this website will take good care of my personal details	0.186	0.840

It is interesting to note that the initial factor analysis in of the hotel data resulted in four factors, while this model contains only two. This factor structure and the loading pattern can be interpreted as follows:

Factor 1 includes all the items belonging to the components Branding, Usability, Company and Products & Services. According to the MoTEC model, the first two components constitute Interface Properties, while the last two refer to a merchant's Competence (which is part of Informational Content). The item loadings on this factor therefore cover aspects referring to a site's graphic design and ease-of-use on the one hand, and information about the company and its offering on the other.

It is interesting to note that this factor regroups components that were conceptually separated in the model. However, data reduction has resulted in a factor that regroups whole components, as opposed to partial components, viz. only some items within a given scale. One way of conceptually combining these components would be to say that, together, they refer to **efficient access to competence-related information**. That is, the interface properties constitute the means by which users navigate to content about the service offering and the service provider. Please note that *integrity* has not been retained as a label in this case, as Factor 1 does not contain any Privacy items.

Factor 2 includes all the items belonging to the components Security and Privacy and refers therefore to **perceived risk**. Again, as in the hotel and retail studies, the combined data analysis confirms that Security and Privacy are not only related to each other theoretically but also in people's minds. Observations during user tests also suggest that users often do not distinguish between security and privacy concerns, as privacy can also be thought of as related to the *security* of personal data. In addition, it is not unusual for websites to publish security and privacy policies in the same section, which may blur the distinction between these two concepts even more.

Another aspect of the model that is confirmed by these results is the distinction between information related to Competence (Company and Products & Services) and Risk (Security and Privacy Security). Most importantly, what the 2-factor model reveals is that Risk-related information is perceived as a separate factor altogether. Indeed, although the MoTEC model included Risk in the Informational Content dimension, it did not say anything about the entirely separate role this component might play.

Figure 5 shows a graphical representation of the items plotted in a space defined by Factor 1 on the x-axis and Factor 2 on the y-axis. It is noteworthy that the risk-related items are all located in the upper-left quadrant. On the other hand, all items referring to efficient access to information are situated in the bottom-right quadrant.

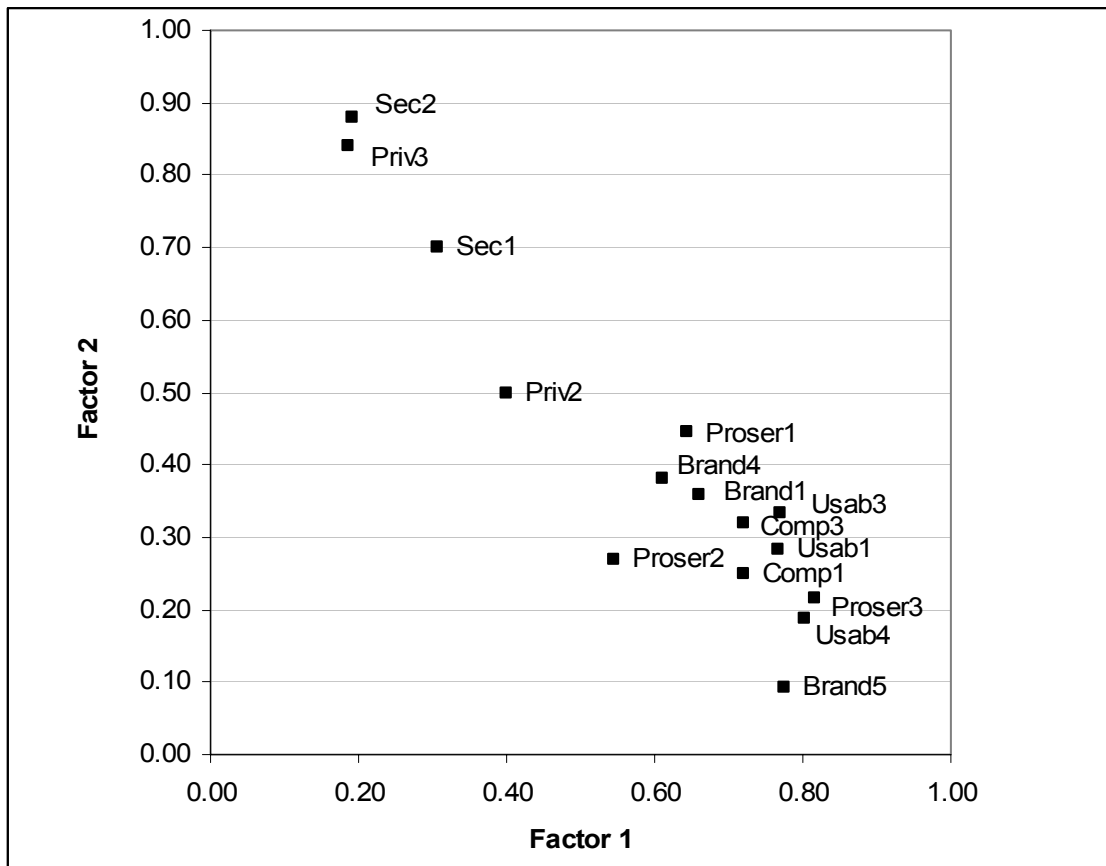


Figure 5 – QuoTEC items in the 2-factor space

Figure 5 also shows that all factor loadings of the questionnaire items are positive. This can be accounted for by the fact that a website can offer more or less efficient access to information and more or less comprehensive information about risk. That is, by definition, these two factors can never be negative.

Although the repartition of items may suggest that they are distributed along one line, we are careful not to conclude that trust is a one-dimensional concept. Indeed, as the initial factor analyses demonstrated, trust can equally well be described by three or four factors. One should not forget that the 2-factor structure only accounts for 61.25% of the variance. As reported above, the initial factor analysis on the combined data, which accounted for 63.69% of the variance, was a 3-factor model. It was only after the exclusion of some items, based on their loading pattern, that a 2-factor model emerged. Thus, it is very unlikely that a one-dimensional model would be able to account for a significant proportion of the variance. The second reason is that the two factors that we interpreted above seem conceptually distinct enough not to be merged. Indeed, it may well be that a website provides efficient access to the products it is selling but inappropriately addresses consumers’ risk-related concerns – and vice versa. Implications of the 2-factor structure for the MoTEC model are discussed at the end of this chapter. As far as the questionnaire is concerned, one could argue that it could be reduced to two items representing the extremes of efficient access to information and perceived risk. However, once again, the results show that it would not be meaningful to exclude the other items as these two factors have too low a coverage of the

total variance to allow this and, in addition, there might be other concepts inside the 2-factor structure that we can neither extract, nor interpret.

It is important to emphasise that the primary aim of the factor analysis in this case was to reduce the number of questions in the questionnaire. This item reduction was successfully carried out, as the number of items went from 23 to 15 with more or less the same proportion of the variance explained. As the questionnaire was the main focus of the factor analysis, it also means that, in order to understand online trust, one should not exclusively focus on the two easily interpretable factors. Given the low coverage of the variance, there is bound to be additional concepts or dimensions that one has to take into account.

5.4.4 Final QuoTEC Questionnaire

A new Scale Reliability test was conducted on the shorter questionnaire. It is noteworthy that the 15-item instrument has a very respectable overall reliability coefficient of 0.9329 (cf. Table 16).

Table 16 – Final Cronbach Alphas for the QuoTEC components

Model Component	Final Cronbach Alpha
Presentation	0.7641
Navigation	0.8856
Company	0.7993
Products & Services	0.7543
Security	0.7204
Privacy	0.5815
Overall	0.9329

Table 17 shows the final QuoTEC questionnaire that now consists of 15 items instead of 23 initially (cf. Section 4.4.2). The remaining questionnaire items are listed per component; however, the order of the questions will have to be randomised when using the questionnaire to gather feedback from users.

Table 17 – Final QuoTEC questionnaire

Questionnaire Items per MoTEC Dimension, Component & Sub-Component
Interface Properties
Branding
1 The site makes a good first impression
2 I find the site's visual design clear and professional
3 The site seems up to date
Usability
4 Access to information is quick and efficient
5 It is easy to find the information one is looking for
6 Navigation is predictable and consistent
Informational Content
Company
7 This company seems to be legitimate and respectable
8 The company seems to be run professionally
Products & Services
9 The information about products/services is complete
10 Prices are transparent
11 The information this site provides is objective and credible
Security
12 The website provides sufficient information about its terms and conditions
13 The website has adequately addressed consumers' concerns about security
Privacy
14 The website has adequately addressed consumers' concerns about privacy
15 I trust this website will take good care of my personal details

5.4.5 Trust Performance Visualisation

The advantage of a two-dimensional structure resides in the fact that each factor can be represented on an axis, which, together, would redefine a space in which individual websites could be plotted. In order to utilise the two dimensions as a visualisation tool, we have first computed averages for each item across respondents, for each of the 10 websites used in the validation studies. Each average was then multiplied by its associated pair of factor loadings displayed in Table 15. In other words, after multiplication, each averaged item produced one value for the x-axis and one value for the y-axis. The last step consisted in averaging all x values and all y values, thus producing coordinates for each of the 10 websites.

Figure 6 show the distribution of the 10 websites on the graph. The thick lines superimposed to the graph represent simulated values for a case where all items had been

given average values. That is, the two thick line create four quadrants, of which the bottom left one represents high perceived risk (= low perceived security) and difficulty in accessing relevant information. Because this is the worst case imaginable, the quadrant has been coloured in grey to highlight badly performing websites.

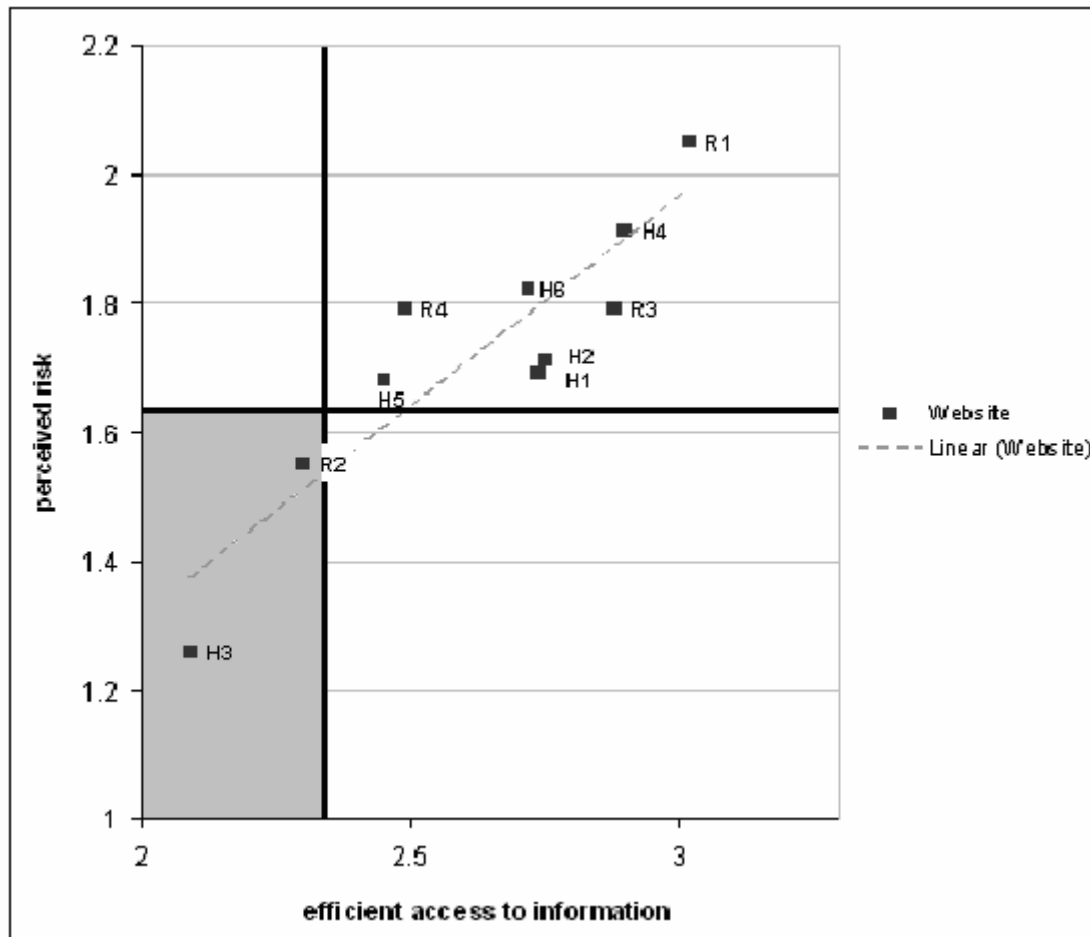


Figure 6 – Trust performance visualisation as a function of efficient access to information and perceived risk

How do these results compare with the trust measures participants had to give on a scale from 1 to 7? To answer this question, the score frequencies were noted for each website. To turn the results from the 7-point scale into a dichotomy of trust-not trust answers, scores 1 to 3 were counted as “trust”, scores of 5 to 7 were counted as “not trust” and neutral scores of 4 were evenly distributed into these two groups. The results are presented per industry and per set of websites. The number in brackets that follows the website label corresponds to the proportion of participants who would trust that website.

The Swiss hotel sites were rated as follows:

$H1 (90.0\%) > H2 (77.5\%) > H3 (60.0\%)$

This sequence confirms our expected trust predictions.

The graph shows H1 and H2 located very closely to another. However, H2 seems to have a slightly higher trust value than H1. This is interesting as it does not correspond

with the expected ranking. H3, on the other hand, scores very low as expected and is located in the grey quadrant. What is remarkable is that, even though H3 is located in the bottom left quadrant, 60% of the respondents would still trust this site.

The Dutch hotel sites were rated as follows:

$H4 (87.5\%) > H5 (72.5\%) = H6 (72.5\%)$

This sequence confirms our prediction that H4 would be rated higher than H5 and H6. However, the pure trust responses show no difference between H5 and H6.

The graph shows indeed that H4 scores better than the other two sites but also indicates that H6 has a higher trust performance than H5.

The computer websites were rated as follows:

$R1 (79.0\%) > R2 (54.0\%)$

This sequence confirms our expected trust predictions.

The graph makes this distinction very clear as R1 gets the best overall trust performance, while R2 is located in the grey quadrant.

The book websites were rated as follows:

$R3 (72.0\%) > R4 (60.0\%)$

This sequence confirms our expected trust predictions.

The graph shows that both websites have the same value on the y-axis, i.e. on perceived risk. However, as R3 scores better on efficient access to information, its overall trust performance is higher.

The comparisons between the proportion of participants who trust a given site and its location on the graph have been made to see whether there was convergent evidence that both methods showed the same results. For the hotel websites, there is a clear match between trust responses and the location of the worst Swiss and the best Dutch sites. However, the performance of the other four sites does not exactly match the trust responses. Regarding the retail websites, both sequences can perfectly be found back in the graph.

In conclusion, the comparisons provide some convergent evidence that the trust visualisation, based on the 2-factor model, reflects people's trust judgements. For cases that did not match the trust responses, two interpretations can be given. First, it may be that the people's answers to the pure trust question do not align with their answers to the questionnaire. This could be because of factors not covered by the questionnaire, such as their familiarity with a particular brand. The second interpretation is that the trust visualisation is not accurate as it is based on a 2-factor model that only accounts for 61.25% of the variance. Of course, mismatches could also be due to a combination of these two factors.

5.4.6 Regression Analysis

Given that we included a pure trust measure in addition to the QuoTEC items, we have enough data to investigate which MoTEC components are the best predictors of trust. This section will present the results of multiple linear regression analyses (step-wise), based on the reduced 15-item questionnaire. As for the factor analysis, we shall first look at the two studies separately and then at the combined data.

One should point out that, statistically, combining data from the two studies might introduce a bias, as data points are generated in clusters. That is, each participant in the hotel study evaluated 6 websites and each participant in the retail study evaluated 4 other websites, which might violate the assumption of independence in the data. However, we tried to minimise this potential bias by counterbalancing the order in which websites had to be evaluated and by collecting a large amount of data.

5.4.6.1 Hotel Websites

The model resulting from the analysis of the hotel data accounts for 63.2% of the variance ($F [2, 117] = 61.899, p < 0.001$). Table 18 shows that only two components, Company and Products & Services, were retained in the equation as significant predictors of consumer trust.

Table 18 – Multiple linear regression for the hotel data

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	0.781	0.321		2.430	0.017
Company	0.858	0.065	0.772	13.205	< 0.001
(Constant)	0.614	0.307		2.001	0.048
Company	0.637	0.084	0.574	7.620	< 0.001
Products & Services	0.273	0.070	0.292	3.882	< 0.001

The resulting linear regression equation reads:

$$\text{Trust} = 0.64 (\text{Company}) + 0.27 (\text{Products \& Services}) + 0.61$$

In Section 5.4.3.2, we tried to account for the fact that the hotel data resulted in one more factor than the model for the retail data. Our explanation had to do with the peculiarity of the hotel industry to the extent that guests will physically be at the hotel and interact with its staff. This hypothesis receives strong support from the regression analysis, as the same two elements that make hotels special are the best predictors of trust in hotel websites. These two factors turn out to be so important that none of the other four MoTEC components is retained in the stepwise analysis.

5.4.6.2 Retail Websites

The model resulting from the analysis of the retail data accounts for 73.2% of the total variance ($F [4, 195] = 114.907, p < 0.001$). Table 19 shows the 6 steps involved in this analysis. An interesting observation that can be made about these results is the transition between steps 5 and 6. Indeed, Branding is one of the best predictors of trust for most of the analysis and ceases to be one the moment Company is introduced.

The resulting linear regression equation reads:

$$\text{Trust} = 0.44 (\text{Privacy}) + 0.30 (\text{Products \& Services}) + 0.24 (\text{Company}) + 0.23 (\text{Usability}) - 0.98$$

Like the equation from the hotel study, this equation also contains Products & Services and Company as good predictors. Interestingly enough, the best predictor of trust turns out to be Privacy. While Branding is no more in the equation, Usability is also retained as a strong predictor variable. One could explain the effect of Usability by the fact that retail websites offer a wide variety of products in different categories. Therefore, locating and selecting a product might be more cumbersome than just specifying what kind of room one would like on a hotel website. The fact that Privacy is the best predictor could be accounted by the fact that customers really pay online (as opposed to a mere booking on most hotel websites) and that they will never meet the people behind the company (as opposed to the hotel staff). Thus, cues about their integrity with respect to risk management need to be picked up directly from the website.

Table 19 – Multiple linear regression for the retail data

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	0.120	0.289		0.415	0.679
Branding	0.983	0.062	0.749	15.929	< 0.001
(Constant)	-0.768	0.270		-2.845	0.005
Branding	0.651	0.066	0.496	9.837	< 0.001
Privacy	0.525	0.062	0.424	8.408	< 0.001
(Constant)	-1.110	0.262		-4.239	< 0.001
Branding	0.349	0.085	0.266	4.120	< 0.001
Privacy	0.468	0.060	0.378	7.835	< 0.001
Products & Services	0.433	0.083	0.327	5.218	< 0.001
(Constant)	-1.085	0.258		-4.202	< 0.001
Branding	0.256	0.091	0.195	2.820	0.005
Privacy	0.450	0.059	0.364	7.609	< 0.001
Products & Services	0.338	0.089	0.255	3.790	< 0.001
Usability	0.202	0.076	0.175	2.658	0.009
(Constant)	-1.047	0.255		-4.104	< 0.001
Branding	0.155	0.098	0.118	1.580	0.116
Privacy	0.429	0.059	0.346	7.273	< 0.001
Products & Services	0.268	0.092	0.203	2.915	0.004
Usability	0.192	0.075	0.167	2.555	0.011
Company	0.192	0.076	0.169	2.534	0.012
(Constant)	-0.976	0.252		-3.873	< 0.001
Privacy	0.443	0.059	0.358	7.580	< 0.001
Products & Services	0.304	0.090	0.230	3.391	0.001
Usability	0.231	0.071	0.201	3.256	0.001
Company	0.241	0.070	0.212	3.465	0.001

5.4.6.3 Combined Data

In order to examine which components are good predictors of consumer trust in general e-commerce websites, we conducted a multiple regression analysis on the combined data, again using the stepwise method. Table 20 shows the four, subsequent

models that were produced. The final model ($F [4, 315] = 163.881, p < 0.001$) explains 67.5% of the total variance.

Table 20 – Multiple linear regression analysis on the combined data

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	0.746	0.206		3.620	< 0.001
Company	0.836	0.042	0.745	19.897	< 0.001
(Constant)	-0.002	0.210		-0.112	0.911
Company	0.651	0.044	0.581	14.677	< 0.001
Privacy	0.381	0.047	0.323	8.172	< 0.001
(Constant)	-0.319	0.207		-1.541	0.124
Company	0.465	0.053	0.414	8.711	< 0.001
Privacy	0.304	0.046	0.258	6.548	< 0.001
Products & Services	0.333	0.058	0.278	5.721	< 0.001
(Constant)	-0.289	0.205		-1.414	0.158
Company	0.434	0.054	0.386	8.052	< 0.001
Privacy	0.280	0.047	0.237	6.000	< 0.001
Products & Services	0.228	0.068	0.191	3.369	0.001
Usability	0.156	0.053	0.154	2.929	0.004

The linear regression equation derived from the analysis therefore reads:

$$\text{Trust} = 0.43 (\text{Company}) + 0.28 (\text{Privacy}) + 0.23 (\text{Products \& Services}) + 0.16 (\text{Usability}) - 0.29$$

As in the factor analysis, the result for the combined data is quite similar to that for the retail websites. Indeed, exactly the same predictor variables are retained in the equation, only in different proportions.

What is remarkable is that Branding and Security have not been found to be significant predictors of trust. One could explain the absence of Branding by the fact that the effect of graphic design is implicit, as it is not an end in itself but a means to attract customers to a site and motivate them to explore it further. As to the Security component, this result is all the more surprising as security concerns are sometimes equated with trust concerns. One explanation for this result could be that visitors to a website do not primarily look for security-related information. Rather, they base their judgement of the vendor's trustworthiness on other factors. It is important to stress that Branding and Security should not be dismissed as being irrelevant. Since the explained variance of the multiple regression analysis is 67.5%, it is hypothesised that these two components do play a non-negligible role nevertheless. Further, this result concerns 10 websites only, so generalisability over websites is not in any way guaranteed.

5.5 Conclusions

In conclusion, the implications of the factor analysis results for the validity of the MoTEC model are twofold.

First, the factor analysis on the combined data provides support for the proposition that the pairs of components defined in the model are indeed related. Indeed, the pairs Usability-Branding and Company-Products & Services are kept together in Factor 1 (of the third iteration). As to the pair Security-Privacy, it is also preserved in Factor 2.

Secondly, the 2-factor structure clearly indicates that that efficient access to competence-related information and information related to perceived risk are statistically distinct factors affecting people's judgement of a vendor's trustworthiness. The present model only list factors likely to affect trust but does not try to attach weights to the different components, nor does it attempt to regroup components across dimensions (cf. Factor 1 that contains both Interface Properties and Informational Content items). At this point, it is important to emphasise the difference between the objectives of HCI research and HCI design. There is no doubt that the 4-, 3- and 2-factor models that were presented in this chapter make a valuable contribution to HCI research. Indeed, our starting point was a list of factors likely to affect trust and now we have a better idea of the *constellation* in which these factors affect trust, for different industries. Thus, these models can be new starting points for HCI research, as they provide additional information.

However, as far as HCI design is concerned, the main product of the factor analysis consists of the revised, 15-item, QuoTEC questionnaire and the trust visualisation tool. The added value of the factor constellation knowledge to HCI designers is less clear. Indeed, in order to design e-commerce systems in which trust-shaping factors are optimised, the same set of guidelines will hold. The model and its derived tools all state the importance of risk-related information, viz., that this kind of information should be comprehensive and well-communicated to the end user. Our conclusion is that HCI designers would benefit more from retaining the MoTEC model presented in Chapter 3, as its structure more easily maps onto the roles and responsibilities of the different actors involved in a design process.

Validation studies involving the QuoTEC and CheckTEC tools, as well as user tests are presented in the next chapter.

CHAPTER 6

Toolbox Validation

Chapter 4 presented a Trust Toolbox consisting of a suite of three tools for trust design and evaluation. The first objective of this chapter was to test the actual contribution of using the CheckTEC checklist, as opposed to using no tool at all. An experimental set-up was created whereby one group of HCI specialists evaluated two websites without the help of a tool, while a matched group evaluated the same two websites using the checklist. The findings indicate that, on average, checklist users find four times as many problems than unguided evaluators, in half the time. In addition, checklist evaluators find about 90% of problems observed in user tests, while that proportion is just over 50% for unguided evaluators. The problems found by the former group also appeared to refer to more MoTEC components. The second objective was to test whether the correspondence between predicted and observed problems was greater for checklist-guided than for unguided evaluators. The third objective was to test to what extent a remote administration of the QuoTEC questionnaire would produce the same results as when it is given to participants straight after a user test. It appeared that the component performance predictions produced in both conditions correlated significantly with each other. However, care must be taken to force questionnaire-only participants to follow the set scenario in a systematic way.

6.1 Introduction

The tools introduced in Chapter 4 have been derived from the MoTEC model, which, in turn, was developed on the basis of literature on trust and user tests. The content of the checklist and the questionnaire can therefore be explicitly traced back to its sources, which maximises the face validity of the tools. However, to this point, we do not have concrete evidence of stronger forms of validity. In simple words, we do not have any evidence that the tools work as such and do contribute to actual HCI design practices. For this reason, this chapter reports a series of studies aimed at testing the actual contribution of the CheckTEC checklist, as well the predictive power of QuoTEC questionnaire results.

6.2 Hypotheses & Approach

The toolbox validation studies reported in this chapter aimed at testing three hypotheses described below.

6.2.1 Hypothesis 1

Hypothesis 1: Evaluators who use the CheckTEC checklist will find more, as well as more varied problems than unguided evaluators.

The first part of the hypothesis refers to the number of predicted problems found by either evaluation method. It was hypothesised that checklist-guided evaluators would find more problems because CheckTEC would force them to consider website attributes that go beyond usability. The second part of the hypothesis refers to the type of trust problems identified. Given the checklist's wide scope, it was hypothesised that the trust problems found by unguided evaluators would refer to fewer MoTEC components than the problems found using the checklist.

Since this hypothesis required a comparison of two expert evaluation studies, we devised the following experimental set-up:

Study 1 consisted of an evaluation of two websites, based solely on 10 evaluators' knowledge and experience. Each evaluator had to report predicted trust problems, i.e. design defects that may cause the consumers not to trust the website. The individual reports were collated and a summarising document was produced.

Study 2 consisted of an evaluation of the same two websites conducted by 10 different experts, this time, using the CheckTEC checklist. Evaluators had to indicate for each CheckTEC item whether it was considered to be a problem or not, by providing a rating on a web-based version of the checklist.

The general set-up for this comparison is presented in Section 6.3, while the results of Study 1 are presented in Section 6.4 and those of Study 2 in Section 6.5. The number of problems identified by each method will be compared in Section 6.6.

6.2.2 Hypothesis 2

Hypothesis 2: The correspondence between predicted and observed problems is greater for the checklist-guided predictions than the unguided predictions.

This hypothesis refers to the overlap of problems reported by target customers in user tests with problems reported in Study 1 and Study 2. It was hypothesised that checklist-guided evaluators would be more accurate in their predictions than unguided evaluators. To test this hypothesis, we needed to conduct user tests in order to collect data from representative users. That is why we devised the following, additional, study:

Study 3 consisted of scenario-based user tests aimed at observing, as opposed to predicting, trust problems. Individual user tests were followed by the administration of the QuoTEC questionnaire. The testing of hypothesis 2 was only concerned with qualitative results in the form of observed trust problems, which can be referred to as **Study 3a** (QuoTEC results, referred to as **Study 3b**, will be discussed in the next section).

The set-up of Study 3 is described in Section 6.7 and the results for the two websites in Section 6.8. The comparison between the problems predicted in Study 1 and Study 2 with observed problems will be presented in Section 6.9.

6.2.3 Hypothesis 3

Hypothesis 3: There is no difference between the questionnaire results in the user tests followed by the questionnaire condition (Study 3b) and the questionnaire-only condition.

This hypothesis refers to the possible effects of the conditions in which the QuoTEC questionnaire is administered to representative users. One condition, exemplified by Study 3, would be to conduct user tests to collect qualitative data and, after the test, ask participants to fill out a questionnaire to collect quantitative data. As this set-up presupposes user tests to be conducted, it is relatively resource-intensive as the facilitator must carefully organise and schedule test sessions and be physically present in all of them. It would, theoretically, be much easier just to distribute the questionnaire, along with a set task, and let participants fill it out in their own time and place. The question is whether the questionnaire results produced without a facilitator are similar to those produced in the more resource-intensive condition. To test the hypothesis that there is no difference between the two conditions, we used the questionnaire results from **Study 3b** and those of the following study:

Study 4 refers to an evaluation of the same two websites solely by means of the questionnaire, i.e. based on the same scenarios as Study 3 but without one-to-one user tests beforehand.

The set-up of this comparative study is described in Section 6.10, while the results of Study 3b and Study 4 are reported in Sections 6.11 and 6.12, respectively. The comparison between the results is made in Section 6.13.

6.2.4 Approach

Figure 7 summarises the four studies introduced in the preceding section.

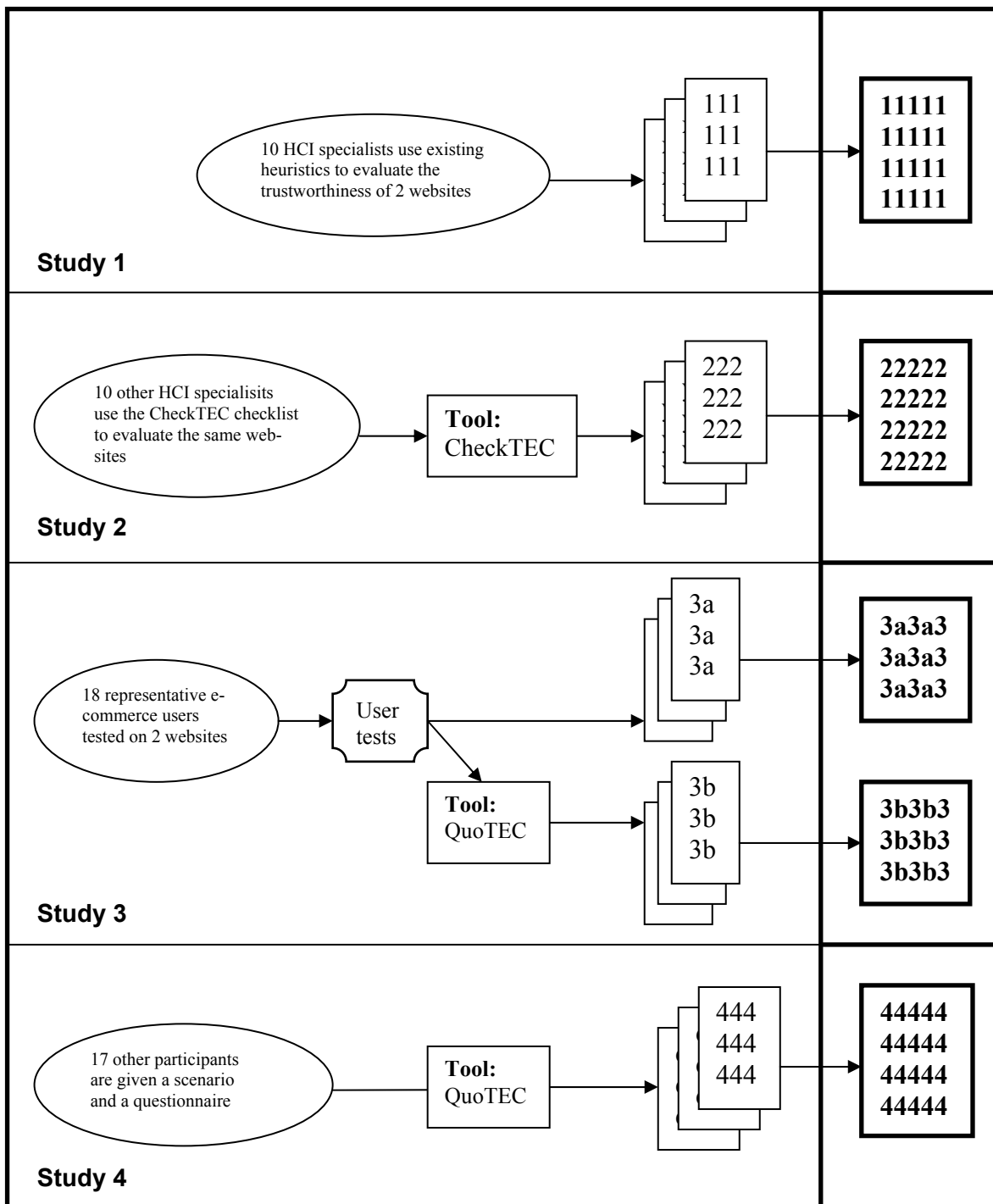


Figure 7 – Toolbox validation approach

6.3 Trust Problems Predicted by Expert Evaluators

The first experiment aimed to demonstrate the actual contribution of the checklist in terms of its ability to predict trust problems. This was done by comparing the performance of unguided evaluators (Study 1, discussed in section 6.4) to that of evaluators who used the checklist to evaluate two specific websites (Study 2, discussed in section 6.5). Such a set-up provides direct qualitative and quantitative support for using the checklist. Our first hypothesis was as follows:

Hypothesis 1: Evaluators who use the CheckTEC checklist will find more, as well as more varied problems than unguided evaluators.

6.3.1 Experimental Design

The design for this study was a between-subjects design with two matched groups. The independent variable was the guidance given to the expert evaluators to perform their task. It had two levels: unguided or checklist-guided. The dependent variable was the number and nature of the reported trust problems. The hypothesis was that expert evaluators in the checklist-guided condition would identify more and more varied problems than those in the unguided condition.

The two conditions were not tested concurrently so that the unguided experts would not be able to have access to the checklist. The checklist-guided condition was tested once all unguided evaluations had been received. In order to get a good idea of how trust problems were identified, it was decided to have 10 evaluators per condition. Nielsen (1993) argues that 10 usability experts should uncover approximately 85% of the usability problems that can be found by evaluators; we hypothesise this to be equally applicable to trust problems. In addition, this would also provide enough sets of data to allow for quantitative inferences. To attain some reliability for results across websites, expert evaluators were asked to review two different websites, one from the retail sector and one from the service industry. It is important to note that we were not interested in comparing the websites *per se* but in the consistency in performance of unguided and checklist-guided experts across sites.

6.3.2 Participants

Given the focus of this study, it was important to have a cohesive group of participants so that differences in performance would mainly be attributed to the different levels of the independent variable. For that reason, we recruited our 20 evaluators from a pool of 1st year and 2nd year postgraduate students, as well as PhD students, all in the area of HCI. All participants had been taught courses on cognition and design, as well as on usability evaluation methods. Thus, we will use the term *experts* to refer to these students, not in the sense of experienced professionals, but in the sense of HCI specialists, rather than mere customers visiting a website. Appendix 2 shows the experience of the evaluators in Studies 1 and 2.

Of course, one limitation of this set-up is that most students did not have extensive field experience evaluating systems, which might have affected their results. It is not unusual, however, to use students in comparative usability studies, although such a sample may lead to limited external validity (Gray & Salzman, 1998). Besides, as the

checklist is intended for HCI practitioners with no or little trust knowledge, this choice of participants is justified.

Since it was important to form two groups that were matched, all participants had to fill in a background questionnaire asking about their experience with HCI (in years), as well as their expertise in expert reviews, web usability and trust issues. Based on these results, 2 groups of 10 participants were formed with an equal distribution of overall expertise. That ensured that the groups were well matched and that their results could be comparable. The unguided condition consisted of evaluators A1 to A10, while the checklist-guided condition consisted of evaluators B1 to B10. All evaluators were given a financial reward for their participation.

6.3.3 Websites

Two websites were selected in this study to test for the reliability of the checklist. Given the international background of the evaluators, both sites had to be in English. To control for the possible effect of origin, the two sites were based in the UK. Also, to control for reputation effects, both sites had to be unknown to the evaluators and therefore non-mainstream. In addition, given this study's focus on trust problems, the two selected websites had to have obvious trust shortcomings. As an example of the service industry, the first website was a flower delivery service. The particularity of such as business is that it deals with non-standard, perishable products where correct and timely delivery is paramount. This site stood out by its crude and dated graphic design, as well as the scarce information it contained about its products and about the flower shop itself (cf. Figure 8). The dialogue in the ordering sequence was also little intuitive, making it difficult to specify a delivery time.

As an example of the retail industry, the second website was a discount fragrance shop (cf. Figure 9). In that case, all products were standard, brand-name perfumes. The site's main selling argument was competitive prices owing to duty-free bulk purchases. This site stood out by its cluttered visual identity and poor usability. For instance, the site did not offer extensive product information, assuming customers would already be familiar with a fragrance before buying it online. Also, it claimed to be "110% secure", although it did not contain any precise information about how security would be guaranteed. The company's overstating its commitment to security is hypothesised to negatively affect its credibility. It is interesting to note that Figure 8 shows another occurrence of the indication "110%", this time in relation with the veracity of their guaranteed low prices. Since the tests were conducted just before Valentine's Day, both websites were using a so-called "falling hearts" theme on their homepages.



Figure 8 – Homepage of the flower website

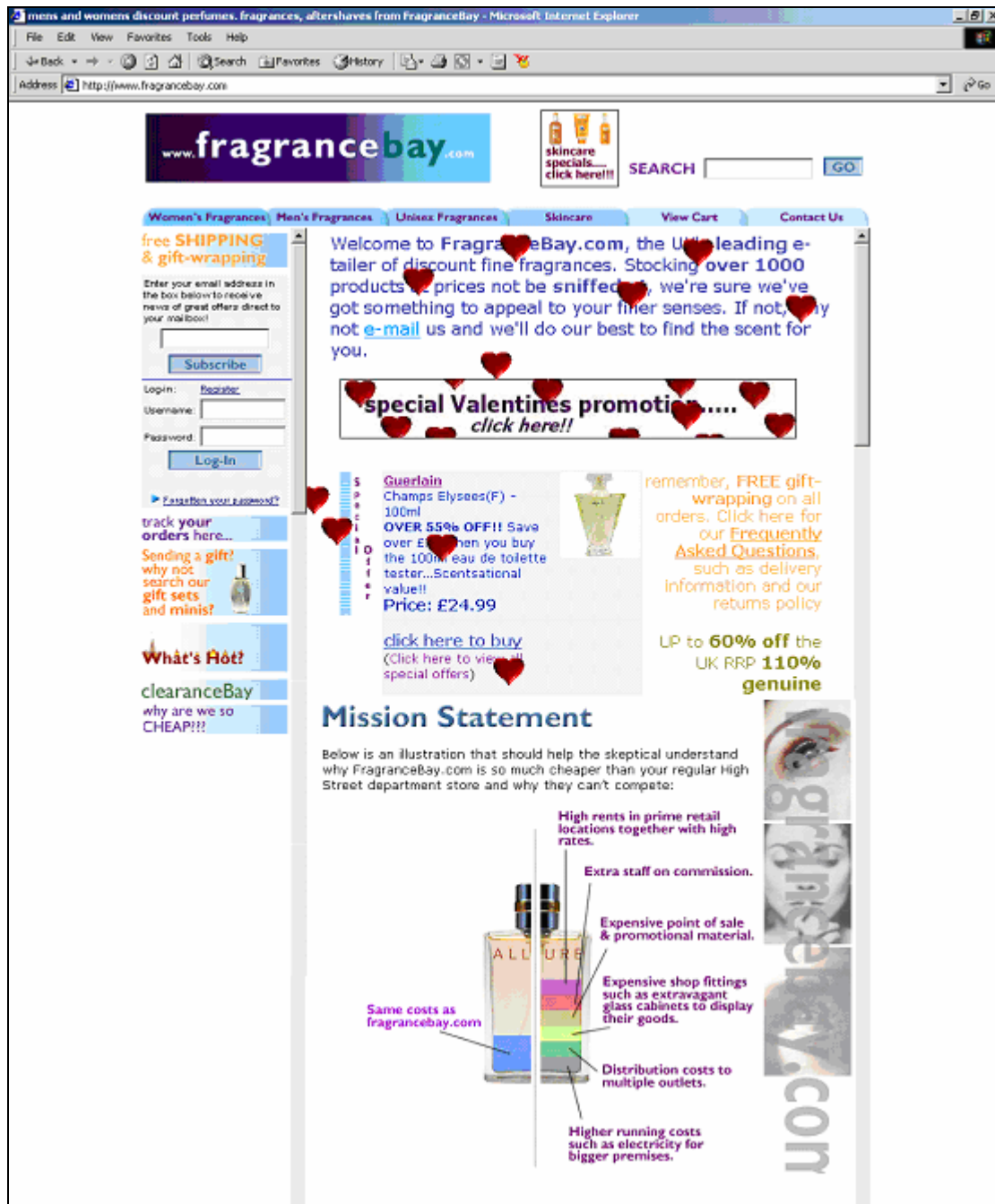


Figure 9 – Homepage of the perfume website

6.4 Study 1: Unguided Expert Evaluations

Study 1 consisted in asking participants to evaluate the two websites without any guidance as to what factors might affect trust.

6.4.1 Procedure

Selected participants received an email message containing their participant number (A1 to A10), as well as a link to a web page. That page contained the instructions they had to follow for their evaluations.

The scenario was as follows: The participants had to play the role of HCI consultants that had been approached by two different clients to evaluate their website. The participants' knowledge about HCI and usability should be used to identify websites elements that were likely to affect user trust, either positively or negatively. The aim of the evaluation was to find out whether potential customers would trust either site enough to buy from it. The instructions further specified that they had to conduct the evaluations solely on the basis of their knowledge and experience. They were neither allowed to use any trust-related resources, nor to conduct user tests.

Evaluators were not given a clear definition of a trust problem to increase the realism of the test situation. That is, a situation was created where a client feels that ill-defined user acceptance issues hinder website usage and transactions.

The trust-affecting factors had to be listed and rated on a scale adapted from Nielsen (1983), where: 1 = trust catastrophe, 2 = major trust problem, 3 = minor trust problem, 4 = cosmetic trust problem and 5 = not a trust problem at all. A template document was available as a download from that page, so that all participants would present their results in a uniform way, i.e. writing the trust problem or booster in the left column and giving it a rating from 1 to 5 in the right column. After they had written one report for each site, participants were required to complete a methods feedback questionnaire that is discussed in Section 6.6.4.

6.4.2 Data Analysis

Each problem (or positive remark) was noted on a small paper note (Post-it™), along with the participant number of the person reporting it and its severity score. Each note was also pre-classified according to what MoTEC component, if any, it would fall into. In a first round of data analysis, several evaluators would only appear on the same note if exactly the same problem was reported. In a second round, problems that referred to different aspects of the same interface or informational element were re-grouped so as to reduce the number of problem notes to a more manageable size. This had the added benefit of emphasising the salience of the problems. Also, problems reported only once that had a severity rating of 4 (cosmetic problem) were not retained for further analysis. The third step involved mapping the reported problems to the items in the checklist, so as to compare problems at the same level of granularity. To avoid a potential bias on the part of the experimenter, the mapping of predicted problems onto checklist items was done by an independent HCI specialist. She was also instructed to make a note of problems that did not fit any existing checklist cate-

gory. Dealing with reported problems in terms of the checklist items was an essential step to be able to compare the performances of the unguided and checklist-guided evaluators.

For each site, a table is provided that shows the number of trust problems per evaluator (A1 to A10), per MoTEC component (cf. Sections 6.9.1 and 6.10.1). The number in brackets after each component refers to the number of *different* problems found in that category. The totals by line therefore refer to the number of problem reports in that category, as one particular problem could have been reported by more than one evaluator. The table also lists the total number of different problems, the total number of problems reported by each evaluator, as well as the total number of problem reports.

6.4.3 Results of Study 1 for the Flower Website

Table 21 shows the number of trust problems found by each evaluator for the flower delivery website.

Table 21 – Study 1-Flowers: Number of problems found per participant

Components	Participants										Tot.
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	
Pre-Interactional Filters (0)	0	0	0	0	0	0	0	0	0	0	0
Branding (4)	1	0	1	2	2	2	1	2	1	2	14
Usability (6)	3	2	2	1	4	2	3	2	3	2	24
Company (4)	1	3	1	0	3	1	1	0	0	2	12
Products & Services (3)	2	0	1	1	3	1	0	1	0	2	11
Security (2)	1	1	0	0	1	0	0	0	0	1	4
Privacy (1)	0	1	0	0	0	0	0	0	0	0	1
Relationship Management (0)	0	0	0	0	0	0	0	0	0	0	0
Total (20)	8	7	5	4	13	6	5	5	4	9	66

Twenty different problems were reported, of which the majority fell into the Usability category. Evaluators especially noted the site's poor legibility, as well as the difficulty of finding relevant information. Selecting and editing items to be purchased also proved problematic. The second most populated category was Branding, with four problems. In general, the site's graphic design was considered to be amateurish and not in line with what people would expect from such as store. As far as the category Company is concerned, evaluators noted the lack of corporate information, as well as the lack of concrete details, such as photographs of the owners and the actual flower shop. Interestingly, some evaluators also suspected the company was being secretive about its business policies by hiding them in a hard-to-read section. In the Products & Services category, the main problems referred to the poor product descriptions and the little informative visual supports. The components referring to Risk, namely Security and Privacy counted two problems and one problem, respectively.

The average number of problems found by one evaluator was 6.6 ($SD = 2.8$). This

means that, on average, each evaluator found 33.0% of all problems reported in this study. In other words, since 20 different problems were found and that 66 problem reports took place, each problem was reported, on average, 3.3 times.

Table 22 shows the problems identified by 3 evaluators or more. The first column displays the number of evaluators that reported the problem, followed by the average severity rating in brackets.

Table 22 – Study 1-Flowers: Most frequently reported problems with their severity ratings

Frequency	Components	Problems
8 (2.38)	Branding	1. Amateur graphic design
7 (2.71)	Usability	2. Poor legibility of text
5 (2.00)	Products & Services	3. Photos of bouquets not informative enough
4 (2.25)	Company	4. Company hiding away essential transaction details
4 (2.50)	Usability	5. Need to input “none” if no message wanted
4 (2.50)	Usability	6. Poor feedback about transaction steps
3 (2.00)	Products & Services	7. Incomplete textual product information
3 (2.33)	Company	8. No corporate information
3 (2.33)	Products & Services	9. Poor integration of the different product categories
3 (3.00)	Usability	10. Inconsistent navigation design

As far as positive remarks are concerned, one participant remarked that the falling heart theme (cf. Figure 8) was an indication that the site was up to date as it was close to Valentine’s Day. In addition, the prominent placement of the phone number and credit card logos was also perceived as trust-inspiring. Another participant noted that the Cancer Charity logo also had a positive influence on his trust in that vendor. Only two evaluators out of 10 stated the information contained on the Privacy & Security page was clear and comforting.

6.4.5 Results of Study 1 for the Perfume Website

Table 23 shows the number of trust problems found by each evaluator for the discount perfume retail website.

On average, each evaluator uncovered 5.2 ($SD = 1.5$) problems in this condition, which amounts to 24.8% of the total problems found. In other words, each problem was reported 2.48 times on average. Out of the 21 different problems identified in this website, 6 referred to the Branding component, as evaluators found that the site’s cheap visual identity did not fit the traditionally glamorous perfume industry. Most problems in Usability referred to inconsistent menu design and inadequate feedback when selecting products.

It is noteworthy that three (non-British) evaluators identified the inability to convert British Pounds into other major currencies as a trust problem. In the problem mapping phase, the independent expert did not find a direct equivalent for the currency problem in the checklist, which points to a clear deficiency of the evaluation tool that

needs to be addressed. One should note, however, that currency issues are mentioned in the Usability component in the MoTEC model. The problem is that no checklist item was produced that mentioned them.

Table 23 – Study 1-Perfumes: Number of problems found per participant

Components	Participants										Tot.
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	
Pre-Interactional Filters (0)	0	0	0	0	0	0	0	0	0	0	0
Branding (6)	1	1	3	1	2	4	3	1	1	0	17
Usability (4)	2	0	0	1	2	1	1	2	2	1	12
Company (4)	1	0	1	1	0	0	0	0	0	1	4
Products & Services (2)	1	0	1	0	2	0	0	2	0	0	6
Security (4)	2	1	0	1	1	0	2	1	1	2	11
Privacy (1)	0	1	0	0	0	0	1	0	0	0	2
Relationship Management (0)	0	0	0	0	0	0	0	0	0	0	0
Total (21)	7	3	5	4	7	5	7	6	4	4	52

Regarding the Company component, contact information was not found to be prominent enough and one evaluator missed photographs of the shop's staff. As in the flower shop, poor product information and little informative photographs were major problems in Products & Services. In Security, it is interesting to note that four different shortcomings were observed, the most notable one being an image that reads "110% Security". This, by itself, was perceived to be suspicious and it was found even more inadequate as it was not linked in any way to a security policy. Regarding Privacy, two experts found that the link to the policy was not prominent enough.

The main results are summarised in Table 24, as it shows the problems reported by at least three expert evaluators. The numbers in brackets refer to the average severity rating for each problem.

Table 24 – Study 1-Perfumes: Most frequently reported problems with their severity ratings

Frequency	Components	Problems
8 (2.38)	Branding	1. Cheap-looking graphic design, does not go with perfume
5 (3.20)	Usability	2. Menu design is inconsistent and gives poor feedback
4 (2.25)	Security	3. "110% Security" sign nor prominent, nor informative
4 (2.75)	Usability	4. Poor feedback about transaction steps
3 (2.33)	Branding	5. Purpose of store not clear from the start
3 (2.33)	Branding	6. Page design does not scale nicely in smaller screens
3 (2.33)	Products & Services	7. Incomplete textual product information
3 (3.00)	Products & Services	8. Photos of perfumes are too small
3 (3.00)	Security	9. Unclear design and content of the delivery policy
3 (3.30)	Usability	10. Not possible to change currency (GBP) into EUR or USD

6.4.6 Study 1: Conclusions across Websites

Unguided evaluations of the Flower and Perfume websites clearly showed that most of the reported trust problems had to do with the sites' poor usability. Evaluators also identified several Branding-related problems in the two sites. To this point, we do not know whether the stress on Usability and Branding faithfully reflects real trust problems or whether it is a function of the evaluators' HCI training. It is also noteworthy that Company and Products & Services information was found to be more problematic in the Flower than in the Perfume website. Inversely, Security-related problems, though reported only once in the Flower website, were reported 11 times in the Perfume website. Privacy problems, however, were noted only once and twice, respectively. The tables also show that not a single problem pertaining to either the Pre-Interactional Filters or Relationship Management was reported in Study 1.

6.5 Study 2: Checklist-Guided Expert Evaluations

Study 2 consisted in having a matched pair of experts evaluate the same two websites as in Study 1, but, instead of relying on their experience or intuition alone, they were required to use the CheckTEC checklist presented in Chapter 4.

6.5.1 Procedure

The checklist-guided expert evaluations were only conducted once all the results from the unguided evaluations had been received. As in Study 1, these evaluators were given instructions remotely and no further interaction about the study or facilitation took place. The 10 participants in this condition were directed to another web page containing their instructions. They had the same scenario as the unguided experts, namely that they had been asked, as HCI professionals, to identify factors likely to damage trust in the websites. They had to do that by indicating to what extent the site complied with each checklist item. The web-based checklist contained trust heuristics in the left column, a pull-down menu to select compliance ratings in the middle and a text box for optional comments on the right (cf. Figure 10).

In case evaluators felt that one particular checklist item was not relevant to the website at hand, they had to give it the neutral rating of "acceptable" and write "not applicable" in the comment box. The checklist also contained a text box at the end asking evaluators to summarise their general impression about how trustworthy the site seemed to be. When submitting the checklist form, the results were automatically received by email.

When possible, please provide some comments to explain your rating. In particular, when you feel one attribute is not relevant, please rate it as "acceptable" and write "NOT APPLICABLE" in the comments box.

Website Attributes	Rating	Comments
1. Pre-interactional filters		
1.1. This industry which this company belongs to is reputable	SELECT: ▾	<input type="text"/>
1.2 This company is known from the offline world or from advertisements	SELECT: ▾	<input type="text"/>
2. Interface Properties		
2.1 Branding		
2.1.1 The purpose of the web site is clear from the start	SELECT: ▾	<input type="text"/>
2.1.2 The graphic design of the website is professional	SELECT: ▾	<input type="text"/>
2.1.3 The colour scheme and graphical elements are appropriate for this kind of website	SELECT: ▾	<input type="text"/>
2.1.4 The homepage incites users to explore the site further	SELECT: ▾	<input type="text"/>
2.1.5 The site pays attention to details, be they	SELECT: ▾	<input type="text"/>

Figure 10 – Screenshot of the online version of CheckTEC

6.5.2 Data Analysis

For each evaluator, the 54 compliance ratings were entered into a spreadsheet. To overcome problems related to the partial use of the compliance scale, a linear transformation was applied to the results prior to analysis. Such a transformation eases data analysis by giving each evaluator's overall average (across all questionnaire items) the same value and the same standard deviation. Guilford (1954) argues that such a transformation addresses the error of leniency, i.e. biased scores due to evaluators consistently rating a particular stimulus too low or too high. This helped us make checklist ratings more comparable across evaluators.

The next phase consisted in identifying checklist items that were not relevant for one particular website, so that these could be removed. This was done by looking at individual results and noting occurrences of the phrase "not applicable" in the comments box. Items that were clearly not relevant but that not necessarily every evaluator had marked as such were disregarded in the analysis. For example, results for Item 3.2.4 ("Sponsored content and advertisements are clearly labelled as such") were not taken into account if a site did not have any sponsored content or advertisements. This helped reduce noise data by focusing on relevant CheckTEC items.

For each evaluator, only relevant items with a score below 3.0 were considered to be problems. That is, items rated as cosmetic problems were not formally counted as problems, which makes the treatment of the checklist data more conservative than that of the unguided condition. The problem analysis per evaluator, rather than per group, was conducted to be consistent and compatible with the unguided condition of Study 1. That ensured that a proper comparison could be made at the end. Of course, overall

averages per checklist items were also computed, to give us an idea of those items considered to be trust problems by most, if not all, evaluators.

A graphical representation of the perceived performance of each of the six components is also presented. For each evaluator, average values of the six components were computed. A component with a score under 3.00 was counted as “problematic”, while a score above 3.00 was recorded as “satisfying”. Values of exactly 3.00 were recorded as half problematic-half satisfying. That allowed us to draw bar graphs emphasising the problematic-satisfying dichotomy by distributing what would have been “grey” scores equally into the “black” and “white” groups. Instead of computing averages per component across participants, we chose to present the results in terms of whether evaluators considered one component to be a problem or not. That is, for each of the six components, the number of “problematic” and “satisfying” scores were counted and later converted to proportional percentages (cf. Sections 6.5.3 and 6.5.4). The rationale for this presentation of the CheckTEC data is that it allows inferring a component’s general performance from the proportion of the population that found it to be problematic.

6.5.3 Results of Study 2 for the Flower Website

Two items that were not applicable were excluded: Item 3.1.4 about sponsored content and advertisement (there was no advertisement) and Item 3.3.6 about the return policy that makes little sense for a bunch of flowers. The remaining 52 items have all been rated as problematic by at least one evaluator, which translates into the total number of observed problems being 52. Table 25 shows the number of problems by means of the checklist.

Table 25 – Study 2-Flowers: Number of problems found per participant

Components	Participants										Tot.
	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	
Pre-Interactional Filters (2)	2	1	0	1	2	0	2	0	2	2	12
Branding (8)	5	4	3	4	2	3	3	1	6	5	36
Usability (11)	1	5	6	3	0	3	2	4	4	5	33
Company (8)	7	6	7	7	6	5	7	5	7	7	64
Products & Services (7)	2	2	2	2	1	0	3	1	5	3	21
Security (7)	5	5	7	5	5	5	6	1	5	5	49
Privacy (6)	2	4	4	1	2	2	2	4	5	2	28
Relationship Management (3)	2	3	3	2	3	2	2	1	3	1	22
Total (52)	26	30	32	25	21	20	27	17	37	30	265

The average number of problems found by one evaluator was 26.5 ($SD = 6.1$). Thus, on average, each evaluator uncovered 51.0 % of all trust problems found in this study. Compared with the unguided condition, it is striking that more than twice as many problems have been identified in this condition and that the average number of problems found per evaluator was approximately four times as high. In addition,

while the average evaluator found one third of the problems observed in the first study, about half of the problems were found in the second study. A more detailed comparison is presented in Section 6.6.

What is striking is that the unguided evaluators only focused on interface and informational elements. That is, they did not pay attention to Pre-interactional Filters and Relationship Management issues, both issues that were found to be problematic in the checklist condition. Besides, checklist-guided evaluators consistently found more problems in each of the remaining categories. What is particularly noteworthy is the number of problems noted in the categories Product & Services, Security and Privacy. This observation is in line with our argument that traditional HCI methods do not focus on the evaluation of perceived value and risk. The inclusion of these dimensions into the checklist thus nicely complements interface-related measures. Table 26 presents the concluding comments made by each of the 10 evaluators in this condition.

Table 26 – Study 2-Flowers: Concluding comments by checklist users

Evaluator	Concluding Comments	Trust Rating
B1	Would trust more with an up-to-date security seal	Low
B2	Trusted it at first, but after completing form noticed that it wasn't trustworthy at all.	Low
B3	Appears to be home-made, quite some info missing or obscure	Low
B4	Looks amateurish, not user-friendly, obscure procedures	Low
B5	Suspicious appearance, lack of SSL protocol	Low
B6	Not really trustworthy, but would try cheap item at first to see how good the service is	Neutral
B7	Seems trustworthy, warn customers that they might be deviations from orders	High
B8	Very trustworthy, thanks to security section, cc logos and company information	High
B9	Trustworthy thanks to third party seal and moderate low risk of flowers as products	High
B10	High trustworthiness	High

The results presented in this table are interesting as trust ratings seem to increase with the number given to each evaluator. Initially, the numbers assigned to evaluators (in each of the matched groups in Study 1 and 2) reflected the amount of experience they had with respect to conducting web usability evaluations. The table suggest that the more evaluators are familiar with the usability evaluation of websites, the more likely they are to be liberal about their trust ratings. Whether this pattern of results is a coincidence or whether it really points to a “personal experience” effect will need to be investigated in future research.

A graphical representation of the results was produced according to the method described in Section 6.5.2. The trust ratings are based on the concluding comments presented above. Figure 11 shows the performance of the individual components. This graph indicates that ease-of-use and company-related information were not considered to be real problems. Branding was found to be inappropriate to half of the evaluators. Poor information about the perfumes and about the company's security and privacy policies explains why more than half of the evaluators did not trust this website.

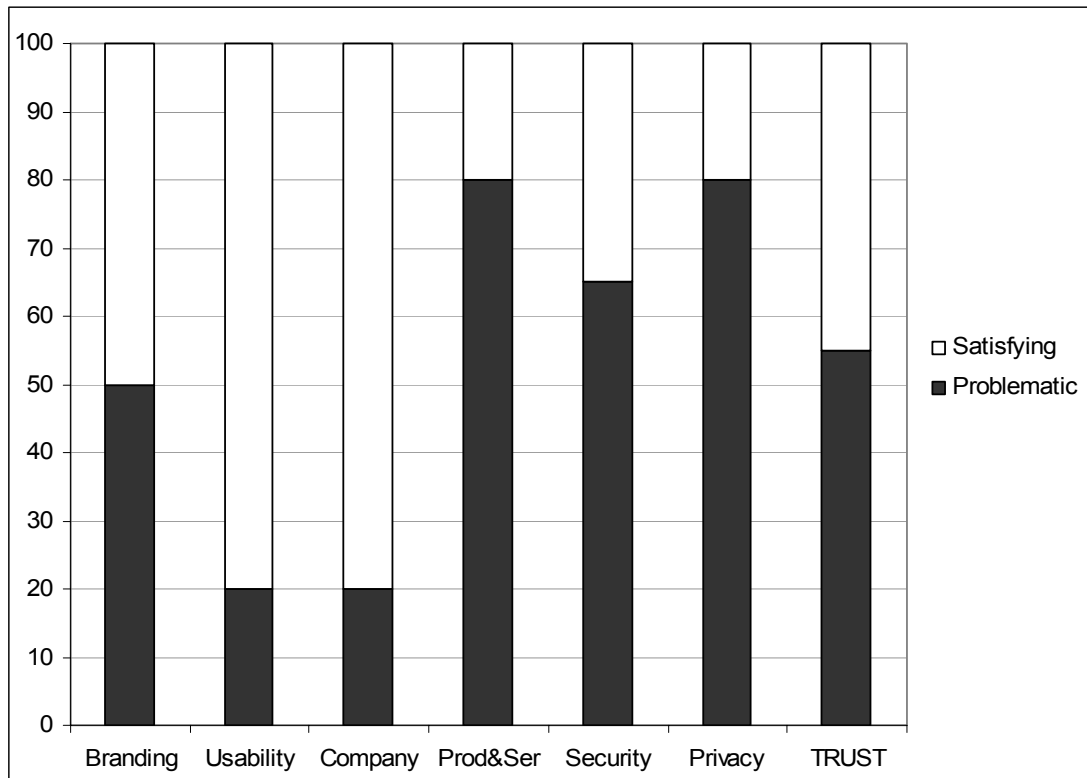


Figure 11 – Study 2-Flowers: Components and trust performances

6.5.4 Results of Study 2 for the Perfume Website

Table 27 shows the number and spread of problems found by the 10 checklist-guided evaluators in Study 2.

Table 27 – Study 2-Flowers: Number of problems found per participant

Components	Participants										Tot.
	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	
Pre-Interactional Filters (2)	2	1	1	1	2	2	1	0	1	1	12
Branding (8)	6	4	4	4	2	5	4	2	5	2	38
Usability (11)	4	2	3	2	3	2	6	0	0	4	26
Company (7)	5	4	2	6	6	6	6	7	4	5	51
Products & Services (7)	2	3	2	2	4	7	3	1	1	4	29
Security (8)	5	4	3	1	4	3	3	5	3	3	34
Privacy (6)	3	2	3	1	2	4	5	2	1	6	29
Relationship Management (3)	1	1	1	1	1	2	0	1	0	0	8
Total (52)	28	21	19	18	24	31	28	18	15	25	227

On average, one evaluator found 22.7 (*SD* = 5.3) problems, which amounts to 43.7 % of all trust problems identified in this study. Compared with the unguided condition

for this website, the average number of problems found is again approximately four times as high, while the proportion also increased from 0.25 to 0.40. A more thorough comparison will be made in Section 6.6.

As in the flower website example, it is noteworthy that a number of Pre-interactional Filters and Relationship Management problems have been noted in this condition. Regarding the Usability and Products & Services components, many more problems were uncovered using the checklist, namely, from 6 to 11 and from 3 to 7, respectively. Again, these results indicate that a number of problems related to the Security and Privacy components were comparatively ignored in Study 1.

Figure 12 offers a graphical representation of these results, based not on the number of problems by component but by the average rating a given component received.

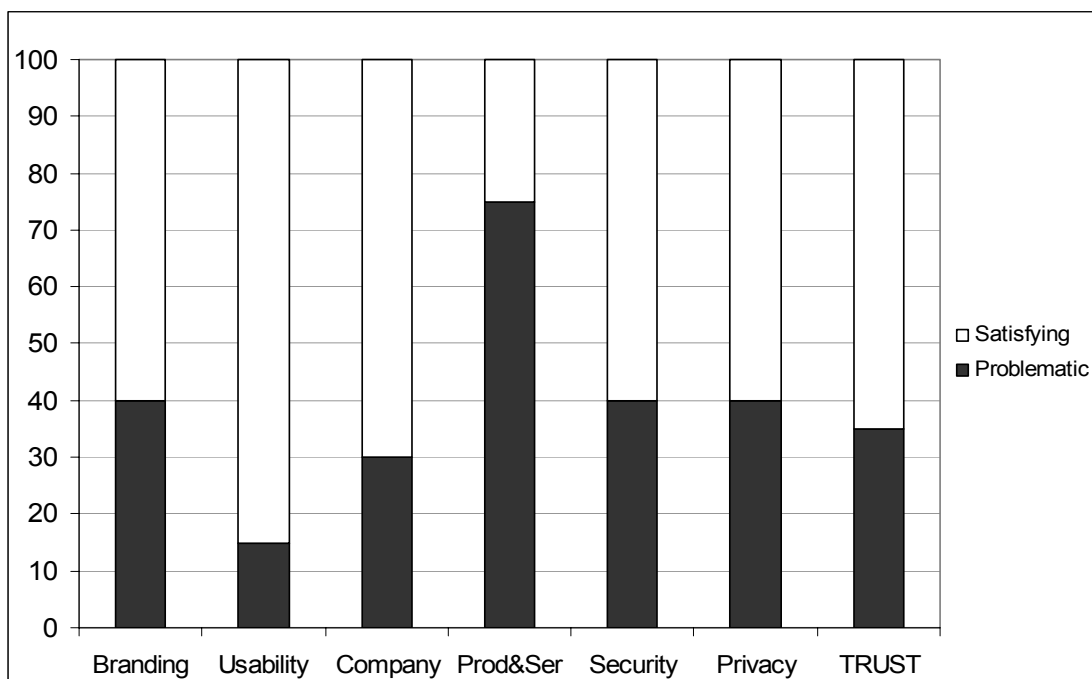


Figure 12 – Study 2-Perfumes: Components and trust performances

The most problematic area seems to be information about perfumes, in particular the short textual description provided and the small, or sometimes even inexistent, photos to illustrate a given fragrance. What the graph also clarifies is that, although 11 Usability problems were noted, only 15% of the evaluators rated the whole component as being problematic. Thus, these usability problems are more likely to have an annoying rather than a detrimental effect on the overall user experience. This distinction would have been lost without the graphical visualisation of the same results.

The graph also shows that 40% of the evaluators considered the graphic design of this website to be inappropriate or little professional. The same amount of evaluators considered Privacy and Security to be problematic. From their concluding comments, it appears that one third of the evaluators found this website not to come across as trustworthy.

6.6 Comparing the Results from Studies 1 and 2

This section summarises the main findings of the expert evaluation studies. The results for the two websites will be compared to investigate to what extent they can be generalised to other business-to-consumer e-commerce websites.

6.6.1 Number of Problems Found

Table 28 provides a summary of the number of problems found for each website in both the unguided and the checklist-guided conditions

Table 28 – Number of problems found by expert evaluations in Studies 1 and 2

Results	Flower Website			Perfume Website		
	Unguided	Checklist	<i>Difference</i>	Unguided	Checklist	<i>Difference</i>
Average number of problems/evaluator	6.6	26.5	+ 19.9	5.2	22.7	+ 17.5
Total number of problems found	20.0	52.0	+ 32.0	21.0	52.0	+ 31.0
Average proportion found/evaluator (%)	33.0	51.0	+ 18.0	24.8	43.7	+ 18.9

This table shows that there were surprisingly few differences between the two websites in terms of the number of problems found with either method. On average, unguided experts only found a quarter to a third of a small set of trust problems. Checklist-guided experts, on the other hand, found about half of a larger set of trust problems. Therefore, the results indicate that checklist-guided evaluations yield 4 times as many trust problems as unguided evaluations. This, in itself, is already a good argument in favour of using the checklist.

A statistical analysis was performed on the data collected in Studies 1 and 2 by comparing the row totals (i.e. the number of problem reports in a given MoTEC category) using paired t-tests. A paired samples t-test on the row totals of Study 1-Flower and Study 2-Flower confirmed that the differences between these two sets of problems were highly significant ($t = -4.386$, $p = 0.002$, 1-tailed). Likewise, the differences between the row totals of Study 1-Perfumes and Study 2-Perfumes were found to be even more highly significant ($t = -5.144$, $p < 0.001$, 1-tailed).

These results clearly confirm the first part of Hypothesis 1, viz. that checklist-guided evaluators find more trust problems than unguided evaluators. The second part of that hypothesis will be discussed in the next section.

To test the reliability of evaluators across the two websites, correlations were computed between the number of problems each one of them found in the first website with the number of problems found in the second. For evaluators A1 to A10 in Study 1, the Pearson correlation coefficient amounts to 0.345 and is not significant. For evaluators B1 to B10 in Study 2, the correlation coefficient is 0.284 and is not significant either. These results indicate that evaluators were not reliably better at finding problems in one site and in the other: Some found more problems in the Flower web-

site, others in the Perfume website. This, of course, does not warrant any conclusions regarding the validity, nor the severity of the problems found. It just points to the fact that that inter-evaluator reliability is very low, as previously observed by Nielsen (1989).

6.6.2 Problem Distribution

The second part of Hypothesis 1 referred to a greater variety of trust problems identified by the checklist-guided evaluators. Variety of problems, in this case, refers to the different MoTEC components which predicted problems correspond to. Table 29 shows the distribution of problems for both websites and for both evaluation methods. The proportion of problems is based on the number of different problems found per component in a given study, for a given website. These numbers are between brackets in Tables 21, 23, 25 and 27.

Table 29 – Proportion (%) of problems per MoTEC component

Components	Flower Website		Perfume Website	
	Unguided	Checklist	Unguided	Checklist
Pre-Interactional Filters	0.0	3.8	0.0	3.8
Branding	20.0	15.4	28.6	15.4
Usability	30.0	21.2	19.0	21.2
Company	20.0	15.4	19.0	13.5
Products & Services	15.0	13.5	9.5	13.5
Security	10.0	13.5	19.0	15.4
Privacy	5.0	11.5	4.8	11.5
Relationship Management	0.0	5.8	0.0	5.8
Total	100.0	100.0	100.0	100.0

What this table makes apparent is that one evaluation method finds a very similar distribution of problems in both websites. On the other hand, the differences between the unguided and the checklist conditions are striking. Indeed, the first observation to make is that the checklist contains items about Pre-interactional Filters and Relationship Management, which both proved problematic in those two websites. Secondly, despite instructions that urged them to look for trust problems, it is surprising that not more evaluators paid attention to risk-related factors, namely to Privacy and Security. Of course, at this point, we are not in a position to differentiate between the quasi absence of risk-related problems or just the fact that they were not noted. The user test study (Study 3) reported in Section 6.7 will show whether unguided evaluators have indeed overlooked these factors or whether there were no significant risk-related shortcomings.

The fact that the distribution of problems in the two checklist cases is almost identical can be accounted by the fact that the checklist contains a fixed number of items per

MoTEC component. Since, for each website, almost all items were rated as problematic by at least one evaluator, the structure of the checklist is reflected in the problem distribution. Indeed, had the checklist contained ten questions about Pre-interactional Filters and only four about Usability, the distribution of problems would have been skewed accordingly. It remains that the checklist forces evaluators to look at website attributes that are not normally part of a pure usability evaluation.

The results in Table 29 can also be considered in terms of variance over the eight categories (rows). Variance is lowest when each category has the same number of problems and highest when all the problems refer to one category. Indeed, if problems are distributed evenly (i.e. 12.5% per category), variance would be zero. If one category gets all the problems, variance would be 35.36. In our case, the standard deviations are as follows:

- Flowers: unguided = 10.69, checklist = 5.56
- Perfumes: unguided = 10.47, checklist = 5.56

The difference between the two conditions for each website amounts to a factor of about two. That means that the variance ratios are close to four, which indicates that the distribution pattern is clearly different for the unguided and checklist conditions.

The observation that the problems identified by unguided evaluators are less varied than those identified in the checklist study can be interpreted in two ways. First, it may be that unguided evaluators did not have a clear idea of what could constitute a trust problem for potential customers. That, in turn, shows that trust issues in e-commerce can only partially be addressed through existing HCI knowledge and that a validated body of knowledge would indeed be beneficial. In that respect, Table 29 demonstrates that the checklist helps uncover potential problems in areas which unguided evaluators would not spontaneously pay attention to.

The second interpretation refers to the fact that the problem distributions in the two unguided evaluations are quite distinct but that they are almost identical in the checklist evaluations. One could interpret the more salient problem pattern in the unguided evaluations as a greater sensitivity to local, i.e. website-related, problems. Accordingly, checklist-guided evaluators might be so focused on checking the site for compliance with *general* trust heuristics that they might be less sensitive to those trust issues that are most relevant to one specific merchant, in one specific industry.

In conclusion, our findings also confirm the second part of Hypothesis 1, viz., that checklist-guided evaluators find more varied trust problems than unguided evaluators. However, the problem distribution patterns in Table 29 might point to checklist evaluations as being less sensitive to a website's specific trust problems.

6.6.3 Time & Satisfaction

Based on the methods feedback results, the average time for completing the evaluation of one website in Study 1 was 68 minutes ($SD = 31$). The checklist feedback results indicate that the average time for evaluating one site in Study 2 was 35 minutes ($SD = 15$). This means that evaluators B1 to B10 found almost four times as many

problems as evaluators A1 to A10 in almost half the time. In other words, the checklist is eight times more efficient in its ability to identify trust problem than solely relying on existing HCI knowledge, intuition and experience.

One should also note the large standard deviation for completion time in the unguided condition, which points to large individual differences in conducting a trust evaluation. It is interesting to see that, in both conditions, Pearson Correlation tests revealed no significant correlations between the amount of time spent evaluating one site and the number of reported trust problems (Unguided: $r = 0.234$; Checklist: $r = 0.355$). This would suggest that experts are just quicker to identify the same problems than less experienced evaluators.

In both conditions, evaluators were asked to evaluate their performance after they had completed their tasks. The question they had to answer was “How well, in your opinion, did you cover trust issues?”. This was to get a subjective report of their satisfaction with their performance in their particular condition. Answers had to be given on a 5-point scale, from “not well at all” to “very well”. Figure 12 clearly shows the difference between the two conditions. Half of the unguided evaluators reported that they saw their performance as average or below average. In the case of the checklist-guided experts, all of them reported their performance as either average or above average. A Wilcoxon test on the subjective performance data revealed that the difference between the two conditions is highly significant ($Z = -2.530$, $p = 0.006$, 1-tailed). These results indicate that a checklist-based evaluation increases subjective performance and satisfaction.

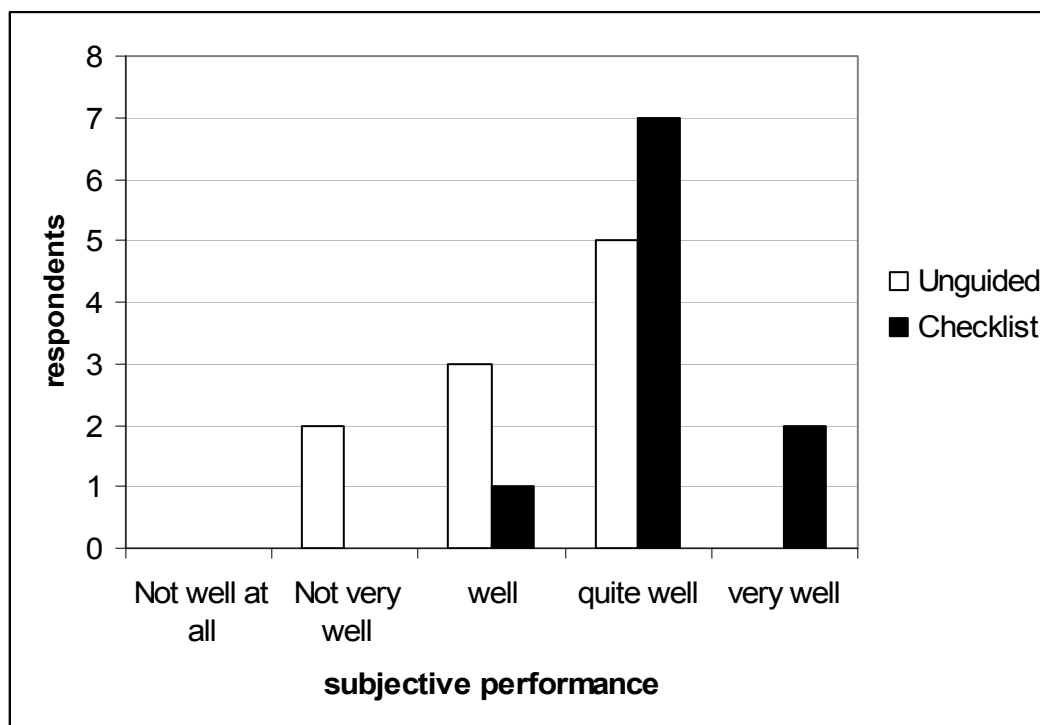


Figure 13 – Reported subjective performance in unguided and checklist studies

6.6.4 Study 1: Methods Feedback

Apart from time and subjective performance, the post-evaluation questionnaire also asked unguided evaluators to report what methods they used to perform their evaluations and what methodological tool would have helped them.

With regards to the tips and tricks participants used to find trust problems in the two websites, four visited the site like an average user and noted down their experiences. Four others reported that they based their evaluations primarily on their own experience with online shopping. One of them indicated that he also used Nielsen's (1993) heuristics as a complementary method, while the remaining two participants reported it as their primary evaluation method. This clearly shows the different approaches HCI practitioners without an explicit trust tool would adopt to diagnose trust problems.

The last question asked participants to indicate what knowledge or tool would have helped them evaluate the trust performance of the two websites (more than one answer per person was possible). Three participants stated that trust heuristics or guidelines would have been particularly helpful. Two people said they would need knowledge about people's perception of risk, while knowledge about criminal cases and a reliability assessment of online payment systems were each mentioned once. Interestingly enough, one participant answered that statistics of website usage, such as a user's click stream throughout a site, would have been useful information. On the methodological side, two people said that user tests would have helped their trust evaluation, whereas another person stated she would have benefited from a second expert opinion. Lastly, one participant also mentioned that a trust questionnaire would have been useful to get trust-specific information directly from users. This survey shows that models like the MoTEC model and provisions like the Trust Toolbox may well meet the needs of HCI practitioners.

6.6.5 Study 2: CheckTEC Feedback

Similarly, all 10 checklist-guided evaluators were also asked to complete a feedback form asking them about items they would add, change or remove from the current checklist.

Regarding possible additions, 3 participants said they would like to see a more explicit question about the first impression made by a website. Since this aspect was covered in Item 2.1.4, it is suggested that it should be rephrased to contain the phrase "first impression". Two people also missed a question relating to a users' situational awareness within a website, i.e. whether they knew where in the site they were. Related to that, one evaluator mentioned the need for a question about users' feeling of control over the site, especially over the payment process. A more general point was the request to provide better explanations of the structure of the checklist and the scales it contains.

As far as changes are concerned, one evaluator argued that the rating scale might not always be appropriate, as most of the questions could be answered by "acceptable" and some questions might just have "yes/no" answers. That person would also like to get an automatically-computed overall score for the website, based on her input. Other

participants claimed that the checklist items were not always phrased very well. For example, the item “The colour scheme and graphical elements are appropriate for this kind of website” does not always have one clear answer. Indeed, the colours can be appropriate for a certain website but the site may look bad nevertheless. Also, some questions were found to be too suggestive, such as “The site contains information about the company’s legal status, e.g. registration with a Chamber of Commerce”. One particular evaluator reports to have looked for these registration details while neglecting other related pieces of information.

With regards to removals, one evaluator said that a lot of questions within a scale could in fact be combined. This would have the benefit of making the checklist shorter and therefore more usable. Two participants suggested that the questions about browser compatibility and broken hyperlinks should be removed, as it is unlikely that evaluators can check these issues thoroughly. Rather, one could ask whether such problems were encountered.

6.7 Studies 3 & 4: Trust Problems Reported by Users

Studies 1 and 2 rank as predictive evaluations, based on expert reviews. Studies 3 and 4 count as empirical evaluations, based on data collected from a representative sample of the target user population. However, one should stress that even empirical evaluations only produce a list of predicted problems. Indeed, since data stems from a small sample of the target population, one can never claim to have identified all existing problems. In addition, usage of the websites is simulated rather than occurring in a real shopping situation. For the sake of the argument, we will refer to user-reported problems as *observed* trust problems, as opposed to *predicted* problems.

Study 3 consisted of user tests (Study 3a), followed by the administration of the QuoTEC questionnaire (Study 3b). Study 4, consisted of the remote administration of the same questionnaire to a matched group of users without user tests beforehand. Study 3 allows us to check to what extent the trust problems predicted by the expert evaluators in Studies 1 and 2 coincide with the problems reported by actual users. We will first deal with Hypothesis 2 (below), as the set-up to test Hypothesis 3 is reported separately in Section 6.10.

Hypothesis 2: The correspondence between predicted and observed problems is greater for the checklist-guided predictions (Study 2) than the unguided predictions (Study 1).

6.7.1 Experimental Design

As with the expert evaluations, the design for this study was a between-subjects design with two matched groups. The independent variable was the method of gathering users’ trust-related feedback. It had two levels: user tests followed by the questionnaire or questionnaire only. The dependent variable was the number and nature of the reported trust problems. Of course, both the number and nature of problems can be established in the user tests, while only the nature of problems can be made apparent in the questionnaire studies. This is because average scores are computed per component, not taking into account concrete problems. Thus, the level of abstraction of iden-

tified problems is higher in the case of questionnaire results.

The same websites were used in the user tests and questionnaire studies as in the expert evaluation studies (Studies 1 and 2). The order in which the websites had to be evaluated was counterbalanced to avoid order effects.

6.7.2 Participants

The participants had to be representative of online shoppers who might be interested in buying from the websites we selected. They were recruited through a posting to the mailing list of the university's international student network. Since the tests were conducted in English, this recruitment strategy ensured that only fluent English speakers responded. In total, 35 people participated in the user feedback studies. Eighteen participants (C1 to C18), 6 of whom were female, took part in the user test and questionnaire study. Seventeen participants (D1 to D17), 9 of whom were female, took part in the questionnaire-only study. They ranged between 24 and 32 years of age. Each participant received a financial reward for their participation.

6.8 Study 3: User Tests

6.8.1 Procedure

After a short introduction, participants were given a written document containing an introduction and instructions to the test session. The introduction set the scene by providing a balanced account of the advantages and the risks inherent to shopping online. It concluded with a sentence stating that, before engaging in a transaction with an online vendor, it was important to determine whether that particular vendor was indeed trustworthy. The aim of the study was to find out what would make users trust or distrust the two selected websites.

The set-up was the same for the two sites. First, participants had to comment on the first impression made by the site. After that, they had about 10 minutes to explore the site while "thinking aloud", i.e. commenting about what they were doing and why. Once they had finished the initial exploration phase, they were asked to carry out one realistic task per site. The thinking aloud method used in these tests was interactive to the extent that the test moderator could intervene to ask participants to clarify their actions. In that respect, it was more akin to that put forward by Dumas and Reddish (1994), for reasons discussed in Boren and Ramey (2000). The interactive think aloud protocol was also reported to be more representative of actual practice than strict verbal protocols (Ericsson & Simon, 1984).

Participants were told to go as far in the ordering process as possible in order to see exactly what stages they had to go through if they wanted to place an order. The maximum time allowed to visit one site was 30 minutes so as to keep it within a reasonably realistic time frame. After they had completed their task on a specific website, participants had to fill in an online version of the revised QuoTEC questionnaire presented in Chapter 5. Questionnaire results correspond to Study 3b and are treated in Section 6.11.

It is important to point out that, for practical reasons, the author acted as the test facilitator in this study. A potential risk of this set-up is that observations might have been affected by the author's knowledge of trust-shaping factors in general and by the results of the expert evaluations in particular. However, to minimise this risk, users were asked to follow the set scenario by thinking aloud and no additional questions were asked, unless they were necessary to understand a particular behaviour.

6.8.2 Data Analysis for Study 3a

Data gathering consisted of noting observed trust problems using a specially designed form. This form contained 6 main sections, one for each model component, as well an additional section for more general remarks. This allowed to pre-classify user comments into appropriate categories, thereby facilitating the analysis of observed trust problems. Appendix 3a shows the raw data for the Flower website and Appendix 3b shows the raw data for the Perfume website.

6.8.3 Results of Study 3a for the Flower Website

The third study about this website consisted of a series of user tests with 18 participants. The scenario they were given was as follows:

“Imagine you're browsing the web to find a company that delivers flowers in the UK. The reason is that you'd like to send flowers to a friend on her birthday, April 16th. Since she works full-time, it is important that the flowers should be delivered after 6pm. After browsing for a while, you end up on the allFLOWERS website. Would you be comfortable ordering your flowers from this site?”

All comments made by users referred to Interface Properties and Informational Content. Regarding Branding, the most frequent remarks had to do with the graphic design found to be very amateurish in general and rather annoying due to the falling hearts feature. An out-of-date reference to Mother's Day, as well as spelling errors were also reported. From the numerous usability problems found, the most serious was the impossibility to enter the delivery time on the same screen as the delivery day. Because of this, most of the participants assumed it was not possible to specify the delivery time at all, which prevented them from completing the set scenario. In fact, it was possible to specify the time of delivery on a later screen. Also, most people expected to see the basket overview page after selecting a product and to be able to checkout from that overview page. However, both functions were not implemented in this way, which made the shopping process little intuitive. These observations are quite different from the rather satisfying rating the Usability component received in the checklist study.

In addition, company information was found to be too hidden and too superficial. For example, several participants would have preferred to be given the name of a contact person and to see photographs of the florists and their shop. Three participants also found the customer testimonials to be biased and unreliable. Regarding the flower arrangements, several people remarked that their description was incomplete. Information about the flowers' origin or the kind of wrapping it would come in would have helped assess the real value of the offering. Two participants also assumed that the

picture that was shown for each arrangement represented the most expensive type of bouquet in that category, which was also found to be misleading. Interestingly, four people remarked that they missed the possibility to create or to specify their own floral arrangement. This is a problem that neither the unguided, nor the checklist-guided evaluators had predicted.

The site's security policy was generally found to be too long and not clear enough. That customers have to fill in their personal details on an unencrypted page of the flower site and their payment details on a third-party website was also found to affect the amount of control they have on the transaction. It is interesting that only 2 out of 18 participants clicked on the Shopsafe seal displayed on the homepage. This turned out to be a real disappointment as the Shopsafe site looked more like a directory of online shops, rather than an organisation genuinely committed to security. Even more worryingly, only one participant clicked on the Which WebTrader seal, only to see that the service had been discontinued and that the seal did not have any value any more.

A systematic comparison of these results with those predicted in Studies 1 and 2 will be made in Section 6.9.

6.8.4 Results of Study 3a for the Perfume Website

The same 18 participants were given the following scenario for the Perfume Website:

“Imagine you're browsing the web to find a company that sells perfumes and cosmetics. The reason is that you're looking for two birthday presents for your friend. She told you that she would be interested in a perfume by Estée Lauder called *Pleasures* and in a moisturizing cream by Lancôme called *Primordiale Intense Nuit*. After browsing for a while, you end up on the FragranceBay website. Would you be comfortable ordering these products from this site?”

Fourteen participants out of 18 noted the site's graphic design to be very cluttered and rather amateurish. For example, the site's logo was perceived to be a banner advertisement by 5 people. Regarding the site's ease-of-use, the majority of participants had problems finding specific items, be it using the navigation menus or using the search engine. The search results page also appeared to be ill-designed as it did not make apparent differences between products from the same brand. In addition, participants were surprised to have to specify the number of items every time they added one to their basket. The default number, they claimed, should be “1”.

It is interesting that the majority of people tested mentioned that there was not a lot of textual and visual information about the different fragrances, while at the same time claiming that it did not really matter as they would only buy online perfumes that they knew from the offline world. Two participants also expected an automatic fragrance advisor, based on a customer's preference profiles. Such an added-value service would certainly help to emulate offline advice and possibly increase the perceived competence of the vendor. Although task-driven participants paid little or no attention to company information, others said they wanted to see an “About Us” section with more detailed information about the company.

As mentioned above, the most significant shortcoming of this site in terms of security was the “110% security” logo on the homepage. Such a claim was dismissed as little credible by half of the participants. The fact that this logo was not clickable and did not provide access to a security policy was seen as an additional problem. The next section will present a comparison between these results and those predicted in Studies 1 and 2.

6.9 Comparing Unguided, Checklist & User Tests Results

Up to this point, we have the results from three different evaluation studies of the same website. How do the results compare? To answer this question, we have first mapped problems identified in the user tests onto appropriate CheckTEc items, whenever possible. The second step involved noting how many test participants has encountered each problem. The frequency a problem occurred was expressed as a percentage so as to be more directly comparable with the predicted problem frequencies for Studies 1 and 2. Appendix 4 shows the predicted and observed problem frequencies for both websites and classified per CheckTEC item.

6.9.1 Comparisons for the Flower Website

Figure 14 shows a Venn diagram summarising the repartition of problems found in the three studies. The surface area of each circle is proportional to the number of problems it contains. Such a presentation allows for a better visualisation of the differences in number of problems identified by each evaluation method.

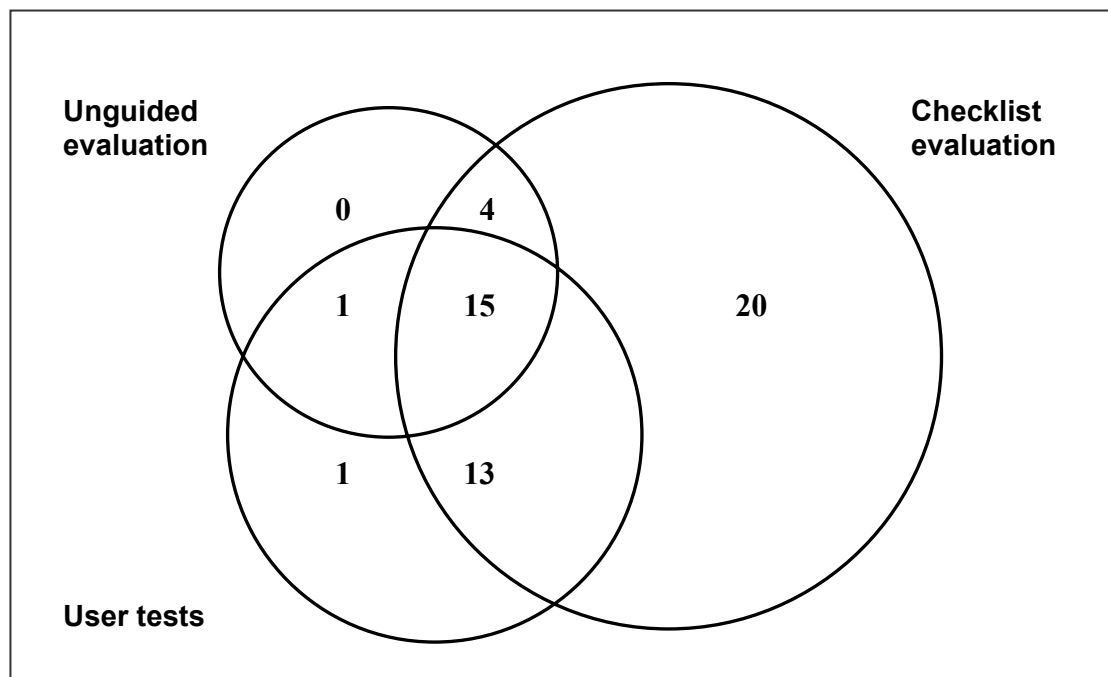


Figure 14 – Number of problems found by the different methods (Flowers)

Thirty “real” problems were observed in the user tests. Out of these problems, 28 (= 93.33%) were correctly predicted by the checklist-guided experts. It is noteworthy that only 16 (= 53.33%) of the problems had been predicted by the unguided experts. Thus, the real contribution of the checklist can be spotted straightaway and consists of 13 additional problems found. This diagram also shows that the unguided experts correctly predicted one problem that the checklist users did not predict: that the overly positive customer testimonials would have a negative effect on their credibility in people’s eyes. There is also one problem that none of the expert groups had predicted: that the lack of floral arrangement customisation was perceived as a problem in the user tests.

What is also striking is the number of apparent false alarms, i.e. factors noted as problems in the checklist condition that were not mentioned in the user tests, nor predicted by the unguided evaluators. Three possible interpretations can be given to account for this observation. Firstly, it could be that these 20 predicted problems only have a marginal influence on people’s feeling of trust and that they should be disregarded. The second is that several checklist items become not applicable if one particular item is not present in the site. For example, if a site does not have a privacy policy, most items in the Privacy scale will be rated as problems, although the lack of such a policy is perceived as only one problem by test participants. Thirdly, another explanation is that most of these issues do have an important role in a real situation where people are genuinely deciding whether they should engage in a transaction. That is, the artificiality of the experimental situation might have suppressed or decreased the influence of a certain number of factors, notably those related to security and privacy. One should also remember that user tests are conducted with a small sample of the target population and that its results, strictly speaking, should also be treated as predictions. In other words, that a predicted problem has not been observed in these user tests does not entail that it would not be a problem to another person. By extension, the main problem of such comparative studies is that one key number will always be missing: the total number of real trust problems.

If we had that number, it would be easier to frame results in terms of detection theory: either a problem is a trust problem or it is not and it is either found or not found. Even without that number, we can still try to interpret our findings in terms of this framework. If people have a low sensitivity to trust problems, the difference between the means of the two distributions (those of trust problems and non-trust problems) is small. That is, people might either miss a trust problem or consider that it is not severe enough to be counted as a problem. However, if one sensitises these people for trust problems, which we did by means of the checklist, the distribution of problems across categories is bound to change in two ways. First, because of a greater awareness of what can constitute a trust problem, the number of found trust problems is likely to increase. Second, the criterion that defines a trust problem can either remain constant or shift to a more lax position. If the criterion shifts towards non-trust problems, even more problems will populate the trust problem category, running the risk that it will also contain some non-trust problems. At this point, one cannot know whether the checklist evaluators are indeed producing false alarms or whether their sensitivity has been increased in such a way that they just identify more of the real trust problems, whereas the sensitivity of the unguided users is still too low to distinguish them from

non-trust problems. Thus, the apparent false alarms might, in fact, be valid trust problems.

6.9.2 Comparisons for the Perfume Website

Figure 15 shows a Venn diagram showing the number of problems uncovered by the different evaluation methods. Again, the surface area of each circle is proportional to the number of problems it contains.

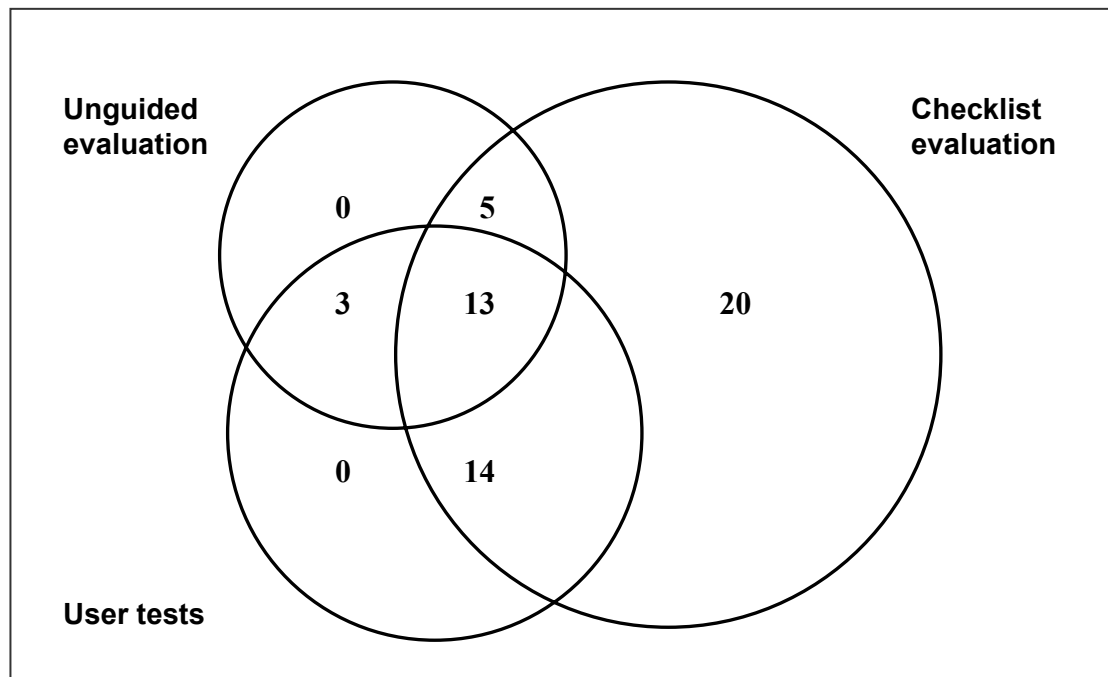


Figure 15 – Number of problems found by the different methods (Perfumes)

These results indicate that 30 different problems were observed in the user tests. Out of these 30, 27 (= 90%) were correctly predicted by the checklist-guided evaluators. The unguided evaluators only predicted 16 (= 53.33%) problems correctly. However, they did predict three problems that were not picked up by the checklist users, namely: (1) the lack of currency conversion, (2) the lack of customer testimonials and (3) the incompleteness of company contact details. The last point is interesting as it is clearly part of the checklist (Item 3.2.1) but it was not rated as problematic by any evaluator. This shows that a scenario-based user test helps investigate to what extent information that is present in a site is easily accessible when users need it.

While the concrete benefit of using the checklist consists of 14 additional problems being uncovered, an important number of apparent false alarms (20) must also be noted. The reasons for these predicted but not observed problems are the same as in the Flower website (cf. Section 6.9.1).

6.9.3 Comparisons across Studies

Table 30 shows the overlap between the problems predicted by the two evaluation

methods and the problems observed in the user tests.

It is interesting to note that the same number of problems was reported by participants in the two user tests. This, of course, is a coincidence, as is the almost identical proportion of problems correctly predicted by either method. It remains that checklist users correctly predicted 90%, or more, of the problems users encountered in the tests. Of course, one should bear in mind the potential false alarm issue discussed above. How to address this problem in future website evaluations will be discussed in the next chapter.

Table 30 – Number of & overlap between predicted and observed problems

Results	Flower Website			Perfume Website		
	Unguided	Checklist	Difference	Unguided	Checklist	Difference
Average number of problems/evaluator	16.00	28.00	+ 12.00	16.00	27.00	+ 11.00
Total number of problems found	30.00	30.00	N/A	30.00	30.00	N/A
Average proportion found/evaluator (%)	53.33	93.33	+ 40.00	53.33	90.00	+ 36.66

Appendix 4 shows how many problems were found by each evaluation method, per MoTEC component and for the two websites. An additional test to investigate the correspondence between predicted and observed problems consists in noting the number of problems found by each method, by component, across websites and exploring correlations between unguided, checklist and user tests results.

Table 31 shows the correlations between the outputs of the three evaluation methods. Please note that for this analysis, only the six main MoTEC components were taken into account (i.e. Pre-interactional Filters and Relationship Management were not included).

Table 31 – Correlation coefficients for unguided, checklist and user tests with significances (N=12)

		Checklist	User tests
Unguided	Pearson Correlation	0.882	0.632
	Sig. (2-tailed)	< 0.001	0.028
Checklist	Pearson Correlation	1	0.810
	Sig. (2-tailed)	.	0.001

The Pearson correlation test revealed that all relationships are highly correlated. For instance, results for the unguided and checklist-guided studies turn out to be the most strongly related ($r = 0.882$, $p < 0.001$, 2-tailed). To explore which of these two methods produce more accurate predictions of observed problems, we can see which data set the user tests results correlate better with. The last column shows that user tests results correlate both with unguided predictions ($r = 0.632$, $p = 0.028$, 2-tailed) and

checklist predictions ($r = 0.810$, $p = 0.001$, 2-tailed). Thus, we can conclude that both expert evaluation methods produce a number of problems by component that highly correlates with the findings from user tests. Since the r value is higher for the checklist condition, one can infer that the relationship between checklist and user tests results is stronger than for the unguided results. The results thus suggest that using the checklist to predict trust problems is likely to produce results that correspond more accurately to results from user tests, compared with predictions from unguided evaluators.

In conclusion, the findings confirm Hypothesis 2, viz. that the correspondence between predicted and observed problems is greater for checklist users than for unguided evaluators.

6.10 Testing Hypothesis 3: Questionnaire Studies

The third hypothesis was as follows:

Hypothesis 3: There is no difference between the questionnaire results in the user tests followed by the questionnaire condition (Study 3b) and the questionnaire without facilitator condition (Study 4).

The results of Study 3b are presented in the next section, while the procedure and results of Study 4 are presented in Section 6.12.

6.11 Study 3b: Questionnaire after User Tests

Since this study deals only with one additional aspect of Study 3, viz., the administration of the questionnaire after the tests, the main experimental set-up for this study will not be repeated here.

One should not that we will only deal with the quantitative results produced by the questionnaire in this section. The results from the online questionnaires were received by e-mail and transferred to a spreadsheet. For each participant, the average value for each of the six MoTEC components was computed automatically. This allowed us to present the results using the same kind of stacked bars graph as in Study 2.

6.11.1 Results of Study 3b for the Flower Website

Figure 16 shows the performance of the six model components, black areas indicating the presence of problems. The graph indicates a consistently low performance of the Flower website across the different components. In general, 40% of the participants would not trust this site, notably because of poor graphic design and ease-of-use, as well as privacy concerns. When compared to the graph produced on the basis of the checklist-guided evaluations, it appears that “pretend” customers of this website are much more forgiving or accepting than HCI evaluators. Indeed, for four out of the six components, the proportion of problems noted by expert ratings was considerably higher. Interestingly, experts overrated the usability of the website, as well as the completeness of company-related information. That the Usability component was

rated lower by test participants may be accounted for by the structured set of tasks they had to complete on the website. That is, they were forced to pay attention to transaction-critical factors, such as delivery day and time.

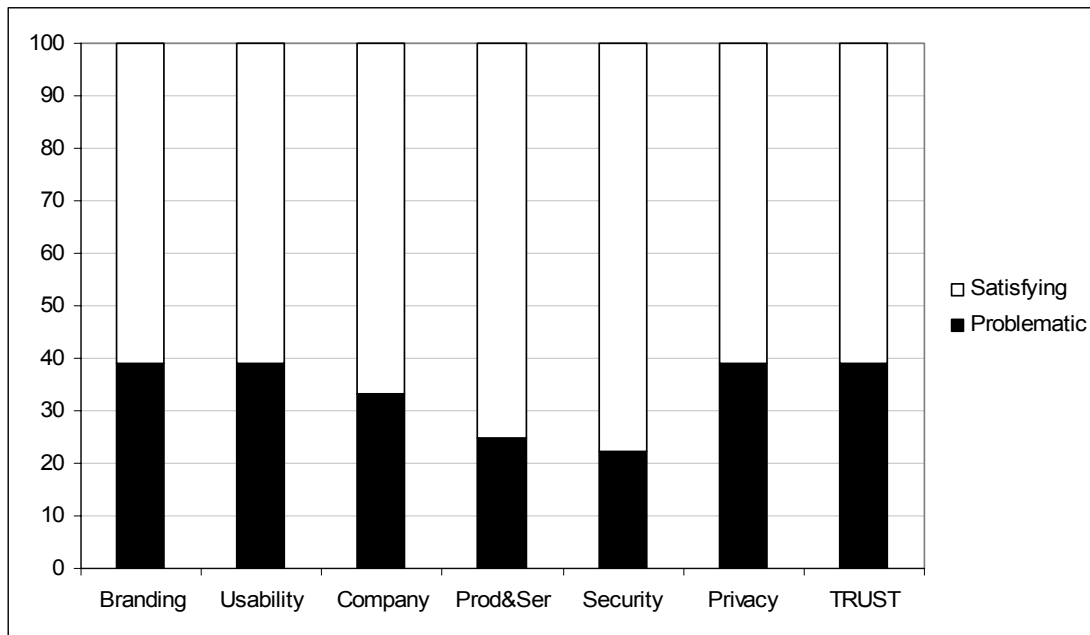


Figure 16 – Study 3b-Flowers: Components and trust performances

6.11.2 Results of Study 3b for the Perfume Website

Figure 17 shows the performance of the different component and the overall trust performance for the perfume website.

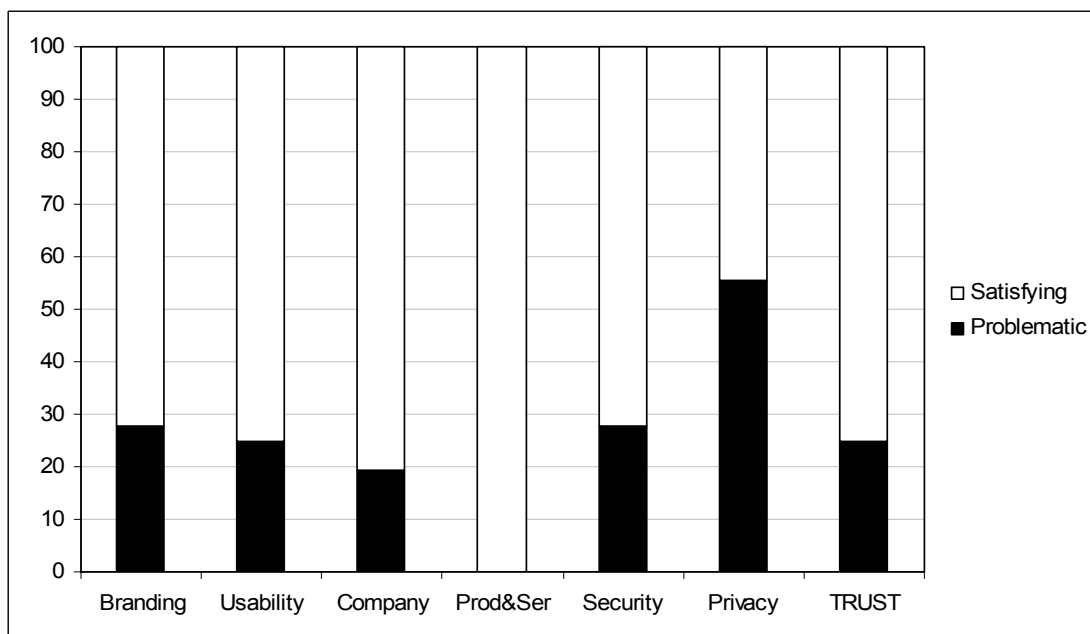


Figure 17 – Study 3b-Perfumes: Components and trust performances

Compared with the graph shown in Section 6.5.4, the most striking difference is that the component that was rated highest by the checklist-guided evaluators is the component rated lowest by actual users. That component is Products & Services, which scored 100% satisfying in the user tests. This can be explained by the fact that, indeed, product information and visuals were scarce but that it did not really matter in this case, as test participants knew what they had to buy (it was specified in the scenario). Also, they claimed that they would only buy perfumes online that they already knew offline, hence the limited need for extensive descriptions. It is likely that a customer that would browse the same website without knowing exactly what product to purchase would react quite differently to this.

Apart from the Privacy component that was more problematic in the tests than what was predicted by checklist users, all components received better ratings by the test participants. This shows, once again, that using the checklist produces more conservative results.

6.12 Study 4: Questionnaire without Facilitator

6.12.1 Procedure

Interested participants had to sign up for the study on a web page. Each participant was then emailed an individual participant number, as well as a hyperlink to a page containing instructions and the scenarios for both sites. Two versions of the instructions page were created to counterbalance the order in which the sites had to be evaluated. The introduction and the scenarios were exactly the same as for the participants of the user tests.

The only difference consisted of the fact that participants carried out their evaluations in their own time and place, without the presence of a facilitator. After each website evaluation, participants had to fill in a questionnaire described in Section 6.8.2. After both evaluations had been completed, they had to submit another online form stating their preferred payment modality for receiving their participation incentive (cash, voucher or bank transfer).

6.12.2 Data Analysis

Since all contact with this group of participants was mediated electronically, the data was gathered using a slightly extended version of the questionnaire used in the previous study. The first part contained exactly the same questions, in the same format and order. The only difference is that the standard questionnaire was complemented by four free-text questions at the end, namely:

- What did you like about this website?
- What did you dislike about this website?
- Would you feel comfortable ordering from this website? Why?
- What could increase your trust in this website?

The rationale for including these four questions was that it would force participants to

reflect on their general trust experience with each website. Data analysis for the standard questionnaire was conducted the same way as in Study 3b, with the same kind of stacked bar graphs as output. The answers to the free text questions were not included in the analysis, as they only fulfilled a “reflective” function in this case.

6.12.3 Results of Study 4 for the Flower Website

Twelve participants were given the same scenario as in Study 3 and were asked to complete the tasks on their own. After that, they had to fill in the same questionnaire. Figure 18 shows the results of the questionnaire-only study for the flower shop.

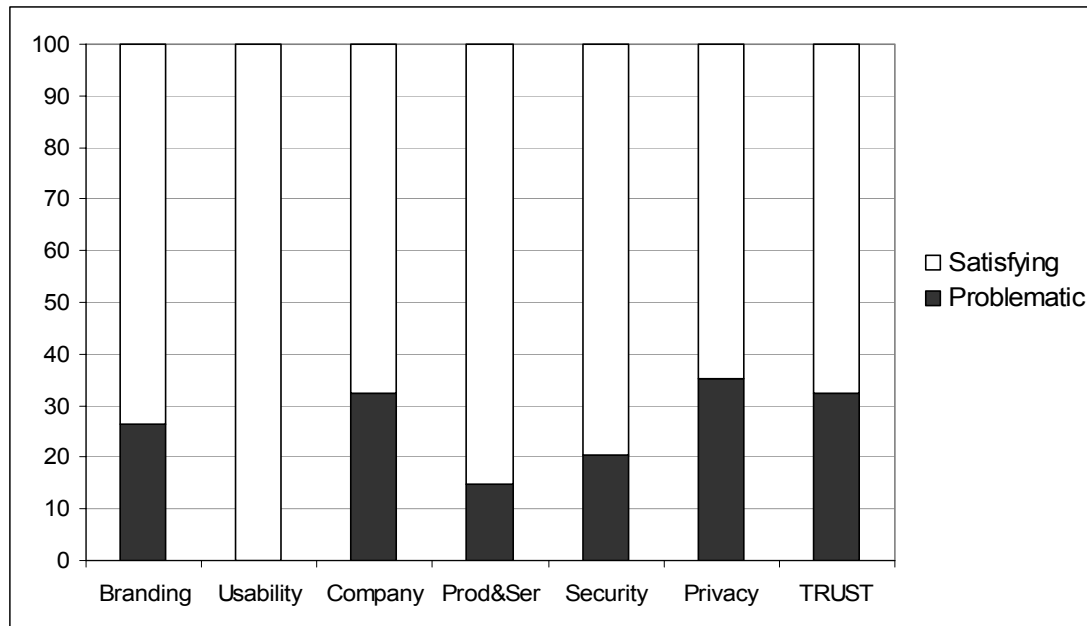


Figure 18 – Study 4-Flowers: Components and trust performances

Compared with the results of the questionnaire after user tests study, two main observations can be made. The first one is the surprising total lack of usability problems noted by this group of participants, while the user tests showed that 40% reported usability as problematic. This result, although it is even more extreme than the one from the checklist-guided experts, can nevertheless be given the same explanation. It is very likely that respondents in this condition invested as little time as possible in following the scenario and completing the questionnaire, as they would receive their financial incentive whatever effort they would put into it.

A quantitative comparison between the questionnaire results of Study 3b and Study 4 was made on the basis of correlations. It is not surprising that, when all components are taken into account, the correlation between the two questionnaire conditions is rather low ($r = 0.197$, not significant). However, when the Usability component is removed, the correlation coefficient increases to 0.819 (significant at the 0.05 level, 1-tailed). This confirms our view that their unsupervised test session led participants to have only a superficial look at the site and to rate usability high as a consequence.

The second observation is that the pattern of the remaining components and the trust measure are remarkably similar, with values fluctuating between 20 and 40%. What

can be concluded from this comparison is that QuoTEC can successfully be used in a questionnaire-only situation to analyse graphic design, competence- and risk-related information. However, in order to get a good indication of a site's usability, it is recommended to conduct one-to-one user tests. An alternative and more cost-effective approach would be a remote usability evaluation set-up where people are required to enter answers to set questions in a software application before being allowed to proceed to the next task. This would force them to pay more attention to the website and be more thorough about their feedback.

6.12.4 Results of Study 4 for the Perfume Website

Figure 19 shows the performance for each component in the questionnaire-only condition. It is noteworthy that the general trend of these results corresponds to the questionnaire results in the user test condition. One notable difference is that almost 20% of the respondents found Products & Services to be problematic. This is more than what was observed above, but still significantly less than what was predicted by the checklist users. Thus, this pattern of results suggests that scenario-based testing, be it with or without facilitator, helps put the website and its offerings into context. That is, website elements that would normally be problematic (e.g. very short product descriptions) are tested to see whether they would indeed pose a problem to prospective customers of this particular site with this particular range of products. One should note that the proportion of respondents who would not trust this site is approximately the same in both questionnaire studies, namely, 25 and 30% respectively.

Compared with the results from Study 3b, one should note that, even with the Products & Services component included, the Pearson correlation coefficient amounts to 0.825 (significant at the 0.05 level, 1-tailed).

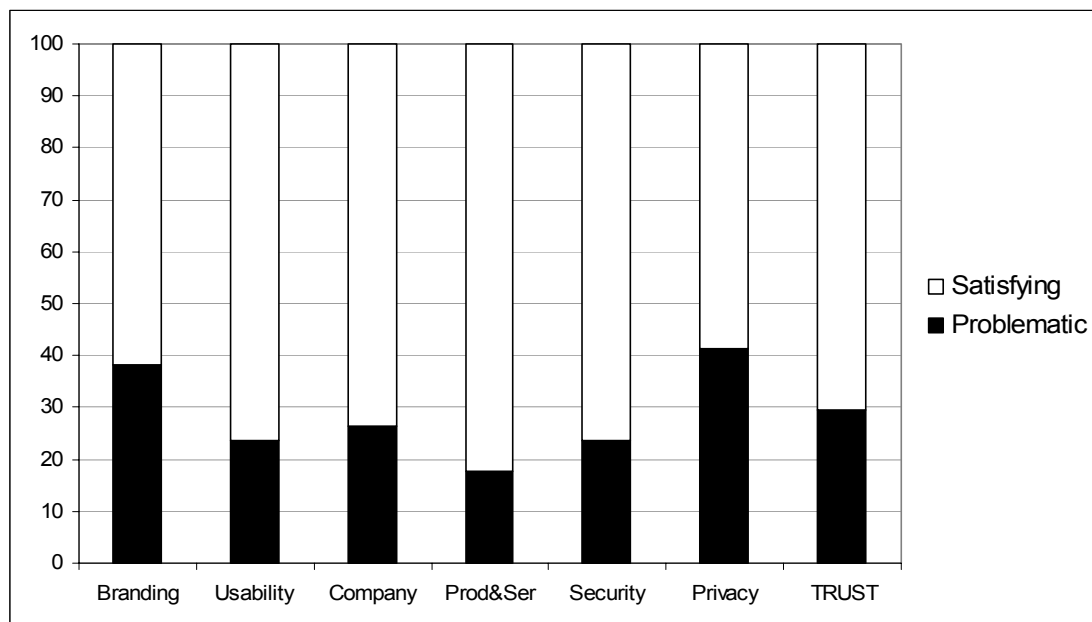


Figure 19 – Component performances based on questionnaires only (Perfumes)

6.13 Comparisons of Questionnaire Results

Regarding the adequacy of the QuoTEC questionnaire, the results suggest that questionnaires administered straight after a user test help visualise the performance of the different components and also the severity of observed trust problems. In that respect, questionnaire results are not mere reflections of problems encountered during a test but can be used as a complementary measure of severity.

The reason we also conducted a questionnaire study without having user tests beforehand was that, in practice, one-to-one tests can be extremely time- and cost-intensive. The argument was that if the same results can be achieved without having user tests requiring an onsite facilitator, this would be a convincing argument for having respondents fill out the questionnaire in their own time and place.

The results indicate that, overall, a similar performance pattern is found in the two questionnaire studies. However, two important exceptions must be noted. First, the Flower website example showed that the Usability component was significantly underrated in the questionnaire-only study. We tried to account for this result by claiming that respondents in that conditions might have been little motivated to invest a lot of time in carrying out the tasks set by the scenario. Therefore, not having encountered transaction-critical problems, they rated that component very high. This problem can be addressed by building in a more interactive dialogue in the remote test set-up. For example, a software application might prompt users to look for certain pieces of information and input the answer into the system. That would help force participants to follow a structured scenario.

The second important exception was the fact that the extremely succinct product descriptions in the perfume website were not considered to be a real problem for test participants, while almost 20% of the questionnaire-only respondents rated the Products & Services component as problematic. One explanation for this difference is that the scenario used in the tests specified which products participants should select. Therefore, they might not have needed extensive product information as they exactly knew what they were looking for. Also, one should remember that most test participants reported that they would only buy a perfume online that they would already know from the offline world. Thus, extensive descriptions are less important when people shop in a goal-directed way and/or for a familiar product. The reason questionnaire-only respondents reacted differently probably was that, once again, they did not follow the set scenario when evaluating this site. Thus, they might have assessed the site, based purely on its objective characteristics (e.g. succinct descriptions) without being immersed in a context that might affect how these characteristics are perceived.

Provided our methodological *caveat* is respected, our findings confirm Hypothesis 3, viz. that there would be no difference between the questionnaire results after user tests and the questionnaire without facilitator condition.

6.14 Conclusions

This chapter described a series of studies, the objective of which was to test the contribution and predictive power of the CheckTEC and QuoTEC tools presented in Chapters 4 and 5. The studies indicate that, compared with unguided evaluators, expert evaluators using the checklist find more problems, need less time and are more satisfied about their performance. On average, checklist-guided evaluators found four times as many problems in half the time. When compared with real trust problems observed in user tests, it appeared that checklist-guided evaluators found almost twice as many problems. An important number of apparent false alarms were also found in the checklist condition. These can be explained both by noise items that are not applicable and by items that are bound to become more salient in a real shopping situation.

Hypothesis 1 could thus be confirmed. Regarding the variety of problems found, checklist users reported more varied problems as shown in Section 6.6. These results confirm our second hypothesis.

Regarding the QuoTEC questionnaire, its application after user tests has shown to offer a complementary view on user comments by providing a graphical way to assess the severity of trust-related problems. The questionnaire studies also indicate that an unsupervised questionnaire administration can produce reliable results, provided respondents are committed to follow a scenario in a structured way. If not, parts of the results might not faithfully reflect the experience of a prospective customer. Correlation between questionnaire results in the two studies confirmed our third hypothesis.

CHAPTER 7

Discussion

This final chapter starts by providing a summary of the different phases and results encountered in this research. The main contribution of this research resides in a global, problem-centred approach that went beyond the traditional boundary of HCI. The MoTEC model, comprehensive in its coverage, is structured around concepts that can easily be related to practice. The validated model-derived tools also make a valuable contribution to HCI practice. Indeed, the CheckTEC checklist and the QuoTEC questionnaire are concrete tools that can be traced back to their theoretical basis. Generally, it was shown that designing the trust experience for B2C e-commerce is above all a multidisciplinary team activity that should continuously be coupled with assessments of perceived trustworthiness. Limitations of this research include: the broad coverage of the model, resulting in a liberal definition of what constitutes a trust problem; the applicability of the different tools; the over-reliance on students as test participants. Future research linking trust, HCI and Marketing is also outlined. Lastly, ethical aspects of designing for perceived trustworthiness are discussed.

7.1 Introduction

This chapter presents a general discussion of the results presented in earlier chapters. We will start with a summary of the approach, phasing and findings reported in this research. The contribution of this work to the discipline of Human-Computer Interaction will be discussed in Section 7.3. Practical implications for HCI research and HCI practice will be exposed in Sections 7.4 and 7.5, respectively. Limitations of this approach and their potential effects on the generalisability of the findings will be considered in Section 7.6. This will be followed by ethical considerations regarding the design of persuasive technology in Section 7.7. Lastly, Section 7.8 will discuss open questions for future research.

7.2 Recapitulation

Chapter 1 started by reviewing the development of electronic forms of commerce. Unlike EDI and Business-to-Business e-commerce, Business-to-Consumer e-commerce was found to be particularly affected by consumer trust concerns. A first analysis indicated that trust concerns were related to concerns regarding security and privacy, the unfamiliarity of some online services, lack of direct interaction with products, salespeople and fellow shoppers and the general low credibility of online information. Human-Computer Interaction was argued to be an effective discipline from which to approach online trust. The objectives of this research were articulated as: (1) To build up substantive knowledge about what makes customers trust e-commerce websites and, (2) To build up and validate methodological knowledge to help practitioners design and evaluate trust-shaping factors in e-commerce websites.

Chapter 2 reviewed literature on trust and risk. Rempel et al.'s (1985) model of trust in romantic relationships was particularly useful at distinguishing between different trust components. Doney and Cannon's (1997) marketing-oriented model of trust in the buyer-seller relationship brought additional trust factors and trust-building processes to the fore. With respect to trust in e-commerce websites, it was noteworthy that several studies looked at individual factors likely to affect trust but, when this research started, none had tried to regroup trust-shaping factors into a model. That provided additional support for our first research objective formulated in Chapter 1. More recent findings are discussed in the next section.

Chapter 3 described how the Model of Trust in E-Commerce (MoTEC) was developed on the basis of the literature reviewed in Chapter 2, in order to address the first objective formulated in Chapter 1. An initial version of the MoTEC model presented in Egger (1998) underwent different iterations to integrate new research findings and observations from user tests. The revised model contains four dimensions: Pre-interactional Filters, Interface Properties, Informational Content and Relationship Management. The model claimed to be descriptive of the different types of factors that may affect a person's judgment of an online vendor's trustworthiness.

Chapter 4 addressed the second objective set in the introductory chapter, viz., the need for methodological knowledge for trust evaluation (diagnosis) and design (prescrip-

tion). The so-called Trust Toolbox contains a suite of tools that have been directly derived from the MoTEC model. The first is GuideTEC, a set of design guidelines that can be used to maximise a site's perceived trustworthiness. The second, CheckTEC, is a checklist to be used by HCI practitioners to conduct expert evaluations. Lastly, QuoTEC is a questionnaire that can be used to get direct feedback about a website's trust performance directly from the target customers.

Chapter 5 reported the application of the QuoTEC questionnaire in evaluations studies in the service and retail industries. The results indicate that there are different trust issues when evaluating a hotel website or a retail website. To get an overall picture of e-commerce sites, data from the two studies were then combined in a single data set and analysed as a whole. The main objective of this chapter was to reduce the number of questionnaire items in QuoTEC, while minimising effects on the explained variance. Using factor analysis, two main underlying factors emerged from the data. The first one was interpreted as "efficient access to information". The second one contained all items referring to the model's Risk component and was thus called "perceived risk". The 2-factor structure was also shown to be valuable to visualise the overall trust performances of the individual websites. However, we also stressed that these 2 factors only account for about 61% of the variance, which indicates that trust is likely to be more than 2-dimensional in nature and that other factors should not be disregarded.

The previous chapter demonstrated the concrete benefit of the CheckTEC checklist. Indeed, compared with unguided evaluators, checklist-guided evaluators found about four times as many problems in half the time. In addition, the correspondence between predicted problems and problems observed in user tests is greater for CheckTEC results than results in the unguided condition. The QuoTEC questionnaire was also found to be an effective remote means to gather representative data about component performances from target users. Some differences in performance between a group who received the questionnaire after user tests and the remote group were attributed to the questionnaire-only group's lack of commitment and thoroughness in following the set scenario. It remains that, administered in a more systematic way, QuoTEC-based evaluations allow HCI practitioners to get valuable feedback from a large number of people over a short period of time.

7.3 Contribution

The research objectives set in Chapter 1 were:

- (1) To build up substantive knowledge about what makes people trust e-commerce websites;**
- (2) To build up and validate methodological knowledge to design and evaluate trust-shaping factors in e-commerce websites.**

How well these objectives have been attained will be assessed by reviewing the types of knowledge produced in this research.

7.3.1 The MoTEC Model

The Model of Trust in E-Commerce presented in Chapter 3 was developed to meet the first objective. Its main contribution lies in the fact that it regroups trust-shaping factors before, during and after the interaction with a website. Such a scope goes beyond traditional HCI, as factors beyond the user interface are also covered by the model. For example, the inclusion of factors related to reputation and customer relationship management, though typically more Marketing-oriented, are nevertheless crucial to the understanding of online trust. Another concrete benefit of the MoTEC model is that its main structure around four main dimensions is relatively simple to understand. Also, care was taken to rename the more Psychologically-oriented components of the initial model in such a way that they are more in line with concepts used and understood by website development and e-commerce management teams.

The model was developed with reference to existing literature on trust. In that respect, its development is traceable in terms of the identification of sources. However, it is not fully traceable in terms of the replicability of the development process. That means that it is very likely that other researchers would identify the same, or similar, trust-shaping factors but their method of model construction might be different.

The MoTEC model can, in principle, be of use to both researchers and practitioners. As far as researchers are concerned, it is important to note that the first objective referred to building up substantive knowledge and not to validating it. This was a conscious choice based on the primary aim of the research, which was to support the work of HCI practitioners by giving them validated methodological knowledge. That is why the model has not directly been tested and generalised in this research. It remains that the QuoTEC applications in Chapter 5 produced results that can be used as a starting point for further research and validation. Implications for HCI research are discussed below, in Section 7.4. Regarding HCI practitioners, it is important to point out that the method by which the model can be used by practitioners has not been addressed.

7.3.2 The Trust Toolbox

The Trust Toolbox introduced in Chapter 4 was developed to meet the second objective. Methodological knowledge, in the form of tools, was thus directly derived from the substantive knowledge contained in the MoTEC model. This direct derivation implies that the contents of the tools can be traced back to literature from Psychology, HCI and Marketing, which gives the tools more theoretical grounding. The output of the validation studies and the current status of the tools are presented next.

7.3.2.1 GuideTEC Guidelines

One should note that the guidelines have not been directly validated in this research. However, a research project by Kirillova (2003), supervised by the author, provided strong experimental evidence for the effectiveness of a subset of the guidelines presented in Section 4.2. That project first analysed consumers' trust concerns about buying insurance online, based on interviews and tests on a selection of online insurance websites. Relevant GuideTEC guidelines were then complemented by insurance-specific guidelines to better address customers concerns. Twenty-one guidelines were

then applied to the redesign of the first two navigational levels of an existing website.

Eight participants were then asked to carry out representative information-seeking tasks on both the original and the redesigned site (counterbalanced) and were asked to fill in the 15-item QuoTEC questionnaire. Performance scores on the six components served as measures to test the effects of the guidelines application. Wilcoxon tests revealed that the perceived performance of the six components in the redesigned prototype was significantly higher. What this study showed is that the application of trust guidelines had a positive effect on perceived trustworthiness, as measured by the six components as well as by a pure trust measure. What this study did not show, however, is *which* specific guideline in a set had *what* effect. In other words, guidelines were not validated individually, but in a group. Although these results look promising, future research will need to validate the guidelines one by one to identify which guidelines have the strongest effect on perceived trustworthiness and, by extension, consumer trust.

7.3.2.2 CheckTEC Checklist

The MoTEC-derived checklist was developed as a tool for expert evaluations. That is, it contains the same trust heuristics as the guidelines, just phrased in a different way and arranged by model component. The validation study in Chapter 6 tested the actual contribution of the CheckTEC checklist as a tool by comparing it to the results produced by a matched group that had to evaluate the same websites without this tool. The results showed that evaluators in the checklist condition identified more than twice as many trust problems as compared with the unguided evaluators. In addition, the proportion of evaluators who found a given problem was consistently higher in the checklist condition. The results also showed that evaluators who used the checklist needed, on average, 35 minutes to evaluate one site, although unguided evaluators needed, on average, 68 minutes. Satisfaction ratings were also found to be significantly higher for checklist users.

The validation studies also compared the number of problems predicted by the expert evaluation groups with the number of problems observed in user tests. The first observation was that unguided evaluators correctly predict about 53% of the observed trust problems, while that number was between 90 and 93% for checklist-guided experts. The second observation was that evaluators who used the checklist also predicted an important number of problems that were not observed in the user tests. Several interpretations have been given to account for this result. It remains that the main problem with comparative studies in HCI is that it is never possible to know the total number of real problems. Indeed, it is important to stress that the problems observed in the user tests, being based only on 18 participants, are themselves predictions of possible problems. Thus, it may well be that the apparent false alarms would in fact be problems for other users, under different conditions.

On the other hand, one could argue that the checklist users might focus too much on the details contained in the trust heuristics and thereby lose a sense of the general experience a site could make on its users. This could of course happen depending on the experience of the evaluator and the *actual* use of the checklist. Indeed, we expect practitioners to use the checklist more informally than in this validation study. For instance, evaluators could only select checklist items (or heuristics) that apply to their

specific case and/or which they need the most guidance for. Also, it is not absolutely necessary to rate the website's compliance with the items on a scale.

The feedback received from the checklist users indicated that some items should be removed, while others could be added or merged. It would be interesting for future research to test more thoroughly the level of abstraction, as well as the phrasing of the different items in order to make the checklist more complete, coherent and fit for purpose.

7.3.2.3 QuoTEC Questionnaire

The MoTEC-derived questionnaire was intended as a tool to gather data about a website's perceived trustworthiness directly from representative customers, over a short period of time. Through the factor analysis applied to the combined QuoTEC results in Chapter 5, the number of questionnaire items could be reduced from 23 to 15, with a minimum loss of explained variance. That analysis also ensured that the questionnaire equally applies to service and retail websites. Another valuable contribution resides in the fact that the factor analysis uncovered two underlying dimensions in the data. A practical implication of these results consists in the opportunity for trust performance visualisation provided by the 2-dimensional space. Plotting the individual websites on that space showed that there was a strong positive relationship between efficient access to information and (the absence of) perceived risk. This provides strong support for our claim that trust can be influenced by seemingly unrelated factors, such as a website's ease-of-use.

7.4 Implications for HCI Research

The first implication of our findings for HCI research concerns the adequacy of the MoTEC model, both in terms of content and predictive power. The third iteration of the factor analysis on the combined data in Chapter 5 uncovered two main underlying constituents of trust. The loading pattern of the resulting model showed that the pairs of components that were predicted to be related in Chapter 3 were indeed kept together. That provided strong evidence for the structure of the model's content. An interesting theoretical implication of the resulting 2-factor structure is the distinction consumers make between the ease with which they can retrieve relevant information from a website and the perceived risk of a transaction. We argued, however, that more factors are necessary to get a comprehensive picture of the factors affecting one's perception of trustworthiness. Indeed, if we refer back to the factor analysis results of the separate studies, we see a 4- and a 3-factor solution that account for a greater proportion of the variance. It is interesting to see that our interpretation of the 3-factor solution almost perfectly coincides with the three factors in the analytic model of trust proposed recently by Corritore et al. (2003) but not yet validated.

Regarding the ability to predict a customer's trust in a website, our findings indicate that the best predictor variables are not the same for all websites. The regression analysis for the hotel websites was remarkably different from that for the retail websites. Thus, it seems that differences between industries can be addressed by changing the weights attached to the model components. An implication of this observation is that we need a better taxonomy of transactional websites so that we can conduct more

systematic comparative studies between different types of sites.

It would be particularly interesting to explore the distinction between low/high involvement on the part of the participant and low/high risk, as described by Briggs (2003). Low involvement refers to situations where people are asked general questions about trust in e-commerce websites or specific questions about a part of a website, without engaging in a deeper assessment of a site, based on all the available trust cues. On the other hand, high involvement refers to people who visit specific websites and who are asked to report what elements makes them trust the sites or not. Our studies all fall in the latter category, as extensive user tests were conducted on existing websites. Briggs (2003) further distinguishes between low and high risk, which refers to the degree of risk (e.g. inconvenience, money, etc.) related to the transaction. To relate this to our results, it may be that the perceived risk of staying in a hotel that is inappropriate is greater than the perceived risk of buying a standard product online.

Another important aspect that Briggs (2003) notes is *framing effects*, i.e. that the way tasks or problems are posed to the participants can have a strong influence on the data collected. What we observed in our studies is that some participants had the tendency to view the test as a usability test rather than a trust study. Although both are related, the focus of the evaluation and the resulting data are clearly different. One implication for research would be to devise a true online shopping task and observe participants without prompting them to focus on trust issues. Indeed, some customers might be relatively unaware of potential risks and would not, naturally, look for specific information or policies. That would place the trust problem in a more naturalistic setting. Fogg et al. (2002) also discusses this idea in the light of Prominence-Interpretation Theory. These authors argue that people will focus their attention on prominent elements of a website, such as graphic design, when they are asked to evaluate it and base their judgement of trustworthiness on these prominent elements.

An assumption of this research was that trust attitude, based on a site's perceived trustworthiness, would lead to trust behaviour. Although we did not test trust behaviour, as no real purchases were made on the websites, the initial assumption remains. The main reason for this is the Theory of Planned Behaviour by Ajzen (1991) which states that attitudes, along with subjective norms and behavioural control, influence behavioural intention. However, it would be interesting for future research to test whether the assumption of the link between trust attitude and trust behaviour in the context of e-commerce stands up.

7.5 Implications for HCI Practice

We stressed the need for concrete tools intended for HCI practitioners from the very beginning. Indeed, instead of focusing on the direct validation of the MoTEC model, the entire approach centred on the development and validation of the Trust Toolbox. One advantage of this approach is that it provides practitioners not only with tools but also with a high-level descriptive model. For instance, the theoretical rationale for using a specific trust design guideline is not made explicit in the otherwise very comprehensive report by the Nielsen Norman Group (2002).

An additional practical implication of this research is that its customer-centred approach emphasised the need for multidisciplinary collaboration. Indeed, if a company wants to launch a B2C e-commerce website, it will be in its interest to have marketers, graphic designers, developers, usability specialists, customer service and management all agree on a common strategy. Indeed, as consumer trust can be affected by so many factors, it is primordial that all parties maximise those trust-shaping factors that come under their responsibility. Ideally, a trust-building strategy should be coordinated by a user/customer experience specialist who would feed back customer comments and concerns directly to the relevant department. The Relationship Management dimension also points to the importance of building trust over time. Therefore, perceived trustworthiness evaluation and redesign should be a continuous process within a company that does business online.

7.6 Limitations

7.6.1 Scope

The main limitation of this research can be imputed to the ambition of the model's coverage. The model tried to include as many trust-shaping factors as possible without trying to scope down the online trust problem to one particular aspect. The high-level model that resulted might therefore seem too general when viewed only from one discipline. Another implication of this global approach to trust is the fact that no single definition of trust in this context was given. Indeed, many factors that were seen to hinder website usage and transactions were considered to be trust problems, although psychologists could object to that liberal definition.

One obvious way to make the high-level model more concrete was the development of the Trust Toolbox. However, these tools primarily deal with Interface Properties and Informational Content aspects, at the expense of Pre-interactional Filters and Relationship Management. This can partly be explained by our present focus on initial trust and also by the practical difficulties involved in manipulating people's attitudes towards a particular company or in manipulating the quality of buyer-seller interactions over time. With respect to the Trust Toolbox in general, one should emphasise that hardly any tool in HCI comes in a "one size fits all" format. A prominent feature of HCI as a discipline is its emphasis that human-computer interactions take place in a particular context to achieve a particular goal. (ACM, 1992). Consequently, changing contexts of use can greatly impact the extent to which one particular tool will be fit-for-purpose. It is therefore up to the individual HCI practitioner to identify situations where a tool can be used unaltered or where it will need to be adjusted to match a given context.

7.6.2 Internal Validity

Are the changes in the dependent variable really due to the independent variables that we identified? The main threats to internal validity in our studies have to do with the possibility of extraneous, or confounding, variables. First, it might be that the MoTEC model and its derived tools missed one crucial factor that had a significant effect on respondent's trust ratings. In particular, it is not possible to control for people's previ-

ous experiences, whatever they may be. A previous interaction with a company or a brand would be an obvious factor that might impact one's trust in its website. Less obvious are cases of brand associations. This would refer to cases where one person has not directly interacted with a company before but where the site bears some resemblance with some other familiar website or just a brand, be it trusted or not. That would be a situation that would bring another factor into the equation that would be hard to make conscious and explicit, let alone quantify.

Given that we tested live e-commerce sites, there is always the possibility that the sites (the stimuli) might have changed from one day to the next, because of the dynamics of web authoring. Since we can never have full control over a live website, participants might have been exposed to slightly different versions of the test objects. This would be akin to a history effect, beyond the control of the researcher. In fact, no changes in the test websites were observed.

A possible maturation effect might also have played a role in the hotel study. Participants were asked to evaluate six different websites, which typically took more than one hour and a half. Because of the length of the test, people might have been less motivated and less alert towards the end of the test. Also, because of the repeated measures design, participants might have remembered questionnaire items from one evaluation to the next. Primed by the questionnaire, they might have evaluated later websites with a greater, albeit less natural, focus on trust issues. We tried to minimise such effects by counterbalancing conditions and by randomising the presentation of the websites.

Lastly, another threat to internal validity could be our selection of test participants. Approximately three quarters of the people tested in the course of this research were university students. This means that the results are based on a relatively homogenous user population in terms of age, socio-economic background and familiarity with technology. Testing such a sample restricts the spectrum of individual differences (cf. Pre-interactional Filters) that may affect one's trust judgements. On the other hand, it allows for better comparisons across studies and conditions.

7.6.3 External Validity

External validity is concerned with the generalisability of the results. As the last point in the preceding section underlined, one question is to what extent our findings generalise to other user population or customer segments. Although students have been prime targets for e-commerce because of their access to and familiarity with technology, increasingly more people are being attracted by the convenience of online shopping. It is, for example, unclear to what extent our findings would apply to an older but technologically more novice user group.

To what extent do the results apply to other settings? As mentioned earlier, Fukuyama (1995) reports cultural differences in terms of propensity to trust. Of course, there are also cultural differences in terms of access to technology. Thus, it may be that our results are culturally biased and that they may not exactly map onto other settings. However, to minimise the effect of culture within our possibilities, we conducted the different studies in three different European countries with a very international pool of

participants.

An interesting question is whether our results can be generalised to other times. Although the basic psychology of trust will certainly not change with time, technology changes at an incredible speed and so does people's attitude to novel means of interaction. In addition, the legal context for online trade might also change in a way that would give customers more rights and more protection. In that case, consumers' risk-related concerns might decrease. However, it is hoped that the basic structure of the MoTEC model was articulated at a high enough level to remain applicable to different flavours of electronically-mediated commerce.

7.6.4 Construct Validity

Construct validity refers to the degree to which one can generalise back to the theoretical construct one started from. In this research, the questions are: Was trust accurately measured? Were the components accurately measured? Regarding the pure measure of trust that we asked respondents, it should be noted that we consciously chose not to define trust to them beforehand. This was to avoid priming them by describing the kinds of risk they might be subjected to when transacting on an untrustworthy website. Of course, this also means that maybe not every respondent had exactly the same concept of what trust is, which might have affected the reliability of the trust ratings. This is related to the framing effects described by Briggs (2003), i.e. that the way trust-related questions or tasks are phrased can have a significant effect on how they will be answered.

One should note that no independent convergent evidence could be produced in this research. With the recent publications of e-commerce trust scales (e.g., Bhattacharjee, 2002 or McKnight et al., 2002), comparative studies can now be conducted to test whether the different measurement instruments all measure the same construct.

7.6.5 Ecological Validity

Ecological validity refers to the similarity of the test situation to the real situation. In this case, was participants' behaviour during the tests similar to that in a real online shopping situation? The main limitation in this research stems from the fact that none of our test participants was really about to make a purchase or a booking on any of the sites that were evaluated. This lack of intrinsic motivation might have affected the quality of the feedback with respect to perceived trustworthiness. In addition, in a normal situation, decisions about whether to shop on one site rather than another are taken alone, without having to rationalise and justify one's decision. This sometimes unconscious process also underlies the principle of impulse buying, where immediately accessible benefits might distract from a thorough assessment of risk. This is also related to the involvement/risk distinction discussed above (cf. Briggs, 2003).

7.7 Future Research

The main endeavour for HCI research would be to validate the MoTEC model by examining more closely the interrelations between the different components. Thus, instead of a static model listing trust-shaping factors, it could describe the dynamics be-

tween, and the different weights attached to, the model components. In particular, as suggested above, it would be interesting to look at different industries to examine the different underlying variables that emerge from the data. As perceived risk was found to be a stable factor in our different studies, closer attention should be paid to people's knowledge of and actual concerns about security and privacy issues. For example, if a person does not know or care about the possible dissemination of his/her personal information, that person could rate a website that lacks privacy information as very trustworthy. Such a pre-disposition could introduce a large variation in trust ratings. In addition, one could investigate the relationship between the objective risk of a transaction (e.g. based on the technological infrastructure) and its perceived risk.

Construct validation should also be addressed by comparing alternative measurements for trust and the model components to the results presented in this research.

Regarding the Trust Toolbox, the most direct follow-up research to this project would be to pay closer attention to the GuideTEC guidelines. Building on the work of Kirillova (2003), isolated or small groups of trust guidelines could be used to develop a prototype website in which specific variables would be manipulated. One could envisage a component-by-component approach aimed at finding out the exact weights prospective customers attach to the different attributes of a website. Such an experimental validation would promote the tested guidelines to the most reliable guideline category mentioned in Chapter 4. It would also be interesting to conduct comparative studies of trust-shaping factors across industries. One preliminary step in that direction was made by Shelat and Egger (2002) who looked at the relative importance of the MoTEC components in the gambling industry. These findings indicate that experienced gamblers are mostly concerned about information, such as company policies, as well as about Relationship Management issues, such as prominent means of contact.

It would also be interesting for future research to focus on the two aspects that were neglected in this project, namely Pre-interactional Filters and Relationship Management. Regarding Pre-interactional Filters, it would be interesting to examine the potential effects of culture and/or age on one's judgement of a website's trustworthiness. Research on how to control for related experiences and brand associations would also be welcome to produce more valid and more reliable data. Regarding the positioning of a new e-commerce website, Marketing could teach HCI a lot about how to advertise a company and its products in such a way that consumers make positive associations with it. Complementing small sample research, as is common in HCI, with large sample research, as is common in Marketing, might also give new insights into how different market segments (or user groups) rate different attributes of a new online company and/or its website.

Since this research dealt only with initial trust, the fourth dimension, Relationship Management, was left aside as it deals with trust development over time. Exactly how trust between a buyer and a seller is maintained over time would be a fascinating topic for further research. In particular, what is the optimal mixture of on- and offline communication for one particular industry and how often should interactions take place? For example, one could research methods used to ensure consistency of a company's communication with one customer across media and distribution channels.

This consistency would be crucial to create an environment of familiarity and genuine client-centeredness - essential ingredients to maintain trust over the long term. Another crucial question is whether and how breaches of trust can be repaired and how long it will take to attain the previous level of trust.

7.9 Ethical Considerations

The work on credibility done by Fogg's (2003) *captology* group at Stanford and the current research both fall into the field of *persuasive technology*. The Stanford group claims to design persuasive technologies that bring about positive changes in the users' attitudes and behaviours. However, the science of persuasion is intimately linked to that of deception. That is why it is important to be aware that designing e-commerce websites in a way that makes them *appear* to be trustworthy can raise serious ethical questions. Although our results are intended to be used by legitimate companies to maximise their perceived trustworthiness, the GuideTEC guidelines could equally well be employed by shady, fly-by-night operators to make a site appear genuine and trustworthy.

Berdichevsky and Neunschwander (1999) discuss the ethical implications of persuasive technology design and propose a list of eight principles which designers should follow. Their "Golden Rule of Persuasion" is as follows:

"The creators of a persuasive technology should never seek to persuade a person or persons of something they themselves would not consent to be persuaded to do."

This principle was directly inspired by the Theory of Justice put forward by the Harvard philosopher John Rawls (1971). Rawls claimed that we should consider ethics from behind a "veil of ignorance", i.e. from the point of view of an "original position of equality". If we did not know who we were (in terms of social status, education, etc.), we would be bound to obey only those rules of ethics that benefit us, no matter who we turn out to be. In other words, we should seek to develop technologies *as if* we did not know on which side of the medium we would end up. That would force us to reflect on the ethical implications of the technology, as we might very well be on the receiving end. Readers interested in the application of the Theory of Justice to usability aspects of HCI are referred to Duquenoy and Thimbleby (1999), who discuss the link between the philosophical concept of justice and HCI design.

Although Berdichevsky and Neunschwander's (1999) main rule of persuasive design makes sense as a guideline, the problem remains that, in practice, enough people do not care about being unethical as long as they can profit from their deceptive designs. The following two examples, published in Egger (2003), illustrate that user experience manipulation happens more often than we think. The first example was encountered during a study into what makes people trust online gambling sites (Shelat & Egger, 2002), in a book by Haywood (2000) entitled *BeatWebCasinos.com*.

Haywood (2000) reports his personal experience browsing casino websites and noticing that several of them featured a *Safebet* trust seal. To find out more about this seal, he clicked on it and, as expected, was taken to the *Safebet* site. The site informed him that the casino he had arrived from was indeed registered with that trusted-third party.

He discovered that *Safebet* is a non-profit, independent organisation that provides certification and dispute resolution services to help regulate the industry and protect the interests of the players. It even has its own team of mathematicians to analyse whether the casinos' odds are fair. In terms of the MoTEC model, the original website had a prominent link that showed endorsement by a trusted third party, which normally helps mitigate risk by a transfer of trust.

However, a simple *whois* search revealed that the person who registered the *safebet.org* domain also registered 51 casino domain names. It turned out that some people created a phoney certification scheme and proudly featured the allegedly independent seal on their own gambling sites. There is no doubt that the average surfer would not have double-checked the legitimacy of the seal and would have been very easily deceived by this little design trick. Let us consider another example.

Imagine a foreign person who wants to work in the United States but only possesses vague information about the green card lottery. She runs a web search and finds the site of the *USA Immigration Services* in Washington D.C. The name certainly sounds blandly official and the site name ends in *.org*. Its logo features the eagle and the American flag. The site's graphic design looks professional and includes the hyper-linked logos of the USA Freedom Corps, The White House and FirstGov (the US e-government portal). There is plenty of information about the green card lottery and even a neatly designed eligibility checking system. Users can fill in a form, pay a processing fee - and hope they will win. Again, things are not as they seem: Unless one reads the terms and conditions, there is no way to know that it is not the website of an official government agency but that of a for-profit intermediary. Participation in the lottery is actually free! Exploiting the target audience's poor familiarity with US institutions, this website misleads people into thinking they are dealing with the Immigration and Naturalization Service.

7.9 Conclusions

In conclusion, one should bear in mind that many strategies can be implemented to increase a website's perceived trustworthiness. It is important to emphasise that a high perceived trustworthiness does not necessarily mean that a company does indeed behave in a trustworthy way – and vice versa. Although some of the knowledge and tools presented in this research could be misused by individuals to make dubious sites appear to be legitimate, the same knowledge can be used to educate online shoppers to make wiser decisions and minimise risks.

BIBLIOGRAPHY

- Abrazhevich, D. (2002). Importance of User-Related Factors in Electronic Payment Systems. In: J.E.J. Prins et al. (Eds.). *Trust in Electronic Commerce: The Role of trust from a Legal, an Organizational and a Technical Point of View*. Kluwer Law International.
- ACM Special Interest Group on Computer-Human Interaction Curriculum Development Group (1992). *ACM SIGCHI Curricula for Human-Computer Interaction*, Technical Report, ACM, New York.
- Ajzen, I. (1991). The Theory of Planned Behaviour. *Organizational Behaviour and Human Decision Processes*, Vol. 50: 179-211.
- Arion, M, J.H. Numan, H. Pitariu & Jorna, R. (1994). Placing Trust in Human-Computer Interaction. *Proc. 7th European Cognitive Ergonomics Conference*, 353-365.
- Arnfeld, A. & Rosbottom, J. (1998). Improving the Availability and Cost-Effectiveness of Guidelines for Guideline Users: Towards a Structured Approach. *Behaviour and Information Technology*, Vol. 17 (3): 135-140.
- Baier, A. (1992). Trust and Antitrust. In: Deigh, J. (Ed.), *Ethics and Personality: Essays in Moral Psychology*, University of Chicago Press.
- Berdichevsky, D. & Neunschwander, E. (1999). Toward an Ethics of Persuasive Technology. *Communications of the ACM*, May 1999, Vol. 42 (5): 51-58.
- Bernstein, T., Bhimani A.B., Schultz, E. & Siegel, C. (1996). *Internet Security for Business*, John Wiley, New York.
- Blythe, M.A., Overbeeke, C.J., Monk, A.F. & Wright, P.C. (2003). *Funology: From Usability to Enjoyment*. Dordrecht: Kluwer.
- Boren, T. & Ramey, J. (2002). Thinking Aloud: Reconciling Theory and Practice, *IEEE Transactions on Professional Communication*, Vol. 43 (3):261-278.

- Boston Consulting Group (1998). *The State of Internet Retailing*. Mimeo, BCG/Shop.org (November).
- Briggs, P. (2003). Making Sense of Trust Research. *Hot Topics, Publication of the Human Oriented Technology Lab*, Carleton University (Canada). Available at: <http://www.carleton.ca/hotlab/hottopics/Articles/Pamelas.html>
- Briggs, P., Burford, B., De Angelli, A., Lynch, P. (2002). Trust in Online Advice. *Social Science Computer Review*, Vol. 20 (3): 321-332.
- Brown, C.M. (1988). *Human-Computer Interface Design Guidelines*. Norwood, NJ: Ablex Publishing Corp.
- Camp, L. J. (2000). *Trust and Risk in Internet Commerce*, The MIT Press.
- Cheskin Research (2000). *Trust in the Wired Americas*. Available at: <http://www.cheskin.com/think/studies/trust2.html>
- Cheskin Research & Studio Archetype (1999). Ecommerce Trust Study. Available at: <http://www.studioarchetype.com/cheskin/>
- Claymon, D. (1998). A Matter of Trust. *Red Herring*, March 1998. Available at: <http://www.redherring.com/mag/issue52/trust.html>
- CommerceNet (1997). *Barriers & Inhibitors to the Widespread Adoption of Internet Commerce*. Available at: www.commerce.net.
- Coombs, C.H., Dawes, R.M. & Tversky, A. (1970). *Mathematical Psychology*. Prentice Hall.
- Corritore, C.L., Kracher, B. & Wiedenbeck, S. (2003). On-line Trust: Concepts, Evolving Themes, a Model. *International Journal of Human-Computer Studies*, Vol. 58: 737-758.
- Cural, S.C. & Judge, T.A. (1995). Measuring Trust between Organizational Boundary Role Persons. *Organizational Behavior and Human Decision Processes*, Vol. 64 (2): 151-170.
- Dasgupta, P. (1988). *Trust as a Commodity*. In: Gambetta, D. (Ed.) (1988). *Op. cit.*
- Davis, F.D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, Vol. 13 (3): 319-340.
- Deutsch, M. (1960). The Effect of Motivational Orientation Upon Trust and Suspicion, *Human Relations*, Vol. 13: 123-139.
- Dix, A., Finlay, J., Abowd, G. & Beale, R. (1993). *Human-Computer Interaction* (2nd ed.). Prentice-Hall Europe.
- Doney, P.M. & Cannon, J.P. (1997). An Examination of the Nature of Trust in the Buyer-Seller Relationship. *Journal of Marketing*, Vol. 51: 35-51.
- Dowell, J. and Long, J. (1989). Towards a Conception for an Engineering Discipline of Human Factors. *Ergonomics*, Vol. 32: 1513-35.

- Dumas, J.S. & Reddish, J.C. (1994). *A Practical Guide to Usability Testing*. Norwood, NJ: Ablex Publishing Corp.
- Duquenoey, P. & Thimbleby, H. (1999). Justice and Design. *Proceedings of the IFIP Conference on Human-Computer Interaction (Interact'99)*, 281-286.
- Egger, F.N. (1998). *Increasing Consumers' Trust in Electronic Commerce through Human Factors Engineering*. Unpublished Master of Science (Ergonomics) Thesis, University of London.
- Egger, F.N. (2000). "Trust Me, I'm an Online Vendor": Towards a Model of Trust for E-Commerce System Design. *Proceedings of CHI 2000 Extended Abstracts*, ACM Press:101-102.
- Egger, F.N. (2001). Affective Design of E-Commerce User Interfaces: How to Maximise Perceived Trustworthiness. In: Helander, M., Khalid, H.M. & Tham (Eds.), *Proceedings of CAHD2001: Conference on Affective Human Factors Design*, Singapore, June 27-29, 2001: 317-324.
- Egger, F.N. (2002). Consumer Trust in E-Commerce: From Psychology to Interaction Design. In: J.E.J. Prins et al. (Eds.). *Trust in Electronic Commerce: The Role of trust from a Legal, an Organizational and a Technical Point of View*. Kluwer Law International.
- Egger, F.N. (2003). Deceptive Technologies: Cash, Ethics & HCI. *SIGCHI Bulletin*, Vol. 35, Issue 2, May-June 2003, ACM Press: 11.
- Egger, F.N. & Groot, B. de (2000). Developing a Model of Trust for Electronic Commerce: An Application to a Permissive Marketing Web Site. *Poster Proceedings of the 9th International World Wide Web Conference*, Amsterdam (The Netherlands), May 15-19, 2000: 92-93, Foretec Seminars Inc.
- Equifax/Harris Consumer Privacy Survey (1996). Archived at:
<http://www.mindspring.com/~mdeeb/equifax/cc/parchive/svry96/docs/intro.html>
- Ericsson, K.A. & Simon, H.A. (1980). Verbal Reports as Data. *Psychological Review*, Vol. 87: 215-251.
- Fischhoff, B., Slovic, C. & Lichtenstein, S. (1977). Knowing with Certainty: The Appropriateness of Extreme Confidence. *Journal of Experimental Psychology: Human Perception of Performance*, Vol. 3: 552-564.
- Fogg, B.J. (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann Publishers.
- Fogg, B.J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P. & Treinen, M. (2001). What Makes Web Sites Credible? A Report on a Large Quantitative Study. *Proceedings of CHI 2001*, ACM Press: 61-68.
- Fogg, B.J., Soohoo, C., Danielsen, D., Marable, L., Stanford, J., & Tauber, E. (2002). *How Do People Evaluate a Web Site's Credibility? Results from a Large Study*. Stanford Persuasive Technology Lab, Stanford University. Available at:
http://www.consumerwebwatch.org/news/report3_credibilityresearch/stanfordPTL_abstr act.htm

- Fogg, B.J. & Tseng, H. (1999). The Elements of Computer Credibility. *Proceedings of CHI 99*, ACM Press: 80-87.
- Fukuyama, F. (1995). *Trust: The Social Virtues and the Creation of Prosperity*, The Free Press, New York.
- Gambetta, D. (Ed.) (1988). (1988). *Trust: Making and Breaking Cooperative Relations*, Basil Blackwell.
- Ganster, C.G. (1987). Worker Control and Well-Being: A Review of Research in the Workplace. In: Sauter, S., Hurrell, I. & Cooper, C. (Eds.) *Job Control and Worker Health*. New York: John Wiley & Sons.
- Garrett, J.J. (2002). *The Elements of User Experience: User-Centered Design for the Web*. New Riders.
- Gefen D. & Straub, D.W. (2000). The Relative Importance of Perceived Ease-of-Use in IS Adoption: A Study of E-Commerce Adoption. *Journal of the Association for Information Systems*, Vol. 1 (8): 1-28.
- Georgia Tech (1998). Georgia Tech Visualisation and Usability WWW Survey (1998). Available on: http://www.cc.gatech.edu/gvu/user_surveys/papers/
- Giddens, A. (1990). *The Consequences of Modernity*. Oxford: Polity Press.
- Glaser, T.R. (1994). The Morality of Trust. Available at: <http://ucsu.colorado.edu/~glasert/archive/5thsem.htm>.
- Godin, S. (1999). *Permission Marketing: Turning strangers into friends and friends into customers*. New York: Simon & Schuster.
- Goleman, D., 1996. *Emotional intelligence: Why it can matter more than IQ*, Bloomsbury, London.
- Good, D. (1988). Individuals, Interpersonal Relations, and Trust. In: Gambetta, D. (Ed.) (1988). *Trust: Making and Breaking Cooperative Relations*, Basil Blackwell.
- Gray, W. D. & Salzman, M. C. (1998). *Damaged Merchandise? A Review of Experiments that Compare Usability Evaluation Methods*. *Human-Computer Interaction*, Vol. 13: 203-261.
- Guilford, J. P. (1954), *Psychometric Methods* (2nd ed.), New York: McGraw-Hill.
- GVU (1998) GVU's 10th WWW User Survey Available at: http://www.gvu.gatech.edu/user_surveys/survey-1998-10/
- Hart, K. (1988). Kinship, Contract and Trust: The Economic Organization of Migrants in an African City Slum. In: Gambetta, D. (Ed.). *Trust: Making and Breaking Cooperative Relations*, 176-193. Oxford: Basil Blackwell.
- Haywood, B. (2000). *BeatWebCasinos.com: The Shrewd Player's Guide to Internet Gambling*. RGE Publishing.

- Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. Sage Publishing.
- Hwang, P. & Burgers, W.P. (1997). Properties of Trust: An Analytical View, *Organisational Behavior and Human Decision Processes*, Vol. 69 (1): 67-73.
- Jarvenpaa, S.L, Tractinsky, N. & Vitale, M. (1999). Consumer Trust in an Internet Store: A Cross-Cultural Validation. *Journal of Computer-Mediated Communication*, Vol. 5 (2). Available at: www.ascusc.org/jcmc/vol5/issue2/jarvenpaa.htm
- Jarvenpaa, S., Tractinsky, N. & Vitale, M. (2000). Consumer Trust in an Internet Store. *Information Technology & Management Journal*, Vol. 1 (1-2): 45-71.
- Jordan, P.W. (2000). *Designing Pleasurable Products*. London: Taylor and Francis.
- Kahneman, D. & Tversky, A. (1973). On the Psychology of Prediction. *Psychological Review*, Vol. 80: 237-51.
- Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgments Under Uncertainty*. Cambridge University Press.
- Keen, P. (1999). *Electronic Commerce Relationships: Trust By Design*. Prentice Hall.
- Kim, J. (1997). Towards the Construction of Customer Interfaces for Cyber Shopping Malls - HCI Research for Electronic Commerce. *Electronic Markets*, Vol. 7 (2): 12-15. Also available at: www.electronicmarkets.org.
- Kim, J. & Moon, J.Y. (1998). Designing Emotional Usability in Customer Interfaces - Trustworthiness of Cyber-banking System Interfaces. *Interacting with Computers*, Vol. 10: 1-29.
- Kirakowski, J. (1994). The use of questionnaire methods for usability assessment. Available at: <http://sumi.ucc.ie/sumipapp.html>
- Kirillova, N. (2003). Usability & Trust in Online Travel Insurance: An Empirical Validation of Design Guidelines. Master's thesis, *User-System Interaction Programme*, Eindhoven University of Technology. ISBN 90-444-0301-X.
- Koller, M. (1988). Risk as a Determinant of Trust. *Basic and Applied Social Psychology*, Vol. 9 (4): 265-276.
- Lee, J. & Moray, N. (1992). Trust and Control Strategies and Allocation of Function in Human-Machine Systems. *Ergonomics*, Vol. 35 (10): 1243-1270.
- Lee, J., Kim, J. & Moon, J. Y. (2000). What Makes Internet Users Visit Cyber Stores Again? Key Design Factors for Customer Loyalty User Experience in E-Commerce. *Proceedings of CHI 2000*; ACM Press: 305-312.
- Lewicki, R.J., McAllister, D.J. & Bies, R.J. (1997). Trust and Distrust: New Relationships and Realities. *Academy of Management Review* (July 1998).
- Lindgaard, G. (1999). Does emotional appeal determine perceived usability of web sites?

- Available at: <http://cyberg.curtin.edu.au/members/papers/49.shtml>
- Lindquist, J.D. (1975). Meaning of Image. *Journal of Retailing*, Vol. 50 (4): 29-38.
- Lohse, G.L. & Spiller, P. (1998). Quantifying the Effect of User Interface Design Features on Cyberstore Traffic and Sales. *Proceedings of CHI 98*, ACM Press: 211-218.
- Long, J.B. & Dowell, J. (1989). Conceptions of the Discipline of HCI: Craft, Applied Science, and Engineering. In: *People and Computers: The Theory and Practice of HCI*. British Computer Society HCI Specialist Group Conference.
- Luchins, A.S. (1942). Mechanization in Problem Solving. *Psychological Monographs*, Vol. 54 (whole issue).
- Luhmann, N. (1988). *Familiarity, Confidence, Trust: Problems and Alternatives*. In: Gambetta, D. (Ed.). *Trust: Making and Breaking Cooperative Relations*, 94-107. Oxford: Basil Blackwell.
- Lumkin, M. (2003). Avatars for Customer Relationship Management. In: Minocha, S. & L. Dawson (Eds.). *Proceedings of Workshop 6: Exploring the Total Customer Experience (TCE): Usability Evaluations of (B2C) E-Commerce Environments*, INTERACT '03, 1-5 September 2003, Zurich (CH).
- Macavinta, C. (2000). Privacy fears raised by DoubleClick database plans. News.com. Available at: <http://news.cnet.com/news/0-1005-202-1531929.html>
- McAllister, D.J (1995). Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. *Academy of Management Journal*, Vol. 38 (1): 24-59.
- McKnight, D.H., Choudhury, V., Kacmar, C. (2000). Trust in E-Commerce Vendors: A Two-Stage Model. *Proc. 21st International Conference on Information Systems*, Brisbane, Queensland, Australia, 532 – 536.
- Meyer, R., Davies, J. & Shoorman, F. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, Vol. 20 (3): 705-734.
- Milgram, S. (1963). Behavioural Study of Obedience. *Journal of Abnormal and Social Psychology*, Vol. 67: 371-8.
- Moorman, C., Deshpande, R., & Zaltman, G. (1993). Factors Affecting Trust in Marketing Relationships, *Journal of Marketing*, 58: 20-38.
- Muir, B.M. (1987). Trust Between Humans and Machines, and the Design of Decision Aids. *International Journal of Man-Machine Studies*, Vol. 27 (5-6): 527-539.
- Muir, B.M. & Moray, N. (1996). Trust in Automation, Part II: Experimental Studies of Trust and Human Intervention in a Process Control Situation. *Ergonomics*, Vol. 39 (3): 429-460.
- Nelson, M.G. (2000). Fast Is No Longer Fast Enough. *Information Week Online*. Available at: <http://www.informationweek.com/789/prweb.htm>
- Nielsen, J. (1993). *Usability Engineering*. Academic Press.

- Nielsen, J. (1999). Reputation Managers are Happening. *Alertbox Column*, Sept. 5, 1999. Available at: <http://www.useit.com/alertbox/990905.html>
- Nielsen, J. (2000). *Designing Web Usability: The Practice of Simplicity*. Indianapolis: New Riders Publishing.
- Nielsen, J. & Molich, R. (1990). Heuristic evaluation of user interfaces. *Proceedings of the CHI'90 Conference on Human Factors in Computing Systems*, 249-256. New York: ACM Press.
- Nielsen Norman Group (2000). Trust: Design Guidelines for E-Commerce User Experience. Available at: <http://www.nngroup.com/reports/ecommerce/trust.html>
- Pavlou, P.A. (2001). Integrating Trust in Electronic Commerce with the Technology Acceptance Model: Model Development and Validation. *Proceedings of the Seventh Americas Conference on Information Systems*: 816-822.
- Popper, K. (1968). *The Logic of Scientific Discovery*. London: Hutchinson.
- Powell, W.W. (1990). Neither Market Nor Hierarchy: Network Forms of Organization. *Research in Organizational Behavior*, Vol. 12: 295-336.
- Preece, J. (1994). *Human-Computer Interaction*. Addison-Wesley.
- Preece, J., Rogers, Y. & Sharp, H. (2002). *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons, Inc.
- Prins, J.E.J., Ribbers, P.M.A, Tilborg, H.C.A van, Veth, A.F.L & Wees, J.G.L. van der (Eds.) (2002). *Trust in Electronic Commerce: The Role of trust from a Legal, an Organizational and a Technical Point of View*. Kluwer Law International.
- Rawls, J. (1972). *A Theory of Justice*. Oxford University Press (Originally published 1971, Harvard University Press).
- Rempel, J.K., Holmes, J.G. & Zanna, M.P. (1985). Trust in Close Relationships. *Journal of Personality and Social Psychology*, Vol. 49 (1): 95-112.
- Riegelsberger, J., Sasse, M.A. & McCarthy, J. (2003). Shiny Happy People Building Trust? Photos on e-Commerce Websites and Consumer Trust. *Proceedings of CHI 2003*, ACM Press: 121-128.
- Robinson, S.L. (1996). Trust and Breach of the Psychological Contract. *Administrative Science Quarterly*, Vol. 41: 574-599.
- Rosenfield, L. & Morville, P. (2002). *Information Architecture for the World Wide Web: Designing Large-Scale Websites*. O'Reilly & Associates.
- Rotter, J.B. (1980). Interpersonal Trust, Trustworthiness, and Gullibility. *American Psychologist*, Vol. 35 (1): 1-7.
- Schellekens, M. & Wees, J.G.L. van der (2002). ADR and ODR in Electronic Commerce. In: Prins et al. (Eds.). *Trust in Electronic Commerce: The Role of trust from a Legal, an Organizational and a Technical Point of View*. Kluwer Law International.

- Shelat, B. & Egger, F.N. (2002). What Makes People Trust Online Gambling Sites? *Proceedings of CHI 2002*, ACM Press: 852-853.
- Spool, J.M. (Ed.) (1999). *Web Site Usability: A Designer's Guide*. San Francisco: Morgan Kaufmann Publishers.
- Stanford, J., Tauber, E., Fogg, B.J. & Marable, L. (2002). Experts vs. Online Consumers: A Comparative Credibility Study of Health and Finance Web Sites. Stanford Persuasive Technology Lab, Stanford University. Available at: <http://credibility.stanford.edu/mostcredible.html>.
- Steinbrück, U., Schaumburg, H., Duda, S. & Kruger, T. (2002). A picture says more than a thousand words: photographs as trust builders in e-commerce websites. *Proceedings of CHI 2002*, ACM Press: 748-749.
- Tan, Y.H & Thoen, W. (1999). Towards a Generic Model of Trust for Electronic Commerce. Proceedings of the 12th Bled International Electronic Conference, Bled, Slovenia, June 7-9, 1999. Vol. 1: 346-359.
- Tversky, A. & Kahnemann, D. (1974). Judgement under Uncertainty: Heuristics and Biases. *Science*, Vol. 125: 1124-1131.
- Tversky, A. & Kahnemann, D. (1980). Causal Schemas in Judgements under Uncertainty. In: Fishbein, M. (Ed.), *Progress in Social Psychology*. Lawrence Erlbaum Associates Inc.
- Wason, P.C. (1968). Reasoning about a Rule. *Quarterly Journal of Experimental Psychology*, Vol. 20: 273-81.

APPENDIX

APPENDIX 1: Abstracts of Papers Produced in this Research

Egger, F.N. (1999). **Human Factors in Electronic Commerce: Making Systems Appealing, Usable & Trustworthy**. Graduate Students Consortium & Educational Symposium, 12th Bled International E-Commerce Conference, June 1999, Bled, Slovenia.

The challenges of electronic commerce (e-commerce) can be summarised in a few words: attract consumers, make them visit the site, establish trust, make them buy and, most importantly, make them come back. Unlike traditional commerce, most of the buyer-seller interaction takes place exclusively through the e-commerce interface. It is therefore imperative that this system be designed with the users/consumers in mind. That is why the general problem of concern in this study is the human-computer interaction (HCI) design of electronic commerce systems. Its specific scope is to identify human factors susceptible to increase the general appeal, usability and trustworthiness of a commercial web site. The working hypothesis in this study is that a user/consumer-centred design approach addressing these factors is more likely to lead to an appealing, usable and trustworthy site than a traditional technology-led approach. The objective of this research is therefore to develop and validate a user/consumer-centred analysis and design method for e-commerce systems.

Egger, F.N. (2000). **"Trust Me, I'm an Online Vendor": Towards a Model of Trust for E-Commerce System Design.** In: G. Szwillus & T. Turner (Eds.): *CHI2000 Extended Abstracts: Conference on Human Factors in Computing Systems*, The Hague (The Netherlands), April 1-6, 2000: 101-102, ACM Press.

Consumers' lack of trust has often been cited as a major barrier to the adoption of electronic commerce (ecommerce). To address this problem, a model of trust was developed that describes what design factors affect consumers' assessment of online vendors' trustworthiness. Six components were identified and regrouped into three categories: Prepurchase Knowledge, Interface Properties and Informational Content. This model also informs the Human-Computer Interaction (HCI) design of e-commerce systems in that its components can be taken as trust-specific high-level user requirements.

Egger, F.N. & Groot, B. de (2000). **Developing a Model of Trust for Electronic Commerce: An Application to a Permissive Marketing Web Site.** *Poster proceedings of the 9th International World-Wide Web Conference*, Amsterdam (The Netherlands), May 15-19, 2000: 92-93, ISBN 1-930792-01-8.

The Internet has notoriously democratised direct access to businesses, putting them only a few mouse-clicks away from consumers. However, research indicates that consumers' lack of trust still constitutes a major psychological barrier to the adoption of new forms of online services. It is therefore imperative to identify factors likely to affect a consumer's perception of an online vendor's trustworthiness. Only then can methods be derived that elicit consumer trust requirements and thereby inform the design of the e-commerce user interface.

Egger, F.N. (2001). **Affective Design of E-Commerce User Interfaces: How to Maximise Perceived Trustworthiness.** In: Helander, M., Khalid, H.M. & Tham (Eds.), *Proceedings of CAHD2001: Conference on Affective Human Factors Design*, Singapore, June 27-29, 2001: 317-324.

Successful e-commerce user experience design depends on a large number of factors. This paper focuses on consumers' acceptance of and trust in an e-commerce system, based on the transaction's value and perceived risk. The model of trust for e-commerce (MoTEC) by Egger (2000) provides a framework making explicit factors likely to affect customer trust. For each model component, design principles are provided, along with more concrete guidelines. It will be shown that the user interface is only one element of the customer experience. Designing for trust therefore requires user experience strategists to look beyond the mere design of the web site and pay attention to more general management and marketing issues

Egger, F.N. & D. Abrazhevich (2001). **Security & Trust: Taking Care of the Human Factor.** *Electronic Payment Systems Observatory Newsletter*, Vol. 9, Joint Research Center of the European Commission, Seville (Spain).

In the e-business chain, the last link that needs to be convinced of the security of an online transaction is the end-user. That is why this article puts forward a user-centred perspective of the problem of trust in online payments, derived from the discipline of Human-Computer Interaction (HCI). We will first offer a general account of e-commerce system design, showing that there is more to trust than only security. The last part gives some recommendations on what can be done to increase consumers' trust.

Shelat, B. & Egger, F.N. (2002). **What Makes People Trust Online Gambling Sites?** *Proceedings of the CHI'02 Conference on Human Factors in Computing Systems*, 852-853. New York: ACM Press.

A validated model of trust was used as a framework for an empirical study to identify on- and offline factors that influence gamblers' perception of an online casino's trustworthiness. The results suggest that the quality with which casinos address gamblers' trust concerns by providing appropriate content is the prime factor. However, designing for trust must be part of a consistent strategy that also involves customer service and usability.

Egger, F.N. (2002). **Consumer Trust in E-Commerce: From Psychology to Interaction Design.** In: J.E.J. Prins et al. (Eds.). *Trust in Electronic Commerce: The Role of trust from a Legal, an Organizational and a Technical Point of View*. Kluwer Law International.

This chapter discusses the issue of trust in business-to-consumer e-commerce. Starting from psychological accounts of trust in romantic and business relationships, the focus will be on trust in electronically-mediated forms of commerce. A model of trust for e-commerce (MoTEC) will be presented as an attempt to classify off- and online factors observed to affect consumers' feelings of trust towards an online vendor. The last part describes how such a model can be applied to the design of e-commerce user interfaces. What is noteworthy is that the issue of trust will be looked at exclusively from the consumer perspective, as opposed to the legal and technology perspectives found in other chapters of this book. The approach adopted in this chapter stems from the discipline of Human-Computer Interaction (HCI), given its concern for end-users and its stress on design knowledge.

Egger, F.N. (2003). **Deceptive Technologies: Cash, Ethics & HCI.** *SIGCHI Bulletin*, Vol. 35, Issue 2, May-June 2003, p.11, ACM Press.

Everyone working on web projects will have noticed how HCI and marketing get increasingly integrated to deliver positive and memorable experiences to users. Since my research has looked at the factors that make people trust e-commerce sites, I've had many opportunities to observe how simple design tricks can affect people's attitude towards a website. Ultimately, my findings will help online businesses implement a communication strategy geared to minimise perceived risks and increase their professionalism.

APPENDIX 2: Background of the evaluators in Studies 1 and 2

Legend

- 1 = not at all familiar/experienced
 2 = not very familiar/experienced
 3 = familiar/experienced
 4 = quite familiar/experienced
 5 = very familiar/experienced

Evaluator labels	General HCI experience	Familiarity with heuristic evaluations	Experience conducting expert reviews	Experience evaluating websites	Experience evaluating e-commerce websites	Experience with trust issues in e-commerce
------------------	------------------------	--	--------------------------------------	--------------------------------	---	--

Study 1

A1	1-2 years	2	3	4	2	1
A2	< 1 year	4	2	1	1	1
A3	< 1 year	3	2	2	2	2
A4	1-2 years	4	3	3	2	2
A5	1-2 years	5	4	3	3	3
A6	1-2 years	5	3	4	4	4
A7	> 5 years	4	2	2	2	3
A8	2-3 years	4	4	4	4	5
A9	3-4 years	3	2	2	2	2
A10	2-3 years	3	3	4	4	5

Study 2

B1	1-2 years	3	2	3	2	1
B2	1-2 years	4	3	4	2	1
B3	3-4 years	5	2	3	3	2
B4	1-2 years	4	4	4	3	3
B5	< 1 year	3	2	3	2	3
B6	2-3 years	4	3	3	3	5
B7	< 1 year	1	1	1	1	1
B8	2-3 years	1	2	1	1	2
B9	3-4 years	3	2	2	1	2
B10	2-3 years	3	2	2	2	5

APPENDIX 3A: Raw data of the user tests for the Flower Website

Appendices 2a and 2b present the raw data from the user tests. The observed problems are classified according to which MoTEC component they refer to. Each problem is followed by the number of the participant(s) who encountered it. To allow for the comparison presented in Appendix 4, the problems have been categorised according to which CheckTEC checklist item they correspond to. The frequency of a given problem is also noted.

Problem Description	Corresponding Checklist Item Frequency	Checklist Item #
Branding		
Splash screen blocks access to information (C4, C7, C12, C18)	C4, C7, C12, C18 = 4	2.1.1
Graphic design is amateurish (C2, C6, C9, C10, C13, C14, C16, C17) Falling hearts are annoying (C2, C4, C5, C6, C8, C9, C11, C12, C13, C14, C16, C17) Colour scheme not harmonious (C4, C14)	C2, C4, C5, C6, C8, C9, C10, C11, C12, C13, C14, C16, C17 = 13	2.1.2
Spelling mistake (C17)	C17 = 1	2.1.6
Some English expressions are difficult to understand (C12)	C12 = 1	2.1.7
Mother's Day info is out of date (C12)	C12 = 1	2.1.8
Usability		
Input field for delivery time expected next to field for delivery day (C1, C4, C5, C7, C8, C9, C11, C15, C16, C17, C18) Bottom navigation tabs are too hidden (C3, C12, C14) No link to checkout screen from the basket overview (C4, C11, C12, C13, C14, C18)	C1, C3, C4, C5, C7, C8, C9, C11, C12, C13, C14, C15, C16, C17, C18 = 15	2.2.3
No link to homepage from order pages (C6, C10)	C6, C10 = 2	2.2.4
Unclear labelling of flower categories (C1, C3, C6, C7, C14, C15) Unclear labelling of the <i>Feedback</i> button (C3) Search engine usage and results are unclear (C5)	C1, C3, C5, C6, C7, C14, C15 = 7	2.2.5
After adding an item to basket, one is not transferred to the basket overview (C5, C6, C8, C10, C14, C16, C17, C18) Confusing labelling of the <i>Add</i> and <i>Checkout</i> functions (C14, C18)	C5, C6, C8, C10, C14, C16, C17, C18 = 8	2.2.8
Delivery date changes when edited from the basket screen (C2, C7, C9, C13, C16)	C2, C7, C9, C13, C16 = 5	2.2.9
Counties/states menu is not dynamically linked to selected country (C2, C5, C10) Need to type "none" if no message required (C5, C7, C13, C17) Removing an item from basket is difficult (C9, C16)	C2, C5, C7, C9, C10, C13, C16, C17 = 8	2.2.10
Some personal information needs to be filled in both in the flower and payment sites (C4)	C4 = 1	2.2.11
No currency conversion	C14 = 1	--

Company

Missing photo of management team (C1, C7, C17) Missing photo of shop and/or van (C1, C7, C17) Contact details not prominent (C2, C7, C8, C14) Missing name of contact person (C15)	C1, C2, C7, C8, C14, C15, C17 = 7	3.1.1
Unclear whether flower shop and bears/balloon shop are the same (C2, C5, C12, C15, C17) An <i>About Us</i> section would be helpful (C5, C13, C18)	C2, C5, C12, C13, C15, C17, C18 = 7	3.1.2
Lack of information regarding affiliation to an interna- tional network of florists (C2, C17)	C2, C17 = 2	3.1.7
Customer testimonials assumed to be biased and unre- liable (C5, C7, C17)	C5, C7, C17 = 3	--

Products & Services

Missing information about flowers' origin (C1) Descriptions are incomplete (C2, C16) Confusing presence of non-flower products (C13)	C1, C2, C13, C16 = 4	3.2.1
Photos of flowers not clear enough (C5, C9, C10, C12, C13, C16) Photos assumed to represent the most expensive bou- quet type (C15, C17)	C5, C9, C10, C12, C13, C15, C16, C17 = 8	3.2.2
No guarantee that the bouquets would indeed match their description (C6, C13, C14)	C6, C13, C14 = 3	3.2.3
Not possible to customise own flower creation	C5, C7, C9, C10 = 4	--

Security

Security policy too hidden (C10)	C10 = 1	3.3.1
Distance selling regulation regarding returning prod- ucts does not apply to flowers (C1) Policies are too long (C1, C15, C16) Policies skipped as "they could write anything any- way" (C2, C18) Little information about the payment intermediary (C4)	C1, C2, C4, C15, C16, C18 = 6	3.3.2
Address form is not encrypted (C4)	C4 = 1	3.3.3
"128 bit technology" could be explained in lay terms (C13, C17)	C13, C17 = 2	3.3.4
No alternative means of payment (C18)	C18 = 1	3.3.5
Shopsafe seal ok but its homepage looks more like a directory: no real commitment to security (C12, C16) WHICH Web Trader seal leads to an announcement that the service has been discontinued (C17)	C12, C16, C17 = 3	3.3.8

Privacy

Policy too hidden (C10)	C10 = 1	3.4.1
Policy is too long (C1, C8, C9)	C1, C8, C9 = 3	3.4.2
Some items in the order form are not relevant C2, C11 <i>Remember my details</i> option is pre-selected (cookie installed) C17	C17 = 3	3.4.6

APPENDIX 3B: Raw data of the user tests for the Perfume Website

Problem Description	Corresponding Checklist Item Frequency	Checklist Item #
Branding		
Too much clutter (C2, C3, C5, C7, C9)	C2, C3, C5, C7, C9 = 5	2.1.1
Unclear use of product images (C1) Notices too many different fonts on homepage (C3, C8, C11) Looks like a tabloid (C5) Doesn't look professional (C6, C7, C10, C14, C17, C18) Some logos are not well designed (C12)	C1, C3, C5, C6, C7, C8, C10, C11, C12, C14, C17, C18 = 12	2.1.2
Design gives no feeling of fragrance (C1) Logo looks like a banner (C3, C12) Graphic design is not appropriate (C7, C8) Looks too American for a European site (C14)	C1, C3, C7, C8, C12, C14 = 6	2.1.3
Unclear introduction (C2)	C2 = 1	2.1.4
Textual layout inaccuracies (C1, C4, C6, C8, C12)	C1, C4, C6, C8, C12 = 5	2.1.5
Usability		
Missed error message feedback because of poor layout (C12)	C12 = 1	2.2.1
Category labels unclear (C3, C9)	C3, C9 = 2	2.2.2
Ordering sequence not logical (C2, C3, C4) Lack of consistency in structure (C5)	C2, C3, C4, C5 = 4	2.2.3
Navigation menu change according to section (C5) Inconsistent use of product type or brand for navigation (C6, C7)	C5, C6, C7 = 3	2.2.4
Confusing left and right hand navigation (C1) Difficult to locate item in search results list (C2, C3, C5, C6, C10, C11, C13) Searching by brand not well supported (C17)	C1, C2, C3, C5, C6, C10, C11, C13, C17 = 9	2.2.5
Need to specify quantity = 1 is not intuitive (C12, C13, C14, C15, C16) Click on brand name to add to cart does not always work (C17)	C12, C13, C14, C15, C16, C17 = 6	2.2.7
Annoying automatic reload after editing item (C4)	C4 = 1	2.2.9
Could not specify EU as shipping destination (C14)	C14 = 1	2.2.10
No currency converter from GBP to EUR (C4)	C4 = 1	--

Company

Not prominent contact details (C5, C8, C16) Does not know if company has offline offices (C7, C13) Suspicious that postal address is a PO box (C12, C17) Misses details and name of a contact person (C15)	C5, C7, C8, C12, C13, C15, C16, C17 = 8	3.1.1
Misses information about where the perfumes come from (C1, C7, C17) Not enough company background (C5, C15)	C1, C5, C7, C15, C17 = 5	3.1.2
Wonders if it's really as successful as it claims (C5)	C5 = 1	3.1.6
Misses information about being part of a known off-line company or organization (C12, C13, C18)	C12, C13, C18 = 3	3.1.7
Misses customer testimonials (C16)	C16 = 1	--

Products & Services

Fragrance brand not always prominent (C2) Product descriptions too short (C7, C16) Some perfumes for men in list for women (C7)	C2, C7, C16, C17 = 4	3.2.1
Photos of perfumes are too small (C1, C2, C3, C15)	C1, C2, C3, C15 = 4	3.2.2
Descriptions are copied from product leaflets and not evaluated by the site (C15)	C15 = 1	3.2.3
Not sure what is content and what is advertisement (C1, C7)	C1, C7 = 2	3.2.4
Suspects that there will be higher shipping costs to compensate low prices (C1)	C1 = 1	3.2.5

Security

"110% Security" logo is misleading (C3, C5, C13, C17) Misses a prominent link to the security policy (C10, C17, C18)	C3, C5, C10, C13, C17, C18 = 6	3.3.1
"110% Security" logo is not linked to a security policy (C3, C4, C13)	C3, C5, C13 = 3	3.3.2
Security information contains too much jargon and is not easy to understand (C3)	C3 = 1	3.3.4
Missed third party seal (C3)	C3 = 1	3.3.8

Privacy

Had problems locating privacy information in Terms & Conditions (C3, C8, C18) Did not find link to privacy policy (C10, C12)	C3, C8, C10, C12, C18 = 5	3.4.1
Assumes that data will be sold anyway (C4) Policy not written in a reader-friendly way (C5)	C4, C5 = 2	3.4.2

APPENDIX 4: Comparative Results for the Unguided, Checklist and User Tests Conditions for the Flower and the Perfume Websites

Legend

- Problems that were predicted by unguided experts and checklist users but **not** in user tests
- Problems that were predicted **in all three conditions**
- Problems that were **only** predicted by checklist users
- Problems that were predicted by checklist users and user tests but **not** by unguided experts

Units

The numbers refer to the proportion of subjects (%) who noted an item as a problem.

			Flower Website			Perfume Website			
	#	checkl. #	unguided	checklist	user tests	unguided	checklist	user tests	
P I F	1	1.1	0	50	0	0	30	0	
	2	1.2	0	70	0	0	80	0	
Branding	3	2.1.1	0	10	22.22	30	10	27.77	
	4	2.1.2	80	60	72.22	80	60	66.66	
	5	2.1.3	30	40	0	80	40	33.33	
	6	2.1.4	0	70	0	0	70	5.55	
	7	2.1.5	40	50	0	20	90	27.77	
	8	2.1.6	0	50	5.55	10	40	0	
	9	2.1.7	20	40	5.55	0	30	0	
	10	2.1.8	0	40	5.55	20	30	0	
	Usability	11	2.2.1	0	40	0	30	10	5.55
		12	2.2.2	70	40	0	0	40	11.11
13		2.2.3	0	40	83.33	0	20	22.22	
14		2.2.4	50	30	11.11	50	40	16.66	
15		2.2.5	40	30	38.88	30	30	50	
16		2.2.6	0	40	0	0	20	0	
17		2.2.7	0	10	0	0	10	33.33	
18		2.2.8	0	10	44.44	0	10	0	
19		2.2.9	10	10	27.77	0	20	5.55	
20		2.2.10	40	50	44.44	0	40	5.55	
21		2.2.11	40	30	5.55	40	20	0	
Company	22	3.1.1	30	20	38.88	30	50	44.44	
	23	3.1.2	50	10	38.88	30	40	27.77	
	24	3.1.3	0	30	0	0	40	0	
	25	3.1.4	not applicable			0	60	0	
	26	3.1.5	0	20	0	0	20	0	
	27	3.1.6	0	60	0	0	20	5.55	
	28	3.1.7	10	20	11.11	0	60	16.66	
	29	3.1.8	0	50	0	0	0	0	

Products & Services	30	3.2.1	0	20	22.22	10	0	22.22
	31	3.2.2	30	60	44.44	0	60	22.22
	32	3.2.3	10	100	16.66	10	90	5.55
	33	3.2.4	0	90	0	0	40	11.11
	34	3.2.5	0	90	16.66	0	70	5.55
	35	3.2.6	0	100	0	0	100	0
	36	3.2.7	0	90	0	0	80	0
	37	3.2.8	40	90	0	10	70	0
Security	38	3.3.1	20	60	5.55	40	40	33.33
	39	3.3.2	0	40	33.33	40	20	16.66
	40	3.3.3	0	70	5.55	0	20	0
	41	3.3.4	0	80	11.11	0	50	5.55
	42	3.3.5	10	80	5.55	20	30	0
	43	3.3.6	not applicable			0	20	0
	44	3.3.7	0	70	0	0	70	0
	45	3.3.8	0	90	16.66	0	90	5.55
Privacy	46	3.4.1	10	50	5.55	20	40	27.77
	47	3.4.2	0	20	16.66	0	30	11.11
	48	3.4.3	0	100	0	0	100	0
	49	3.4.4	0	20	0	0	30	0
	50	3.4.5	0	60	0	0	60	0
	51	3.4.6	0	30	16.66	0	30	0
Rel. Man.	52	4.1	0	40	0	0	30	0
	53	4.2	0	80	0	0	30	0
	54	4.3	0	100	0	0	40	0
	Not in checklist		20	0	16.66	20	0	5.55
	Not in checklist		0	0	22.22	30	0	5.55

SUMMARY

Business-to-consumer electronic commerce on the Internet has revolutionised the purchase of products and services by giving consumers round the clock access to worldwide providers. However, B2C e-commerce has also shown to be associated with a myriad of factors hindering adoption and usage by private customers. Such factors include concerns regarding security and privacy, the unfamiliarity of some online services, lack of direct interaction with products, salespeople and fellow shoppers and the generally low credibility of online information. These factors were collectively defined as “trust issues”, as they refer to a purchase decision customers have to make in a situation of uncertainty and risk.

The first objective of this research was to build up substantive knowledge about which specific factors make customers trust e-commerce websites. The second objective was to build up and validate methodological knowledge in the form of tools that HCI practitioners can use to design and evaluate trust-shaping factors in e-commerce websites. On the basis of literature on trust and e-commerce surveys, a first model of trust in e-commerce (MoTEC) was developed. Through user tests, the initial model was refined to increase its descriptive power. The final MoTEC model contains four main dimensions, containing components and sub-components. It is structured as follows:

1. **Pre-interactional Filters** refer to factors that may affect a person’s trust in an online vendor before accessing its website. They are composed of *User Psychology* and *Pre-purchase Knowledge*.
2. **Interface Properties** refer to surface cues in the user interface, namely graphic design and ease of use. The corresponding components are *Branding* and *Usability*.
3. **Informational Content** refers to the different types of information contained in the website. There are two main types of information, each having two sub-components: *Competence*, containing information about Company and Products & Services; and *Risk*, containing information about Security and Privacy.
4. **Relationship Management** refers to interactions with the company over time, both before and after a purchase (*Pre-purchase* and *Post-purchase Interactions*).

The MoTEC model was then used to derive a Trust Toolbox containing a suite of concrete tools for designers. The first tool was called GuideTEC and was a set of trust design principles and guidelines. The second, ChekTEC, was a checklist evaluators can use to diagnose the trust performance of a website. Thirdly, QuoTEC was a questionnaire that can be administered to representative users, either after a user test with a facilitator or on its own.

The QuoTEC questionnaire was used to collect user feedback in two different studies. The first dealt with the service industry and examined user reactions to six hotel websites in Switzerland and in the Netherlands. The second dealt with the retail industry and examined user reactions to two computer websites and two online bookstores in the United Kingdom. The main objective of this double study was to reduce the number of items in the questionnaire. The original set of 23 items was reduced to 15 items, while keeping the effect of the reduction on the explained variance minimal. These studies also uncovered differences in the factors underlying and predicting trust in the two industries. Trust in hotel websites was found to be best predicted by the components Company and Products & Services. On the other hand, trust in retail websites was best predicted by the components Privacy, Products & Services, Company and Usability. This difference was accounted for by the fact that hotel guests will physically stay at the hotel and interact with its staff and, often, only make a booking, while retail customers actually buy goods online from a company they will only interact with online.

A validation study demonstrated that evaluators using the CheckTEC checklist found four times as many problems as unguided evaluators, in half the time. Also, checklist-guided evaluators paid attention to a greater range of factors than unguided ones, who mostly noted factors related to Branding and Usability. Compared with the results from user tests, checklist-guided experts correctly predicted about 90% of all observed problems. The QuoTEC questionnaire was also tested to compare the results produced after its administration after a user test, with a facilitator, and those produced in a remote evaluation set-up, without a facilitator. The findings indicate some differences in the results that were mostly due to the questionnaire-only participants not systematically following the set scenarios. Building in some controls in the remote evaluation would address this issue by forcing participants to evaluate websites more thoroughly, which would increase the reliability of the results.

Given the differences observed between the hotel and the retail websites, the MoTEC model should be applied to more varied types of industries and validated for each of them individually. This would show which constellation of factors are the most important in each case. As the GuideTEC guidelines were only indirectly validated through the checklist items, future research should examine the effect each guideline makes on perceived trustworthiness. In conclusion, concrete examples illustrate the ethical implications of making websites appear to be trustworthy.

SAMENVATTING

De ontwikkeling van op particulieren gerichte verkoop via het Internet heeft een revolutie betekend op gebied van de handel in produkten en diensten, door de consumenten 24 uur per dag toegang te bieden tot leveranciers overal ter wereld. Op particulieren gerichte e-commerce wordt echter ook geassocieerd met een groot aantal factoren die de acceptatie en het gebruik door eindklanten bemoeilijken, waaronder de bezorgdheid over veiligheid en privacy, de onbekendheid van sommige online diensten, het gebrek aan directe interactie met produkten, verkopers en andere klanten en de over het algemeen verminderde geloofwaardigheid van online informatie. Deze factoren zijn samengebracht onder de noemer “vertrouwensproblematiek” omdat ze stuk voor stuk te maken hebben een aankoopbeslissing die klanten moeten maken in een onzekere en risicovolle situatie.

Het eerste doel van dit onderzoek bestond uit het opbouwen van fundamentele kennis over welke specifieke factoren ervoor zorgen dat klanten vertrouwen hebben in bepaalde e-commerce websites. Het tweede doel was gericht op het opbouwen en valideren van methodologische kennis in de vorm van tools voor mensen die Human-Computer Interaction in het veld toepassen en die kunnen worden gebruikt voor het ontwerpen en evalueren van vertrouwensgerelateerde factoren in e-commerce websites. Op basis van literatuur over vertrouwen en e-commerce werd een eerste model voor vertrouwen in e-commerce (MoTEC) ontwikkeld. Door middel van gebruikerstesten werd dit eerste model verfijnd om het de werkelijkheid beter te laten beschrijven. Het uiteindelijke MoTEC-model bestaat uit 4 dimensies, die componenten en sub-componenten omvatten in de volgende structuur:

1. **Pre-interactiefilters** verwijzen naar factoren die a priori iemands vertrouwen in een online verkoper kunnen beïnvloeden, dus voordat deze de website gezien heeft, en die bestaan uit Gebruikerspsychologie en achtergrondkennis.
2. **Interface eigenschappen** verwijzen naar oppervlakte-elementen in de gebruikersinterface, namelijk grafisch design en gebruiksgemak. De bijbehorende componenten zijn *Merkpositionering* and *Gebruiksvriendelijkheid*.
3. **Informatieve Inhoud** verwijst naar de verschillende soorten informatie die op de website worden aangeboden. Er zijn twee hoofdsoorten informatie, die elk twee subcomponenten bevatten: *Competentie*, die informatie bevat over Bedrijf en Produkten & Diensten; en *Risico*, met informatie over Veiligheid en Privacy.

4. Relatiemanagement verwijst naar interacties met het bedrijf in de loop van de tijd, zowel vóór als na aankoop (*Vóór-aankoop-* en *Na-aankoopinteractie*).

Het MoTEC model werd vervolgens gebruikt om een vertrouwentoolbox voor ontwerpers te ontwikkelen. De eerste tool heet GuideTEC en bestaat uit een aantal ontwerpprincipes en richtlijnen voor vertrouwen. De tweede, CheckTEC, is een checklist die website evaluatoren kunnen gebruiken om de mate van vertrouwen in een website te kunnen vaststellen. De derde tool, QuoTEC, is een vragenlijst die aan een representatieve gebruikersgroep kan worden uitgereikt, niet alleen na een gebruikerstest met begeleider, maar ook zonder begeleider als een op zichzelf staande tool.

De QuoTEC vragenlijst werd vervolgens gebruikt in twee verschillende onderzoeken om gebruikersfeedback te vergaren. Het eerste onderzoek richtte zich op de dienstensector en onderzocht de gebruikersreacties op zes hotelwebsites in Zwitserland en Nederland. Het tweede was gericht op de detailhandel en onderzocht de reacties van gebruikers op twee computerwebsites en twee online boekhandels in Groot-Brittannië. Het hoofddoel van dit dubbelonderzoek was het verminderen van het aantal vragen op de vragenlijst QuoTEC. Het oorspronkelijke aantal van 23 vragen werd hierdoor teruggebracht naar 15, terwijl het effect hiervan op de verklaarde variantie minimaal gehouden werd. Deze onderzoeken lieten ook verschillen zien in de factoren die vertrouwen voorspellen en beïnvloeden. Vertrouwen in hotelwebsites werd vooral voorspeld aan de hand van de componenten Bedrijf en Producten & Diensten. Daarentegen werd vertrouwen in de detailhandel voornamelijk beïnvloed door de componenten Privacy, Producten & Diensten, Bedrijf en Gebruiksvriendelijkheid. Dit verschil wordt verklaard door het feit dat hotelgasten vaak alleen de boeking online doen, daarna echt in het hotel zullen verblijven en daarbij contact zullen hebben met het hotelpersoneel, terwijl klanten die online goederen kopen van een detailhandelaar, alleen maar online contact met dat bedrijf zullen hebben.

Een validatieonderzoek toonde aan dat evaluatoren die CheckTEC gebruikten, vier keer zo veel problemen vonden als evaluatoren zonder leidraad, en dat in de helft van de tijd. Ook lieten CheckTEC-evaluatoren op een groter aantal factoren dan de evaluatoren zonder CheckTEC, die voornamelijk factoren noteerden die verband hielden met Merkpositionering en Gebruiksvriendelijkheid. Vergeleken met de resultaten van de gebruikerstesten, voorspelden de CheckTEC-evaluatoren ongeveer 90% van alle waargenomen problemen. Er werd ook een test uitgevoerd om de resultaten te vergelijken tussen een situatie waarbij CheckTEC werd uitgedeeld na een gebruikerstest met begeleider en een evaluatiestudie op afstand, zonder begeleider. De bevindingen laten enkele verschillen zien in de resultaten, die voornamelijk te wijten zijn aan het feit dat de deelnemers in de onbegeleide situatie niet systematisch de voorgeschreven scenario's volgden. Dit probleem kan worden aangepakt door een aantal controles in te bouwen die de deelnemers in de onbegeleide situatie dwingen de websites grondiger te onderzoeken, hetgeen de betrouwbaarheid van de resultaten zal vergroten.

Vanwege de verschillen die tussen de evaluaties van de hotel- en de detailhandelwebsites naar voren kwamen zou het MoTEC model op meerdere sectoren in de industrie moeten worden toegepast en onafhankelijk voor iedere sector moeten worden gevalideerd. Dit zal laten zien welke samenhang van factoren het belangrijkste zijn voor elke sector. Aangezien de GuideTEC richtlijnen alleen maar indirect werden gevalideerd door middel van de punten op de checklist, zou toekomstige studie het effect moeten onderzoeken dat dat elk van de richtlijnen heeft op waargenomen vertrouwenswaardigheid. Ten slotte laten concrete voorbeelden de ethische implicaties zien van het betrouwbaar laten *lijken* van websites.

CURRICULUM VITÆ

- 29 Sept. 1975 Born in Lausanne (CH)
- 1991-1994 Gymnase de CESSRIVE, Lausanne (CH)
→ Swiss Federal Maturity & Baccalaureate (Latin-English)
- 1994-1997 City University, London (UK)
→ BSc (Hons) Psychology & Philosophy
- 1997-1998 University College London (UK)
→ MSc Human-Computer Interaction
- 1999-2003 J.F. Schouten School for User-System Interaction Research
Department of Technology Management
Eindhoven University of Technology (NL)
→ PhD Human-Computer Interaction
- Since Sept. 2003 HCI & user experience consultant
Geneva (CH)
- Updates: <http://www.ecommuse.com>
Contact: egger@ecommuse.com