

ISSN 0265-1335

Volume 23 Number 4 2006

International Marketing Review

**Issues and advances in
international marketing research**

Guest Editors: Rudolf R. Sinkovics and
Thomas Salzberger



www.emeraldinsight.com



International Marketing Review

ISSN 0265-1335

Volume 23
Number 4
2006

Issues and advances in international marketing research

Guest Editors

Rudolf R. Sinkovics and Thomas Salzberger

Access this journal online _____	347
Editorial advisory board _____	348
Guest editorial _____	349
The psychometric properties of eTail quality: an international investigation across product categories <i>Albert Caruana and Michael T. Ewing</i> _____	353
An illustration of modeling moderating variables in cross-national studies <i>Rosanna Garcia and Destan Kandemir</i> _____	371
Reconsidering the problem of data equivalence in international marketing research: contrasting approaches based on CFA and the Rasch model for measurement <i>Thomas Salzberger and Rudolf R. Sinkovics</i> _____	390
Researching the socio-cultural context: putting social representations theory into action <i>Elfriede Penz</i> _____	418

CONTENTS

Access this journal electronically

The current and past volumes of this journal are available at:

www.emeraldinsight.com/0265-1335.htm

You can also search more than 100 additional Emerald journals in Emerald Fulltext (www.emeraldinsight.com/ft) and Emerald Management Xtra (www.emeraldinsight.com/emx)

See page following contents for full details of what your access includes.



CONTENTS
continued

**Employing information communication technologies
to enhance qualitative international marketing
enquiry**

Sharon Loane, Jim Bell and Rod McNaughton _____ **438**

Access this journal online

www.emeraldinsight.com/imr.htm



As a subscriber to this journal, you can benefit from instant, electronic access to this title via Emerald Fulltext or Emerald Management Xtra. Your access includes a variety of features that increase the value of your journal subscription.

How to access this journal electronically

To benefit from electronic access to this journal, please contact support@emeraldinsight.com. A set of login details will then be provided to you. Should you wish to access via IP, please provide these details in your e-mail. Once registration is completed, your institution will have instant access to all articles through the journal's Table of Contents page at www.emeraldinsight.com/0265-1335.htm. More information about the journal is also available at www.emeraldinsight.com/imr.htm.

Our liberal institution-wide licence allows everyone within your institution to access your journal electronically, making your subscription more cost-effective. Our web site has been designed to provide you with a comprehensive, simple system that needs only minimum administration. Access is available via IP authentication or username and password.

Key features of Emerald electronic journals

Automatic permission to make up to 25 copies of individual articles

This facility can be used for training purposes, course notes, seminars etc. This only applies to articles of which Emerald owns copyright. For further details visit www.emeraldinsight.com/copyright

Online publishing and archiving

As well as current volumes of the journal, you can also gain access to past volumes on the internet via Emerald Fulltext or Emerald Management Xtra. You can browse or search these databases for relevant articles.

Key readings

This feature provides abstracts of related articles chosen by the journal editor, selected to provide readers with current awareness of interesting articles from other publications in the field.

Non-article content

Material in our journals such as product information, industry trends, company news, conferences, etc. is available online and can be accessed by users.

Reference linking

Direct links from the journal article references to abstracts of the most influential articles cited. Where possible, this link is to the full text of the article.

E-mail an article

Allows users to e-mail links to relevant and interesting articles to another computer for later use, reference or printing purposes.

Structured abstracts

Emerald structured abstracts provide consistent, clear and informative summaries of the content of the articles, allowing faster evaluation of papers.

Additional complimentary services available

Your access includes a variety of features that add to the functionality and value of your journal subscription:

E-mail alert services

These services allow you to be kept up to date with the latest additions to the journal via e-mail, as soon as new material enters the database. Further information about the services available can be found at www.emeraldinsight.com/alerts

Connections

An online meeting place for the research community where researchers present their own work and interests and seek other researchers for future projects. Register yourself or search our database of researchers at www.emeraldinsight.com/connections

Emerald online training services

You can also access this journal online. Visit www.emeraldinsight.com/training and take an Emerald online tour to help you get the most from your subscription.

Choice of access

Electronic access to this journal is available via a number of channels. Our web site www.emeraldinsight.com is the recommended means of electronic access, as it provides fully searchable and value added access to the complete content of the journal. However, you can also access and search the article content of this journal through the following journal delivery services:

EBSCOHost Electronic Journals Service
ejournals.ebsco.com

Informatics J-Gate
www.j-gate.informindia.co.in

Ingenta
www.ingenta.com

Minerva Electronic Online Services
www.minerva.at

OCLC FirstSearch
www.oclc.org/firstsearch

SilverLinker
www.ovid.com

SwetsWise
www.swetswise.com

Emerald Customer Support

For customer support and technical help contact:

E-mail support@emeraldinsight.com

Web www.emeraldinsight.com/customercharter

Tel +44 (0) 1274 785278

Fax +44 (0) 1274 785204

EDITORIAL ADVISORY BOARD

David Ballantyne

University of Otago, New Zealand

Professor Jean J. Boddewyn

The City University of New York, USA

Dr Marylyn Carrigan

University of Birmingham, UK

Professor Tefvic Dalgic

University of Texas at Dallas, USA

Professor Adamantios Diamantopoulos

University of Vienna, Austria

Professor Krzysztof Fonfara

Poznań University of Economics, Poland

Dr Debra Harker

University of the Sunshine Coast, Queensland,
Australia

Dr Michael Harker

University of the Sunshine Coast, Queensland,
Australia

Professor Nigel J. Holden

Kassel International Management School, Germany

Dr Gavin Jack

University of Leicester, UK

Professor Constantine S. Katsikeas

Leeds University, UK

Sam Okoroafo

University of Toledo, USA

Professor Stan Paliwoda

University of Strathclyde, UK

Dr Michael Jay Polonsky

Victoria University, Melbourne, Australia

Professor K.N. Rajendran

The University of Northern Iowa, USA

Professor Ilkka Ronkainen

Georgetown University, USA

Professor Saeed Samiee

University of Tulsa, USA

Professor Bodo B. Schlegelmilch

Wirtschaftsuniversität, Wien, Austria

Professor Vern Terpstra

University of Michigan, USA

Professor Sandra Vandermerwe

Imperial College, University of London, UK

EDITORIAL REVIEW PANEL

Professor John W. Cadogan

Loughborough University, UK

Professor Paul Chao

Eastern Michigan University, USA

Dr Irvine Clarke III

James Madison University, USA

Dr John B. Ford

Old Dominion University, Norfolk, USA

Professor E. Kaynak

Pennsylvania State University, USA

Dr Fredric Kropp

Monterey Institute of International Studies, USA

Professor Leonidas Leonidou

University of Cyprus, Cyprus

Professor Hans Muhlbacher

University of Innsbruck, Austria

Dr Helen Perks

UMIST, UK

Professor Nigel Piercy

University of Warwick, UK

Dr Ronald Savitt

Portland, Oregon, USA

Aviv Shoham

University of Haifa, Israel

Dr K. Sivakumar

Lehigh University, Pennsylvania, USA

Dr Anne Souchon

Loughborough University, UK

Professor Chris Styles

The University of Sydney, Australia

Dr Isabelle Szmigin

University of Birmingham, UK

Professor Peter G.P. Walters

The Hong Kong Polytechnic University,
Hong Kong

About the Guest Editors Rudolf R. Sinkovics is Senior Lecturer in International Business at The University of Manchester, Manchester Business School, UK. His research centres on interorganisational governance, the role of ICT in firm internationalisation and research methods in international business. His work has appeared in the *Journal of International Business Studies*, *Journal of Business Research*, *International Business Review*, and other journals. He received his PhD from Vienna University of Economics and Business Administration, Austria, in 1998.

Thomas Salzberger is Assistant Professor in Marketing Management at the Vienna University of Economics and Business Administration, Vienna, Austria. His main areas of research focus on market research, methodology and measurement in marketing. His work has appeared in the *Journal of Advertising*, *Australasian Marketing Journal*, *Rasch Measurement Transactions*, and other journals. He received his PhD from Vienna University of Economics and Business Administration, Austria, in 1998.

Introduction to the special issue on “Issues and advances in international marketing research”

“Issues and advances in international marketing research” is the title of this special issue of *IMR*. With an unprecedented proliferation of research methodologies, processes and procedures over the last couple of years, some readers might find that there is not much room for further advancement.

In fact, this theme has been selected because, despite progress and methodological advancements, liberalising markets and technology impacts prompt for even more methodological and conceptual rigour when conducting research across borders. It has been argued that market dynamics and global competitive pressures create new challenges for researchers and companies (e.g. Craig and Douglas, 2001) and call for a diversification of research methods (Nakata and Huang, 2005). International marketing research is challenged in two significant ways. Firstly, it needs to fulfil the intrinsic quest for extending and also transcending boundaries, which is characteristic of any scientific research. The process of internationalisation and globalisation, however, represents a powerful external trigger specifically for cross-national research. Secondly, international research entails additional methodological problems due to the increased complexity making it exceptionally prone to errors and misleading conclusions. These challenges persist, irrespective of paradigmatic views, whether following a qualitative or quantitative perspective.

The call for papers solicited topics including reliability and validity issues in international marketing research, safeguarding equivalence in cross-national/cross-cultural research, construct measurement issues, development of frameworks for construct measurement, demonstration of appropriate modelling techniques, innovative qualitative designs, techniques or experiments and applications of tools or techniques for cross-cultural studies from other disciplines. The call for papers prompted submission of 19 papers. This somewhat modest response was surprising to us, but may indicate that the next wave of methodological advancements and methodological pluralism is yet to come. Most of the contributions were following quantitative methodology and while companies are said to direct an increasing share of their market research budgets to qualitative designs, the submissions to this special issue did not

mirror this development. Additionally, our invitation for comparative ethnography was not taken up, and we believe that there is still much room for the application of tools and techniques such as simulations. In all, “advancement in international marketing research” appears to be incremental, focusing on established methodologies, rather than radical, in terms of pushing us towards fundamentally new research avenues.

Of the papers received, 13 passed a first screening by the guest editors and were subsequently subjected to a more thorough review process, strictly following the *IMR* procedures (2-3 double-blind peer-reviews). One of these papers is authored by the guest editors, yet was as a matter of course also subjected to the review process. The review process resulted in five manuscripts accepted for publication in this special issue. All manuscripts include empirical foundations to varying degrees.

The first paper of this special issue by Albert Caruana and Michael T. Ewing looks at the psychometric properties of eTail quality across different product categories in three countries. The testing procedures for conceptual and psychometric equivalence are a useful showcase for equivalence testing. This paper also makes a strong argument in favour of replication and extension research in International Marketing Research. While scholars agree that building on established work is worthwhile and scientifically meritorious, published papers in marketing journals provide scant evidence. This paper is a replication with extension that examines the applicability of online service quality. It tests whether earlier US based findings can be generalised to other countries.

Rosanna Garcia and Destan Kandemir explore how moderation can and should be modeled in cross-national/cultural contexts. They build on a multi-national study of consumer involvement to demonstrate appropriate methods for modelling different types of moderation. Specifically, they demonstrate how to identify form and strength moderators. Form moderators are modelled using multiplicative interactions while strength moderators are modelled using multi-group SEM. The authors provide a valuable research framework which should help to utilise their approach with cross-country datasets.

The guest editors to this special issue, Thomas Salzberger and Rudolf Sinkovics, are “Reconsidering the problem of data equivalence in international marketing research”. The paper contrasts approaches of equivalence testing based on CFA and the Rasch model for measurement by reanalysing a cross-country dataset on technophobia. Data equivalence is closely connected to the measurement theory employed. While most international marketing research builds on classical test theory, recently, approaches based on item response theory, or latent trait theory (LTT), have been suggested as a promising alternative to this standard paradigm (Schaffer and Riordan, 2003; Singh, 2004). Given very attractive features of this paradigm, which is well adopted in other disciplines, the authors introduce the Rasch (1960) model, which is a special class of LTT models, to international marketing research.

Elfriede Penz, in her paper “Researching the socio-cultural context: putting social representations theory into action” demonstrates the value of boundary-spanning of academic disciplines. She applies the social representations theory, which originates in the social psychology literature, to the context of international marketing research. This is important because the marketing community will not necessarily be familiar with the reasoning of this branch of the literature, built around some French authors. It is also innovative because the social representations theory is utilised to study what members of a certain culture think of objects and products, which values they associate

with these and which norms they follow. In practical research terms, the contribution is that social representations go beyond the traditional confines of groups and/or multidimensional research methodologies with individual views and perspectives.

Finally, the paper by Sharon Loane, Jim Bell and Rod McNaughton, posits that new information communication technologies (ICT), particularly the Internet, can significantly improve the robustness of qualitative and mixed-method international marketing research. The authors describe and evaluate the application of ICT in a cross-national study of rapidly internationalising small firms. They develop a conceptual framework and outline how the appropriate integration of ICT into the research process can help to refine sample identification and selection procedures, improve response rates and encourage greater respondent “buy-in” to depth interviews.

Acknowledgements

Special thanks go to the *IMR* editor, who took the initiative to make this special issue happen. Furthermore, we are very grateful to those colleagues who were prepared to invest time and effort to review the papers submitted to this special issue. Their contribution has been vital to the development of the papers and to the overall coherence of this special issue. The reviewers are listed in alphabetical order.

Vic Adamowicz, University of Alberta, Canada
David Ballantyne, University of Otago, New Zealand
John Cadogan, Loughborough University, UK
Charles C. Cui, Manchester Business School, UK
Z. Seyda Deligonul, St. John Fisher College, USA
Arne Floh, Wirtschaftsuniversität Vienna, Austria
Andrea Hemetsberger, University of Innsbruck, Austria
Margaret Hogg, Lancaster University, UK
Kevin Iyk N. Ibeh, University of Strathclyde, UK
Kathy Keeling, Manchester Business School, UK
Daekwan Kim, Florida State University, USA
Erich Kirchler, University of Vienna, Austria
Larry Lockshin, University of South Australia, Australia
Frank McDonald, University of Hull, UK
Julie Pallant, Swinburne University, Melbourne, Australia
Saeed Samiee, University of Tulsa, USA
Arnold Schuh, Wirtschaftsuniversität Vienna, Austria
Nitish Singh, California State University, Chico, USA
Anne Souchon, Loughborough University, UK
Chris Styles, University of Sydney, Australia
Alan Tennant, Leeds University, UK
Denis Towill, Cardiff University, UK
Mary Wolfenbarger, California State University Long Beach, USA

Rudolf R. Sinkovics

Manchester Business School, UK, and

Thomas Salzberger

Vienna University of Economics and Business Administration, Austria

References

- Craig, C.S. and Douglas, S.P. (2001), "Conducting international marketing research in the twenty-first century", *International Marketing Review*, Vol. 18 No. 1, pp. 80-90.
- Nakata, C. and Huang, Y. (2005), "Progress and promise: the last decade of international marketing research", *Journal of Business Research*, Vol. 58 No. 5, pp. 611-8.
- Rasch, G. (1960), *Probabilistic Models for Some Intelligence and Attainment Tests*, Danish Institute for Educational Research, Copenhagen.
- Schaffer, B.S. and Riordan, C.M. (2003), "A review of cross-cultural methodologies for organizational research: a best-practices approach", *Organizational Research Methods*, Vol. 6 No. 2, pp. 169-215.
- Singh, J. (2004), "Tackling measurement problems with item response theory: principles, characteristics, and assessment, with an illustrative example", *Journal of Business Research*, Vol. 57 No. 2, pp. 184-208.



The psychometric properties of eTail quality

Properties of
eTail quality

An international investigation across product categories

Albert Caruana

Centre for Communication Technology, University of Malta, Msida, Malta, and

Michael T. Ewing

Department of Marketing, Monash University, Melbourne, Australia

353

Received July 2005

Revised January 2006

Accepted February 2006

Abstract

Purpose – The links between quality and customer satisfaction, retention and loyalty in offline retail settings are well established. It therefore seems fair to posit that quality will also be a determinant of online retailer success. This assumption motivated Wolfinbarger and Gilly to develop a scale for the measurement of “eTail quality”. The paper’s purpose is to validate that scale.

Design/methodology/approach – The study investigates the equivalence of the eTailQ scale across different product categories in three countries. Tests for both conceptual and psychometric equivalence are conducted.

Findings – Findings suggest that eTailQ exhibits conceptual equivalence. It also exhibits psychometric equivalence by providing acceptable levels of reliability, variance extracted and both discriminant and nomological validity.

Originality/value – eTailQ can be used with confidence by online goods and service retailers in English speaking countries outside the USA.

Keywords Australia, Malta, South Africa, Internet shopping, Customer loyalty, Electronic commerce

Paper type Research paper

The critical role of service quality in offline retailing is well documented in the literature (Brady and Cronin, 2001; Dabholkar *et al.*, 1996; Parasuraman *et al.*, 1985, 1988; Rust and Oliver, 1994). However, until recently, little was known about retail quality in online environments. Wolfinbarger and Gilly’s (2003) conceptualisation and measurement of online retail service quality (eTailQ) was based on a sample of respondents in the USA and explicitly excluded participants who purchased services, such as financial services (Wolfinbarger and Gilly, 2003, p. 189). However, eTailQ’s generalizability is currently restricted to one country and it is questionable whether service offerings are also covered. However, globalisation is a powerfully homogenizing influence, as consumers in many countries receive the same intense communications, and global brands become ubiquitous. As Deshpandé (1999) points out, many of the most interesting marketing problems are global, not local. By its very nature, the Internet facilitates global retailing. In a world where services are becoming increasingly salient and are potentially easier to provide online than tangible products, it is useful to have a measurement instrument that can assess online service quality for both tangibles and intangibles. An important consideration in cross-cultural research



is whether the same measures might be used in global contexts and whether these require local adaptation. To their credit, marketing scholars have responded with a growing body of literature focusing on cross-national research topics (Durvasula *et al.*, 1993; Netemeyer *et al.*, 1991). However, less attention has been paid to replications involving cross-cultural measurement issues (Mullen, 1995; Singh, 1995). This is particularly true for Internet and e-Business research, where cross-cultural theory generalizability and managerial applicability are both salient, given the inherent “borderless” nature of the Internet.

It is perhaps paradoxical that while most academics would claim to accept the importance of replication as a worthwhile pursuit (Collins, 1985; Rosenthal and Rosnow, 1984) a look at published papers in the main marketing journals provides scant evidence of this. The bias in favour of “original” theory-driven research is widespread. Hubbard and Armstrong (1994) report that of over a thousand papers sampled from three major marketing journals, none were replications. Less than 2 per cent were extensions and they consumed but 1 per cent of the journal space. Similar results are reported in accounting, economics, finance, management and marketing journals (Hubbard and Vetter, 1991, 1996). The pursuit of originality has meant that the marketing “literature is largely made up of uncorroborated, fragmented, ‘one-off’ results” (Hubbard and Lindsay, 2002, p. 382). While innovative theory driven research is not to be discouraged, the marketing discipline may be suffering from an excess of it. This paper is a replication with extension that examines the applicability of a new measure of online service quality. It tests whether results reported by the authors from the USA can be generalized to other countries. We argue that looking for sameness through replication is as useful a pursuit as looking for differences. Looking for sameness provides the basis for generalizability that may often be in doubt. For a journal that styles itself as international and a special issue that focuses on the international nature of market research, we argue that a replication study is akin to what fuel is to an engine. As Easley *et al.* (1994, p. 205) point out, “Research is not only a creative process, it is a discipline. Some concertos are best understood by being played more than once”. We therefore take up Wolfinbarger and Gilly’s (2003, pp. 196-7) recent call for “future research [that] investigates how characteristics of various product categories affect the importance of the four [eTailQ] factors in predicting quality”. While replication principally addresses validity issues (Carman, 1990) we consider several additional factors. First, we take a more global contextual perspective and utilize measurement equivalence to test the instrument cross-nationally. This is significant given the cross-cultural nature of the Internet and the need for internationally applicable theory development. Second, while Wolfinbarger and Gilly (2003) looked at tangible products, we extend the application by examining two tangible offerings (books in South Africa and groceries in Malta) and an intangible category (shares in Australia). Third, we undertake a detailed analysis of the individual eTailQ dimensions. Rather than just consider the effect of the overall online quality construct, on each of Perceived Value, Overall Satisfaction and Loyalty we look at the effect of each of the four dimensions of online quality on each of the three constructs. We further extend the nomological net to include perceived value; a construct not looked at by Wolfinbarger and Gilly (2003). Theoretically, Perceived Value is treated as an antecedent or moderating variable to quality (Caruana *et al.*, 2000b) while Overall Satisfaction and Loyalty are consequences.

The paper is set out as follows. First, we revisit the case for replication in the social sciences, and position the present study within Berthon *et al.*'s (2002) replicative research framework. We then review the literature underpinning retail service quality – in both offline and online settings, concluding with a description of Wolfenbarger and Gilly's (2003) eTailQ scale (the focal construct in this study). Next, we explore the methodological case for establishing measurement equivalence and ground our six hypotheses in appropriate theory. The data collection procedure is then described, data are analysed and results discussed. The paper concludes with implications, study limitations and future research directions.

The role of replication in advancing knowledge

The role of replication in the social sciences has received considerable attention of late, but as Madden *et al.* (1995) lament, articles on the subject in marketing generally begin with the observation that few replication studies are being reported in the literature and conclude with a call for more replication research. For example, Hubbard and Armstrong (1994) found that, compared with other business disciplines, marketing journals appear least inclined to publish replicative research. This is somewhat of a paradox given that Kane (1984) regards replicability as the cornerstone of scientific research, and that Rosenthal and Rosnow (1984) contend that replicability is almost universally accepted as the most important criterion of genuine scientific knowledge. As Berthon *et al.* (2002) note, the underlying reasons for this are unclear and thus worthy of more scholarly attention. Is there a lack of outlets for this type of research, despite the editor of one leading journal stating that replication was sought for and encouraged (Kinneer, 1992), and another implying that space in core journals could be devoted to the reporting of replication results (Monroe, 1992)? Is replicative research in marketing not rewarded? Are there theoretical problems? Is it simply a case of neglect? Or have marketing researchers lacked a framework to conceptualize, structure and guide replicative research efforts?

Assuming the latter, we need a conceptual framework of possible replication with extension combinations. Even at their most basic level, "straight" replications of an earlier study also involve an extension as, at a minimum, the time dimension is different. Yet, time can and does impact earlier results. A good example is the Haire (1950) classic "shopping list" study, the results of which used to be duly featured in any textbook dealing with projective techniques in marketing research. Of course, wide exposure in textbooks provides legitimacy and broad acceptance of generalizability. However, later straight replication of the research, again in the USA, provided contrary results (Webster and von Pechmann, 1970).

We use Berthon *et al.*'s (2002) framework to position the replicative component of this paper. Berthon *et al.* (2002, 2003) argue that research is an epistemological process which occupies a conceptual space defined by four primary dimensions: problem or phenomenon, theory, method and context. Research space (\mathfrak{R}) can thus be defined:

$$\mathfrak{R} =_{\text{def}} (P \cdot T \cdot M \cdot C)$$

Where P represents the research phenomenon or problem, T the theory, M the method and C the context. For simplicity, let T_r designate a replication of theory from one project to the next, and T_g as the generation or change of theory from one project to the next; similarly for method ($M_{r,g}$), context ($C_{r,g}$) and problem ($P_{r,g}$). On any one

dimension, change or generation can be of two main types: incremental ($_{gi}$) or comprehensive (extensive) ($_{ge}$). A particular research strategy is thus[1]:

$$\mathfrak{R} = \text{fn}(P \cdot T \cdot M \cdot C)$$

where $P: P \in \{r, gi, ge\}$, $T: T \in \{r, gi, ge\}$, $M: M \in \{r, gi, ge\}$, $C: C \in \{r, gi, ge\}$ Holding the problem dimension constant, potential research space is defined by three dimensions: theory, method and context. Research can generate knowledge through eight possible combinations: $T_r M_r C_r$, $T_r M_r C_g$, $T_r M_g C_r$, $T_g M_r C_r$, $T_r M_g C_g$, $T_g M_r C_g$, $T_g M_g C_r$, $T_g M_g C_g$. In terms of the eight research strategies the following points are pertinent. First, $T_r M_r C_r$ corresponds to a pure replicative strategy where all three dimensions are held constant, implying a research strategy of zero degrees freedom. Second, $T_g M_g C_g$ correspond to the pure generative strategy, where there is change on all dimensions – i.e. a research strategy of three degrees of freedom. Third, the remaining six strategies, comprise the region of extension. Specifically, $T_r M_r C_g$, $T_r M_r C_r$, $T_g M_r C_r$ contain only one changed parameter (one degree of freedom – df), while $T_r M_g C_g$, $T_g M_r C_g$ and $T_g M_g C_r$ contain two (two degrees of freedom). This paper takes an incrementally new method (investigating conceptual equivalence in addition to psychometric equivalence) and generatively new, multiple contexts (shares in Australia; groceries in Malta; books in South Africa) for the eTailQ instrument (Wolfenbarger and Gilly, 2003) that seeks to capture the online retail service quality construct. It is, therefore, an example of a $T_r M_{gi} C_{ge}$ replication with 2 df.

Retail service quality

Online service quality is a more recent development of offline service quality. Definitions of service quality revolve around the notion that it results from the comparison customers make between their expectation about a service and their perception of how the service was actually performed (Lehtinen and Lehtinen, 1982; Lewis and Booms, 1983; Grönroos, 1984, Parasuraman *et al.*, 1985). Grönroos (1984) conceptualizes two dimensions of service quality, viz. “Technical quality” or what is received by the customer, and “functional quality” or how a service is provided. The latter is perceived as the more critical and is concerned with the psychological interaction taking place during the exchange transaction. Rust and Oliver (1994) suggest a three component measure consisting of service product, delivery and environment where the first two components, respectively, reflect the earlier conceptualizations of technical and functional quality put forward by Grönroos (1984).

Qualitative research by Parasuraman *et al.* (1985) identified ten service quality factors, their so-called “gap model” and the development of SERVQUAL as a measure of the service quality gap between expected service and perceived service. Later quantitative research (Parasuraman *et al.*, 1988) reduced the expected ten factors to five: tangibles, reliability, responsiveness, assurance and empathy. SERVQUAL has undergone innumerable replications in the USA and internationally, with noteworthy variations in the factor structure (Carman, 1990; Babaakus and Boller, 1992; Winsted, 1997). Parasuraman *et al.* (1994, p. 211) concede that replication results support:

... the possibility of a three-dimensional structure wherein responsiveness, assurance and empathy meld into a single factor, thereby reducing the number of SERVQUAL factors from five to three.

The instruments developed to measure service quality have been primarily generic. However, Dabholkar *et al.* (1996) propose a new multi-level model and measure, specifically of retail service quality, consisting of five primary dimensions termed: physical aspect, reliability, personal interaction, problem solving and policy with the first three factors, each having two sub-dimensions. The physical aspects dimension is said to be broader than tangibles in SERVQUAL in that, besides the appearance of the outlet, it includes the convenience offered via layout and physical facilities. However, the reliability factor is similar to SERVQUAL reliability and includes: promises and doing it right, as separate sub-dimensions while personal interaction that includes two sub dimensions: inspiring confidence and courteous/helpful, is similar to responsiveness in the 1994 version of SERVQUAL. Specific to the retail situations are probably the two additional dimensions of problem solving and policy of the store. The former underlines the increased importance of service recovery and involves the handling of returns, exchange and complaints while the latter focuses on opening hours, credit cards and parking.

The factor structures of the various instruments considered are summarised in Table I. While factor naming is by no means an exact science, it appears that instruments can be grouped into two main streams. One stream coming from Lehtinen and Lehtinen (1982); Grönroos (1984) and Rust and Oliver (1994) emphasises three main dimensions of service quality, namely: Service product; Service delivery and Service environment. The Parasuraman, Zeithamal and Berry stream emphasises Tangibles, Reliability and Responsiveness. Tangibles appears as the common denominator across both streams while the additional dimensions in Dabholkar *et al.* (1996) possibly result from a specific focus on retail service quality, rather than service quality generally. A final major development has been the instrument proposed by Brady and Cronin (2001). They conceptualise service quality as consisting of the three dimensions proposed by Rust and Oliver (1994) that each consists of sub-dimensions as in Dabholkar *et al.* (1996) and are in turn made up of items from the reliability, responsiveness and empathy dimensions of SERVQUAL.

Since, e-commerce is mostly conducted without human contact, the instruments that have been developed to measure service quality may not inappropriate to use to determine online service quality. Four main instruments can be identified in this area, namely: WEBQUAL (Lociacono, Watson and Goodhue, 2000), WebQual (Barnes and Vidgen, 2002), eTailQ (Wolfenbarger and Gilly, 2003) and more recently E-S-Qual (Parasuraman *et al.*, 2005). WEBQUAL with its 12 dimensions has an internet interface design focus to help web site designers improve their interaction with users. It is therefore less relevant to the measurement of online service quality (Zeithaml *et al.*, 2002). WebQual (Barnes and Vidgen, 2002) is a 22-item instrument capturing three aspects (and five factors), namely: web site usability (site design and usability), information quality and service interaction quality (Trust and Empathy). WebQual has an information systems focus and the psychometric testing of the instrument is not extensive. Conversely, the eTailQ instrument (Wolfenbarger and Gilly, 2003) was arguably the first psychometrically robust instrument that focused specifically on online service quality. The instrument consists of 14-items that result in four factors, labelled: web site design (five-items), Customer service (three-items), Fulfilment/Reliability (three-items) and Security/Privacy (three-items). Wolfenbarger and Gilly (2003, p. 193) define web site design as all elements of the consumer's

Table I.
Factor structures of the
various instruments
considered

Author/s	Dimensions						
Parasuraman <i>et al.</i> (1988)	Reliability	Responsiveness	Assurance	Empathy	Tangibles	Tangibles	Policy
Parasuraman <i>et al.</i> (1994)	Service product	Responsiveness	Assurance	Empathy	Service environment	Physical aspects	Problem solving
Rust and Oliver (1994)	Service delivery	Personal interaction	Assurance	Empathy	Physical outcome	Physical outcome	Policy
Dabholkar <i>et al.</i> (1996)	Reliability	Personal interaction	Assurance	Empathy	Physical outcome	Physical outcome	Policy
Brady and Cronin (2001)	Interaction quality	Personal interaction	Assurance	Empathy	Physical outcome	Physical outcome	Policy

experience of the web site. It includes navigation, information search, order processing, appropriate personalization and product selection. Customer service refers to responsive, helpful, willing service that responds to customer inquiries quickly. Fulfilment/reliability involves the accurate display and description of a product so that what customers receive is what they thought they ordered, and the delivery of the right product within the time frame promised. Finally, security/privacy refers to security of credit card payments and privacy of shared information. More recently Parasuraman *et al.* (2005) have followed accepted psychometrical procedures to develop E-S-Qual and provide an alternative means of measuring the online service quality. E-S-Qual is a 22-item instrument that also identifies four dimensions, namely: efficiency, fulfilment, system availability and privacy. In addition, a subscale of 11 items has been developed focusing on measuring customers' assessment of the quality of the service recovery provided by online service providers. The dimensions of Fulfilment and Privacy are common to both instruments – Table II. Data for the psychometric testing of both eTailQ and E-S-Qual was collected in the USA. We focus our study on the eTailQ instrument because it was the first available instrument that met acceptable psychometric criteria. We seek to establish whether the construct and scale are international in scope and hold across product categories.

Establishing equivalence

Cross cultural equivalence has been conceptualised in many different ways. Indeed, Johnston (1998) identifies 52 types of equivalence, some of which are poorly defined with much overlap in what they describe. We make use of the distinction put forward by Cheung (1999) who discusses two types of equivalence: conceptual equivalence and psychometric equivalence. Conceptual equivalence refers to similarity in meaning of items of an instrument across countries and aims to ensure that the instrument is measuring the same construct. Respondents need to agree on the item loadings and factor structure of an instrument and therefore the instrument can be said to “mean” the same thing. Because conceptual equivalence is a prerequisite to making meaningful comparisons across respondent groups (Cheung and Rensvold, 1998; Reise *et al.*, 1993), the majority of existing studies initially focus on establishing the conceptual equivalence of an instrument (Facteau and Craig, 2001). In contrast, psychometric equivalence indicates that the different respondents reply to the instrument in the same way. This is evidenced by equivalent levels of reliability, variance, range of ratings, mean level of ratings, and intercorrelations among factors. In cross-cultural research, conceptual equivalence is a prerequisite, as it provides a basis for assuming that the construct of interest has an identical meaning in the different cultures and allows for meaningful discussion of the construct across cultures. However, an absence of psychometric equivalence is not necessarily negative. It may indicate meaningful differences between respondents from the different cultures.

Author/s	Dimensions					
Wolfenbarger and Gilly (2003)	Web site design	Privacy/security	Fulfilment/reliability	Customer service		
Parasuraman <i>et al.</i> (2005)		Privacy	Fulfilment		System availability	efficiency

Table II.
Factor structures of the instruments measuring online service quality

The most commonly used procedure for testing equivalence (or invariance) is multigroup confirmatory factor analysis. In seeking evidence of multi-group equivalence we are interested in whether the eTailQ construct holds across different products/cultures (shares in Australia; groceries in Malta; books in South Africa). Jöreskog (1971) recommended that tests of equivalence commence with a test of the equality of covariance structures across groups. Rejection of this condition would result in increasingly restrictive hypotheses testing aimed at identifying the source of non-invariance. However, research has shown cases where this basic omnibus test is held or not supported, yet subsequent tests show or do not show invariance (Byre, 1998, p. 261). In testing the eTailQ measure across product categories/countries we therefore propose three hypotheses that consider conceptual equivalence:

- H1a.* The four underlying eTailQ factors will hold across all three product categories – invariance in the number of factors.
- H1b.* The pattern of factor loadings across the three product categories are equivalent – invariance of factor loading pattern.
- H1c.* Structural relations among the four factors of the construct are equivalent – invariance of factor variance and covariances.

We subsequently assess the psychometric properties of the scale by considering a further three hypotheses concerning psychometric equivalence:

- H2a.* The reliability of the scale is high and consistent across product categories.
- H2b.* The scale exhibits discriminant validity across product categories.
- H2c.* The scale exhibits nomological validity across product categories.

Data collection

The eTailQ instrument (Wolfenbarger and Gilly, 2003) was used to measure the four quality constructs of fulfillment/reliability, customer service, web site design and privacy/security (three-, three-, five- and three-items, respectively). Measures for perceived value, overall satisfaction and customer loyalty were also included to enable testing for nomological validity. Perceived value is measured using the five-item instrument by Dobbs *et al.* (1991); Overall satisfaction and Customer loyalty is measured on four- and six-item scales, respectively, both developed by Oliver (1996). A number of classificatory demographic variables were also added. All items were measured using seven-point scales (1 – strongly disagree; 7 – strongly agree). All instruments exhibited strong psychometric properties with acceptable reliability and reported validity. Arrangements were made with three internet vendors. All three countries use English as their main commercial language and all customers are bilingual in the English language, eliminating the need to complicate matters via translations. The companies for the three product categories agreed to send emails to their online customers on their database requesting them to visit a webpage and complete the questionnaire. A total of 1,165, 692 and 210 replies was collected for the books, shares and groceries businesses, respectively.

Analysis

Preliminary single group analysis

eTailQ data for each product category sample was first analysed separately to confirm the appropriateness of the four factor model suggested by Wolfenbarger and Gilly (2003). Their analysis indicated that such a model provided the best confirmatory factor analysis (CFA) fit statistics. Analysis of the four factor model that provides the best fit indices indicates common problems across the models for the three product categories with large modification indices for correlated errors between items 11 and 10 and between items 9 and 8 (TD11,10; TD9,8). Freeing these parameters improved model fit but additional modification indices were small and further fitting was not pursued as it risked over-specification. This model was treated as the baseline model. Results shown in Table III provide support for a four factor model in all three product/country groups.

Tests of conceptual equivalence

The identified baseline model was used to test for conceptual equivalence. Results indicate that the initial model is consistent with one that is well fitting. A look at the indices of fit shows a root mean square error of approximation (RMSEA) of 0.07, and a comparative fit index (CFI) of 0.96 confirming that eTailQ is most appropriately described by a four-factor model. We therefore test increasingly restrictive models in terms of the pattern of factor loadings and invariance in factor variance/covariance. An assessment of the fit of the different models has, until recently, relied on an assessment of the change in χ^2 ($\Delta\chi^2$) and degrees of freedom (Δdf) of the model relative to the first model. However, $\Delta\chi^2$ as a test of invariance among groups is affected by sample size and model specifications. On the basis of 768,000 CFA model estimations, Cheung and Rensvold (2003) show that ΔCFI is a more suitable statistic for testing between-group invariances of CFA models. They argue that a value of ΔCFI smaller than or equal to -0.01 indicates that the null hypothesis of invariance should not be rejected. Values between -0.01 and -0.02 should be viewed with increasing suspicion, while values greater than -0.02 indicate a likely difference between the constrained and unconstrained models. Inspection of Table IV provides fit indices in rows 1, 2, and 3 that support *H1a* invariance in the number of factors, *H1b*: invariance in the factor loading pattern and *H1c* invariance of factor variance and covariances, respectively.

Tests of psychometric equivalence

The first tests undertaken examined reliability and variance extracted. The reliabilities of the different dimensions all exceeded the recommended 0.7 level (Nunnally, 1978). The weakest score obtained was for Fulfilment/reliability in the groceries sample, which provided a value of 0.71. For all three samples, the average variance extracted comfortably exceeded the 0.5 level (Hair *et al.*, 1992, p. 450), except for the Fulfilment/reliability dimension in the groceries sample where a value of 0.44 was obtained. These results provide support for *H2a* that the reliability of scale is high and consistent across product categories investigated.

Evidence for discriminant validity is provided from several sources. The first test adopted the procedure suggested by Anderson and Gerbing (1988) and the data for eTailQ and Perceived Value were tested with a four-factor and a five-factor model. The χ^2 tests for the pair of models in all three samples reveal statistically significant

Table III.
Comparison of various
eTailQ factor models
across three product
categories

	χ^2	df	NFI	CFI	RMR	GFI	AGFI
<i>Books</i>							
Four factors; one second order factor	758	73	0.93	0.93	0.12	0.91	0.88
<i>Four factors</i>	742	71	0.93	0.94	0.12	0.92	0.88
<i>Four factors with modification</i>	530	69	0.95	0.95	0.11	0.94	0.91
Three factors (fulfilment/reliability + web design, privacy/security, customer service)	1339	74	0.88	0.89	0.14	0.86	0.80
Two factors(fulfilment/reliability + web design + privacy/security, customer service)	2883	76	0.77	0.77	0.17	0.74	0.64
One factor	4011	77	0.68	0.68	0.17	0.67	0.55
<i>Shares</i>							
Four factors; one second order factor	618	73	0.92	0.93	0.09	0.89	0.84
<i>Four factors</i>	614	71	0.93	0.93	0.09	0.89	0.83
<i>Four factors with modification</i>	304	69	0.96	0.97	0.06	0.94	0.91
Three factors (fulfilment/reliability + web design, privacy/security, customer service)	1028	74	0.89	0.89	0.12	0.82	0.75
Two factors(fulfilment/reliability + web design + privacy/security, customer service)	2155	0.76	0.77	0.77	0.13	0.69	0.57
One factor	2855	77	0.7	0.7	0.16	0.63	0.49
<i>Groceries</i>							
Four factors; one second order factor	170	73	0.9	0.94	0.1	0.89	0.84
<i>Four factors</i>	170	71	0.9	0.94	0.1	0.89	0.84
<i>Four factors with modification</i>	132	69	0.93	0.96	0.09	0.91	0.87
Three factors (fulfilment/reliability + web design, privacy/security, customer service)	218	0.74	0.88	0.91	0.11	0.87	0.81
Two factors (fulfilment/reliability + web design + privacy/security, customer service)	814	76	0.69	0.72	0.24	0.63	0.49
One factor	778	77	0.62	0.65	0.17	0.64	0.51

improvement in model fit. Additionally, the more stringent test suggested by Fornell and Larcker (1981) for discriminant validity was used. The findings reported in Table V, show that in two of the three categories considered, the average variance extracted for the web site design dimension did not exceed the squared correlations with the remaining factors. This does not in itself point to an absence of discriminant validity. The results provide support for *H2b* that the scale exhibits discriminant validity across product categories investigated.

Wolfenbarger and Gilly (2003, p. 192) and Table V also report a similar issue with the web site design dimension but do not suggest possible remedial action. The results from the original study and this study do indicate that the web site design items could benefit from further reassessment and enhancement. Interestingly, the results in Table V show that the web site design scores are lower in the case of the service category (online shares). This could either mean that web site design issues are less salient among these customers who are likely to have done their research beforehand and use the site primarily to effect transactions. Or, that the issues being tapped by the items that capture the web site design construct may need to be reconsidered.

To investigate nomological validity, the four dimensions of eTailQ were treated as a set of predictor variables and on the basis of existing literature, perceived value, Overall satisfaction and loyalty were considered as criterion variables – enabling the data to be tested with canonical correlations. Results in Table VI provide canonical correlations ranging from 0.706 to 0.744 that are significantly different from zero at the 0.001 level by the chi square test. The canonical correlation provides an estimate of the strength of the relationship between the predictive and the criterion variables. The magnitude of the redundancy index in canonical correlation indicates how much variance in one set of variables is shared by variability in the other set (Stewart and Lowe, 1968). The results also show that both sets of variables explain approximately one third of the variance in each of the three constructs in each of the three samples.

	χ^2	df	NFI	RMSEA	CFI
Validity of four factor structure	998	207	0.95	0.07	0.96
Invariant pattern of factor loadings	1,072	228	0.95	0.07	0.96
Invariant factor variance/covariances	1,184	247	0.94	0.07	0.95
Invariant error variance covariance ^a	2,703	281	0.87	0.11	0.88

Note: ^aThe equality of error variance and covariance is considered to be the least important constraint to test (Bentler, 1995) but included for completion purposes

Table IV.
Simultaneous tests for
conceptual equivalence

	Shares				Books				Groceries			
	1	2	3	4	1	2	3	4	1	2	3	4
1. Fulfilment/reliability	0.82				0.55				0.44			
2. Customer service	0.53	0.78			0.45	0.77			0.35	0.64		
3. Web site design	0.72	0.69	0.51		0.40	0.55	0.54		0.44	0.52	0.53	
4. Security/privacy	0.40	0.28	0.48	0.81	0.27	0.31	0.45	0.79	0.21	0.24	0.37	0.85

Note: All correlations significant at $p < 0.01$

Table V.
Squared correlations
among eTailQ and
perceived value with
average variance
extracted on diagonal

Table VI.
Results of canonical analysis showing the effects of eTailQ dimensions on perceived value, overall satisfaction and loyalty

	Std. canonical weights		Canonical loadings		Canonical cross-loadings		
	Books	Shares	Books	Shares	Books	Shares	Groceries
<i>Independent var.</i>							
Fulfil/reliability	0.265	0.183	0.780	0.849	0.556	0.632	0.474
Customer service	0.373	0.363	0.866	0.874	0.617	0.651	0.563
Web site design	0.433	0.476	0.875	0.936	0.624	0.697	0.641
Security/privacy	0.130	0.116	0.699	0.700	0.498	0.521	0.548
Redundancy coefficient	0.332	0.395					
<i>Dependent var.</i>							
Perceived value	0.425	0.566	0.775	0.901	0.553	0.671	0.604
Overall satisfaction	0.580	0.319	0.852	0.752	0.607	0.559	0.473
Loyalty	0.231	0.312	0.763	0.798	0.544	0.594	0.605
Redundancy coefficient					0.323	0.372	0.318
Canonical correlation					0.713	0.706	0.744
Wilk's λ					0.427	0.426	0.467
χ^2					986.59	149.35	585.80
df					12	12	12
P(χ^2)					0.000	0.000	0.000

The relative importance of a variable in each set of variables is indicated by the standardized canonical weights extracted for each variable, their canonical loadings and canonical cross-loadings. These coefficients are similar in the three samples indicating that web site design appears to have the most effect and security/privacy the least. An examination of the canonical loadings that show how much variance each dimension or construct shares with other variables in the same set indicate that these are all strong. Canonical cross-loadings reflect the correlation to variables between sets. These are all above 0.490 providing support for a significant relationship between eTailQ dimensions and the constructs of perceived value, overall satisfaction and loyalty for all three product categories. These results support *H2c* that the scale exhibits nomological validity across the product categories investigated.

Conclusions

The results suggest that the eTailQ construct and measure does generalise across product categories and countries. The instrument exhibits both conceptual and psychometrical equivalence. Conceptually, there are four common factors across the three different product categories taken from three different countries. Product categories exhibit the same pattern of item factor loadings making across product comparisons possible. They also exhibit invariant factor variances and covariances indicating that respondents for the different product categories use an equivalent range of the scale and the relations between dimensions do not differ by product category. Psychometrically, the instrument exhibits acceptable levels of reliability and variance extracted and provides acceptable levels of both discriminant and nomological validity. The instrument can be used with confidence – by retailers selling tangible goods and by in service industries, and it can also be used in English speaking countries outside the USA

Theory-building and managerial implications

This paper's contribution to both scholars and practitioners can best be assessed by way of comparing what has been achieved in one well conceptualised $T_r M_{gi} C_{ge}$ replication (Berthon *et al.*, 2002), with the lengthy and *ad hoc* process in which SERVQUAL has evolved. In the 17 years since SERVQUAL was originally published (Parasuraman *et al.*, 1988), it has been replicated more frequently than any other instrument in marketing's short history. For the most part, the replication studies considered either a single new product category, or were conducted in one new country. However, despite the lack of a replicative research framework, both SERVQUAL and the extant literature on service quality measurement did evolve – it just took an inordinate amount of time. For example, following Cronin and Taylor's (1992) criticism (as well as in response to other replication-based criticisms), Parasuraman, Zeithaml and Berry published a revised scale in 1994 – which in turn was criticised by Caruana *et al.* (2000a) and challenged by Brady and Cronin (2001). This raised the obvious question as to whether it is possible for eTailQ to evolve more quickly – particularly given the time-compressed Internet environment. To the best of our knowledge, ours is the first comprehensive replication of eTailQ. By substantially confirming the scale's reliability, validity and generalizability, we have hopefully paved the way for future research to progress more quickly and in a more structured and directed manner. Similarly, managers can now use the scale with confidence and focus their energy on each of the four dimensions of eTail quality as they are shown to have an effect on

perceived value, satisfaction and loyalty that in turn are known to result in positive word of mouth and customer retention.

Limitations

Common with all survey data collected online it is difficult to ascertain the randomness of the sample. We relied on the database of customers of collaborating firms in the three different countries. The small size of the country in the case of Malta meant that the groceries sample is relatively smaller in absolute terms when compared to the other two data collections. Second, the three countries chosen are English speaking, or ones where English is widely understood and used. It is possible that countries that make no or little use of English may conceptualize the construct differently. Moreover, the need for translation of the items requires a further test of equivalence in terms of item meaning. Third, conceptual and psychometric equivalence are prerequisites for, but do not necessarily ensure, complete measurement equivalence. As in most situations, further testing of other aspects of equivalence can be undertaken. Fourth, it is of course possible to check for other types of equivalence besides those tested for in this study. Thus, if one seeks to conduct a comparison of means across countries it is necessary to undertake configural, metric and scalar invariance tests (Steenkamp and Baumgartner 1998). Although complete invariance is not often encountered at least configural equivalence needs to be achieved for there to be potential transferability of a construct (Salzberger *et al.*, 1999). Such testing would provide enhanced support for cross-cultural equivalence. Finally, it is worth noting that Wolfinbarger and Gilly (2003, p. 193, Figure 1) treat their online service quality construct as a reflective first order and formative second order specification, rather than as a completely formative construct. This may point to possible model misspecification (Jarvis *et al.*, 2003).

Future research opportunities

While English has become the language of business, there is an argument for undertaking further testing of the instruments where the predominant language of a potential customer base may not be English. While the instrument has performed well, it would also be useful to extend the testing of the instrument beyond one service offering to others that may include banking, insurance and so on. Given the results of the discriminant validity test it appears that the items constituting web site design require further scrutiny, as the findings obtained for shares and books in this study echo the results reported by Wolfinbarger and Gilly (2003) who also indicate a similar problem with this dimension. A useful framework that may lend itself well in the reconsideration and improvement to the items that make up the web site design dimension is the 7C's model for internet customer interface design by Rayport and Jaworski (2003, p. 151). Of the five items that currently constitute the web site design dimension, two tap into content (good selection, depth of information), one into communication (does not waste my time); one into commerce (quick and easy to complete transactions) and one into customization (right level of personalization). This leaves a further three elements of the 7C's framework consisting of context (in terms of layout and design); connection (degree to which a site is linked to other sites) and community (the way a site enables user to user communication) that do not appear to have been tapped. Future research can consider adding questions to subsequently identify a set of items that can better capture the web site design dimension.

Note

1. More comprehensively, one might specify $P_e: P_e \in (0,1)$, $T_e: T_e \in (0,1)$, $M_e: M_e \in (0,1)$, $C_e: C_e \in (0,1)$, where P_e , T_e , M_e , and C_e are the indices of extension for a particular dimension.

References

- Anderson, J.C. and Gerbing, D.W. (1988), "Structural equation modelling in practice: a review and recommended two-step approach", *Psychological Bulletin*, Vol. 103, pp. 411-23.
- Babaakus, E. and Boller, G.W. (1992), "An empirical assessment of the SERVQUAL scale", *Journal of Business Research*, Vol. 24 No. 3, pp. 253-68.
- Barnes, S. and Vidgen, R. (2002), "An integrative approach to the assessment of e-commerce quality", *Journal of Electronic Commerce Research*, Vol. 3 No. 3, pp. 114-27.
- Bentler, P.M. (1995), *EQS: Structural Equations Program Manual*, Multivariate Software Inc., Encino, CA.
- Brady, M.K. and Cronin, J.J. (2001), "Some new thoughts on conceptualizing perceived service quality: a hierarchical approach", *Journal of Marketing*, Vol. 65, pp. 34-49.
- Berthon, P., Pitt, L.F., Ewing, M.T. and Carr, C. (2002), "Potential research space in MIS: a framework for replication, extension and generation", *Information Systems Research*, Vol. 13 No. 8, pp. 416-27.
- Berthon, P., Ewing, M.T., Pitt, L.F. and Berthon, J-P. (2003), "Ad copy? Reframing replicative research in advertising", *International Journal of Advertising*, Vol. 22 No. 4, pp. 511-30.
- Byre, B.M. (1998), *Structural Equation Modeling with Lisrel, Prelis and Simplis: Basic Concepts, Applications and Programming*, Lawrence Erlbaum Associates Publishers, Hillsdale, NJ.
- Carman, J.M. (1990), "Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions", *Journal of Retailing*, Vol. 66 No. 1, pp. 33-55.
- Caruana, A., Ewing, M.T. and Ramaseshan, B. (2000a), "Assessment of the three-column format SERVQUAL: an experimental approach", *Journal of Business Research*, Vol. 49 No. 1, pp. 57-65.
- Caruana, A., Money, A. and Berthon, P. (2000b), "Service quality and satisfaction – the moderating role of value", *European Journal of Marketing*, Vol. 34 Nos 11/12, pp. 1338-52.
- Cheung, G.W. (1999), "Multifaceted conceptions of self-other ratings disagreement", *Personnel Psychology*, Vol. 52 No. 1, pp. 1-36.
- Cheung, G.W. and Rensvold, R.B. (1998), "Testing factorial invariance across groups: a reconceptualization and proposed new method", *Journal of Management*, Vol. 25 No. 1, pp. 1-27.
- Collins, H.M. (1985), *Changing Order: Replication and Induction in Scientific Practice*, Sage, Beverly Hills, CA.
- Cronin, J.J. and Taylor, S.A. (1992), "Measuring service quality: a re-examination and extension", *Journal of Marketing*, Vol. 56, pp. 55-68.
- Dabholkar, P.A., Thorpe, D.I. and Rentz, J.O. (1996), "A measure of service quality for retail stores: scale development and validation", *Journal of the Academy of Marketing Science*, Vol. 24 No. 1, pp. 3-16.
- Deshpandé, R. (1999), "Introduction: 'foreseeing' marketing", *Journal of Marketing*, Vol. 63, pp. 164-7.
- Dobbs, W.B., Monroe, K.B. and Grewal, D. (1991), "Effects of price, brand and store information on buyers' product evaluations", *Journal of Marketing Research*, Vol. 28 No. 3, pp. 307-19.

- Durvasula, S., Andrews, J.C., Lysonski, S. and Netemeyer, R.G. (1993), "Assessing the cross-national applicability of consumer behaviour models: a model of attitude toward advertising in general", *Journal of Consumer Research*, Vol. 19, pp. 626-36.
- Easley, R.W., Madden, C.S. and Dunn, M.G. (1994), "A review and extension of the role of replication in achrol", in Achrol, R. and Mitchell, A. (Eds), *Enhancing Knowledge Development in Marketing: 1994 AMA Educators Proceedings 5*, American Marketing Association, Chicago, IL, p. 205.
- Facteau, J.D. and Craig, S.B. (2001), "Are performance appraisal ratings from different rating sources comparable?", *Journal of Applied Psychology*, Vol. 86 No. 2, pp. 215-27.
- Fornell, C. and Larcker, F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18, pp. 39-50.
- Grönross, C. (1984), "Service quality model and its market implications", *European Journal of Marketing*, Vol. 18 No. 4, pp. 36-44.
- Hair, J.R., Anderson, R.E., Tatham, R.L. and Black, W.C. (1992), *Multivariate Data Analysis*, 3rd ed., Macmillan, New York, NY.
- Haire, M. (1950), "Projective techniques in marketing research", *Journal of Marketing*, Vol. 14, pp. 649-56.
- Hubbard, R. and Armstrong, J.S. (1994), "Replication and extensions in marketing: rarely published but quite contrary", *International Journal of Research in Marketing*, Vol. 11, pp. 233-48.
- Hubbard, R. and Lindsay, R.M. (2002), "How the emphasis on 'original' empirical marketing research impedes knowledge development", *Marketing Theory*, Vol. 2 No. 4, pp. 381-402.
- Hubbard, R. and Vetter, D.E. (1991), "An empirical comparison of published replication research in accounting, economics, finance, management and marketing", *Journal of Business Research*, Vol. 35, pp. 153-64.
- Hubbard, R. and Vetter, D.E. (1996), "Replications in the finance literature: an empirical study", *Quarterly Journal of Business and Economics*, Vol. 30 No. 4, pp. 70-80.
- Jarvis, C.B., Mackenzie, S.B. and Podsakoff, P.M. (2003), "A critical review of construct indicators and measurement model misspecification in marketing and consumer research", *Journal of Consumer Research*, Vol. 30, pp. 199-218.
- Jöreskog, K.G. (1971), "Simultaneous factor analysis in several populations", *Psychometrika*, Vol. 36, pp. 409-26.
- Kane, E.J. (1984), "Why journal editors should encourage the replication of applied econometric research", *Quarterly Journal of Business and Economics*, Vol. 23, pp. 3-8.
- Kinnear, T.C. (1992), "From the editor", *Journal of Marketing*, Vol. 56, pp. 1-3.
- Lehtinen, U. and Lehtinen, J.R. (1982), "Service quality – a study of dimension", unpublished working paper, Service Management Institute, Helsinki, pp. 439-60.
- Lewis, R.C. and Booms, B.H. (1983), "The marketing aspects of service quality", in Berry, L.L., Shostasck, G. and Upah, G. (Eds), *Emerging Perspectives in Services Marketing*, Marketing Association, Chicago, IL, pp. 99-107.
- Madden, C.S., Easley, R.W. and Dunn, M.G. (1995), "How journal editors view replication research", *Journal of Advertising*, Vol. 14 No. 4, pp. 77-87.
- Monroe, K. (1992), "Replications in consumer research", editorials in *Journal of Consumer Research*, Vol. 19 No 1, June, and No 2, September.
- Mullen, M.R. (1995), "Diagnosing measurement equivalence in cross-national research", *Journal of International Business Studies*, pp. 573-96, 3rd Quarter.

-
- Netemeyer, R.G., Durvasula, S. and Lichtenstein, D.R. (1991), "A cross-national assessment of the reliability and validity of the CETSCALE", *Journal of Marketing Research*, Vol. 28, pp. 320-7.
- Nunnally, J.C. (1978), *Psychometric Theory*, 2nd ed., McGraw Hill, New York, NY.
- Oliver, R.L. (1996), *A Satisfaction Behavioral Perspective on the Consumer*, McGraw Hill, New York, NY.
- Parasuraman, A., Zeithamal, V.A. and Berry, L.L. (1985), "A conceptual model of service quality and its implications for future research", *Journal of Marketing*, Vol. 49, pp. 41-50.
- Parasuraman, A., Zeithamal, V.A. and Berry, L.L. (1988), "A SERVQUAL: multiple-item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994), "Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria", *Journal of Retailing*, Vol. 70 No. 3, pp. 201-30.
- Parasuraman, A., Zeithamal, V.A. and Malhotra, A. (2005), "E-S qual: a multiple item scale for assessing electronic service quality", *Journal of Service Research*, Vol. 7 No. 3, pp. 213-33.
- Rayport, J.F. and Jaworski, B.J. (2003), *Introduction to E-Commerce*, 2nd ed., Vol. 2, McGraw-Hill, Singapore.
- Reise, S.P., Widaman, K.F. and Pugh, R.H. (1993), "Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance", *Psychological Bulletin*, Vol. 114, pp. 552-66.
- Rosenthal, R. and Rosnow, R.L. (1984), *Essentials of Behavioral Research: Methods and Data Analysis*, McGraw-Hill, New York, NY.
- Rust, R.T. and Oliver, R.L. (1994), "Service quality: insights and managerial implications from the frontier", in Rust, R.T. and Oliver, R.L. (Eds), *Service Quality: New Directions in Theory and Practice*, Sage, Thousand Oaks, CA, pp. 1-19.
- Salzberger, T., Sinkovics, R.R. and Schlegelmilch, B.B. (1999), "Data equivalence in cross-cultural research: a comparison of classical test theory and latent trait theory based approaches", *Australasian Marketing Journal*, Vol. 7 No. 2, pp. 23-38.
- Singh, J. (1995), "Measurement issues in cross-national research", *Journal of International Business Studies*, pp. 597-619, 3rd Quarter.
- Steenkamp, J-B.E.M. and Baumgartner, H. (1998), "Assessing measurement invariance in cross-national consumer research", *Journal of Consumer Research*, Vol. 25, pp. 78-90.
- Stewart, D.K. and Lowe, W.A. (1968), "A general canonical correlation index", *Psychological Bulletin*, Vol. 70, pp. 160-3.
- Webster, F.E. Jr and von Pechmann, F. (1970), "A replication of the 'shopping list' study", *Journal of Marketing*, Vol. 34, pp. 61-3.
- Winsted, K.F. (1997), "The service experience in two cultures: a behavioural perspective", *Journal of Retailing*, Vol. 73 No. 3, pp. 337-60.
- Wolfmbarger, M. and Gilly, M.C. (2003), "eTailQ: dimensionalizing, measuring and predicting etail quality", *Journal of Retailing*, Vol. 79, pp. 183-98.
- Zeithaml, V.A., Parasuraman, A. and Malhotra, A. (2002), "Service quality delivery through web sites: a critical review of extant knowledge", *Journal of the Academy of Marketing Science*, Vol. 30 No. 4, pp. 362-75.

Further reading

- Chase, J.E. (1970), "Normative criteria for scientific publications", *American Sociologist*, Vol. 5, pp. 262-5.
- Cronin, J.J. and Taylor, S.A. (1994), "ASERVPERF versus SERVQUAL: reconciling performance based and perception based-minus expectations measurements of service quality", *Journal of Marketing*, Vol. 58, pp. 125-31.
- Easley, R.W., Madden, C.S. and Dunn, M.G. (2000), "Conducting marketing science: the role of replications in the research process", *Journal of Business Research*, Vol. 48 No. 1, pp. 83-92.
- Johnson, T.P. (1998), "Approaches to equivalence in cross-cultural and cross national survey research", *ZUMA-Nachrichten Spezial*, January, pp. 1-40.
- Vandenberg, R.J. and Lance, C.E. (2000), "A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research", *Organizational Research Methods*, Vol. 3, pp. 4-69.

Corresponding author

Michael T. Ewing can be contacted at: michael.ewing@buseco.monash.edu.au



An illustration of modeling moderating variables in cross-national studies

Rosanna Garcia

Northeastern University, Boston, Massachusetts, USA, and

Destan Kandemir

*Department of Marketing, Albers School of Business and Economics,
Seattle University, Seattle, Washington, USA*

Modeling
moderating
variables

371

Received September 2005
Revised February 2006
Accepted March 2006

Abstract

Purpose – This paper seeks to explore how moderation can and should be modeled in cross-national/cultural contexts. A multi-national study of consumer involvement is utilized to demonstrate proper methods for modeling the different types of moderation.

Design/methodology/approach – Using data from a consumer survey regarding wine purchasing preferences conducted in Australia, New Zealand and the USA, the paper demonstrates how to identify moderators of form and of strength. A form moderator is modeled using multiplicative interactions while a strength moderator is modeled using multi-group analyses in structural equation modeling (SEM). Differences in consumers across the three countries are examined from the results.

Findings – This study suggests that search behavior is positively influenced by involvement in New Zealand and the USA but not in Australia. It also shows that perceived risk of occasion decreases involvement in all three countries, while partial support for the positive effects of importance of tradition on involvement is found. Furthermore, “perceived risk of occasion,” identified as a moderator of form, is found to significantly moderate the relationship between importance of tradition and involvement in the US sample only. Finally, the results demonstrate significant differences across the three samples in relationships among importance of tradition, perceived risk of occasion, involvement, and search behavior, indicating that the country variable has significant moderator effects.

Originality/value – Understanding form vs strength moderation is important when evaluating multi-national/cultural differences so that proper methodology can be utilized. This paper provides international marketing researchers with guidelines on how to model interactions and multi-group comparisons using SEM.

Keywords Cross-cultural studies, Linear structure equation modelling, Wines, New Zealand, Australia, United States of America

Paper type Research paper

Introduction

When conducting cross-national studies it is imperative that the appropriate research design be followed when gathering data in order to ensure “comparability between data collected in different cultural contexts” (Douglas and Craig, 1983, p. 131). Yet, once data collection is complete, data evaluation for the different contexts must also be appropriate. One method routinely used by international researchers for modeling



The authors would like to acknowledge the support of Northeastern University's Institute of Global Innovation Management on this project as well as two anonymous reviewers.

cross-national/cross-cultural contexts is to use “country” as a moderator (e.g. Calantone *et al.*, 1996; Song and Parry, 1997; Song and Xie, 2000). However, moderation has been defined in numerous ways by researchers across many different disciplines (Carte and Russell, 2003) and has been approached using different methodologies.

International marketing researchers typically have used three different approaches for modeling moderating effects: multi-group analysis (Ford *et al.*, 1997; Homburg *et al.*, 2005), multiplicative multiple regression (MMR) analysis (Buvik and Andersen, 2002; Kim and Oh 2002; Song and Xie, 2000) or analysis of variance (Vanhonacker and Pan, 1997). In cross-national studies, organizational researchers frequently utilize multi-group analyses in structural equation modeling (SEM) (Calantone and Zhao, 2001; Granzin and Painter, 2001; Keillor *et al.*, 2004) or MMR with dummy variables, whereas consumer marketing researchers frequently utilize analysis of variance (Chan, 1996; Tansey *et al.*, 1997). Based on these observations, we raise the following questions: Under what circumstances is one methodology more appropriate over another? Should organizational researchers follow one method, yet, consumer marketing researchers another when considering cross-national contexts? A second observation in reviewing the international marketing research literature is that when “country” is the primary moderator under consideration, contingencies are often not considered. This is presumably because the methodology to model two moderators in the same study can become very complex. This raises the second question: How can contingencies be added to cross-national studies with minimal complexity?

We use the context of a cross-national study regarding product involvement conducted with wine consumers in Australia, New Zealand and the USA to address these questions. We introduce the Andrews *et al.* (1990) framework of consumer involvement in search behavior to illustrate appropriate methodological approaches to moderation in this cross-national context. We provide a definition for moderator and summarize the different methodological approaches for modeling moderation before establishing and testing our hypotheses.

Identifying moderators

Baron and Kenny’s (1986, p. 1174) seminal definition of a moderator is a variable Z that:

... affects the direction and/or strength of the relationship between an independent or predictor variable [X] and a dependent or criterion variable [Y].

Furthermore, Sharma *et al.* (1981) (also see Prescott 1986) propose that there are two common types of moderator variables, strength and form, and these should be operationalized using different methodologies. Each variable type requires different analytic techniques for modeling moderation.

Baron and Kenny (1986) suggest different analytic techniques for modeling moderating effects based on measurement level of the independent variables. Although well-recognized methodological guidelines, they are worth reviewing from a cross-national perspective. In this discussion, we assume one moderator, “country” one independent variable and one dependent variable. Four cases are presented by Baron and Kenny (1986).

In case 1, the moderator and the independent variable are both categorical variables. In this situation, it is easy to identify the moderator, “country” as a categorical variable. The appropriate method in this case is an analysis of variance design

(ANOVA/ANCOVA/MANOVA) where moderation is indicated by the interaction. However, ANOVA and MANOVA assume homogeneity of variance across all cells. ANCOVA additionally assumes a linear relationship between the covariate and the dependent variable, homogeneity of regression coefficients, and independence between the independent variable and the covariate (Cohen and Cohen, 1983). These restrictions are not always valid in cross-national studies. ANOVA is also known to produce coefficient estimates that are biased and inefficient for latent variables, thus, making interpretation difficult (Ping, 1995). Correctly interpreting the findings of ANOVA in cross-national studies requires the testing of homogeneity on several levels.

In case 2, the moderator is a categorical variable and the independent variable is a continuous variable. This is the most common occurrence in cross-national/cross-cultural studies in international marketing research. Under these circumstances, one method to use is correlational tests; however, this type of test assumes homogeneity of variance of the independent variable at each level of the moderator (Baron and Kenny, 1986). An alternative method is to use MMR following the guidelines of Cohen and Cohen (1983), among others. When using MMR, it is imperative to include the direct effects of the independent variable and the moderator as well as the interaction variable. However, introducing all three variables adds multicollinearity to the model. One method of reducing multicollinearity is to mean center the data (Ping, 1995). MMR also requires measurement equivalence across samples prior to interpreting the data, particularly in cross-national studies. Thus, multi-group SEM becomes practical in cross-national studies since differential measurement error in the independent variables can be adjusted for (Myers *et al.*, 2000). As seen in the extant literature, SEM is frequently utilized when country is the primary moderator.

In case 3, the moderator is a continuous variable and the independent variable is a categorical variable. If “country” is the primary moderator under consideration, it will never be continuous, and thus, case 3 is not valid under our assumptions in this study. For completeness, it is suggested that in case 3 linear relationships between the independent variable and the dependent variable can be tested using MMR. When the relationship is not linear other regression procedures as described by Cohen and Cohen (1983) are required.

Similarly, case 4, where both the moderator and independent variable are continuous, is not a circumstance that would occur if “country” is the moderator of interest. However, we will consider this case in order to address our second question, which pertains how to model contingencies in cross-national studies. In the case where both the moderator and the independent variable are continuous, MMR is a sufficient method to follow as long as the measurement error in the moderator and/or independent variable is considered following Kenny and Judd (1984), or a similar method. MMR also suffers from complexity in interpreting the results. This can be demonstrated by looking at the regression model:

$$Y = (\beta_0 + \beta_2 Z) + (\beta_1 + \beta_3 Z)X + \varepsilon,$$

where Z represents the moderator:

The coefficient β_3 represents the interaction. For a unit change in Z , the slope of the X - Y relationship is moderated (changed) by amount β_3 (Judd *et al.*, 1995, p. 451).

Yet, this equation can also be written as:

$$Y = (\beta_0 + \beta_1 X) + (\beta_2 + \beta_3 X)Z + \varepsilon,$$

where X becomes the moderator. Either is correct, but interpretation drastically changes.

Alternatively to Baron and Kenny (1986), Sharma *et al.* (1981) propose categorizing types of moderator variables based on the relationships between the independent variable (predictor) and the dependent variable (criterion) with the proposed moderator. Moderators of strength, also called homologizer variables, influence the strength of the relationship between a predictor variable, X , and a criterion variable, Y . Strength is measured by the predictive validity coefficient of the model (R^2). Homologizers are the best modeled using a multi-group analysis. Moderators of form, also called pure or *quasi* moderators, modify the form of the relationship between the predictor variable, X , and the criterion variable, Y . This alteration of form is demonstrated by revisiting the equation:

$$Y = (\beta_0 + \beta_2 Z) + (\beta_1 + \beta_3 Z)X + \varepsilon$$

where Z is the “form” moderator affecting the slope of the X - Y relationship. Moderators of form are best modeled using multiplicative interaction effects. Sharma *et al.* (1981) suggest a three-step process for determining if a moderator is one of form or of strength (Figure 1). For cross-national studies, this implies that multi-group analyses model moderation of strength, whereas multiplicative interaction effects model moderation of form.

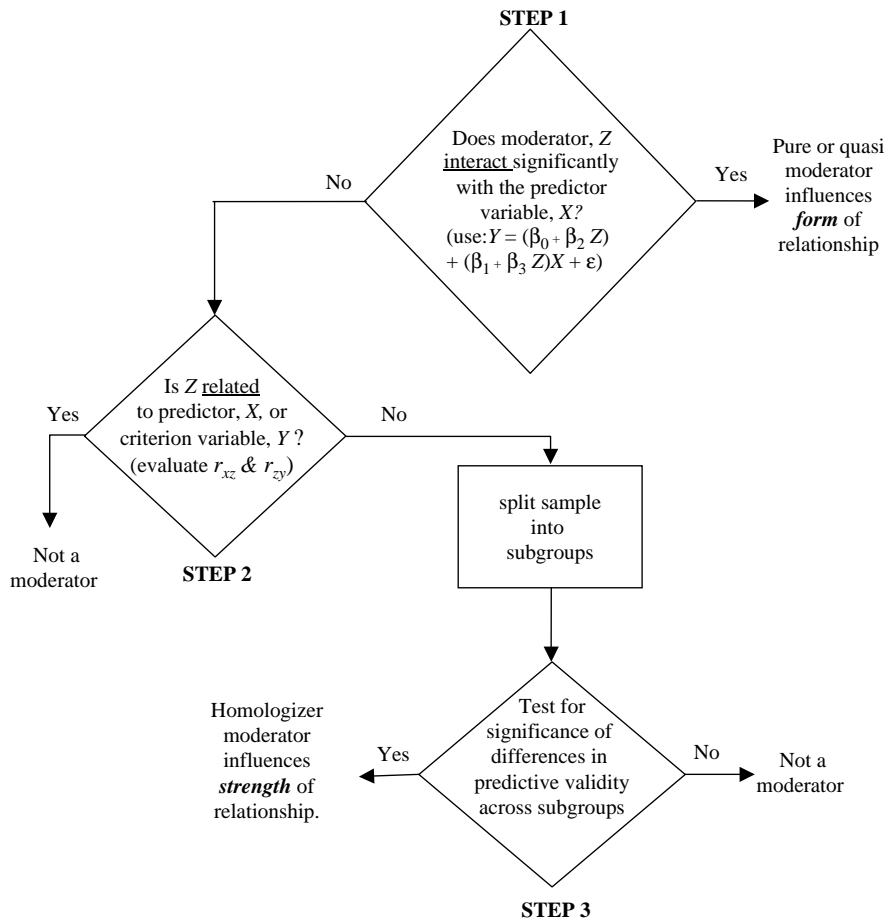
The Baron and Kenny’s (1986) approach maintains consistency with Sharma *et al.* (1981) as long as form moderators are also continuous variables and strength moderators are categorical variables. Sauer and Dick (1993, p. 637) provide a rectification for an discrepancies by contending:

... while it is possible to statistically evaluate whether a variable is functioning as a mediator, a moderator, or a hybrid, the overriding concern should be whether the theory being tested supports a moderator or mediator role for the variable in question. That is, theory should be used to define the functional form of the model.

Thus, although in some cases country may not be a strength moderator, it should be modeled using multi-group analysis following the guidelines of Myers *et al.* (2000).

We provide one caveat in regards to modeling continuous variables as strength moderators. This requires discretizing the variables, which results in a loss of information (Russell and Bobko, 1992). Additionally, researchers may choose division points that yield equal sample sizes in the groups, although this technique may obscure the interaction effects (Rigdon *et al.*, 1998) and may result in insufficient sample sizes. This transformation may also lead to interpretation errors if the interacting variables are measured with error (Rigdon *et al.*, 1998), another common problem in cross-national studies.

In the remainder of this paper, we introduce the involvement model to provide an example of how moderation impacts the relationships between factors. Moderation techniques are then demonstrated by modeling “form moderation” with multiplicative interaction as demonstrated by Ping (1995) and by modeling “strength moderation” with a multi-group analysis using SEM as demonstrated by Bentler (1995).



Source: Based on Sharma *et al.* (1981)

Figure 1. Process for identifying form vs strength moderators

Conceptual framework and hypotheses

To ground our methodological analyses, we utilize the Andrews *et al.*'s (1990) framework for the conceptualization of the involvement construct. This nomological network theorizes relationships among involvement, its antecedents, its consequences, and possible moderators. Antecedents to involvement are grouped into "personal/enduring" and "situational/decision" factors, while "search behavior" is identified as an outcome of involvement based on this framework.

In the marketing context, involvement evaluates a customer's internal state of arousal in response to external stimuli relating to products. The relevance of involvement in the purchase of wine has been supported empirically (Lockshin *et al.*, 1997; Quester and Smart, 1998) and has been shown to be a relevant segmentation variable for product planning across nations (Aurifeille *et al.*, 2002). Specifically, Atkin *et al.* (2005) found that in making wine purchasing decisions, Australians had higher involvement than did New Zealanders or Americans. The universality of the concept of

involvement and its use in international segmentation research for wine make it an ideal variable to investigate in a cross-national study.

Personal factors

Andrews *et al.* (1990) theorize that personal needs, goals and characteristics of the individual purchasing a product will be related to involvement. Internal state of arousal toward a product, advertisement, or other stimuli is driven by personal traits of the consumer (Petty and Cacioppo, 1986). One of the personal characteristics that Andrews *et al.* (1990) focus on is the importance assigned to an object by the individual. Accordingly, we model “importance of tradition” as a personal factor in this study. An important tradition associated with wine consumption is the opening of the bottle. The ritual of opening a bottle of wine often involves viewing the bottle label, listening for the pop of the cork, smelling the cork, swirling and tasting the wine. Many people smell the cork even though this signifies little about the quality of the wine. Yet, it is a procedure that many people follow because it has become part of the tradition of drinking wine. We suggest that wine consumers who assign greater importance to the tradition of opening a bottle of wine will be more likely to have greater internal state of arousal (involvement) to wine purchasing. Thus:

H1. Consumer involvement is positively influenced by importance of tradition.

Situational factors

Andrews *et al.* (1990) suggest that “perceived risk of decision” is a situational factor that affects involvement. Inherent risk associated with a product may become salient to the consumer when interacting with the product in some manner (i.e. deciding which brand to purchase, evaluating the product attributes, etc.) (Bettman, 1973; Dholakia, 2001). Jacoby and Kaplan (1972) identified five risk dimensions in purchasing behavior including psychological, financial, performance, physical, and social risk. Our focus is on psychological risk, which is anxiety or psychological discomfort arising from anticipated post-behavioral reactions such as worry or regret (Perugini and Bagozzi, 1999). Because product choice can significantly differ when choosing a wine to drink at home from choosing a wine for a formal dinner with a boss or other superiors (Hall and Lockshin, 2000), we look at the “perceived risk” provoked when choosing a wine for a formal dinner party with friends or business associates. In this public situation, consumers who perceive this type of purchasing occasion as having possible negative outcomes will have greater involvement in choosing a suitable wine, while others who perceive little threat in this type of purchasing occasion will expend little involvement in the choice. When risk of mispurchase is high, consumers are more likely to be involved (Laurent and Kapferer, 1985). Thus:

H2. Consumer involvement is positively influenced by the perceived risk of a purchasing occasion.

Search behaviors

It is generally accepted that high levels of involvement will result in extensive information search before a purchase decision is made (Beatty and Smith, 1987; Kassarian, 1981). Kassarian (1981) theorizes that individual differences make some consumers more interested, concerned, or involved in the purchase decision process

regardless of the product or situation. Beatty and Smith (1987) found empirical support that involvement is positively related to external search. Likewise, Andrews *et al.* (1990) suggest that involvement is a predictor of increased search and shopping behavior. With regards to wine purchasing, we suggest that involved consumers will conduct more external search behaviors, such as consulting personnel from the retail store, restaurant and winery when making purchase decisions. Thus:

H3. Search behavior is positively influenced by consumer involvement.

Interaction between personal factors and situation factors

The notion that the two types of antecedents to involvement, personal and situational factors, interact together to influence involvement has been widely accepted in the extant literature (Bloch and Richins, 1983; Celsi and Olson, 1988; Richins *et al.*, 1992). Bloch and Richins (1983) and Beatty and Smith (1987), among others, propose an interactive model where personal factors interact with situational factors to amplify a variable's relationship with involvement. Richins *et al.* (1992) suggest that when a product is generally important to a consumer, the stakes associated with its purchase are heightened and involvement is magnified. We have hypothesized a positive relationship between importance of tradition and involvement. However, this relationship can be amplified under certain situational factors. We suggest one such situation is the purchasing of wine for formal occasions and thus, model "perceived risk of occasion" as a form moderator. In this situation, elevated perceived risk of the situation causes tradition to become more salient since others join in the ceremonial activities of opening the wine. Thus:

H4. An interaction effect between importance of tradition and perceived risk magnifies the relationship between tradition and involvement.

Moderating effect of country

Several studies have empirically shown that involvement differs across countries/cultures (Aurifeille *et al.*, 2002; Goldsmith *et al.*, 1998; Quester and Smart, 1998). Goldsmith *et al.* (1998) found that French wine consumers had greater involvement than did American or German wine consumers. Similarly, Atkin *et al.* (2005) found Australians wine consumers had higher involvement than Americans. We, thus, expect that the proposed model will differ across countries/cultures, where the country variable is a strength moderator. Thus:

H5. "Country" will moderate the relationships as proposed in the involvement model, thus, variant relationships will be seen across countries.

Research design and data collection

To determine the interrelationships in the proposed conceptual model, we collected data from wine consumers in Australia, New Zealand and the USA. Respondents were subscribers to wine-related e-newsletters. These subscribers were asked to complete a web-based survey on their purchasing preferences regarding wine. We obtained 1,098 respondents from Australia, 385 from New Zealand, and 1,203 respondents from the USA (see Atkin *et al.*, 2005 for details).

Analysis*Measures*

The three measures for involvement were borrowed from Lockshin *et al.* (1997). The remainder of the measures was developed for this study following the guidelines of Churchill (1979). The items used as shown in the Appendix. The survey was developed with the support of industry partners from Napa Valley, California, the USA and Adelaide, Australia. The measures had previously been used in industry-related consumer research studies, and thus, had been tested under many different conditions. For this study, the measures were pre-tested with 100 respondents in the US and 25 respondents in Australia. In order to make comparisons across populations, we asked the Australian pre-test respondents to note any inconsistencies with local phraseology even though English is the common language of all the surveys. Each country had a separate web site, which allowed us to customize the survey to country specifics. These customizations were minor and were related to questions dealing with local currencies that were not utilized in this study.

Construct validity

We tested the scales for dimensionality, reliability and validity using confirmatory factor analysis (CFA) (Bentler, 1995) before assessing the hypothesized relationships shown in Figure 2. Furthermore, it is necessary to identify equivalent phenomena when conducting cross-national research (Calantone *et al.*, 1996).

A separate CFA that involved raw data as input was performed for each of the three countries (Anderson and Gerbing, 1988). All constructs in the three samples were composed of identical items. Table I displays the results of the CFA. Consistent with Anderson and Gerbing's (1988) suggestion for purifying the measurement model, we eliminated items that loaded on multiple constructs or had low item-to-construct loadings. The loadings of items on their respective factors were highly significant ($p < 0.01$). Further, we examined the reliabilities. As shown in the Appendix, the composite reliabilities ranged from 0.73 to 0.92 across the three country samples.

We next followed the procedure recommended by Bagozzi and Yi (1988) to evaluate the fit of the measurement model for each of the three countries. The fit indices of the CFA models, including Bentler-Bonett's nonnormed fit index (NNFI), comparative fit index (CFI), and Bollen's fit index (IFI), ranged between 0.937 and 0.970 for each of the three countries (Table I). Furthermore, the root mean square of error approximation (RMSEA) values were found to be 0.053, 0.063, and 0.065 for the USA, Australian, and New Zealand samples, respectively. Overall, these fit indices suggest that the CFA models for all three countries fit the data adequately (Bentler, 1995; Bollen, 1989).

Finally, we established discriminant validity of the factors using the procedures suggested by Fornell and Larcker (1981). The results showed that the average variance extracted by the measure of each factor is larger than the squared correlation of that factor's measure with all measures of other factors in the model. These values suggest that all factors in the three measurement models possess strong discriminant validity. Aggregating all aspects of the model evaluation, we conclude that the measurements models are clean with evidence for unidimensionality, convergent validity, reliability and discriminant validity.

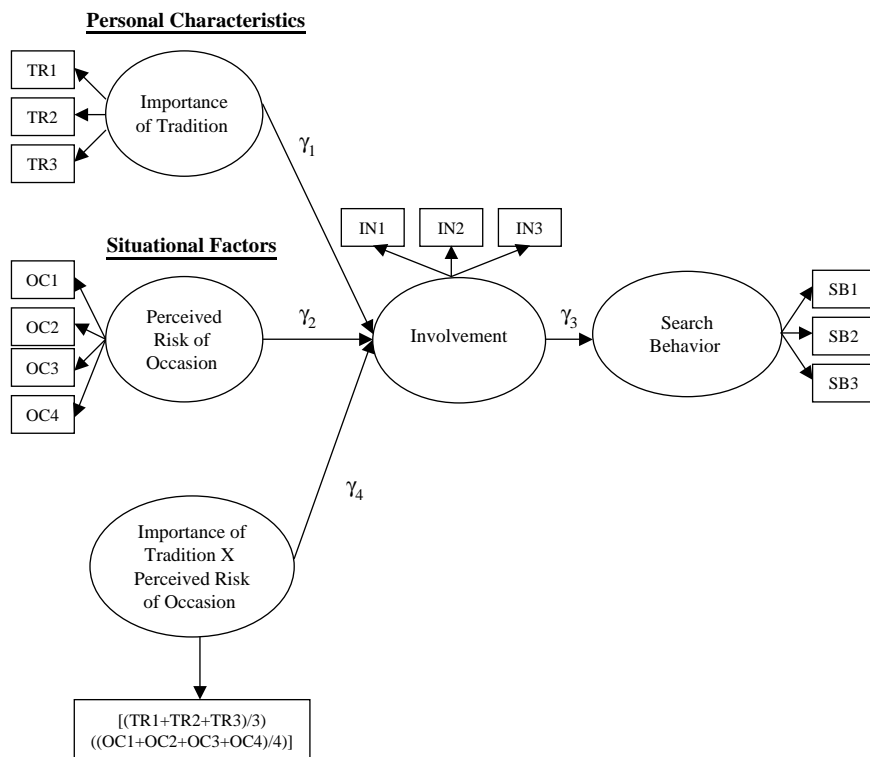


Figure 2. Hypothesized model of involvement in search behavior

Measurement invariance

Measurement invariance addresses the question whether the same models hold across different populations/cultures/countries. As illustrated in Table II, we tested for measurement invariance across the three countries following the method suggested by Myers *et al.* (2000) and examined the moderation effects of form and strength following the methods suggested by Ping (1995) and Bentler (1995).

A test of configural variance (Model 1) produced $\chi^2 = 567.783$ with 177 degrees of freedom (df) as shown in Table III. CFI was found to be 0.963 and RMSEA was 0.034. On the basis of these fit indices, the four-factor model is supported in all groups. These findings reveal that the three countries exhibit the same simple factor structure, and the same baseline model can be used for each sample population.

The second model (Model 2) examines whether or not the factor covariances are equal across the three countries. This was accomplished by constraining the φ s to be equal across groups. The results in Table III show that the chi-square difference between the simple structure model (Model 1) and the equal variance model (Model 2) was 11.183 with 12 df. This is not significant at $p < 0.05$, which indicates that the factor correlations are invariant across the three countries.

In Model 3, the factor loadings (λ s) for the three countries were set to be equal for all questionnaire items. This produced $\chi^2 = 614.668$, with $df = 203$, $CFI = 0.961$, and $RMSEA = 0.033$ (Table III). Accordingly, there was a significant difference in

	USA	Item loadings Australia	New Zealand
<i>Importance of tradition (TR)</i>			
TR1	0.81	0.85	0.85
TR2	0.75	0.73	0.77
TR3	0.62	0.58	0.62
Average variance extracted (percent)	53.2	53.0	56.9
Highest shared variance (percent)	1.0	0.5	3.0
<i>Perceived risk of occasion (OC)</i>			
OC1	0.79	0.79	0.79
OC2	0.91	0.88	0.84
OC3	0.91	0.90	0.91
OC4	0.84	0.83	0.77
Average variance extracted (percent)	74.6	72.4	68.7
Highest shared variance (percent)	11.0	16.0	8.0
<i>Involvement (IN)</i>			
IN1	0.73	0.77	0.67
IN2	0.88	0.80	0.88
IN3	0.56	0.76	0.63
Average variance extracted (percent)	53.4	60.0	53.9
Highest shared variance (percent)	11.0	16.0	8.0
<i>Search behavior (SB)</i>			
SB1	0.59	0.68	0.60
SB2	0.66	0.68	0.67
SB3	0.82	0.81	0.85
Average variance extracted (percent)	48.4	52.8	51.0
Highest shared variance (percent)	2.0	2.0	3.0
<i>t</i> -value range	16.13-34.56	15.36-31.33	9.17-18.34
Model fit statistics			
χ^2	205.301	239.709	122.772
Degrees of freedom	59	59	59
<i>p</i> -value	0.00	0.00	0.00
Bentler-Bonett nonnormed fit index (NNFI)	0.960	0.945	0.937
Comparative fit index (CFI)	0.970	0.959	0.952
Bollen fit index (IFI)	0.970	0.959	0.953
Root mean square of error (RMSEA)	0.053	0.063	0.065

Table I.
Results of the CFA,
country specific analyses

chi-square values between the constrained model (Model 3) and the unconstrained model (Model 1) ($\Delta\chi^2 = 46.885$, $\Delta df = 26$, $p < 0.05$). Consequently, we tested Model 4 in which both the factor correlations (φ s) and the factor loadings (λ s) were constrained to be equal. The difference between the model in which the loadings were constrained to be equal (Model 3) and the model in which both the loadings and the factor correlations were constrained to be equal (Model 4) was not significant ($\Delta\chi^2 = 12.081$, $\Delta df = 12$, $p < 0.05$). These two results indicate that although the factor structure is consistent across the three countries, the latent constructs are composed differently with respect to the measured variables. For instance, the LM test indicated that 4 of the 13 factor loadings were variant ($p < 0.05$) between the USA and Australian samples and the USA and New Zealand samples. However, all factor loadings were invariant between the Australian and New Zealand samples. In addition, the factor loading from each indicator to its respective construct was highly significant ($p < 0.01$).

Analysis	Model	Purpose
Measurement model (Bentler, 1995)	Perform confirmatory factor analysis (CFA) for the USA, Australia, and New Zealand individually	Test construct validity. Eliminate factors with low loadings or loadings on multiple constructs
Measurement invariance (Myers <i>et al.</i> , 2000)	Perform multi-group CFA (Model 1) for the USA, Australia, and New Zealand simultaneously in which no constraints are imposed across groups	Test for configural invariance across countries
	Perform multi-group CFA (Model 2) by constraining the correlations among the factors in three groups	Test nomological validity
	Perform multi-group CFA (Model 3) by constraining the factor loadings across three groups	Establish metric equivalence
	Perform multi-group CFA (Model 4) by constraining both the factor correlations and the factor loadings across three groups	Examine the composition of constructs with respect to measured variables across countries
Hypothesis testing (Ping, 1995)	Test hypothesized model with individual country sample including interaction effects	Evaluate form moderation
Hypothesis testing (Bentler, 1995)	Test hypothesized model by performing multi-group structural model for the USA, Australia, and New Zealand simultaneously	Examine structural equivalence across countries and evaluate strength moderation

Table II.
Data analysis

When we aggregated the results yielded from:

- comparing Model 3 to Model 1;
- comparing Model 3 to Model 4; and
- examining the LM test outcomes in Model 3, we concluded that the measurement models were invariant across the USA, Australian, and New Zealand samples (Calantone *et al.*, 1996).

Moderation of form or strength

We next modeled the moderators by evaluating whether the factors modify the strength or the form of the relationships using the three-step process suggested by Sharma *et al.* (1981) (Figure 1). We first examined “perceived risk of occasion”. Step 1 determines whether a

Model	χ^2	df	<i>p</i> -value	CFI	RMSEA
Model 1: Multi-group – unconstrained	567.783	177	0.00	0.963	0.034
Model 2: Multi-group – φ constrained	578.966	189	0.00	0.963	0.033
Model 3: Multi-group – λ constrained	614.668	203	0.00	0.961	0.033
Model 4: Multi-group – φ, λ constrained	626.749	215	0.00	0.961	0.032

Table III.
Measurement invariance tests through a CFA constrained at several levels

significant interaction is present between the hypothesized moderator (perceived risk of occasion) and the predictor variable (importance of tradition) with the criterion variable (involvement) using moderated multiple regression analysis. In this case, we found β_3 , the interaction coefficient, to be significant. If β_3 is not equal to zero, Z is said to moderate the relationship between X and Y . However, Carte and Russell (2003) caution that it is not enough to solely evaluate β_3 . In addition, the F -statistic [1] examining the difference in R^2 must be calculated. An F -statistic significantly greater than 1.00 leads to a rejection of H_0 : $\Delta R^2 = 0$. For “perceived risk of occasion” the F -statistic is 4.368, which is significantly greater than 1, and β_3 is 0.125, $p < 0.06$. Thus, it can be concluded that “perceived risk of occasion” is a form moderator. Based on our earlier discussion, the SEM indicant product guidelines of Ping (1995) were used to test interactions between these two latent independent variables. This method allows modeling of errors and avoids some of the problems associated with other methodologies.

We also hypothesized the moderating effect of country across the relationships in the model. Using the Sharma *et al.* (1981) procedure, in step 1 we found that country does not interact with either of the predictor variables, importance of tradition nor perceived risk of occasion. Step 2 revealed that country was related to each predictor, thus, country is not a moderator but an antecedent to involvement. However, Sauer and Dick (1993) contend that theory should be used to define the functional form of the model, and as such we model country as a strength moderator using multi-group analysis.

Interaction effect estimation in SEM

To estimate the effects of an interaction variable identified as “importance of tradition*perceived risk of occasion” in SEM, we utilized a two-step estimation approach suggested by Ping (1995) for modeling form moderators. This approach involves calculating the loading and error variance for the single indicator of the latent product using measurement model parameter estimates, and then fixing the loading and error variance at their calculated values in the structural model. We next explain the steps involved in the interaction analysis.

The first step was to center the variables, importance of tradition, perceived risk of occasion, involvement, and search behavior by replacing their raw scores with deviation scores. After centering the variables we summed the indicators of each latent variable to form the latent product. The summated scores for importance of tradition and perceived risk of occasion were multiplied to form the single indicator of the latent product ($TR^*OC = [(TR1 + TR2 + TR3)/3]*[(OC1 + OC2 + OC3 + OC4)/4]$).

The third step involved estimating the measurement model for each country. As per Ping (1995), the λ , θ , and φ values of the main effect latent variables from this analysis were used to compute λ_{tr*oc} and $\theta_{e_{tr*oc}}$ values for the indicator of the latent product. These values are used in the multiplicative model in which the latent product (importance of tradition*perceived risk of occasion) is included. The loading for the indicator of the product latent variable (λ_{tr*oc}) and its error variance ($\theta_{e_{tr*oc}}$) are calculated by the following equations (Ping, 1995):

$$\lambda_{tr*oc} = \Gamma_{tr} * \Gamma_{oc}$$

$$\theta_{e_{tr*oc}} = \Gamma_{tr}^2 \text{Var}(TR)\theta_{oc} + \Gamma_{oc}^2 \text{Var}(OC)\theta_{tr} + \theta_{tr} * \theta_{oc}$$

As noted by Anderson and Gerbing (1988), the unidimensionality of the latent variables must be verified based on estimates from the measurement model when

fixing certain parameter values in a structural model. Based on the measurement model results for each of the three countries, it was concluded that the latent main effect variables (importance of tradition and perceived risk of occasion) were reasonably unidimensional. Finally, we tested the multiplicative model using the λ_{tr*oc} and θ_{tr*oc} of the latent product variable. Next we demonstrate our results derived from the interaction analysis.

SEM analyses and hypothesis testing

As shown in Figure 2, the hypothesized model was estimated for each of the three countries individually. The US model converged well (as shown in Table IV). We found importance of tradition to be positively associated with involvement ($\gamma_1 = 0.153$, $t = 3.605$), supporting *H1*. Perceived risk of occasion had a significant, yet a negative impact on involvement ($\gamma_2 = -0.440$, $t = -6.995$) although we had hypothesized a positive relationship; thus *H2* was not supported. The path from involvement to search behavior was significant ($\gamma_3 = 0.152$, $t = 3.543$), supporting *H3*. Finally, the product term related to perceived risk of occasion and importance of tradition was positively associated with involvement ($\gamma_4 = 0.205$, $t = 2.362$) supporting *H4*.

The model for the Australian sample converged well (Table IV). Again, a negative relationship was observed for *H2* ($\gamma_2 = -0.407$, $t = -5.008$) so it was not supported. None of the other hypotheses were significant, and thus, are not supported. The New Zealand model also converged well (Table IV); two of the four paths were significant. However, perceived risk of occasion was found to be negatively associated with involvement ($\gamma_2 = -0.305$, $t = -2.229$), thus *H2* was not supported. The path from involvement to search behavior was significant ($\gamma_3 = 0.158$, $t = 2.017$), supporting *H3*.

Multi-group model

Next we evaluated the moderating effect of the strength moderator, country, by using a multi-group comparison method of EQS version 6.1 (Bentler, 1995). The objective of multi-group simultaneous path analysis is to determine whether the path coefficients for the relationships between importance of tradition, perceived risk of occasion, involvement and search behavior were equal across the three countries. We first constrained one path to be invariant across the three countries and then freely estimated this path. The LM method in EQS was used to determine which paths were different. The LM tests showed that the tradition-involvement path was variant for Australia (AUS) when compared to the USA (US) and New Zealand (NZ); the US and NZ paths were invariant. For the perceived risk of occasion-involvement path NZ was variant compared to US and AUS, which showed invariance between these two latter countries. Only for AUS was the involvement-search behavior path insignificant; for the US and NZ it was highly significant and invariant. As demonstrated in Table IV, the interaction term (importance of tradition * perceived risk of occasion) was only significant for the US; the paths for AUS and NZ were insignificant and invariant. Thus, *H5* was partially supported.

Discussion

When evaluating moderating effects it is important to understand which methodology is most appropriate for the effect being studied. This is especially true in cross-national

Table IV.
Results of structural
equation model analysis
of individual country
models

Path coefficients	Hypothesis	Standardized loading (<i>t</i> -value)				Multi-group
		USA	Australia	New Zealand		
TR → IN (γ_1)	H1	0.153 (3.605)**	0.055 (0.852) ^{NS}	0.199 (0.563) ^{NS}	US = NZ** > AUS ^{NS}	
OC → IN (γ_2)	H2	-0.440 (-6.995)**	-0.407 (-5.008)**	-0.305 (-2.229)*	US = AUS** > NZ**	
IN → SB (γ_3)	H3	0.152 (3.543)**	0.058 (1.301) ^{NS}	0.158 (2.017)*	US = NZ** > AUS ^{NS}	
TR*OC → IN (γ_4)	H4	0.205 (2.362)*	-0.006 (-0.048) ^{NS}	-0.077 (-0.183) ^{NS}	US = NZ** > AUS = NZ ^{NS}	
Model fit statistics						
χ^2		232.805	270.692	144.531	676.440	
Degrees of freedom		72	72	72	234	
<i>p</i> -value		0.00	0.00	0.00	0.00	
Bentler-Bonett nonnormed fit index (NNFI)		0.959	0.944	0.934	0.952	
Comparative fit index (CFI)		0.968	0.955	0.948	0.959	
Bollen fit index (IFI)		0.968	0.956	0.949	0.959	
Root mean square of error (RMSEA)		0.050	0.060	0.062	0.032	

Notes: ^{NS} = not significant; * *p* = 0.05; ** *p* = 0.01; TR: importance of tradition; IN: involvement; OC: perceived risk of occasion; SB: search behavior

research where moderation examination is common. Accordingly, this study illustrates a three-step process for modeling moderation in international marketing research. First, we establish measurement invariance across populations/countries/ cultures. Next we identify potential moderators of strength or form. Finally, we extend upon existing methodological foundations by demonstrating how to model multiplicative interactions using a single indicator of the latent product (product of main effect variables) with the use of a multi-group analysis in SEM.

Adhering to these analytic techniques, we utilized the Andrews *et al.* (1990) conceptualization for the involvement construct to investigate the differences in search behavior for wine across three countries. We demonstrated that involvement does vary across countries as it positively influences search behavior in the USA and New Zealand but not in Australia. Although we had expected a positive relationship between perceived risk of occasion and involvement, it was negative for all three countries. This indicates an inverse relationship between perceived risk and involvement. Knowledge may be a moderating factor in this case. High involvement is often associated with higher knowledge about a product category (Celsi and Olson, 1988), which is common with wine. Therefore, more involved consumers would actually not be affected by purchasing a bottle of wine for a formal party, whereas less involved consumers would find the situation intimidating. Additionally, we looked at the interaction effects between importance of tradition and perceived risk of an occasion (described as choosing a bottle of wine for a formal dinner party), where perceived risk of occasion is modeled as a form moderator. This interaction effect was only significant for the USA, implying that the perceived risk of the occasion interacts with tradition to amplify its relationship with involvement contingent upon country.

We draw upon the suggestions of Baron and Kenny (1986), Sharma *et al.* (1981) and Sauer and Dick (1993) to demonstrate that the analytical techniques chosen to model moderation should be based upon measurement level, type of moderation being investigated and theory. We suggest that it is important for cross-national studies to use multi-group analysis when evaluating the differences in relationships across countries. Following the guidelines of Calantone *et al.* (1996) ensures that the researcher evaluates measurement invariance across sample populations and builds a model that can be easily interpreted. Dummy variables and MMRs can easily lead to misinterpretation of the results, whereas ANOVA can lead to ignoring variances in measures across different populations/countries/cultures. The unit of analysis should not lead researchers to favor one methodology over another; instead this decision should be guided by the type of moderator being investigated, strength or form (Sharma *et al.*, 1981).

We also encourage cross-national studies to include contingencies in order to advance our knowledge base in international marketing research. Ping (1995) provides a model that can easily incorporate multiplicative interactions into a multi-group international study. We demonstrated how this can be done and found significant differences across the three countries studied regarding the interaction between importance of tradition and perceived risk of occasion. Investigating multiple moderators in a single study is important to advance our knowledge in cross-national/culture studies.

Note

$$1. \quad F_{(df_{>multi} - df_{add}, N - df_{multi} - 1)} = \frac{\Delta R^2 / df_{multi} - df_{add}}{(1 - R_{multi}^2) / (N - df_{multi} - 1)};$$

where R_{multi}^2 refers to $Y = a + b_1X + b_2Z + b_3XZ$ and R_{add}^2 refers to $Y = a + b_1X + b_2Z$.

References

- Anderson, J.C. and Gerbing, D.W. (1988), "Structural equation modeling in practice: a review and recommended two-step approach", *Psychological Bulletin*, Vol. 103 No. 3, pp. 411-23.
- Andrews, C., Durvasula, S. and Akhter, S. (1990), "A framework for conceptualizing and measuring the involvement construct in advertising research", *Journal of Advertising*, Vol. 19 No. 4, pp. 27-40.
- Atkin, T., Garcia, R. and Lockshin, L. (2005), "A multinational study of the diffusion of a discontinuous innovation", working paper, Institute of Global Innovation Management, Northeastern University, Boston, MA.
- Aurifeille, J.M., Quester, G.P., Lockshin, L. and Spawton, T. (2002), "Global vs international involvement-based segmentation", *International Marketing Review*, Vol. 19 No. 4, pp. 369-86.
- Bagozzi, R.P. and Yi, Y. (1988), "On the evaluation of structural equation models", *Journal of the Academy of Marketing Science*, Vol. 16 No. 1, pp. 74-94.
- Baron, R.M. and Kenny, D.A. (1986), "The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations", *Journal of Personality and Social Psychology*, Vol. 51 No. 6, pp. 1173-82.
- Beatty, S.E. and Smith, S.M. (1987), "External search effort: an investigation across several products categories", *Journal of Consumer Research*, Vol. 14 No. 1, pp. 83-95.
- Bentler, P.M. (1995), *EQS Structural Equations Program Manual*, Multivariate Software, Encino, CA.
- Bettman, J.R. (1973), "Perceived risk and its components: a model and empirical test", *Journal of Marketing Research*, Vol. 10 No. 2, pp. 184-90.
- Bloch, P.H. and Richins, M.L. (1983), "A theoretical model for study of product importance perceptions", *Journal of Marketing*, Vol. 47 No. 3, pp. 69-81.
- Bollen, K.A. (1989), *Structural Equations with Latent Variables*, Wiley, New York, NY.
- Buvik, A. and Andersen, O. (2002), "The impact of vertical coordination on ex-post transaction costs in domestic and international buyer-seller relationships", *Journal of International Marketing*, Vol. 10 No. 1, pp. 1-24.
- Calantone, R.J. and Zhao, Y.S. (2001), "Joint ventures in China: a comparative study of Japanese, Korean, and US partners", *Journal of International Marketing*, Vol. 9 No. 1, pp. 1-23.
- Calantone, R.J., Schmidt, J.B. and Song, M.X. (1996), "Controllable factors of new product success: a cross national comparison", *Marketing Science*, Vol. 15 No. 4, pp. 341-58.
- Carte, T.A. and Russell, C.J. (2003), "In pursuit of moderation: nine common errors and their solutions", *MIS Quarterly*, Vol. 27 No. 3, pp. 479-501.
- Celsi, R. and Olson, J. (1988), "The role of involvement in attention and comprehension processes", *Journal of Consumer Research*, Vol. 15 No. 2, pp. 210-24.

-
- Chan, T.S. (1996), "Concerns for environmental issues and consumer purchase preferences: a two-country study", *Journal of International Consumer Marketing*, Vol. 9 No. 1, pp. 43-55.
- Churchill, G.A. Jr (1979), "A paradigm for developing better measures of marketing constructs", *Journal of Marketing Research*, Vol. 16 No. 1, pp. 64-73.
- Cohen, J. and Cohen, P. (1983), *Applied Multiple Regression/Correlation Analysis for the Behavior Sciences*, 2nd ed., Erlbaum, Hillsdale, NJ.
- Dholakia, U.M. (2001), "A motivational process model of product involvement and consumer risk perception", *European Journal of Marketing*, Vol. 35 Nos 11/12, pp. 1340-60.
- Douglas, S. and Craig, S. (1983), *International Marketing Research*, Prentice-Hall, Englewood Cliffs, NJ.
- Ford, J.B., LaTour, M.S. and Honeycutt, E.D. Jr (1997), "An examination of the cross-cultural female response to offensive sex role", *International Marketing Review*, Vol. 14 No. 6, pp. 409-23.
- Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18 No. 1, pp. 39-50.
- Goldsmith, R.E., d'Hauteville, F. and Flynn, L.R. (1998), "Theory and measurement of consumer innovativeness", *European Journal of Marketing*, Vol. 32 Nos 3/4, pp. 340-53.
- Granzin, K.L. and Painter, J.J. (2001), "Motivational influences on 'buy domestic' purchasing: marketing management implications from a study of two nations", *Journal of International Marketing*, Vol. 9 No. 2, pp. 73-94.
- Hall, J. and Lockshin, L. (2000), "Using means-end chains for analyzing occasions not buyers", *Australian Marketing Journal*, Vol. 8 No. 1, pp. 45-54.
- Homburg, C., Kuester, S., Beutin, N. and Menon, A. (2005), "Determinants of customer benefits in business-to-business markets: a cross-cultural comparison", *Journal of International Marketing*, Vol. 13 No. 3, pp. 1-31.
- Jacoby, J. and Kaplan, L. (1972), "The components of perceived risk", in Venkatesan, M. (Ed.), *Proceedings of the 3rd Annual Conference Association for Consumer Research Champaign*, pp. 382-93.
- Judd, C.M., McClelland, G.H. and Culhane, S.E. (1995), "Continuing issues in the everyday analysis of psychological data", *Annual Review of Psychology*, Vol. 46 No. 1, pp. 433-65.
- Kassarjian, H. (1981), "Low involvement: a second look", in Hunt, H.K. (Ed.), *Advances in Consumer Research*, Association for Consumer Research, Ann Arbor, MI.
- Keillor, B.D., Hult, G.T.M. and Kandemir, D. (2004), "A study of the service encounter in eight countries", *Journal of International Marketing*, Vol. 12 No. 1, pp. 9-35.
- Kenny, D.A. and Judd, C.M. (1984), "Estimating the nonlinear and interactive effects of latent variables", *Psychological Bulletin*, Vol. 96 No. 2, pp. 201-10.
- Kim, K. and Oh, C. (2002), "On distributor commitment in marketing channels for industrial products: contrast between the United States and Japan", *Journal of International Marketing*, Vol. 10 No. 1, pp. 72-97.
- Laurent, G. and Kapferer, J.N. (1985), "Measuring consumer involvement profiles", *Journal of Marketing Research*, Vol. 22 No. 1, pp. 41-53.
- Lockshin, L.S., Spawton, A.L. and MacIntosh, G. (1997), "Using product, brand, and purchasing involvement for retail segmentation", *Journal of Retailing and Consumer Services*, Vol. 4 No. 3, pp. 171-83.

- Myers, M.B., Calantone, R.J., Page, T.J. and Taylor, C.R. (2000), "An application of multiple-group causal models in assessing cross-cultural measurement equivalence", *Journal of International Marketing*, Vol. 8 No. 4, pp. 108-21.
- Perugini, M. and Bagozzi, R.P. (1999), "The role of desires and anticipated emotions in goal-directed behaviors: expanding and deepening the theory of planned behavior", working paper, University of Michigan, Ann Arbor, MI.
- Petty, R. and Cacioppo, J. (1986), *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, Springer-Verlag, New York, NY.
- Ping, R.A. Jr (1995), "A parsimonious estimating technique for interaction and quadratic latent variables", *Journal of Marketing Research*, Vol. 32 No. 3, pp. 336-47.
- Prescott, J.E. (1986), "Environments as moderators of the relationship between strategy and performance", *Academy of Management Journal*, Vol. 29 No. 2, pp. 329-46.
- Quester, P.G. and Smart, J. (1998), "The influence of consumption situation and product involvement over consumers' use of product attribute", *Journal of Consumer Marketing*, Vol. 15 No. 3, pp. 220-38.
- Richins, M.L., Bloch, P.H. and McQuarrie, E.F. (1992), "How enduring and situational involvement combine to create involvement responses", *Journal of Consumer Psychology*, Vol. 1 No. 2, pp. 143-53.
- Rigdon, E., Schumacker, R.E. and Wothke, W. (1998), "A comparative review of interaction and nonlinear modeling", in Schumacker, R.E. and Marcoulides, G.A. (Eds), *Interaction and Nonlinear Effects in Structural Equation Modeling*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Russell, C.I. and Bobko, P. (1992), "Moderated regression analysis and likert scales: too coarse for comfort", *Journal of Applied Psychology*, Vol. 77 No. 3, pp. 336-42.
- Sauer, P.L. and Dick, A. (1993), "Using moderator variables in structural equation models", *Advances in Consumer Research*, Vol. 20 No. 1, pp. 637-40.
- Sharma, S., Durand, R.M. and Gur-Arie, O. (1981), "Identification and analysis of moderator variables", *Journal of Marketing Research*, Vol. 18 No. 3, pp. 291-300.
- Song, M.X. and Parry, M.E. (1997), "A cross national comparative study of new product development processes", *Journal of Marketing*, Vol. 61 No. 2, pp. 1-18.
- Song, M.X. and Xie, J. (2000), "Does innovativeness moderate the relationship between cross-functional integration and product performance?", *Journal of International Marketing*, Vol. 8 No. 4, pp. 61-89.
- Tansey, R., Hyman, M.R., Zinkhan, G.M. and Chowdury, J. (1997), "An advertising test of the work ethic in the UK and the US", *Journal of International Consumer Marketing*, Vol. 9 No. 3, pp. 57-77.
- Vanhonacker, W.R. and Pan, Y. (1997), "The impact of national culture, business scope, and geographic location on joint venture operations in China", *Journal of International Marketing*, Vol. 5 No. 3, pp. 11-30.

Further reading

- Lockshin, L.S., Quester, P. and Spawton, T. (2001), "Segmentation by involvement or nationality for global retailing: a cross-national comparative study of wine shopping behaviours", *Journal of Wine Research*, Vol. 12 No. 3, pp. 223-36.

	USA	Composite reliabilities Australia	New Zealand
Importance of Tradition (scale: not important to important)	0.77	0.77	0.80
1. The tradition of opening wine sealed with a cork (uncorking the bottle and sniffing the cork)			
2. The sound of the cork "pop" when opening a bottle			
3. The ritual of opening wine (i.e. the presentation and first taste at a restaurant)			
Perceived Risk of Occasion (scale: strongly disagree to strongly agree)	0.92	0.91	0.90
4. I find choosing the right wine for a formal party daunting because of the abundance of options available			
5. I find choosing the right wine for a formal party a bit intimidating as I might make a mistake			
6. I find choosing the right wine for a formal party a difficult task			
7. I find choosing the right wine for a formal party threatening because of how others will judge my selection			
Involvement (scale: strongly disagree to strongly agree)	0.77	0.82	0.77
8. I have a strong interest in wine			
9. Wine is important to me in my lifestyle			
10. Drinking wine gives me pleasure			
Search Behavior (scale: strongly disagree to strongly agree)	0.73	0.77	0.75
11. Consult the restaurant menu recommendations			
12. Consult with store personnel			
13. Consult with restaurant/winery personnel (server, sommelier, etc.)			

Notes: ^aThese questions were prefaced with: "We have a few additional questions we would like to ask you regarding the role of wine purchases in your lifestyle." Details of this wine choice survey are available upon request from the first author

Table AI.
Measurement items and reliabilities^a

About the authors

Rosanna Garcia is the Assistant Professor of marketing at Northeastern University, Boston, MA. She received her PhD in Marketing from Michigan State University. Her research interests are in new product development and complex systems. She has published in the *Journal of Product Innovation Management*, *Decision Sciences* and other journals. Rosanna Garcia is the corresponding author and can be contacted at: r.garcia@neu.edu

Destan Kandemir is Visiting Professor of marketing at Seattle University. She earned her PhD in Marketing and International Business from Michigan State University. She has published articles in the *Journal of Management*, *Journal of International Marketing*, and others. Her research interests are in market-oriented knowledge management and global alliance management.

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: www.emeraldinsight.com/reprints



Reconsidering the problem of data equivalence in international marketing research

Contrasting approaches based on CFA and the Rasch model for measurement

Thomas Salzberger

*Vienna University of Economics and Business Administration,
Vienna, Austria, and*

Rudolf R. Sinkovics

Manchester Business School, UK

Abstract

Purpose – The paper investigates the suitability of the Rasch model for establishing data equivalence. The results based on a real data set are contrasted with findings from standard procedures based on CFA methods.

Design/methodology/approach – Sinkovics *et al.*'s data on technophobia was used and re-evaluated using both classical test theory (CTT) (multiple-group structural equations modelling) and Rasch measurement theory.

Findings – Data equivalence in particular and measurement in general cannot be addressed without reference to theory. While both procedures can be considered best practice approaches within their respective theoretical foundation of measurement, the Rasch model provides some theoretical virtues. Measurement derived from data that fit the Rasch model seems to be approximated by classical procedures reasonably well. However, the reverse is not necessarily true.

Practical implications – The more widespread application of Rasch models would lead to a stronger justification of measurement, in particular, in cross-cultural studies but also whenever measures of individual respondents are of interest.

Originality/value – Measurement models outside the framework of CTT are still scarce exceptions in marketing research.

Keywords Data analysis, Set theory, Measurement, Cross-cultural studies

Paper type Research paper

1. Introduction and purpose

Advances in International Marketing research rely on cross-country comparisons and require sound conceptualisations and empirical generalisations. Meaningful comparisons are possible only if data are derived from equivalent constructs and associated measures. Issues of reliability, validity as well as dimensional consistency need to be addressed (Davis *et al.*, 1981; van de Vijver and Leung, 1997) in order to avoid misinterpretation of results and fallacious conclusions for managerial decision making.

On the one hand, advancements in statistical and computational methodologies and procedures (Tabachnick and Fidell, 2001) have helped solve methodological problems. On the other hand, applied researchers and practitioners are faced with an accumulation



of competing frameworks and methods for analysing cross-national datasets (van de Vijver and Leung, 1997). Scholars have looked at measurement equivalence issues from both a qualitative and quantitative perspective (Mullen, 1995; Salzberger *et al.*, 1999; Singh, 1995). There is also a heated discussion around scale development frameworks and the nature of operationalisations, incited by Rossiter, 2002 C-OAR-SE paradigm (Diamantopoulos, 2005; Rossiter, 2005). Empirically, multi-group structural equations modelling (Steenkamp and Baumgartner, 1998), which is grounded in classical test theory (CTT), and to a lesser degree generalizability theory as an extension of CTT (Finn and Ujwal, 2005; Rentz, 1987) have gained popularity.

The main purpose of this paper is to examine the use and applicability of an alternative measurement approach, Rasch measurement theory (RMT) which goes back to the Danish statistician and measurement pioneer Rasch (1980). Although measurement issues are listed high on the agenda of international marketing researchers, and despite enjoying an increasing popularity in disciplines such as psychology or rehabilitation medicine this particular theory has not yet received the attention in International Marketing it possibly deserves.

In order to examine the usefulness and potential contribution of RMT we examine the construct “technophobia” in a multinational context. We employ both the popular multigroup confirmatory factor analysis approach (MG-CFA) and the Rasch methodology to a dataset of more than 900 respondents from the UK, Mexico and Austria. Technophobia has been introduced by Sinkovics *et al.* (2002) to assess a negative, anxious (phobic) attitude of consumers towards certain innovative products, which results in their being less open to these products, feeling uncomfortable when using them and disregarding technological benefits related to their use. The reluctance of consumers to buy new technology-driven products represents a key threat to a successful and fast diffusion of the market. In the context of international marketing, the awareness of different levels of technophobia can affect market entry decisions and may help design marketing actions to overcome possibly substantial levels of technophobia.

The present study differs from previous research in two significant ways. Firstly, we transcend mostly conceptual discussions regarding advantages of one approach over the other (Ewing *et al.*, 2005; Fan, 1998; Salzberger *et al.*, 1999; Singh, 2004) and address their empirical application. Rather than advocating a particular approach a priori, we attempt to compare approaches empirically. The multigroup CFA approach has been widely used in marketing and international business (Steenkamp and Baumgartner, 1998) and is considered the state-of-the-art methodology in international research. The alternative measurement approach based on Rasch (1980) is still largely underutilized, however, the methodology builds upon a more fundamental definition of measurement. Secondly, rather than using simulated data (Meade and Lautenschlager, 2004; Salzberger *et al.*, 1999), we present a real set of multi-national data. This serves the need of many marketing practitioners and researchers alike to experience the methodology “in real life” rather than in controlled laboratory settings.

2. Problem of data equivalence

International research almost inevitably crosses cultural boundaries. Admittedly, nationality is by no means an ideal definition of culture. However, for the present purpose, nationality serves as an acceptable approximation of cultural affiliation. We do not investigate the sub-cultural level within a nation but we are aware of the fact

that intra-cultural differences might be as relevant as cross-national factors. The point is when multiple populations are to be compared, the issue of data equivalence (Salzberger *et al.*, 1999), also referred to as measurement invariance (Vandenberg, 2002), emerges as a relevant matter of investigation. Data are equivalent across populations when measures bear the same meaning. Therefore, data equivalence is the prerequisite of comparability. Disregarding the question of data equivalence may lead to wrong conclusions. Firstly, observed differences in measures might be attributed to substantial differences between populations although the observed differences are solely caused by differential response behaviour and not by real differences in the latent variables. Secondly, true differences might be masked by differential response behaviour and remain undetected. Consequently, in mean comparisons, both type one and type two errors are increased in an uncontrollable way.

The most prominent case of multiple populations pertains to cross-cultural data. The need to establish and test for cross-cultural equivalence has become accepted in intercultural marketing research. However, there is no final consensus on a specific methodological approach to determine data equivalence. Since, data equivalence refers to differences in quantity as a result of differences in quality, we have to deal with both qualitative and quantitative issues (Salzberger *et al.*, 1999). A typical cross-cultural study entails the application of a scale that is translated into a different language. If the wording of a translated version of an item has a slightly different meaning leading to, say, a higher manifest score compared to the original version given the same latent score, then the qualitative difference in the meaning of the item causes a quantitatively different response depending on the language group. Consequently, every effort has to be taken to ensure comparability of the data during preparation of the instrument (e.g. appropriate translation techniques), administration of the data collection (e.g. comparable setting of interviewing respondents) and so forth. It should be noted, however, that we finally need to empirically verify the actual equivalence of the data.

While there is agreement on the necessity of addressing data equivalence explicitly and quantitatively rather than merely examining the administration and the design of the study in a qualitative fashion, there are different approaches to actually test for data equivalence. There is a good reason why this issue cannot be resolved easily. Dealing with quantitative differences in the measures of different groups can, by definition, not be disentangled from the very issue of measurement itself. Firstly, we have to decide on what theory of measurement we want to rely on, before we can analyse potentially different functioning of items.

3. Conceptual foundations

3.1 Approaches to test for data equivalence

Currently, there are two approaches identified as best practices for assessing equivalence (Reise *et al.*, 1993; Schaffer and Riordan, 2003). Within the paradigm of CTT (Churchill, 1979; Lord and Novick, 1968), the MG-CFA lends itself to testing data equivalence. Introductions into the MG-CFA approach are widely available both within the marketing domain and other disciplines (Cheung and Rensvold, 1998; Steenkamp and Baumgartner, 1998; Vandenberg, 2002). As for the international marketing research community, Steenkamp and Baumgartner, 1998 contribution enjoys a particularly high diffusion and the MG-CFA has become an established approach in marketing research over the recent years.

The alternative approach is based on item response theory (IRT), however IRT applications remain scarce exceptions. The reason is the slow penetration of the measurement theory itself. Embretson and Reise (2000) provide an excellent general introduction to IRT. We, therefore, concentrate on the most important issues and, in particular, focus on those aspects wherein the methods differ substantially.

3.2 Assessing data equivalence by multigroup confirmatory factor analysis

Under CTT, the observed response x_i is a linear combination of a true score and error. In the factor analytic model this translates to an observed response x_i explained by j latent variables (ξ_j) multiplied by the factor loadings λ_{ij} , an intercept τ_i , and an error term δ_I (Meade and Lautenschlager, 2004). Even in the case of multidimensional constructs, we usually relate one item uniquely to one factor (or latent variable). Without considering a person index, in a unidimensional model, the formula then simplifies to:

$$X_i = \tau_i + \lambda_i \xi + \delta_I \quad (1)$$

In one-group studies we usually disregard the intercept term τ_i because it is a constant across respondents and has no impact on (co-)variances. However, if more than one group is considered, differences in item intercept values for the same item in different groups do have an impact on the groups' means. Thus, it is absolutely essential to model item mean vectors and to address this issue in the investigation of data equivalence.

The investigation of data equivalence is carried out by evaluating models that are increasingly stringent (Steenkamp and Baumgartner, 1998). In the baseline model, for all groups considered, the same structure is imposed, i.e. all items are allocated to the same factors and the remaining loadings (non-salient loadings) are fixed to zero. It should be noted that the set of items need not necessarily be the same across groups, even though standard software such as LISREL (Jöreskog and Dag, 2003) require an identical set of items. Baumgartner and Steenkamp (1998) showed that imaginary variables with means of zero, variances of one and covariances of zero with all other variables can be introduced to balance an unequal number of items. The model has to be evaluated following the usual recommendations for judging CFA models. Provided fit is satisfactory, configural invariance is said to hold. Subsequently, constraints are imposed on the loading parameters, specifying the metric invariance model, which establishes a common metric across groups. The decrease in fit can be easily evaluated by a chi-square-difference test with degrees of freedom equal to the difference in the degrees of freedom of the two nested models. While metric invariance may hold for some items, it might not for others. The modification indices point out which items' loadings should be estimated unconstrained. The final model implies partial metric equivalence. For mean comparisons, the origin of the scale of the latent variable has to be defined in the same way for all groups. This requires constraints on the item intercept estimates for those items, which are metrically invariant. This model is termed the scalar invariance model.

Further tests of invariance of error variances and factor (co-)variances can be based on scalar invariance but are not essential for establishing data equivalence. Particularly interesting in the context of intercultural comparisons are constraints on the means of the latent variables. Such a model in comparison with the scalar invariance model can be used to test the hypothesis of equal means across groups.

3.3 Assessing data equivalence based on item response theory

Models based on CTT refer to aggregate statistics like variances, covariances and means. In contrast, in IRT, the manifest response is modelled directly in a probabilistic way. In other words, the model refers to the probability of responding positively, i.e. agreeing, coded as 1, as opposed to disagreeing, coded as 0, in a dichotomous item. In the Rasch (1980) model for dichotomous data (equation (2)), the response depends on person characteristics, covered by the person parameter β_v , as well as item characteristics, operationalised by the item parameter (δ_i). In a general context of measurement, we suggest using the neutral term item location parameter. This parameter has its closest parallel in the intercept parameter in CFA. However, there are fundamental differences. Firstly, the item location parameter is placed on the same scale as the person parameter. Consequently, item and person parameters can be compared directly. Secondly, a reasonable range of item locations in a scale is essential for determining the fit of the data to the model. Moreover, the hierarchy of items helps better understand the construct and adds to construct validity. So, we are always interested in the item location parameters while we usually ignore the intercept parameters in single group factor analysis. The one-parameter logistic model, also known as the Rasch model, is confined to this type of item parameter. We will concentrate on this model because it has some unique features which make it very attractive for measurement in social sciences.

More comprehensive IRT models (see Embretson and Reise, 2000 for an overview) introduce a further item parameter a_i , the item discrimination parameter. However, since varying item discrimination is incompatible with properties, which are important for measurement (primarily specific objectivity), we concentrate on the Rasch (1980) model (equation (2)). In this model, the a_i parameter is dropped from the equation tantamount to a discrimination parameter of 1 for all items.

$$P(X_{vi} = 1) = \frac{e^{(\beta_v - \delta_i)}}{[1 + e^{(\beta_v - \delta_i)}]} \quad (2)$$

x_{vi} , response of person v to item i ; β_v , person location parameter; δ_i , item location parameter (endorseability).

For polytomous data, the generalisation of the model is straightforward. Since, there are no assumptions about the scale level, in particular, the response scales are not assumed to be interval scales, polytomous items are characterised by a set of threshold parameters stating the boundaries of adjacent categories. For example, in a five-point rating scale type item, the first threshold tells us where the probability of the first category is equal to the probability of selecting the second. With m categories, we therefore need to estimate a set of $m - 1$ threshold parameters for each item. The threshold parameters can be constrained to be equal across items (rating scale model, Andrich, 1978) or be estimated independently for each item (partial credit model, Masters, 1982). Equation (3) states the formula of the general Rasch model for polytomous data (Andrich, 1988).

$$P(a_{vi} = x | \beta_v, \tau_{ij}, j = 1..m, 0 < x \leq m) = \frac{e^{\left(\sum_{j=1}^x - \tau_{ij}\right) + x \cdot (\beta_v - \delta_i)}}{\gamma} \quad (3)$$

with:

$$\gamma = 1 + \sum_{k=1}^m e^{\left(\sum_{j=1}^k -\tau_{ij} \right) + k \cdot (\beta_v - \delta_i)}$$

a_{vi} , answer of person v to item I (item score); β_v , person v location parameter; δ_i , item I location parameter; τ_{ij} , threshold parameter j of item i ; m maximum score, i.e. number of categories

The most important merit of the Rasch model is the feature of specific objectivity (Rasch, 1961, 1977). In essence it says that the item parameter estimates and the person parameter estimates are independent from one another, provided the data fit the model. In other words, the model has to be invariant against all possible groupings of respondents. The invariance property is a defining feature of the class of Rasch models. The relationship between invariance as a model feature and as a property of the data is crucial. According to Fan (1998):

... [t]he invariance property (...) makes it theoretically possible to solve some important measurement problems (...) such as (...) test equating and computerized adaptive testing.

However, the more fundamental question is whether measurement has been achieved, at all. Fan (1998) is concerned that the invariance property has been little explored empirically. He seems to suggest that invariance is a property IRT models are said to deliver but that may not hold in reality questioning the value of the model. From a Rasch perspective, the measurement model requires the invariance property of the data in order to provide measures that are comparable, i.e. measures that are on the same scale. Thus, the invariance property is highly important for measurement. If invariance does not hold in the data, the data lack a fundamental property required for measurement.

A test whether the model remains invariant for different cultural groups is therefore only a special case of testing the data model fit. If, however, an item has a different meaning for respondents from different cultures, then the item has a different location or may not even fit at all in one or more groups. Such a non-invariant item is said to display item differential functioning (DIF). A formal test of DIF can be based on the residuals in different groups. In a pooled data set, the mean of the residuals (i.e. the difference between the expected item score and the actual score) is zero. If there is no DIF, this also holds in subgroups, except for random variations. In case of DIF, the residuals deviate systematically from zero, i.e. the mean is positive in one group but negative in another group. A two-way ANOVA lends itself to test the difference for significance. One factor is the group, while the second is the class interval along the latent scale of parameters (Andrich *et al.*, 2003a, b).

In the absence of any DIF, full scalar equivalence holds and data equivalence is given. Like in the CFA approach, the Rasch model also allows for partial equivalence. An item displaying DIF can be split into several versions for each group. That way, a group specific item location parameter is estimated for each item affected by DIF. Moreover, it is very easy to retain an item for one group but discard the item for another. Splitting an item and subsequent deletion of the item in one culture can be carried out conveniently in standard software like RUMM 2020 (Andrich *et al.*, 2003b) whereas the equivalent in structural equation modeling software requires a new set up of the input matrices (Baumgartner and Steenkamp, 1998).

3.4 *Contrasting the approaches*

3.4.1 *Theoretical comparison.* Both, the CFA based approach and the Rasch based method to test for data equivalence are appropriate within their respective context. Thus, the evaluation of the competing approaches is best based on the underlying measurement theories. Ewing *et al.* (2005) undertook a comprehensive theoretical comparison and concluded that the Rasch model provides the more powerful basis of measurement in the social sciences. Singh (2004) contrasted CTT and a more general IRT model and pointed out that there are several issues that differ substantially. For example, IRT aims at establishing a broad bandwidth of the instrument, i.e. the scale should include items of widely varying locations. According to Singh (2004) this comes at the expense of reducing fidelity. Singh refrains from favouring either methodological framework but views them as complementary.

Regarding the prerequisites data have to meet, the Rasch model offers some interesting advantages. It can easily be applied to all possible scale formats. Dichotomous data as well as polytomous data or any combination of items with different response formats can all be treated in the same way. In particular, there are no assumptions made about the scale level. In contrast, in CFA we usually do assume a metric scale even though we know that this is extremely doubtful. A further advantage of the Rasch model is the independence of the distribution of respondents, which need not be normal or meet any other predefined shape. The stringency of the Rasch model referring to item discrimination favours CFA, which allows for item specific discrimination. It should be noted that more complex IRT models do incorporate discrimination parameters. In the Rasch model, discrimination is constrained to be equal across items for theoretical and philosophical reasons (Ewing *et al.*, 2005).

An evident benefit of the MG-CFA is its embedding in the standard measurement theory in marketing research - irrespective of possible theoretical virtues of the Rasch approach. If one wants to remain within the paradigm of CTT, the MG-CFA approach is appropriate and it will certainly prosper in marketing research. No clear differences can be found for sample sizes. In contrast to more complicated IRT models, the Rasch model works with about the minimum number of respondents usually recommended for CFA studies. The software to estimate the models has become more user friendly in both cases, so we cannot see any reason to favour either method simply because of the user-friendliness of the software.

Another issue is the handling of items displaying DIF. Accounting for partial invariance by allowing the item to vary across cultures can be done easily in both cases. Sometimes, however, an item may fit on one culture but misfit in another. In a MG-CFA study such an item is typically discarded even though Baumgartner and Steenkamp, 1998 show a way to overcome this problem. The Rasch model offers a different resort. Since, the input to the Rasch model is the raw data rather than aggregated statistics like covariance matrices, missing data represents no substantial problem. While missing data do increase standard errors of parameter estimates, the estimation is not affected, in principle. That is why an item parameter may easily be estimated for one group while the same item is discarded for other groups. Consequently, the inclusion of culture-specific items (emics) is easier with the Rasch model, both from a conceptual and operational point of view.

3.4.2 *Empirical comparison.* Only the application to empirical data can shed light on how the theoretical differences bear on the conclusions drawn from analyses of data

equivalence. Ewing *et al.* (2005) demonstrate how the Rasch model can be applied to a multi-group set of real data. However, no comparative analysis with the traditional CFA approach has been undertaken. Nevertheless, the study shows that Rasch analysis is powerful in analysing typical marketing research data.

Salzberger *et al.* (1999) compared Rasch analysis and the MG-CFA approach using a simulated data set. The study, which is mainly a conceptual contribution, illustrates how the different models work in a situation where one subset of items is affected by the same additive bias in one group while another subset of items is invariant across groups. Since, bias is always relative, either subset of items may be considered to be invariant with the other subset being biased. In this case, the MG-CFA approach is very unlikely to recover this fact while the Rasch analysis typically reveals this fact. The reason is that the MG-CFA approach relies on one particular item for definition of the latent scale while the Rasch model defines the scale origin as the mean of all item locations. An important conclusion from this study is the fact that the statistical analysis of data equivalence should always be complemented by substantial content-related considerations.

A very comprehensive comparative study is provided by Meade and Lautenschlager (2004). The authors compare the MG-CFA method with an IRT model that allows for different item discrimination, i.e. a non-Rasch model. Based on a series of simulated data sets varying sample sizes and the type of DIF (location parameter DIF and item slope DIF, respectively), Meade and Lautenschlager (2004) conclude that the IRT approach is “somewhat better suited for detecting differences when they were known to exist”. They also emphasize that “CFA methods . . . were inadequate at detecting items with differences in only b parameters.” and “CFA methods were also largely inadequate at detecting differences in item a parameters.” This seems to be a strong case in favour of the IRT approach. However, as the authors concede in their discussion of limitations, the way the data is simulated is crucial. Meade and Lautenschlager (2004) used IRT-based software (Baker, 1994). This implies that the responses are reflecting a non-linear relationship of the latent variable and the manifest responses. Consequently, a basic assumption of the CFA approach, namely that there is a linear relationship, is violated suggesting the CFA approach is inappropriate in the first place. Still, the conclusions of Meade and Lautenschlager (2004) should be considered valid. The reason is that real data, if they are suitable for measurement, should be non-linear because responses are bounded between a limited number of response categories and item locations (or item intercepts) are varying between items. Thus, Meade and Lautenschlager (2004) illustrate how the CFA model behaves with reasonable data.

The issue of data generation also indicates that there actually can be no “theory-free” comparison of approaches which differ substantially in their foundation. One can either draw conclusions purely on a theoretical basis or one can refer to data. If the data are real data, one does not know which items, if any, are affected by DIF. Then, different results cannot be interpreted without reference to theoretical considerations. If the data are simulated, one has to decide according to which model the data are generated. So, the theoretical input creeps into the study at this stage. However, data are generated, one has to argue for the model chosen. Thereby, one postulates how data should look like and how they should be structured. However, this claim is tantamount to favouring one model over the other, what in turn is a theoretical decision.

4. The empirical study

4.1 Purpose of the comparison

The present study refers to a comparison of a Rasch based approach and the MG-CFA approach to test for data equivalence in a set of real data. Naturally, practical marketing research is concerned with real data. We never know if data are equivalent or where DIF occurs. We even do not know a priori that measurement has been achieved at all. Data may contain so much error that we have to reject the notion of quantification. Although rarely considered possible, a variable may not even exist in quantity, then measurement ceases to be meaningful altogether.

Simulation studies like those conducted by Meade and Lautenschlager (2004) or Salzberger *et al.* (1999) as well as applications of Rasch models to cross-cultural data (Ewing *et al.*, 2005) and intra-cultural data demonstrate that Rasch and IRT methods do represent an interesting option of conceptualising measurement. Based on theoretical considerations, we frame the following expectations. We refrain from calling them hypotheses simply because there is no unambiguous theoretical argument but often pros and cons.

Expectation E1.

In general, we expect more items to fit the CFA model compared to the Rasch model.

It is often argued that the Rasch model is more demanding than the traditional test theory. In particular, items are required to be equally discriminating under the Rasch model. This should lead to more misfits when analysing data using the Rasch model. In particular, items deviating strongly from the mean discrimination (factor loading), should misfit the Rasch model. However, if the items vary substantially in their location, then floor and ceiling effects can lead to reduced item-covariances. Then the items may even fit better under the Rasch model. Since, the items have not been generated with an eye on maximising item locations, we do not expect this effect to play a substantial role, though.

Expectation E2.

In the analysis of data equivalence, we expect a similarity of items lacking scalar invariance in the MG-CFA approach and displaying DIF in the Rasch approach.

Item intercepts and the item locations are related parameters. If items are non-invariant, item intercepts should differ between groups and items should display DIF. In general, the item intercept parameters and the item locations should be inversely correlated. The harder an item is agreed with, the larger the item location and the smaller the item intercept because the manifest score is smaller compared to an easier item given the same latent score.

Expectation E3.

For items lacking fit across groups, we expect some to fit in at least one group but not in another. Therefore, under the Rasch model, the number of items in the final scale should be larger, all other things being equal.

This expectation differs from the two before. In the Rasch model it is very easy to retain an item for one group but discard it for another. The question is whether there are such items in the scale, which is a substantial problem rather than a methodological

one. Nevertheless, the inclusion of items unique to one group may balance the effect of expectation E1.

4.2 *Conceptual foundation of technophobia*

The Anglo-American literature offers a multitude of conceptual foundations for technophobia, particularly pertaining to synonyms such as techno stress (Brod, 1984), cyberphobia (Price and Ridgeway, 1983), computer aversion (Meier, 1985) or computer anxiety (Raub, 1982). The findings however, are largely restricted to “computer”-phobia, as computers were used as anchor products. Scholars have argued that computerphobia and technophobia relate to the same latent variable (Rosen and Weil 1990a, b). However, in view of potential generalization problems in the international context, Sinkovics *et al.*, 2002 established a generic technophobia scale that is applicable to a variety of products and services. The scale is exemplified by referring to automated teller machines.

Their instrument is deemed to represent negative psychological reactions towards technology, which can arise in various forms and intensity. Hereby the term ‘phobia’ is not used in a strict medical sense, relating to the results of the exposure to a feared situation (often demonstrated in symptoms such as sweating, tremors, flushing, etc.), but the notion of phobia implies rational (Röglin, 1994) and – what is even more – irrational psychological aspects to the anxiety (Jaufmann, 1991). The authors derive a three-dimensional factor-structure for technophobia (Sinkovics *et al.*, 2002). The first factor relates to “personal failure”, i.e. issues describing problems, frustrations, and failures when using sophisticated or innovative machines, the second factor represents issues which elicit the ambiguity between human and machine interaction, i.e. fears about machines dominating interactions. Lastly, the third factor is related to “convenience” issues, when using machines. The three factors correlate between 0.45 and 0.59.

4.3 *Data*

To illustrate the procedures for equivalence testing, a subset of data originally collected for a large multi-country survey of consumers was used (Sinkovics *et al.*, 2002). The original study established a measure for the concept of “technophobia” and comprised additional measures such as “innovativeness” (Hirschman, 1980; Price and Ridgeway, 1983). The items used five-category Likert scales. Sinkovics *et al.*, 2002 developed the “technophobia” scale and found reasonably well reliability scores and indications for validity, following exploratory and confirmatory multi-group structural equations modelling procedures. In terms of controlling for equivalence (Craig and Douglas, 2005), different sampling frames were employed. In Great Britain, the sample was drawn from four metropolitan areas, in Mexico, a student sample was taken and in Austria a quota sample was drawn which was representative for the adult Austrian population. Quota descriptors included age, gender and occupation.

4.4 *Descriptive results*

Our analysis builds on data (total $n = 927$) from the UK ($n = 278$), Mexico ($n = 200$) and Austria ($n = 449$) in a deliberate attempt to maximize some of the outcomes of the methodological comparison. Sample sizes were equally high in all three countries, while at the same time we were dealing with fundamentally different cultural

environments. English, Spanish, and German languages were represented in the data and different perspectives of the technophobia phenomenon were expected.

Given the sampling procedures described above, there existed a slight bias in terms of younger age group ($F = 8.577$, $df = 2$, $p < 0.001$). Another potential bias was introduced by the fact that data collection was mainly administered in urban regions. However, these effects were not considered to threaten the methodological direction of our research, quite contrastingly, since these were structural biases of the samples, consistent across countries, the comparison between CFA and Rasch was deemed to be even more interesting.

While card ownership is generally wide-spread in the countries under scrutiny, it is significantly higher in the UK-sample (93.1 percent have ATM cards) than in the Austrian and Mexican sample with 75.9 and 73.0 percent, respectively. The same pattern prevails for the frequency of card usage. The differences are considered to be the result of unequal stages of technological development and product diffusion patterns for automated teller machines.

4.5 Results of the MG-CFA analysis

The MG-CFA was conducted following the scheme recommended by Steenkamp and Baumgartner (1998) using Lisrel 8.54 (Jöreskog and Dag, 2003). The starting point was a multi-group model without any constraints across groups other than the mere factorial structure, i.e. a unidimensional model. We attempted to keep the scale unidimensional for the sake of parsimony. However, if the data had indicated that this notion is not viable, we would have switched to a multidimensional model.

In assessing fit, we mainly focussed on the RMSEA and the χ^2 -statistic. The first model included all 30 items. The fit of the model was highly unsatisfactory. Subsequently, the modification indices were used to identify items the error terms of which show significant covariation. In most cases these items were similar in content. Consequently, one of them was deleted. This procedure was carried out iteratively until a model was derived that fitted well. Finally, a set of six items turned out to fit in all three groups under scrutiny. This model established configural invariance. It acts as the baseline model against the more stringent model of metric invariance is tested.

The next step consisted in imposing equality constraints on the loading parameters across groups. The resulting model formalises full metric equivalence across groups ensuring the same unit of measurement prevails in all three countries. The difference in the χ^2 -statistics of the configural and the metric invariance model can be evaluated by the χ^2 -difference test with degrees of freedom equal to the difference in degrees of freedom of both models. The decrease in fit of the metric invariance model turned out to be non-significant implying full metric invariance to hold true in the three samples considered. All tests were evaluated on the 1 percent-level for type one error.

Subsequently, the item intercept estimates were constrained to be equal across groups in the full scalar invariance model. These further constraints resulted in a significant decrease in model fit compared to the metric invariance model indicating that some items are affected by non-invariance. Again, the modification indices of the intercept parameters showed us where we had to lift constraints. For one item a unique item intercept had to be estimated in each group while another item required a separate estimate in the Austrian sample but worked well with a common estimate for the UK and Mexico.

In summary, the MG-CFA approach established a relatively high degree of invariance given the diverse nature of the countries. Metric invariance prevails fully four items are scalar invariant. It should be noted, though, that 24 items were discarded and only six items retained. A set of six items should be sufficiently precise for most applications and it is certainly highly economical concerning the expenses during data collection. However, the question arises whether six items out of 30 adequately represent the construct. When looking at the content of those items that were retained and of those that were discarded, it is striking that the retained items are primarily about the ease of using ATMs (worry about making mistakes, easy to learn, time-consuming, etc.) while only one item (I do not trust ATMs with my money.) is emotionally coloured.

Based on the model of partial scalar invariance, further constraints can be imposed, which are not necessary for data equivalence in terms of mean comparability but provide interesting information. First of all, a latent mean test can be conducted by requiring the latent group means of all three samples to be equal. While the latent means differed significantly between all three groups, a subsequent test of equal latent means of the UK and the Mexican sample was insignificant suggesting that technophobia is significantly smaller in Austria versus the UK and Mexico with no significant difference between the latter countries.

Invariance was further extended by forcing item loadings to be equal across items in addition to equal loadings across groups. This is tantamount to equal item discrimination across items. A model of partial scalar equivalence with equal item discrimination for three items could be established. Alternatively, error variances were selected to be equal across groups. Some equality constraints of this type had to be lifted but finally a partial scalar invariance model with equal error variances fitted the data well. These further tests of invariance demonstrate that a high degree of equivalence prevails in the data clearly exceeding the absolutely necessary level of comparability (Table I).

4.6 Results of the Rasch analysis

The Rasch analysis is based on the partial credit model (Andrich, 1988; Masters, 1982), i.e. each item is assumed to have its own set of distances between category thresholds. The additional advantage of the Rasch approach is that the threshold ordering of these items can be empirically examined. Although constructed to imply an increasing level of the latent dimension, in practice the response categories may not work properly and the threshold estimates may take any order suggesting violations of the hypothesis that the response data are ordinal. In the present study threshold disordering did occur. A *post hoc* remedy to cope with this issue is collapsing adjacent categories. For this reason, in the current data set, all items were rescored by retaining the first category (fully disagree), collapsing the second category (disagree) and the middle category, and by collapsing the remaining categories of agree and fully agree.

The assessment of data equivalence using the Rasch model features some similarities but also exhibits some differences. Since, all items are required to discriminate equally, metric equivalence can only be achieved fully. Scalar invariance, however, can be partial insofar as an item may display differential item functioning (DIF) requiring a split of the item between the groups. An important difference lies in the fact that the Rasch analysis starts with a full invariance model. Consequently,

Table I.
Assessment of data
equivalence with the
MG-CFA approach

Type of invariance	Model/constraints	RMSEA	χ^2 (df)	<i>p</i> value	$\Delta\chi^2$ (df)	<i>p</i> value
Starting model		0.115	4057.30 (1131)	<0.0001	–	–
Deletion of items due to significant covariation of error terms	No constraints across groups					
Configural invariance	No constraints across groups	0.046	43.11 (27)	0.025	–	–
Full metric invariance	Constraints on loadings (λ_g)	0.043	58.22 (37)	0.015	15.11 (10)	0.128
Full scalar invariance	Constraints on item intercepts (τ_{ig})	0.093	154.16 (47)	<0.0001	95.94 (10)	<0.0001
Setting free item intercept parameters						
Partial scalar invariance	Constraints on some item intercepts (τ_{ig})	0.046	71.35 (44)	0.006	13.13 (7)	0.069
Latent mean test	Constraints on all latent means	0.068	107.87 (46)	<0.0001	36.52 (2)	<0.0001
Latent mean test Austria versus (the UK and Mexico)	Constraints on latent means in the UK and Mexico	0.046	73.25 (45)	0.005	1.9 (1)	0.168
Full metric invariance plus equal discrimination of three items	Additional constraints on loadings across three items (λ_i)	0.041	58.88 (39)	0.021	0.66 (2)	0.719
Partial scalar invariance plus equal discrimination of three items	Additional constraints on loadings across three items (λ_i)	0.041	72.19 (46)	0.008	0.84 (2)	0.657
Partial scalar invariance plus some equal error variances	Additional constraints on error variances (for two items for all three countries; for two items for Austria and the UK; for two items for Austria and Mexico)	0.041	89.09 (52)	0.001	17.73 (8)	0.0233

the item parameters are estimated from pooled data. The overall fit of the model can be assessed by a χ^2 -test that compares, for each item, the expected scores based on probabilities with actual scores based on proportions in several score groups, added up over all items (Andrich *et al.*, 2003a). The test can also be used for the evaluation of an individual item. Basically, misfit of an item can have two reasons. Either the item misfits in at least one of the groups or the item exhibits DIF. Therefore, at each step, item fit statistics have to be examined carefully as well as the test for DIF. Besides, the person separation index was monitored. This statistic is similar to classical reliability. It is bounded between zero and one. High values mean that the items discriminate between the persons. Small values (below 0.85) are problematic because the test of fit loses power. All analyses were conducted using RUMM 2020 (Andrich *et al.*, 2003b).

The analysis was carried out iteratively. At each step, only one change to the model was undertaken in order to disentangle the effects of misfit of different items. Either an item was split up because of DIF or an item was discarded (for all countries or, after a split up, only for one country). The final model comprised a total of 13 items. However, only four of these were invariant across all three groups. Four other items were invariant across two groups with one case where the item misfitted in the remaining group. Two items had to be split up for each group due to DIF. Two items fitted only in two groups and exhibited DIF, one item fitted only in Austria. In summary, 12 items are available for Austrian respondents, whereas the scale comprises 11 items for respondents from the UK and Mexico, respectively.

Since, the test of DIF based on the analysis of variance may suffer from the unequal sample sizes, the results were scrutinized carefully. Firstly, unequal sample sizes imply different power of detecting DIF involving a particular country. Secondly, the sum of squares cannot be allocated unequivocally to the main effects and the interaction term when the design is not orthogonal. To address the first issue, we screened the DIF findings. In principle, the power of the test of DIF is expected to be higher when Austria is involved because it has the largest sample size. However, this does not result in more items showing DIF for Austria in our example. There are three items that are interesting in this respect because they fit in all three countries. In one case, there is a common estimate for the UK and Mexico (in line with higher power for differences against Austria), in another case there is a common estimate for Austria and Mexico, and finally, in the third case there is a common estimate for Austria and the UK. So, it seems that the issue of unequal power does not bear on the outcome in the Rasch analysis. In the traditional CFA analysis, in one case, there is a separate estimate for Austria and a common one for Mexico and the UK.

To examine the second potential threat to the DIF analyses, a random subsample of the data set was drawn with 194 respondents from Mexico and 200 from Austria and the UK, respectively. Since, the DIF diagnosis did not depart in any way from the findings based on the complete data set, we consider the findings tenable. Apart from the DIF analyses, the sample size also has an impact on fit statistics. When assessing the overall fit in Rasch models, the sample size and the person separation index have to be considered particularly. In the present data set, 817 respondents are available for fit assessment. The remaining respondents display extreme scores and, therefore, do not qualify for fit analyses. In the context of Rasch models, 817 represent a large number. Moreover, the person separation index amounts to 0.89 implying excellent power of the test of fit. This leads to a very sensitive χ^2 -statistic. We, therefore, focussed on

the incremental improvement of fit when deleting or modifying items at each step. Another indication of fit was the fit statistics for each individual item. Even the worst fitting item exceeded the 1 percent level. Furthermore, RUMM provides the option of calculating adjusted χ^2 -statistics based on a different sample size. Boomsma and Hoogland (2001) mention $n = 200$ as the minimum sample size when discussing robustness of structural equation modeling against small sample sizes. If we consider $n = 200$ as the minimum sample per country, which is also a reasonable minimum sample size in a Rasch analysis, we get a total sample size of 600. With 600 respondents the χ^2 -statistic amounts to 113.07 ($df = 88, p = 0.037$). However, greater confidence can be derived by taking random samples of 600 respondents and repeatedly calculating fit statistics like in a bootstrapping approach. With a mean χ^2 of 126.35 ($df = 88, p = 0.005$) from 30 runs the model fit appears to be marginal. Analysing each country separately, the data fit the model satisfactorily (Austria $\chi^2 = 72.63, df = 48, p = 0.012$; UK $\chi^2 = 66.90, df = 44, p = 0.015$; Mexico $\chi^2 = 53.90, df = 44, p = 0.145$). For these reasons we deem the final model tenable (Table II).

Figure 1 shows the main results of the Rasch analysis. It depicts the person locations (displayed on top) against the item threshold locations (displayed downwards), which are placed on the same dimension. The difference between Austria and the UK and Mexico, respectively, is about one logit unit and statistically significant. The bell-shaped curve shows the amount of information the instrument provides for respondents depending on the level of technophobia. The more information we have about a person, the smaller is the standard error of the person location tantamount to higher measurement precision. The distribution of the respondents is shifted to the left, implying that for many persons the items are relatively hard to agree with. Consequently, the scale is more sensitive in the area of moderate to severe technophobia. However, that is exactly what the scale is supposed to be. We do not want to differentiate between respondents of negligible degrees of technophobia.

4.7 Comparison of the results

The comparison of the results gained from either approach shows some striking parallels (see Table III). Each approach suggests four items, which are fully equivalent. Three of these are reported invariant (i.e. v07, v27, and v29) in both analyses. Another item (v18) displays a relatively small but significant additive bias in the Austrian sample against the other two samples whereas the Rasch analysis could not find any indication of DIF (the ANOVA DIF test is insignificant with $p = 0.32$). Item 14 is particularly interesting since it turns out to be fully invariant in the MG-CFA approach. In contrast, in the Rasch model it fails to meet a reasonable level of fit and also shows strong indication of DIF. The item is special insofar as it is one of the few reverse coded items. Such items are sometimes problematic since they are prone to confuse respondents. In many cases people do not simply respond in a “reverse” way. Instead, they favour extreme categories and make less use of the finer distinctions in between. In the Rasch model this may lead – quite reasonably – to misfit while in the CFA approach the more “pronounced” responses may even enhance the fit of the item. This may have occurred in the present study. Another point is whether the perceived easiness of learning how to use an ATM is actually a good indicator of technophobia.

Type of invariance	Model	Fit statistics/ANOVA statistics		Person separation index
		χ^2 (df)	<i>p</i> value	
Full scalar invariance	Common parameter estimation across all groups	2,107.93 (270)	<0.000001	0.93
Deleting items or splitting up items	Some items split up because of DIF	153.96 (88)	0.000018	0.89
Partial scalar invariance	Analysis of variance of latent means	$F = 18.73$ (6,912)	<0.000001	
Latent mean test	Analysis of variance of latent means, <i>post hoc</i> test	Scheffé test: Austria – 1.584, the UK – 0.505, Mexico – 0.381	Austria vs the UK $p < 0.001$ Austria vs Mexico $p < 0.001$ the UK vs Mexico $p < 0.717$	

Table II. Assessment of data equivalence with the Rasch partial credit model

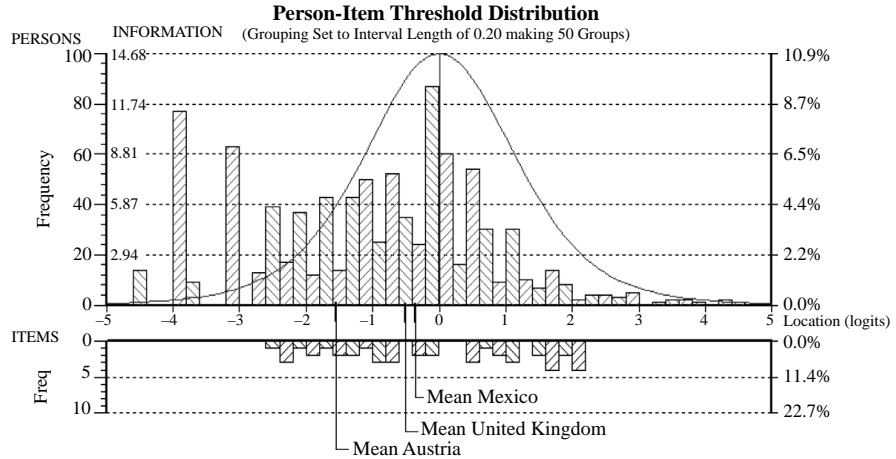


Figure 1.
Rasch Person measures
and item threshold
parameters

Code	Item wording	MG-CFA (item intercepts) ^a			Rasch analysis (item locations) ^b		
		Austria	The UK	Mexico	Austria	The UK	Mexico
v07	Using ATMs is time-consuming	1.93	1.93	1.93	0.393	0.393	0.393
v18	I dont trust ATMs with my money	2.13	2.28	2.28	-0.215	-0.215	-0.215
v27	I find ATMs instructions confusing	2.13	2.13	2.13	0.01	0.01	0.01
v29	I feel confident that I could teach someone how to use an ATM (reverse coding)	3.74	3.74	3.74	-0.267	-0.267	-0.267
v21	I wish I were more adept at using ATMs				0.289	-0.437	
v26	Machines should not handle people's money transactions					-0.203	-0.203
v13	I refuse to use ATMs				-0.196		0.425
v01	I feel some anxiety when I approach an ATM				0.736	-0.091	-0.411
v06	I worry about making mistakes when using ATMs	2.00	2.53	2.91	-0.067	-0.864	-1.273
v17	Thinking about ATMs makes me nervous				1.347		
v09	It takes me a long time to complete bank transactions when using an ATM				0.684	0.079	0.079
v08	ATMs agitate me				0.528	-0.111	0.528
v23	ATMs seem very complicated				0.224	0.224	-0.582
v14	It is easy to learn how to use ATMs (reverse coding)	3.93	3.93	3.93			
	Reliability (Cronbach's α)	0.76	0.83	0.72	0.875	0.884	0.879

Notes: ^aThe parameter estimates from the partial scalar invariance model are shown here without any further equality constraints on loadings across items or error variances; ^bIn the table cells the Rasch overall item locations are stated. These locations are the mean of the two thresholds that are estimated for each item. The mean of all item thresholds over all items is zero by definition

Table III.
Results of assessing data
equivalence

In fact, people might be intelligent and clever and still reject the idea of “machines handling money transactions”.

By both approaches, item v06 is reported to be an indicator of technophobia in all three countries requiring additive correction for DIF. According to the MG-CFA and the Rasch approach, the item is easiest to agree with in Mexico, followed by the UK and finally Austria, representing another parallel between the approaches.

Interestingly, the Rasch methodology reveals eight other items that are at least in one of the countries valid indicators. Four of these items are equivalent, i.e. free of DIF, for two groups providing a stronger link between the countries and enhancing comparability.

Looking at the item intercept estimates in the MG-CFA and the item locations from the Rasch analysis, the same rank order appears for the almost completely invariant items retained under both models. Item v29 is the easiest item, followed by v18, v27, and v07. The metric correlation of all intercept estimates and item locations from common results amounts to -0.47 ($p < 0.01$).

Regarding the comparison of the means of all respondents between the countries, the two approaches come to the same conclusion. Respondents from Austria are less technophobic than people from the UK or Mexico. On the individual level, the correlation of the Rasch person measures and the latent variable estimates from the CFA is 0.85 across all countries. Within the three different groups the correlation coefficient are even higher with $r = 0.89$ in case of Austria and Mexico, and $r = 0.87$ for respondents from the UK. This seems to suggest that Rasch measures can be used interchangeably with CFA derived scores. However, such a conclusion would be premature for at least three reasons. Firstly, Ewing *et al.* (2005) pointed out that the traditional method based on CFA might be an approximation if, and only if the data fit a Rasch model. The Rasch model provides evidence about the data justifying measurement that lies outside the potential of CFA. Secondly, even a correlation as high as 0.89 means that almost 21 percent of the variance is not shared. This issue will be addressed below when we investigate further evidence of validity based on correlations. Thirdly, an important difference between the approaches considered is the non-linearity of the relationship between the measure and the raw score in the Rasch model and the linear relationship in case of CFA. Since, the raw score is transformed in a non-linear way in the Rasch model but in a linear way in CFA, a non-linear relationship prevails between the Rasch measure and the CFA based score. In particular, the non-linearity becomes the most prominent at the extremes. Incidentally, this implies that it is theoretically impossible that both approaches simultaneously yield linear, interval scaled measures. In the present study, only few respondents possess a high degree of technophobia. Consequently, the non-linearity at the upper end plays no substantial role. At the lower end the non-linearity is clearly visible (see Figure 2 for the relationship in the Austrian sample, the figures for the UK and Mexico look very similar). However, the majority of the respondents are located in the central area where the relationship is virtually linear.

4.8 Testing the expectations

In the theoretical section, we raised three expectations (E1-E3) regarding potential differences in the outcome of the analyses. E1 claimed that more items would fit

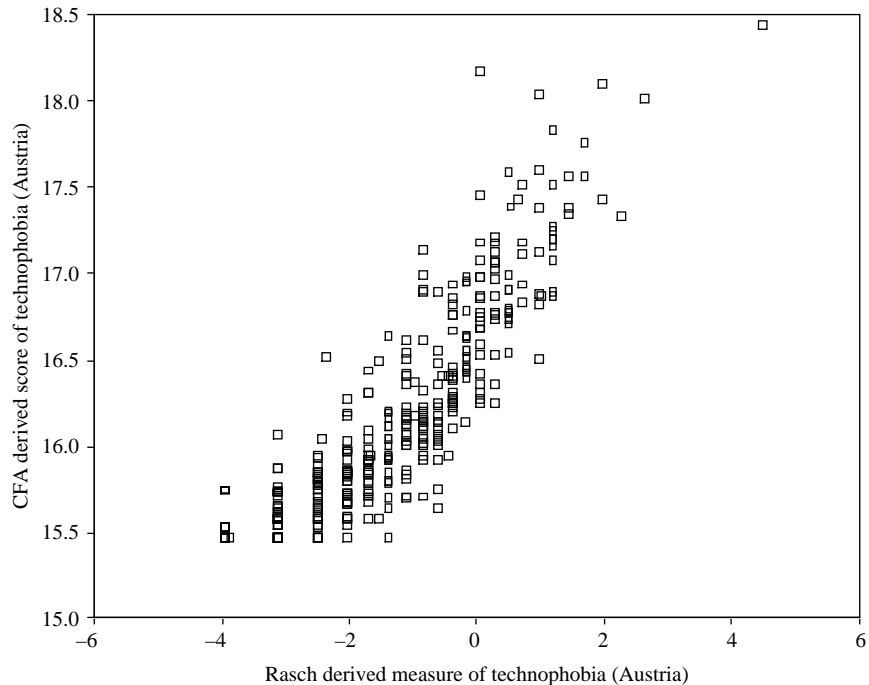


Figure 2.
Relationship of Rasch
derived measures and
CFA based scores of
technophobia

the CFA model compared to the Rasch model. This expectation has to be rejected. In fact, more items, if only partially (see also E3), are retained in the Rasch analysis. While the traditional solution based on CFA has to get by with six items, the Rasch derived scales comprise 11 (the UK and Mexico) and 12 items (Austria), respectively. In turn, the Rasch model rejects one rather doubtful item, which features even scalar equivalence in the MG-CFA. In total, the Rasch model seems to provide a more justified and precise person measure compared to the CFA even though it does not make a difference on the aggregate level of group means. In terms of item content, the set of items retained in the Rasch model covers the emotional component (anxiety, agitation, nervousness, intimidation) much more comprehensively than the CFA based scale. The reason why the Rasch model fits even better and why, all in all, both approaches display significant parallels are twofold. Firstly, the items show a high degree of equivalence and allow for additional constraints on the loadings across items as well as on the error variances. This is not so much a cross-cultural issue but a general feature of the data that explains why the data fit the Rasch model reasonably well. Secondly, the items have been generated following the classical paradigm that favours high inter-item-correlations. In particular, items were not generated aiming at substantial differences in item intercepts (this is referred to as 'bandwidth' issue by Singh, 2004). In contrast, the Rasch (and also the IRT) philosophy requires sufficient variation in item locations for establishing construct validity (Ewing *et al.*, 2005; Singh, 2004) and

providing enough information across a wide range of the scale. The technophobia scale is certainly at the lower end of the acceptable range of item locations. This also favours the CFA (see Salzberger *et al.*, 1999 for reasons why differences in item intercepts may also have an impact on loadings).

The second expectation E2 stated that there would be a similarity of items lacking scalar invariance in CFA and displaying DIF in Rasch analysis. The results support E2 in principle. Firstly, the reverse is true, i.e. those items that meet scalar invariance requirements do not exhibit DIF. Secondly, one item (v06) shows the same pattern of relative biases in both analyses.

The third expectation, according to which some items lacking fit should fit the Rasch model in some of the groups. E3 is strongly supported. Eight items are part of the Rasch derived instrument but are discarded during the MG-CFA.

4.9 Further evidence of validity

Notwithstanding the somewhat different understanding of construct validity depending on the theoretical foundation of measurement, the CFA as well as the Rasch analysis support the construct validity of the technophobia scale as well as its cross-cultural validity as far as the three countries under scrutiny are concerned. Criterion-based validity provides further evidence of validity. Consequently, the technophobia scores from CFA and from the Rasch analysis were correlated with the “use innovativeness” measure, as established by Price and Ridgeway (Hirschman, 1980; Price and Ridgeway, 1983). The construct deals with the use of previously adopted products in novel ways and encompasses five factors: creativity/curiosity, risk preferences, voluntary simplicity, creative reuse, and multiple use potential. The first four were used in the present study. An index across these dimensions was also calculated and named use innovativeness. The scales were analysed exclusively based on CFA.

Considering the technophobia measure based on MG-CFA within the country groups, nine correlations are significant (Table IV). There are some noteworthy differences between the countries. For example, risk preferences are not significantly correlated with technophobia in the UK but correlate -0.40 in Mexico.

Since, the Rasch derived measure and the CFA based score are highly correlated, it is not surprising that very similar patterns of correlations occur if the Rasch measure (Table V) is used instead of the CFA score. Looking at the nine correlations that are significant in both cases, seven are higher with the Rasch score, one is equally high and only one correlation is lower when the Rasch measure is used. In turn, three correlations are non-significant in both cases, one is lower in case of Rasch, two are lower when the CFA is employed. Finally, one correlation is non-significant with CFA but significant with Rasch. In summary, eight correlations are closer to 1 (significant correlations) or closer to 0 (non-significant correlations) with Rasch. Only three correlations are closer to 1 (significant correlations) or closer to 0 (non-significant correlations) with CFA. The findings cannot be generalised but it seems that the Rasch measures have got higher precision and accuracy sharpening the correlations accordingly.

In any case, the findings confirm the assumption that higher levels of innovativeness go with lesser levels of technophobia.

Table IV.
Correlations “use
innovativeness” and
“technophobia” based on
MG-CFA

Pearson correlation with technophobia score from MG-CFA analysis	Creativity, curiosity		Risk preferences		Voluntary simplicity		Creative reuse		Use innovativeness	
	$r(n)$	\hat{p} value	$r(n)$	\hat{p} value	$r(n)$	\hat{p} value	$r(n)$	\hat{p} value	$r(n)$	\hat{p} value
The United Kingdom	0.04 (262)	0.58	-0.10 (263)	0.10	-0.29** (266)	<0.01	-0.25** (265)	<0.01	-0.19** (255)	<0.01
Mexico	0.05 (168)	0.52	-0.40** (175)	<0.01	-0.28** (176)	<0.01	-0.24** (177)	<0.01	-0.19* (166)	0.02
Austria	0.08 (438)	0.12	-0.12** (438)	0.01	-0.07 (438)	0.12	-0.10* (438)	0.05	-0.04 (438)	0.40
Pooled data	0.08* (868)	0.02	-0.09** (876)	0.01	-0.25** (880)	<0.01	-0.18** (880)	<0.01	-0.10** (859)	<0.01

Notes: *, **Correlation is significant at the 0.05, 0.01 level (two-tailed), respectively

Pearson correlation with technophobia score from MG-CFA analysis	Creativity, curiosity		Risk preferences		Voluntary simplicity		Creative reuse		Use innovativeness	
	$r(n)$	p value	$r(n)$	p value	$r(n)$	p value	$r(n)$	p value	$r(n)$	p value
The United Kingdom	-0.01 (269)	0.92	-0.09 (270)	0.14	-0.37*** (273)	<0.01	-0.28*** (272)	<0.01	-0.26*** (262)	<0.01
Mexico	-0.06 (179)	0.40	-0.42*** (189)	<0.01	-0.29*** (190)	<0.01	-0.24*** (192)	<0.01	-0.26*** (166)	<0.01
Austria	0.12* (449)	0.02	-0.13*** (449)	0.01	-0.08 (449)	0.10	-0.10* (449)	0.05	-0.02 (449)	0.72
Pooled data	0.07* (897)	0.03	-0.17*** (908)	<0.01	-0.24*** (912)	<0.01	-0.19*** (913)	<0.01	-0.13*** (887)	<0.01

Notes: * , ** Correlation is significant at the 0.05, 0.01 level (two-tailed), respectively

Table V. Correlations “Use innovativeness” and “technophobia” based on Rasch analysis

5. Summary and implications

The Rasch model and the measurement theory the model is based on, is a promising alternative to the standard approach to measurement rooted in the classical measurement theory. As Ewing *et al.* (2005) have pointed out, RMT can be seen as a more comprehensive framework compared to the classical factor analytic approach. The Rasch model justifies the computation of the raw score that is calculated in either approach in the same way (except for different weighting of the item scores). While the classical theory regards the scores as linear measures, the Rasch model acknowledges the non-linearity and transforms the raw scores into a linear, interval-scaled measure by a logistic function. From this it follows that the Rasch measures and the CFA measures cannot concurrently be linear.

Where the data do not fit the Rasch model, the theoretical foundation of measurement, as it is naturally and routinely asked for in modern physical science, is missing. Resorting to CTT or to more complicated IRT models may or may not lead to a solution. However, we should realise that this sort of measurement is not compatible with the notion of quantification held in the physical sciences. One might argue measurement in the social sciences is harder to achieve. Indeed, this seems to be the case but does it really exempt us from being rigorous and allow us to be more speculative? In fact, one would hardly object to the proposition that numerals without quantitative meaning do not qualify for any sort of statistical computations. Hence, science commits us to provide empirical evidence that numerals reflect quantity.

Where the data do fit the Rasch model, we can be confident that the underlying variable is quantitative and that measurement has been achieved. Of course, this applies within the limits of unavoidable statistical error affecting any sort of empirical hypothesis testing. However, even in this case, the level of manifest item scores is only ordinal. Consequently, the raw data do not meet the assumptions for factor analytic procedures, at least not for those based on covariances or Pearson correlations. On the other hand, the vast majority of empirical research in marketing is based on CTT. It would be overdone to claim that all these findings are invalid. Despite the fundamental philosophical differences between CTT and Rasch measurement, CTT can be seen as an approximation to Rasch measurement – provided the data fit the Rasch model. This implies that many findings would remain valid. What we would “lose” – or rather identify as being fallacious – are those findings where our asserted measures do not reflect quantity. Consequently, when applying the Rasch model there is nothing to lose but much to be won.

In more practical terms, the Rasch model offers further interesting advantages. It can be applied to any number of response categories with all possible combinations within one instrument. The manifest responses are assumed to be ordinal and need not be interval-scaled. The distribution of the respondents can take any shape without endangering parameter estimation, in principle. A fundamental difference between the paradigms of RMT and CTT is the necessity of variation in the distribution of items, i.e. the bandwidth of the scale. In the classical paradigm, typically no attention is paid to the issue of bandwidth, whereas the Rasch model asks for markedly different item locations. For many constructs it may be challenging to generate items that potentially do provide such variation.

In our comparative study of a scale measuring technophobia was investigated in three countries (the UK, Austria, and Mexico). The results of the MG-CFA and the Rasch approach were very similar to a large extent. Contrary to what we expected, the Rasch model retained more items than the classical approach. Still, we claim that, in general, the CFA would keep more items – although wrongly from a Rasch point of view. The reason is the higher flexibility of CFA in terms of item discrimination. In our case, the data meet stringent levels of equivalence. Metric invariance holds for all items in the final scale and half of the loadings can even be constrained across items. Thus, it is not surprising that the data fit the Rasch model reasonably well. On the other hand, the Rasch model eliminated one item that seems to be doubtful but nevertheless reaches even scalar invariance in the MG-CFA. The ease with which the Rasch model handles missing data allows us to split items affected by misfit in one of the groups and discard the responses in that group only. Consequently, more items, specific to only one or two groups can be retained enhancing the precision of the person measures. More pronounced correlations with the external construct of innovativeness support this conclusion.

The use of real data may be seen as an inherent limitation of the present study but also as an advantage because only real data allow for investigating the behaviour of the models in a real world situation of true responses rather than generated responses following a particular model. We claim that a conclusive evaluation of the methods discussed requires reference to theory. If one still wants to draw conclusions purely from empirical applications, one study is certainly insufficient. A meta-study would help identify persistent patterns. Unfortunately, up to now, applications of the Rasch model in single culture studies are scarce, let alone in multi-cultural settings with a full comparison of the MG-CFA and the Rasch approach.

In terms of validating scales based on the Rasch approach, only the adoption of the Rasch philosophy by the researcher, in particular during item generation, may help fully exploit the potential of the Rasch model. Significant variation of the item location would, in all likelihood, cause more prominent differences in scale validation between CFA and the Rasch model. The justification of linear, interval-scaled measures of latent constructs in marketing would certainly benefit from a more widespread application of Rasch analyses, in particular in cross-cultural studies but also whenever the measures of individuals are of interest.

References

- Andrich, D. (1978), "A rating formulation for ordered response categories", *Psychometrika*, Vol. 43 No. 4, pp. 561-73.
- Andrich, D. (1988), "A general form of Rasch's extended logistic model for partial credit scoring", *Applied Measurement in Education*, Vol. 1 No. 4, pp. 363-78.
- Andrich, D., Sheridan, B.S. and Luo, G. (2003a), "Displaying the Rumm2020 analysis", working paper, RUMM Laboratory, Perth.
- Andrich, D., Sheridan, B.S. and Luo, G. (2003b), *Rumm2020: Rasch unidimensional Measurement Models*, RUMM Laboratory, Perth.
- Baker, F.B. (1994), *Genir V: Computer Program for Generating Item Response Theory Data*, University of Wisconsin, Laboratory of Experimental Design, Wisconsin, MA.

-
- Baumgartner, H. and Steenkamp, J-B.E.M. (1998), "Multi-group latent variable models for varying numbers of items and factors with cross-national and longitudinal applications", *Marketing Letters*, Vol. 9 No. 1, pp. 21-35.
- Boomsma, A. and Hoogland, J.J. (2001), "The robustness of Lisrel modeling revisited", in Cudeck, R., Dau Toit, S. and Sörbom, D. (Eds), *Structural Equation Modeling: Present and Future, a Festschrift in Honor of Karl Jöreskog*, SSI Scientific Software International, Lincolnwood, IL, pp. 139-68.
- Brod, C. (1984), *Technostress: The Human Cost of the Computer Revolution*, Addison-Wesley, Reading, MA.
- Cheung, G.W. and Rensvold, R.B. (1998), "Cross-cultural comparisons using non-invariant measurement items", *Applied Behavioral Science Review*, Vol. 6 No. 1, pp. 93-110.
- Churchill, G.A. (1979), "A paradigm for developing better measures of marketing constructs", *Journal of Marketing Research*, Vol. 16 No. 1, pp. 64-73.
- Craig, C.S. and Douglas, S.P. (2005), *International Marketing Research*, 3rd ed., Wiley, Chichester.
- Davis, H.L., Douglas, S.P. and Silk, A.J. (1981), "Measure unreliability: a hidden threat to cross-national marketing research?", *Journal of Marketing*, Vol. 45 No. 1, pp. 98-109.
- Diamantopoulos, A. (2005), "The C-OAR-SE procedure for scale development in marketing: a comment", *International Journal of Research in Marketing*, Vol. 22 No. 1, pp. 1-9.
- Embretson, S.E. and Reise, S.P. (2000), *Item Response Theory for Psychologists*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Ewing, M.T., Thomas, S. and Sinkovics, R.R. (2005), "An alternate approach to assessing cross-cultural measurement equivalence in advertising research", *Journal of Advertising*, Vol. 34 No. 1, pp. 17-36.
- Fan, X. (1998), "Item response theory and classical test theory: an empirical comparison of their item/person", *Educational and Psychological Measurement*, Vol. 58 No. 3, p. 357.
- Finn, A. and Ujwal, K. (2005), "How fine is C-OAR-SE? A generalizability theory perspective on Rossiter's procedure", *International Journal of Research in Marketing*, Vol. 22 No. 1, pp. 11-21.
- Hirschman, E.C. (1980), "Innovativeness novelty seeking, and consumer creativity", *Journal of Consumer Research*, Vol. 7 No. 3, pp. 283-95.
- Jaufmann, D. (1991), *Einstellungen Zum Technischen Fortschritt: Technikakzeptanz Im Nationalen und Internationalen Vergleich*, Campus Verlag, Frankfurt/Main.
- Jöreskog, K.G. and Dag, S. (2003), *Lisrel 8.54*, Scientific Software International Inc., Chicago, IL.
- Lord, F.M. and Novick, M.R. (1968), *Statistical Theories of Mental Test Scores*, The Addison-Wesley Series in Behavioral Science: Quantitative Methods, Addison-Wesley, Reading, MA.
- Masters, G.N. (1982), "A Rasch model for partial credit scoring", *Psychometrika*, Vol. 47 No. 2, pp. 149-74.
- Meade, A.W. and Lautenschlager, G.J. (2004), "A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance", *Organizational Research Methods*, Vol. 7 No. 4, pp. 361-88.
- Meier, S.T. (1985), "Computer aversion", *Computers in Human Behavior*, Vol. 1 No. 2, pp. 171-9.
- Mullen, M.R. (1995), "Diagnosing measurement equivalence in cross-national research", *Journal of International Business Studies*, Vol. 26 No. 3, pp. 573-96.

-
- Price, L.L. and Ridgeway, N.M. (1983), "Development of a scale to measure use innovativeness", in Richard, P.B. and Tybout, A.M. (Eds), *Advances in Consumer Research*, Vol. 10, Association for Consumer Research, Ann Arbor, MI, pp. 679-84.
- Rasch, G. (1961), "On general laws and the meaning of measurement in psychology", paper presented at Berkeley Symposium on Mathematical Statistics and Theory of Probability, University of California Press, Berkeley, CA, Vol. IV, pp. 321-33.
- Rasch, G. (1977), "On specific objectivity. An attempt at formalizing the request for generality and validity of scientific statements", *Danish Yearbook of Philosophy*, Vol. 14, pp. 58-94.
- Rasch, G. (1980), *Probabilistic Models for Some Intelligence and Attainment Tests*, Mesa Press, Chicago, IL, reprint of 1960, Danish Institute of Educational Research.
- Raub, A.C. (1982), "Correlates of computer anxiety in college students", dissertation, University of Pennsylvania, Philadelphia, PA.
- Reise, S.P., Widaman, K.F. and Pugh, R.H. (1993), "Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance", *Psychological Bulletin*, Vol. 114 No. 3, pp. 552-66.
- Rentz, J.O. (1987), "Generalizability theory: a comprehensive method for assessing and improving the dependability of marketing measures", *Journal of Marketing Research*, Vol. 24 No. 1, pp. 19-28.
- Röglin, H.C. (1994), *Technikängste und Wie Man Damit Umgeht*, VDI-Verlag, Düsseldorf.
- Rosen, L.D. and Weil, M.M. (1990a), "Computers classroom instruction, and the computerphobic university student", *Collegiate Microcomputer*, Vol. 8 No. 4, pp. 275-83.
- Rosen, L.D. and Weil, M.M. (1990b), "Myths and realities of computerphobia", *Anxiety Research*, No. 3, pp. 175-91.
- Rossiter, J.R. (2002), "The C-OAR-SE procedure for scale development in marketing", *International Journal of Research in Marketing*, Vol. 19 No. 4, pp. 305-35.
- Rossiter, J.R. (2005), "Reminder: a horse is a horse", *International Journal of Research in Marketing*, Vol. 22 No. 1, pp. 23-5.
- Salzberger, T., Sinkovics, R.R. and Schlegelmilch, B.B. (1999), "Data equivalence in cross-cultural research: a comparison of classical test theory and latent trait theory based approaches", *Australasian Marketing Journal*, Vol. 7 No. 2, pp. 23-38.
- Schaffer, B.S. and Riordan, C.M. (2003), "A review of cross-cultural methodologies for organizational research: a best-practices approach", *Organizational Research Methods*, Vol. 6 No. 2, pp. 169-215.
- Singh, J. (1995), "Measurement issues in cross-national research", *Journal of International Business Studies*, Vol. 26 No. 3, pp. 597-619.
- Singh, J. (2004), "Tackling measurement problems with item response theory: principles, characteristics, and assessment, with an illustrative example", *Journal of Business Research*, Vol. 57 No. 2, pp. 184-208.
- Sinkovics, R.R., Stöttinger, B., Schlegelmilch, B.B. and Ram, S. (2002), "Reluctance to use technology-related products: development of a technophobia scale", *Thunderbird International Business Review*, Vol. 44 No. 4, pp. 477-94.
- Steenkamp, J-B. and Baumgartner, H. (1998), "Assessing measurement invariance in cross-national consumer research", *Journal of Consumer Research*, Vol. 25 No. 1, pp. 78-90.
- Tabachnick, B.G. and Fidell, L.S. (2001), *Using Multivariate Statistics*, 4th ed., Harper Collins College Publishers, New York, NY.

Vandenberg, R.J. (2002), "Toward a further understanding of and improvement in measurement invariance methods and procedures", *Organizational Research Methods*, Vol. 5 No. 2, pp. 139-58.

van de Vijver, F. and Leung, K. (1997), *Methods and Data Analysis for Cross-Cultural Research*, 1st ed., Sage, Thousand Oaks, CA.

Appendix

Code	Item retained in the final MG-CFA model	Item retained in at least one group in the final Rasch model	Item wording
v01		X	I feel some anxiety when I approach an ATM
v02			I prefer to have people handle my bank activities than to use an ATM
v03			ATMs are fun to use
v04			I feel comfortable when using ATMs
v05			I want to learn more about using ATMs
v06	X	X	I worry about making mistakes when using ATMs
v07	X	X	Using ATMs is time-consuming
v08		X	ATMs agitate me
v09		X	It takes me a long time to complete bank transactions when using an ATM
v10			I think most people know how to use ATMs better than I
v11			I resent that ATMs are becoming so prevalent in our daily lives
v12			I can conduct my bank transactions without using an ATM
v13		X	I refuse to use ATMs
v14	X		It is easy to learn how to use ATMs
v15			I feel frustrated when I use an ATM
v16			I feel inadequate about my ability to use ATMs
v17		X	Thinking about ATMs makes me nervous
v18	X	X	I do not trust ATMs with my money
v19			ATMs make things too complicated
v20		X	ATMs are intimidated
v21			I wish I were more adept at using ATMs
v22			ATMs make bank transactions easier
v23		X	ATMs seem very complicated

Table A1.
Item pool of the
technophobia scale

(continued)

Code	Item retained in the final MG-CFA model	Item retained in at least one group in the final Rasch model	Item wording
v24			I like that ATMs are so convenient
v25			I feel more confident dealing with a human teller than an ATM
v26		X	Machines should not handle people's money transactions
v27	X	X	I find ATMs instructions confusing
v28			I have no fear of ATMs
v29	X	X	I feel confident that I could teach someone how to use an ATM
v30			I do not go to the bank after lobby and drive-thru teller hours

Source: Sinkovics *et al.* (2002)

Table AI.

Corresponding author

Thomas Salzberger can be contacted at: thomas.salzberger@wu-wien.ac.at



Researching the socio-cultural context: putting social representations theory into action

Elfriede Penz

*Institute of International Marketing and Management,
University of Economics and Business Administration, Vienna, Austria
Department of Marketing, Management School, Lancaster University, UK*

Abstract

Purpose – This paper aims at contributing to the conceptual and methodological advancement of international marketing research.

Design/methodology/approach – The theory of social representations is utilized to study what the representatives of a certain culture think of objects and products; which values they associate with these, which norms they follow, and in general how they view the world.

Findings – Triangulation was employed to facilitate researchers' access to the social representations approach. Six focus group discussions were conducted combined with a free association instrument answered by a sample of 250 respondents. The empirical study began from a social constructionist perspective. Findings showed that several conceptual and functional aspects of a technological innovation (m-commerce services) are idiosyncratic to particular groups, reflecting the group's societal affiliation and position.

Research implications/limitations – A translation problem prevails when dealing with specific words from cross-cultural research and the similarity of concepts must be considered when translating free associations in the course of international research. Other methodologies were only conceptually presented but not empirically used. Visual approaches such as pictorial instruments or ethnographical tools should be applied in future research.

Originality/value – The results provide a solid basis for international marketing research and suggestions for expanding the current study into international marketing research are provided.

Keywords Marketing, Research, Marketing theory, Postmodernism

Paper type Research paper

Introduction

A shift in marketing research has occurred from explaining differences between countries to the identification of trans-national similarities (Sheth and Parvatiyar, 2001). Similarities across countries allow standardized marketing and economies of scale. At the same time, individualization and migration have led to heterogeneous societies, making it increasingly important to study heterogeneity within countries.

Comparing phenomena and examining differences and similarities in behavior and response patterns across countries is vital in international marketing (Craig and Douglas, 2001). Research within this field treats the matter by comparing contexts, seeking to establish equivalence of concepts and research instruments. This procedure follows a world view in which culture matters ("emic" versus "etic" way of thinking) (Luna and Gupta, 2001). The lack of familiarity with cultures and environmental aspects is a major issue, which reduces the comparability of phenomena. A conceptual



framework that delivers rich cultural analysis is “ethnoconsumerism” (Venkatesh, 1995). It regards consumer behavior from the point of view of the subjects of the study, i.e. as a social or cultural group. In this respect, ethnoconsumerism differs from traditional comparative or cross-cultural approaches. It enables a deep understanding of how knowledge is constructed from a culture’s point of view. Compared to other comparative studies, it does not make assumptions on the data-level but focuses on phenomena at a sociocultural level. In other words, individuals are perceived “as products of their cultures and . . . as their cultures change, so do they” (Venkatesh, 1995, p. 30). Studying individuals as part of a culture, a social group and also equipped with individual characteristics enables a holistic view of them as consumers.

It appears necessary to understand what the representatives of a certain culture think of objects and products; which values they associate with these, which norms they follow, and in general how they view the world. The theory of social representations allows for studying these issues on a social and cultural level. By looking at how individuals group together through shared beliefs, emotions, lifestyles, and consumption practices, it is able to deal with the fuzziness and heterogeneity of today’s markets and postmodern consumer cultures (Elliott, 1999). Individuals who share a common vision of life are networks of societal micro groups (Cova, 1997; Cova and Cova, 2002; Featherstone, 1991). The theory of social representations views perceptions, ideas, and representations as processed collectively and as social creations (Elliott, 1999; Forgas, 1981).

This paper introduces social representation theory which contributes to the conceptual and methodological advancement of international marketing research. The contributions that are looked at in this paper relate to first, views of groups (group versus individual level) and second, tools and techniques to study the complex and fragmented structures of postmodern societies (Craig and Douglas, 2001; Jovchelovitch, 2001; Philogène, 2001). To facilitate access to the social representations approach for international marketing researchers, the methodological section includes an empirical example dealing with the perception and use of mobile phone services.

Looking beyond the confines of international marketing research

The beginning of this section outlines what social representation theory is. The following section focuses on the newness of social representations, i.e. to stress the social and cultural context as important influences on the elaboration of objects and situations, particularly in times of changes. Subsequently the contribution of social representation theory to international marketing research is discussed, followed by a section on innovative research methods. International marketing research is linked to social representation strategies and a typology of possible alternative avenues is outlined. Eventually, some practical applications of social representation theory are presented.

The theory of social representations

Social representations are forms of common sense knowledge (Moscovici, 1981, 1984) that refers to what people think they know of social objects or situations (Stewart and Lacassagne, 2005). Social representations are a bundle of organized judgments, attitudes and information with a hierarchical structure that a social group creates with respect to a social object or situation (Abric, 1996a). They are collective elaborations of a social object or situation, which are relevant for a social group. They are constructed in daily life by

individuals by communication and behavior (Wagner *et al.*, 1999), and serve as future guideline for people's communication and behavior (Moscovici, 1963). The theory of social representations is considered as a framework "of concepts and ideas to study psychosocial phenomena in modern societies" (Wagner *et al.*, 1999, p. 95).

Social representations are formed from new knowledge and ideas, which are integrated into already existing representations by means of "anchoring" and "objectification" (Abric, 1996b; Moscovici, 1984; Roland-Lévy, 2001). Anchoring is the process by which new knowledge, ideas, etc. are adopted by a social group. If the new knowledge, idea, etc. fits an existing categorization scheme, it will be integrated and will henceforth update the underlying categorization scheme. Objectification refers to the process by which abstract ideas or concepts achieve a concrete form, either as object or picture. This process echoes the importance of images in postmodern consumer cultures which can be made more explicit by studying the iconic aspect of social representations (Elliott, 1999).

Social representations comprise organized elements (Abric, 1996a, b). Two different types of elements of a social representation are distinguished: first, central elements form the nucleus of a social group's social representations. The nucleus aims at organizing the social group's ideas, generating meanings for the ideas, and operates as a normative constraint for the social group. The nucleus is assumed to be rather stable over time and changes less quickly than peripheral elements. The second type, peripheral elements may vary from individual to individual and reflect constant fluctuation due to ongoing changes and re-interpretation by consumers (Abric, 1993).

Taking a new look at groups: the social and cultural context in the study of social representations

Any form of knowledge is supposed to be bound to its social and cultural context of production (Jovchelovitch, 2001). For example, what a group of teenagers knows about and expects of the latest computer games may be related to their way of interacting with peers, their access to media and consumption patterns, and to parental perceptions and influences. If societal conditions change so too does knowledge. Viewing consumer behavior from a postmodern angle means acknowledging permanent change (Elliott, 1999) on a collective level (Wagner, 1989). Cultural groups develop certain beliefs and values, which are expressed by shared common-sense representations (Stewart and Lacassagne, 2005). The attention to the "social representations-glue" (Markus and Plaut, 2001, p. 186), which is not visible and becomes evident only through processes of making sense of unfamiliar or changing phenomena, is very much evident in contemporary consumer research's quest to identify the underlying, shared meanings of goods. Therefore, in order to understand consumers' behavior it is vital to understand the social and cultural context and the consequences that societal change has for people's knowledge (Jovchelovitch, 2001).

Collective elaborations of a social object by a community lead to their respective representation of an object and consequently result in a social reality (Moscovici, 1963; Wagner *et al.*, 1999). Subject and object are not seen as functionally separate; an object "... is what it is because it is in part regarded by the person or the group as an extension of their behavior" (Moscovici, 1973, p. xi). Representations and interpretations are acquired differently by groups, either via formal communication such as newspapers, or via informal conversations in a café. Bauer and Gaskell (1999)

call this process cultivation within groups and emphasize that the nature of sender-receiver relations may lead to a plurality of representations of a single issue. For example, the information consumer-friendly organizations provide or what is discussed by people in informal get-togethers, as compared to what companies tell their customers may vary and result in diverse social representations.

Above and beyond this view of social representations as socially shared dimensions, once social representations are formed they have an impact on individual behavior. Within applied social representations research, responses consist of individual opinions, attitudes, etc. taking inter-individual differences into account. It is through the application of different data-analysis methods pertinent to the social representation approach, that the organizing principles (common to the groups of individuals) are pieced together (Doise *et al.*, 1993).

Contributions of social representation theory to international marketing research

The social representations approach overcomes the view of culture as a static and bounded entity and understands it as a place of diversity in meanings and practices of diverse groups. Cultural contexts, for example, can be described “in terms of the nature and form of social representations that can be chronically accessible or easily activated” (Markus and Plaut, 2001, p. 188).

As suggested by Bauer and Gaskell (1999), the analysis of groups at times of a “crossover of their cultural projects,” i.e. when they face the same challenges, is promising. This will result in group-specific reactions to the concerns. By taking into account the different cultural projects that arise due to underlying differences, such as thoughts, languages, or values, social representations help: to understand underlying motivations and structures, and in comparing and understanding different meanings of social objects within independent social and cultural groups (Bergman, 1998; Stewart and Lacassagne, 2005).

Traditional segmentation strategies are primarily based on individual socio-demographic, attitudinal, or psycho-graphic characteristics. However, this falls short because the same goods can have different meanings for different consumer groups at the same time. These perceptions are idiosyncratic for them and even more important against the backdrop of postmodern consumer cultures, which are difficult to grasp with traditional marketing research techniques (Elliott, 1999). Traditional research usually takes an individual perspective, following that any variations in opinions or values between individuals within the same social group points to inter-individual differences. However, this view does not sufficiently facilitate the attainment of profound insight into social behavior and phenomena (Flick, 1995). Social contexts contribute to the overall picture that consumers have of relevant objects. Thus, the view that individuals have of technological innovations cannot be measured by extracting only individual information exempt from the social context.

Related to segmentation is the sampling procedure. Choosing the right sample is regarded as a vital element in international marketing research and recommendations maintain that it should be aligned with research objectives. Useful recommendations are given by Reynolds *et al.* (2003), who discuss the issue of representative sampling in great detail. Contextual factors, such as GNP, unemployment rate, cultural values, and the degree of acculturation or urbanization help in establishing sample equivalence across cultures (Sin *et al.*, 1999). Individual characteristics such as age, gender or

education help in establishing equivalent samples within cultures (Usunier, 1998). Equivalence in sampling aims at obtaining comparable samples across and within cultures (Malhotra *et al.*, 1996).

In contrast to taxonomic groups, which are the result of segmentation, social or natural groups (i.e. milieus) usually have a “common project,” shared ideology, a collective memory, and/or a common history and future., respectively, (*Weltanschauung*, see Bauer and Gaskell, 1999). Events in the past usually influence the meaning of social objects (Wagner *et al.*, 1999). Brands have particular meanings to consumers because they were used in their childhood. In order to obtain a natural group with respect to the topic of research, social and micro-cultural groups should be selected. Social groups that have a shared history, worldview or are exposed collectively to media or other communicative sources are self-referential and share a collective memory (Bauer and Gaskell, 1999; Gaskell, 2000). In practice, this means that it is necessary to explore the collective and shared understanding of concepts within selected natural groups before actually starting comparative marketing research.

Research techniques to study complex and fragmented socio-cultural structures

The social representations approach faces observable “climate” changes in research by turning towards more contextual and qualitative issues as compared to usual, cross-cultural research (Philogène and Deaux, 2001).

Craig and Douglas (2001, p. 85) recommend “to create and make imaginative and thoughtful use of new approaches to understand the changing marketplace”. Researchers should avoid single method bias in data collection (Sin *et al.*, 1999), for instance by using different methods or conceptualizations of the problem (Jick, 1979; Scandura and Williams, 2000). Triangulation consists of adding various dimensions to a project, i.e. having multiple sources of data (diverse respondent groups) or having a team of researchers. The multi-method approach is part of triangulation and points to the application of various methods within a project (Potter, 1996).

Peterson (2005) assumes that a high proportion of variance in consumer behavior studies remains unexplained due to an “instant” construction of responses in self-report studies. Consumers construct their response when answering a question and thereby “are using the questionnaire to decide what their “attitudes” are” (Peterson, 2005, p. 349). Consequently, methods other than self-reports are suggested to overcome these limitations. These methods can include experimental designs in which researchers control predictor variables and assess the effect on the dependent variable, collecting non-verbal data, observational or the use of responses generated by pictures or words. The understanding of people’s behavior at different cultural levels can be improved by applying projective techniques, elicitation tasks, and qualitative methods (Craig and Douglas, 2001). The seemingly vast variety of behaviors and symbolic meanings calls for diverse techniques capable of unraveling unconscious, non-rational but emotional and changing meanings within fragmented consumer groups (Elliott, 1999).

Studying social representations means running into the duality between individual and social. Investigations focus on the social context in which consumers learn, perceive, and develop rather than individual attitudinal dimensions (Moscovici, 1961, 1963). Research methods need to be at hand that allow the capturing of various aspects that “transcend the confines of more traditional . . . research” (Wagner *et al.*, 1999, p. 101). In contrast to traditional approaches, the social representations approach

comes with particular methodological strategies that seamlessly connect theory and empirical applications (Philogène and Deaux, 2001).

In the following, research techniques pertinent to the social representation theory are presented to study complex and fragmented socio-cultural structures. Table I contains international marketing research issues and connects them to the social representation approach. The strategies by Bauer and Gaskell (1999) can be seen as a typology of possible alternative avenues. The table illustrates what is done in social representations research and provides a comparison to concepts of equivalence in traditional international marketing research.

When studying the adoption of new mobile phone services, individuals' views are based, for example, on how important technology is in general or what aspects are highlighted within a social group. Being embedded in a group context where technological innovations are welcomed is likely to produce a greater sense of openness towards new mobile phone services compared to contexts in which people view new technology more in terms of concerns and unfamiliarity.

The social representations approach has produced numerous methodologies that overcome the limitations of traditional research. One issue in international marketing research is for example conceptual equivalence, i.e. what meaning constructs have for respondents. By studying images, metaphors and free associations, the content and process of social representations can be assessed and the meanings of constructs in relation to the social and cultural settings and changes made explicit. In particular, studying the core and peripheral elements of social representations and revealing the anchoring and objectification processes tells us about the socially shared metasytem that influences thoughts and actions (Philogène, 2001). In their framework, Bauer and Gaskell (1999) as well as Stewart and Lacassagne (2005), build upon these methodologies and recommend studying images, metaphors, and free associations, which allows assessment of the function (anchoring, objectification) and organization (center and periphery) of social representations within a specific social group. The analysis of social representations within a defined social group combines different data sources in multi-method analyses. According to the different research aims, observational field research is recommended for behavioral habits, questionnaires for individual perception, and document analysis for formal communication (Moscovici, 1984).

As Bauer and Gaskell (1999, p. 178) point out, the:

... life-cycle of social representations is somewhere between the elusiveness of the flow of consciousness and the "long duree" of mentalities.

Changes in the structure and functions of social representations due to changes in the socio-economic or political environment can be observed by means of comparative and longitudinal analysis. In the course of empirical research researchers should avoid "social engineering" (Bauer and Gaskell, 1999, p. 179) and creating social realities for respective social groups based on common assumptions regarding differences (Bergman, 1998). While this in general reflects a call for "neutral" research, cases exist where more subjective research is appropriate. A personal perspective may enable the researcher to "critically see through their subjectivity" (Reason, 1998, p. 280); however they are aware of potential distortions.

Table I.
International marketing
research issues and social
representation strategies

<p>International marketing research issues Social representations strategies (Bauer and Gaskell, 1999; Stewart and Lacassagne, 2005) Content and process of social representations Meaning of constructs</p>	<p>What is done in social representations research? Study of images, metaphors, free associations Assess the content of social representations for social groups Anchoring/objectification</p>	<p>Relation to equivalence issues in traditional international marketing research (e.g. Salzberger <i>et al.</i>, 1999)</p>	<p>Illustration of social representations approach</p>
<p>Crossovers of cultural projects and diverging coping strategies Culture as context Identification of ideas, etc.</p>	<p>Study of images, metaphors, free associations Assess the function and organization of social representations for social groups Core/periphery Specification of appropriate social segments Search for a shared history, world view, collective exposure to mass media, events in the past or common future</p>	<p>Equivalence of research topic: Functional equivalence</p>	<p>Focus-group discussion Four stimulus situations "m-commerce" Elicit free associations Evaluation of associations Function and content of social representations (analysis of correspondence) Polarity index</p>
<p>Selection of natural groups Sampling design Traditional segmentation strategies</p>	<p>Core/periphery Specification of appropriate social segments Search for a shared history, world view, collective exposure to mass media, events in the past or common future</p>	<p>Equivalence of research units/sample: Definition Selection</p>	<p>Before data collection Results from focus-group (underlying themes) Representative socio-demographics (gender, education) Micro-cultural factors (urbanization, age) After data collection Content of associations Usage patterns Search for interactive information in collected data Media outputs on new products Information given by market</p>
<p>Cultivation within the social group Communication patterns Survey method</p>	<p>Highly process oriented Can include both analysis of informal and formal communication relations between sender and receivers</p>	<p>Equivalence of research administration: Timing Interaction</p>	<p>Search for interactive information in collected data Media outputs on new products Information given by market (<i>continued</i>)</p>

Combination of different data sources	According to different research aims, e.g.	Equivalence of research methods	Focus-group discussion Elicit free associations
Multi-method analysis	Observatory field research (behavioral habits) Questionnaires (individual perception) Document analysis (formal communication) Use of longitudinal data Comparative analysis	Stimuli Data collection	Media reports Advertising
Dynamic nature of social representations	Observe changes in the structure and the functions of social representations	Equivalence of research administration and research methods: Timing Stimuli	
Neutral researcher	Avoid "social engineering"	Data collection	No interventions but thorough documentation of findings
Self-reference criterion	No virtually created social realities	Equivalence of data handling Response translation Response categories	

Table I.

The last column of Table I illustrates the empirical study which is presented below in the methodological part of this paper. The application of social representation theory appears particularly well suited for investigations within contexts of considerable degree of change, since in times of change debate increases among social groups, existing social representations are often re-evaluated (Furnham, 2001), and the meanings of consumer goods are socially re-constructed and re-evaluated (Elliott, 1999). Within international marketing there is considerable change with respect to innovations. Examples of studies that utilize the social representation approach are studies on payment innovations, such as the introduction of the e-purse (Penz *et al.*, 2004) or the introduction of the Euro in European countries (Meier and Kirchler, 1998).

Methodology

To facilitate marketing researchers' access to the social representations approach, this section includes an empirical example dealing with the perception and use of m (mobile)-commerce based phone services. M-commerce means electronic commerce made through mobile devices (M-commerce, 2005) and reflects a technological change that occurs both within a society and internationally.

Studies found that mobile devices were among the most quickly adopted consumer products (Rupp and Smith, 2002). There is already a high penetration of mobile phones across Europe (for details see, e.g. Hayward, 2000). Mainly because of different standards, the USA has lagged behind Europe in terms of both the penetration of mobile phones and the early adoption of m-commerce opportunities. Whereas in Europe text messaging and WAP technology were developed rather early, the USA is further ahead in the development of internet-based technology applicable to the wireless environment (Hayward, 2000). The key advantage of m-commerce over e-commerce is the location function (Yeo and Huang, 2003). Main applications for consumers are: gaming and entertainment, information, billing (mobile payment mechanisms), safety and tracking (Hayward, 2000; Yeo and Huang, 2003), although the latter has reaped criticism (Dye, 2005).

Traditional research focused on consumer adoption by using the technology acceptance model (Bruner *et al.*, 2005) or tried to identify factors that affect the likelihood of transacting business by means of mobile phones (Magura, 2003). An international study utilized focus-group discussions to gain insight into values and limitations of m-commerce services (Jarvenpaa *et al.*, 2003). Yet none of these studies focused on the view that groups with a common social context have of technological innovations and most used methods that primarily addressed individual characteristics of consumers.

Goal and assumptions

The goal of the study was to analyze social representations of consumer groups with respect to technological innovations. In particular, the collectively shared perceptions, ideas, and representations of m-commerce services were assessed and the results interpreted as social creations (Elliott, 1999; Forgas, 1981). The application of the theory of social representations complements the aforementioned discussion of its contribution to the conceptual advancement of international marketing research.

The assumptions are that consumers group together according to their *Weltanschauung*, which reflects their social and cultural context and goes beyond

socio-demographic descriptions. The social representations of m-commerce services are expected to be idiosyncratic to these natural groups and allow understanding the function and organization of the services.

The following methodology was applied: in the beginning, focus-group discussions were conducted to understand the respondents' view on mobile communication in general, their usage patterns, and ideas about future m-commerce services. This was followed by the application of an instrument that evoked associations and evaluations of four stimulus situations of m-commerce services. Themes that had emerged in the focus groups served as a background for a refined analysis within sub-groups. The study uses the criteria age, usage behavior, and level of urbanization as a starting point for defining micro cultures (Steenkamp, 2001), and the emerging, group specific common themes as ways to assess and structure social representations.

Method

Participants. The selection of focus group participants is based on dividing the social space into social strata and functions[1] and representations. While social strata are usually known from the outset (age, gender, income), representations are unknown and unfold only through appropriate research techniques. The results represent the corpus, i.e. a finite collection of material which the researcher continues to work with (Gaskell, 2000).

The six focus groups comprise four to five participants homogenous in terms of social strata information (age, gender, occupation, and mobile phone usage) and heterogeneous with respect to the socio-cultural context (reflected in lifestyle, preferences, etc.). Discussions were conducted with individuals of any age who did not own mobile-phones, and with the following age groups: people between 14 and 19 years, students (between 20 and 25 years old), employees (between 34 and 50 years old), executive employees and entrepreneurs (25-55 years old), and retirees (between 55 and 79 years old).

Instruments. To obtain a holistic picture of consumers' perceptions regarding technological innovation presented by mobile communication, triangulation, and more specifically, a multi-method approach were used. Focus group discussions are used as a valid method to learn more about underlying ideas and views of consumers coming from different social contexts on socially relevant phenomena and new products and services (Craig and Douglas, 2001). They prove useful because within groups there is an ongoing sharing and negotiating of realities and thus social interactions are reflected more accurately than in individual-based research techniques (Gaskell, 2000).

In order to incorporate social representations ideology and step beyond the traditional research focus, the measurement tool comprised an associative network (de Rosa, 1995; de Rosa and Kirchler, 2001; Vergès and Bastounis, 2001), which served as a means to identify the concepts' latent evaluative components. The technique deals with spontaneous associations by asking respondents to write down ideas, pictures, and thoughts that arise when they think of the stimulus word(s). The resulting associations are first listed and then evaluated by the same respondents as positive, negative, or neutral (de Rosa, 1995; de Rosa and Kirchler, 2001). In the present study, four anchor-situations (1-4, as follows) were designed to assist the participants in coming up with ideas. The situations were long enough to offer an idea about what the service would be and short enough not to influence their reactions. The design was

pre-tested and a few changes were made in order to make the initial stimulus situations as objective as possible. Feedback from the participants confirmed the usefulness of the situations. The sequential order of the situations varied systematically in order to increase the validity of the measurement.

- (1) The first situation focused on information services. Provided was a brief account about a person traveling to a foreign city: the person is looking for appropriate accommodation and asks for information via mobile phone.
- (2) The second situation is about a person, who – chronically suffering from a specific health problem – is involved in an accident. The service that the person uses informs the nearing ambulance in advance about the appropriate treatment.
- (3) The third situation deals with a person waiting for a bus who receives a message from an entertainment service that a potential “dream-partner” is nearby.
- (4) Finally to exemplify billing services a person is described who is informed that a call would cost less by moving a few meters further down the road where there is a connection to a less expensive tower. This service builds on tower-specific pricing.

Procedure. The focus-group discussions took place in private settings in urban and rural environments. Two trained research assistants helped in the data collection. The discussions were tape-recorded and subsequently transcribed. After each discussion the interviewer, the moderator, and the project leader met to summarize the process and collect observational data about the respective groups.

Collection of associations and evaluations began in October and was completed in December 2000. A non-probability quota sample technique was applied, representative of the general population of Austria, providing for a geographical mix of rural, urban, and metropolitan areas (Churchill and Iacobucci, 2002). The quota was based on age, gender, and level of education in the respective areas and a total number of 250 interviews were personally conducted by two interviewers. Participants ranged from 15 to 83 years old, and were, on average, 38.35 years old. Of these, 56.2 percent were female and 43.8 percent were male. Most of the respondents lived in urban (60.4 percent), 39.6 percent in a rural area. The majority of participants owned a mobile phone (79.4 percent), only 20.6 percent did not. About half of participants had jobs (53.4 percent), while the remaining 46.6 percent were pupils, students, housewives, or retirees. Within the representative sample, sub-groups are identified to share group-specific, common perspectives of daily life and thus have access to a collective memory.

Results

Emerging themes. Three broad issues come out of the focus-group discussions, namely:

- (1) communication;
- (2) technological innovation; and
- (3) social reality (Figure 1).

Mobile communication (1) is viewed as a means of achieving success by being reachable and able to make calls at all times. It is perceived as enabling more and

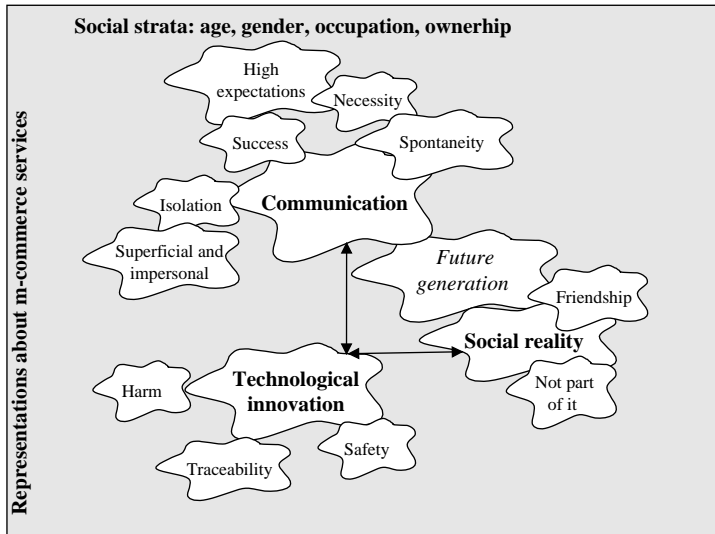


Figure 1.
Themes emerging from
focus-group discussions of
m-services

spontaneous contacts and is regarded as a necessity for some groups (e.g. young people, employees). These views go along with high expectations regarding future possibilities of communication. Concerns and fears regarding shorter, more superficial and impersonal contacts are raised by people without a mobile phone. Isolation seemed to be a possible consequence of mobile phone use.

Some groups are concerned about health and safety issues that derive from overly technically-driven innovations (2). Other concerns deal with the protection of data and one's traceability and visibility. However, the idea that mobile phones might help ill or injured people is accepted by many, previously skeptical participants. There is a call for better and more sophisticated applications.

Finally, the influence of social reality (3) becomes obvious in the discussions. Young people in particular have extensive knowledge of the possible services that mobile phones can deliver. This kind of expertise contributes to their self-esteem, as has been observed in the discussions. Knowing more and more precisely boosts their status among their peers. The self-named "future generation" believes in the positive effects of mobile communication in general and future location-based services in particular. Their vision includes intensifying friendships based on the same provider as those with the same provider save money communicating with one another. The most threatening issue for this group is to be unable to respond to calls or to miss calls from friends and parents. In contrast, for retirees, the idea of mobile communication is rejected simply because it had never been a part of their lives.

Content of social representations (anchoring and objectification). A total of 1,767 associations were gathered by the free association technique on the basis of $n = 250$ respondents. Of these associations:

- 512 words were associated with the story about information-service
- 421 with the story about safety-services; and

- 447 with the story about entertainment-services
- 387 with the story about billing-services.

In general, 36.62 percent of the associations were different, indicating a moderate level of stereotypical views and a rather heterogeneous representation regarding all of the services (Table II).

The frequencies of the associations were the basis for a correspondence analysis. The analysis yielded two factors explaining 35.3 percent and 25.8 percent of variance, respectively. A third dimension explained another 19.5 percent, which was not included in the interpretation since dimension one and two already explained 61.1 percent of the variance. The overall spatial variation (total inertia) was 0.29, indicating that the correlation between row points (associations) and column points (stimulus situations, age, gender, residence, job, mobile-phone ownership) was moderate (Figure 2).

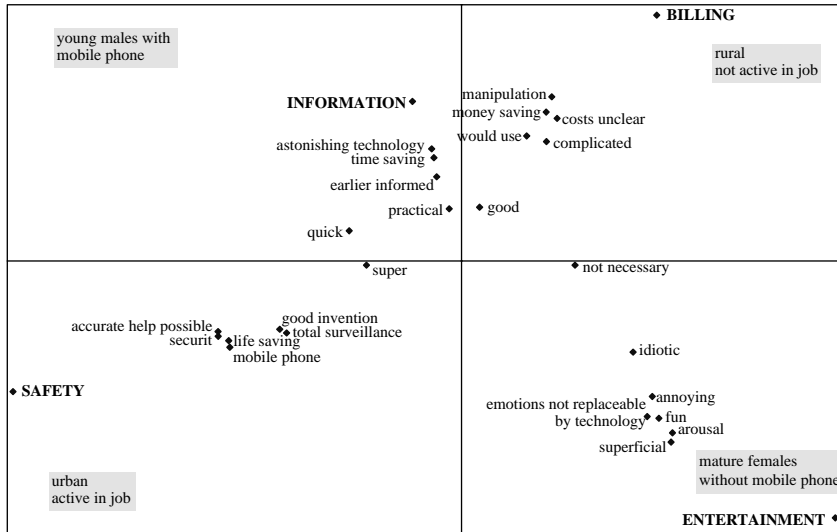
The four m-services were anchored differently and are characterized by different content. The nature of the associations differed in terms of evaluation and content: while information services revealed more positive and rational associations among young males, entertainment services were usually rejected by 35 + females based on negative, emotional aspects. Obviously, the latter group does not consider mobile phones or related services as necessary and thus can hardly be convinced of their advantages. Especially when it comes to more emotional themes such as finding a partner or being entertained, the resistance of the 35 + females seemed to be an expression of the threat imposed (e.g. “communication will decrease” or “emotions will be replaced by technology”) and their coping mechanisms. They rejected the service as superficial and idiotic, but they agreed that the described situation had a certain potential in terms of arousal and could be fun.

A second distinction was found between safety and billing services. A more positive attitude was found with regard to the safety service while billing services were highly rejected. People living in cities basically perceived the safety service as a good invention. They were also aware of total surveillance but seemed to accept this for safety sake. People from rural environment and those not actively employed, directed their attention on the money-saving aspect of the service. The technology seemed to be complicated to them, nevertheless they would use it.

Indexes of polarity and neutrality were calculated separately for the four stimulus situations [2] (de Rosa, 1995; de Rosa and Kirchler, 2001). An analysis of variance revealed that the index of polarity ($F(3,210) = 14.54; p < 0.001$) and the index of neutrality ($F(3,210) = 3.18; p < 0.025$) differed significantly among the four situations.

Table II.
Frequencies of associated words regarding four m-services

	Frequency	Different associations	Percentage
Information	512	235	45.90
Safety	421	146	34.68
Entertainment	447	203	45.41
Billing	387	175	45.22
Total	1,767	647	36.62



Note: Row points (associations) and column points (m-services) that contribute significantly to one or both dimensions are displayed (Matiasko, Dobrov, and Bronner, 1994).

Figure 2. Content of social representations and diverging/converging coping mechanisms

In general, safety and information services were evaluated more positively than billing. Entertainment was the least positively valued service. Additionally, inter-group effects were analyzed, resulting in significant differences in the evaluations (polarity index) between rural and urban sub-samples ($F(1,70) = 5.48; p < 0.05$) and people who own mobile-phones and those who don't ($F(1,70) = 4.07; p < 0.05$). The rural sub-group was more positive than the urban group. Regarding mobile-phone ownership, those with one were more positive about the services (Table III).

Function and organization of social representations (core and peripheral elements). In line with the theory of social representations, the four services in question were analyzed for their core and peripheral elements. According to Abric (1993; 1996b), words which were mentioned most frequently and appeared right at the beginning in the association process can be considered elements belonging to the core. Peripheral elements were mentioned less frequently or later in the association process, and stand for ideas, attitudes, evaluations, etc. that enable the stability of the core. They comprise

	Information services			Safety services			Entertainment services			Billing services		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
Urban	0.56	0.52	59	0.75	0.52	59	0.23	0.76	59	0.50	0.64	59
Rural	0.71	0.50	37	0.76	0.60	37	0.35	0.77	37	0.53	0.73	37
Mobile	0.65	0.49	72	0.73	0.57	72	0.28	0.75	72	0.57	0.64	72
No mobile	0.53	0.58	24	0.82	0.49	24	0.26	0.82	24	0.34	0.75	24
TOTAL	0.62	0.52	96	0.76	0.55	96	0.28	0.76	96	0.51	0.67	96

Table III. Index of polarity regarding the four m-services

those elements that have changed over time or are inconsistent among social groups (Vergès and Bastounis, 2001).

Information services in general were perceived as practical, super, as a good idea that helps to save time. Not really necessary, services were related to convenience (core elements). Peripheral elements add to the understanding of information services' social representation. There were two types of elements: one type reflects the admiration for what technology makes possible (astounding technology, progressive, futuristic, informed earlier) while the other type consists of dubious notions such as expensive, complicated, probably not full information. Both types were mentioned at the same time and were found in the periphery of the organization of information services.

In the core of safety services were expressions such as super, life saving, good invention, safety, and the remark that in this particular case it would be positive. More critical remarks appear solely in the periphery (e.g. total surveillance, unrealistic) indicating that such a service is not without critique, but is essentially important (accurate help possible).

Entertainment services were essentially idiotic, not necessary, unrealistic, superficial, but can be arousing and also as super. Some detailed references to other concepts (such as personal ads) or links to one's own usage (also not using it) were revealed in peripheral zones, which points to more disagreement about the value of this particular service. Further elements that occurred in the peripheral zones were fun, good, but also more critical ones such as horror, emotions are not replaceable by technology, communication is lacking, and a general loss of contact.

Finally, core elements of billing services indicated the function of money saving, but also not necessary or complicated and idiotic. Other core elements were good, super, practical, and reference to actual usage. The peripheral elements lead to the conclusion that certain questions still need to be answered before making an ultimate judgment (costs unclear, depends on price advantage), indicating a particular openness (good idea, interesting, why not) toward the service.

Discussion

Culture is sometimes regarded as a complex phenomenon, perhaps too complex to be captured by only a few dimensions (Steenkamp, 2001). Traditional methodologies in international marketing research usually follow a positivistic and rather "etic" (or "imposed etic") view of cultural phenomena. The proposed theory of social representations views reality as epistemologically socially constructed, and helps to develop an "emic"-view on phenomena. It overcomes the positivist limitation that assumes homogenous meanings of goods within taxonomic market segments (Elliott, 1999). By means of the social representation approach it is possible to observe how social groups make sense of new elements by adding them to the existing meaning system (anchoring). Furthermore, it is possible to identify the rules and conventions that determine behavioral patterns (objectification) (Harré 1998). The current empirical study began from a social constructionist perspective, allowing the social context to shape consumers' perceptions of product innovations meaningfully. This view further emphasizes the importance of research on a micro cultural level.

Because market segments are no longer helpful in predicting consumer behavior from a postmodern perspective, selecting appropriate samples arose as an issue. In this paper, it was proposed that selecting natural rather than nominal sample groups helps

in overcoming this limitation. Collective memory and shared history as well as cultivation form groups' consumption practices. Significant sampling criteria were gained by focusing on group discussions, such as the label "future generation," which was identified as a common way for young consumers to identify themselves.

The social representations approach makes available a variety of methodologies for overcoming limitations often imposed by traditional survey methods, such as construction bias and individual-level focus. By definition, social representations reflect the group level and "allow" different instruments to reveal the group's metasystem. Of the available methods, this study applied the technique of free associations to understand the concept and function of the proposed m-commerce services. The results reflect deep-seated differences among socio-cultural groups.

Several conceptual, functional and emotional aspects of m-commerce services were identified. The concept of time revealed insights into the respondents' societal configurations. People, who want to save time, and who think that consuming requires time, have different concepts of time, thus forming natural market segments. The mobile phone as functional commodity is not essential. It is its value to symbolize one's time resources and reflecting one's position within a society (Featherstone, 1991). All of the m-commerce services were found to be not really necessary, but at the same time, helpful in daily life. Bruner *et al.* (2005) found that while perceived usefulness (utilitarian aspect) contributed to consumer acceptance of Internet devices, what contributed more was the "fun" attribute (hedonic aspect). In the case of m-commerce services, considering the utilitarian value went along with positive expressions. Material goods are communicators and their original use is increasingly difficult to interpret. Mobile phones signify pleasure or freedom, which thus determines people's significance within specific social groups (Featherstone, 1991).

Emotional ties are reflected in the study's results: young respondents in the focus-groups mentioned a feeling of "connectedness" that is fostered by mobile phones. Older respondents and those opposing mobile phones argued that personal ties would disintegrate and personal contact would be lost. The loss of personal contact and the fear that emotions could be replaced by technology, etc. were important themes that derived from the functional analysis of associations. M-commerce services add to the various forms – "disorders" – of freedom that consumers already experience and they could decrease in value for consumers. To be successful, m-commerce services need to assure the various groups that they enable social ties rather than destroy them.

Based on the theory of social representations (Bauer and Gaskell, 1999), seven different strategies (Table I) form a typology of possible research avenues to international marketing research. In the same vein, yet more recently, Stewart and Lacassagne (2005) confirmed the usefulness of social representations to identify cultural differences. This paper contributes to conceptual and methodological advancement by including the concept of ethnoconsumerism (Venkatesh, 1995) and following Wagner *et al.* (1999, p. 101), who proposes to "take into account the complexity of the phenomenon under investigation and pay attention to the context and the diversity of voices".

Suggestions for expanding the current study into the area of international marketing research are provided. In order to find out how new products are integrated into the existing social and cognitive opinion system, collective associative network, assisted by similarity analysis, is proposed. This technique is applied to analyze the

configuration of social representations, particularly linkages with product innovations and the connecting socio-economic environment. Respondents link existing dimensions and socioeconomic concepts by drawing connecting lines. For the purposes of analysis, the number of connected word pairs are counted and put into graphs of similarity (Vergès, 1987, 1996; Vergès and Bastounis, 2001).

Research on a more global level does not necessarily mean carrying out the same procedure applied in a particular cultural setting. The richness of data must not be forfeited for universalisation of phenomena. Splitting goals of research into description, comparison, and generalization, Venkatesh (1995) warns of the danger of losing the richness of data by moving from description to generalization, i.e. from a focus on understanding to a focus on universalizing. Comparative research without losing quality is possible by using a single cultural setting and making comparisons with other cultures on the basis of secondary data. Recent studies on various cultures (Asian, such as Japan, Korea, India and China; and emerging eastern European markets) suggest, as Venkatesh (1995, p. 42) put it:

... that there are deep-seated differences in the way consumer trends are unfolding, and, therefore, that superficial similarities may hide real cultural differences.

Useful comparisons with other cultures ideally employ triangulation and, for example, as suggested by Venkatesh (1995), textual information instead of applying uni-cultural concepts to cross-cultural settings.

Taken as a whole, this paper aims at contributing to the development of sound methodologies in international marketing. The results of the empirical study, such as the identified, collectively constructed themes of time, emotional connectedness, and different functions of material goods, served as an illustration of the proposed approach and confirmed its usability and contribution to international marketing research. Yet some limitations remain. A translation problem prevails when dealing with specific words from cross-cultural research. The similarity of concepts must be considered when translating free associations in the course of international research. Also, the variety of methods is limited in this paper. Corresponding to the need of fragmented consumer cultures for projective techniques, visual approaches such as pictorial instruments or ethnographical tools should be applied in future research.

Notes

1. "Social strata and functions" refer to the range of external variables that are controllable for material and participant selection (Gaskell, 2000).
2. The polarity index is the relation between the difference of positively and negatively evaluated associations and the total number of associations. The polarity index ranges from +1 to -1; higher values expressing a positive evaluation of the stimulus word. The neutrality index expresses the relative frequency of neutral associations; the higher it is, the more neutral the respective stimulus word is evaluated.

References

- Abric, J.-C. (1993), "Central system, peripheral system: their functions and roles in the dynamics of social representations", *Papers on Social Representations*, Vol. 2 No. 2, pp. 75-8.
- Abric, J.-C. (1996a), "De l'importance des représentations sociales dans les problèmes de l'exclusion sociale", *Exclusion Sociale*, pp. 11-17, Insertion et Prévention, Erès, Saint-Agne.

- Abric, J.-C. (1996b), "Specific processes of social representations", *Papers on Social Representations*, Vol. 5 No. 1, pp. 77-80.
- Bauer, M.W. and Gaskell, G. (1999), "Towards a paradigm for research on social representations", *Journal for the Theory of Social Behaviour*, Vol. 29 No. 2, pp. 163-86.
- Bergman, M.M. (1998), "Social representations as the mother of all behavioral predispositions? The relations between social representations, attitudes, and values", *Papers on Social Representations*, Vol. 7 Nos 1/2, pp. 77-83.
- Bruner, I., Gordon, C. and Kumar, A. (2005), "Explaining consumer acceptance of handheld internet devices", *Journal of Business Research*, Vol. 58 No. 5, pp. 553-8.
- Churchill, G.A. and Iacobucci, D. (2002), *Marketing Research: Methodological Foundations*, South-Western, Mason, OH.
- Cova, B. (1997), "Community and consumption: towards a definition of the "linking value" of product or services", *European Journal of Marketing*, Vol. 31 Nos 3/4, pp. 297-316.
- Cova, B. and Cova, V. (2002), "Tribal marketing: the tribalisation of society and its impact on the conduct of marketing", *European Journal of Marketing*, Vol. 36 Nos 5/6, pp. 595-620.
- Craig, C.S. and Douglas, S.P. (2001), "Conducting international marketing research in the twenty-first century", *International Marketing Review*, Vol. 18 No. 1, pp. 80-90.
- de Rosa, A.S. (1995), "Le 'réseau d'associations' comme méthode d'étude dans la recherche sur les représentations sociales", *Cahiers Internationaux de Psychologie Sociale*.
- de Rosa, A.S. and Kirchler, E. (2001), "Ambiguous images in advertising: an application of the associative network", in Roland-Lévy, C., Kirchler, E., Penz, E. and Gray, C. (Eds), *Everyday Representations of the Economy*, WUV-Verlag, Wien, pp. 49-65.
- Doise, W., Clémence, A. and Lorenzi-Cioldi, F. (1993), *The Quantitative Analysis of Social Representations*, Harvester Wheatsheaf, Hemel Hempstead.
- Dye, P. (2005), "Under surveillance", *Precision Marketing*, pp. 23-4, July.
- Elliott, R. (1999), "Symbolic meaning and postmodern consumer culture", in Brownlie, D., Saren, M., Wensley, R. and Whittington, R. (Eds), *Rethinking Marketing*, Sage, London, pp. 112-25.
- Featherstone, M. (1991), *Consumer culture & postmodernism*, Sage, London.
- Flick, U. (Ed.) (1995), *Psychologie des Sozialen*, Rowohlt Taschenbuch Verlag, Hamburg.
- Forgas, J.P. (Ed.) (1981), *Social Cognition: Perspectives on Everyday Understanding*, Academic Press, London.
- Furnham, A. (2001), "Social representations of welfare and economic inequality", in Roland-Lévy, C., Kirchler, E., Penz, E. and Gray, C. (Eds), *Everyday Representations of the Economy*, WUV-Verlag, Wien, pp. 113-36.
- Gaskell, G. (2000), "Individual and group interviewing", in Bauer, M.W. and Gaskell, G. (Eds), *Qualitative Researching with Text, Image and Sound: A Practical Handbook*, Sage, London, pp. 38-56.
- Harré, R. (1998), "The epistemology of social representations", in Flick, U. (Ed.), *The Psychology of the Social*, University Press, Cambridge, MA, pp. 129-37.
- Hayward, C. (2005), "The outlook for mCommerce: technology and applications to 2005", Business Insights Ltd, London.
- Jarvenpaa, S.L., Lang, K.R., Takeda, Y. and Tuunainen, V.K. (2003), "Mobile commerce at crossroads", *Association for Computing Machinery. Communications of the ACM*, Vol. 46 No. 12, pp. 41-4.
- Jick, T.D. (1979), "Mixing qualitative and quantitative methods: triangulation in action", *Administrative Science Quarterly*, Vol. 24 No. 4, pp. 602-11.

- Jovchelovitch, S. (2001), "Social representations, public life, and social construction", in Deaux, K. and Philogène, G. (Eds), *Representations of the Social*, Blackwell, Oxford, pp. 165-82.
- Luna, D. and Gupta, S.F. (2001), "An integrative framework for cross-cultural consumer behavior", *International Marketing Review*, Vol. 18 No. 1, p. 45.
- Magura, B. (2003), "What hooks m-commerce customers?", *MIT Sloan Management Review*, Spring, p. 9.
- Malhotra, N.K., Agarwal, J. and Peterson, M. (1996), "Methodological issues in cross-cultural marketing research: a state-of-the-art review", *International Marketing Review*, Vol. 13 No. 5, pp. 7-43.
- Markus, H.R. and Plaut, V.C. (2001), "Social representations: catching a good idea", in Deaux, K. and Philogène, G. (Eds), *Representations of the Social*, Blackwell, London, pp. 183-9.
- M-commerce (2005), available at: <http://en.wikipedia.org/wiki/M-commerce>
- Meier, K. and Kirchler, E. (1998), "Social representations of the euro in Austria", *Journal of Economic Psychology*, Vol. 19, pp. 755-74.
- Moscovici, S. (1961), *La Psychoanalyse: Son Image et Son Public*, Universitaires de France, Paris.
- Moscovici, S. (1963), "Attitudes and opinions", *Annual Review of Psychology*, Vol. 14, pp. 231-60.
- Moscovici, S. (1973), "Foreword", in Herzlich (Ed.), *Health and Illness*, Academic Press, London.
- Moscovici, S. (1981), "On social representations", *Social Cognition: Perspectives on Everyday Understanding*, Academic Press, London.
- Moscovici, S. (1984), "The phenomenon of social representations", in Farr, R. and Moscovici, S. (Eds), *Social Representations*, Cambridge University Press, Cambridge, pp. 3-69.
- Penz, E., Meier-Pesti, K. and Kirchler, E. (2004), "It's practical, but no more controllable: social representations of the electronic purse in Austria", *Journal of Economic Psychology*, Vol. 25 No. 6, pp. 771-87.
- Peterson, R.A. (2005), "Response construction in consumer behavior research", *Journal of Business Research*, Vol. 58, pp. 348-53.
- Philogène, G. (2001), "A theory of methods", in Deaux, K. and Philogène, G. (Eds), *Representations of the Social*, Blackwell, Oxford, pp. 39-41.
- Philogène, G. and Deaux, K. (2001), "Introduction", in Deaux, K. and Philogène, G. (Eds), *Representations of the Social*, Blackwell, Oxford, pp. 3-7.
- Potter, W.J. (1996), *An Analysis of Thinking and Research about Qualitative Methods*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Reason, P. (1998), "Three approaches to participative inquiry", in Denzin, N.K. and Lincoln, Y.S. (Eds), *Strategies of Qualitative Inquiry*, Sage, Thousand Oaks, CA, pp. 261-91.
- Reynolds, N.L., Simintiras, A.C. and Diamantopoulos, A. (2003), "Theoretical justification of sampling choices in international marketing research: key issues and guidelines for researchers", *Journal of International Business Studies*, Vol. 34 No. 1, pp. 80-9.
- Roland-Lévy, C. (2001), "Financial behavior of households and social representations", in Roland-Lévy, C., Kirchler, E., Penz, E. and Gray, C. (Eds), *Everyday Representations of the Economy*, WUV-Verlag, Wien, pp. 67-91.
- Rupp, W.T. and Smith, A.D. (2002), "Mobile commerce: new revenue machine or black hole?", *Business Horizons*, Vol. 45 No. 4, pp. 26-9.
- Salzberger, T., Sinkovics, R.R. and Schlegelmilch, B.B. (1999), "Data equivalence in cross-cultural research: a comparison of classical test theory and latent trait theory based approaches", *Australasian Marketing Journal*, Vol. 7 No. 2, pp. 23-38.

- Scandura, T.A. and Williams, E.A. (2000), "Research methodology in management: current practices, trends, and implications for future research", *Academy of Management Journal*, Vol. 43 No. 6, pp. 1248-64.
- Sheth, J.N. and Parvatiyar, A. (2001), "The antecedents and consequences of integrated global marketing", *International Marketing Review*, Vol. 18 No. 1.
- Sin, L.Y.M., Cheung, G.W.H. and Lee, R. (1999), "Methodology in cross-cultural consumer research: a review and critical assessment", *Journal of International Consumer Marketing*, Vol. 11 No. 4, pp. 75-96.
- Steenkamp, J-B.E.M. (2001), "The role of national culture in international marketing research", *International Marketing Review*, Vol. 18 No. 1, p. 30.
- Stewart, I. and Lacassagne, M.-F. (2005), "Social representations as a diagnostic tool for identifying cultural and other group differences", *Psychology & Marketing*, Vol. 22 No. 9, pp. 721-38.
- Usunier, J.-C. (1998), *International and Cross-Cultural Management Research*, Sage, London.
- Venkatesh, A. (1995), "Ethnoconsumerism: a new paradigm to study cultural and cross-cultural consumer behavior", in Costa, J.A. and Bamossy, G.J. (Eds), *Marketing in a Multicultural World: Ethnicity, Nationalism, and Cultural Identity*, Sage, Thousand Oaks, pp. 26-67.
- Vergès, P. (1987), "A social cognitive approach to economic representations", in Doise, W. and Moscovici, S. (Eds), *Current Issues in European Social Psychology*, Cambridge University Press, Cambridge, pp. 271-304.
- Vergès, P. (1996), "A method of analysis by questionnaires to know social representations of the economy", paper presented at the IAREP, Paris.
- Vergès, P. and Bastounis, M. (2001), "Towards the investigation of social representations of the economy: research methods and techniques", in Roland-Lévy, C., Kirchner, E., Penz, E. and Gray, C. (Eds), *Everyday Representations of the Economy*, WUV-Verlag, Wien, pp. 19-48.
- Wagner, W. (1989), "Social representation and beyond: brute facts, symbolic coping and domesticated worlds", *Culture & Psychology*, Vol. 4 No. 3, pp. 297-329.
- Wagner, W., Duveen, G., Farr, R., Jovchelovitch, S., Lorenzi-Cioldi, F., Marková, I. and Rose, D. (1999), "Theory and method of social representations", *Asian Journal of Social Psychology*, Vol. 2, pp. 95-125.
- Yeo, J. and Huang, W. (2003), "Mobile e-commerce outlook", *International Journal of Information Technology & Decision Making*, Vol. 2 No. 2, pp. 313-32.

Further reading

- Matiasko, W., Dobrov, I. and Bronner, R. (1994), "anwendung der korrespondenzanalyse in der imageforschung", *Marketing Zeitschrift für Forschung und Praxis (ZFP)*, Vol. 1, pp. 42-54.

Corresponding author

Elfriede Penz can be contacted at: elfriede.penz@wu-wien.ac.at



Employing information communication technologies to enhance qualitative international marketing enquiry

Sharon Loane and Jim Bell
University of Ulster, Northern Ireland, UK, and
Rod McNaughton
University of Waterloo, Waterloo, Canada

Abstract

Purpose – Extant international marketing enquiry has been widely criticised for lacking scope and ambition. Typically, empirical investigations have involved single market studies employing quantitative methods and survey techniques. Consequently, researchers have been challenged to embrace greater methodological pluralism and broaden their geographical perspectives. This contribution posits that new information communication technologies (ICT), particularly the internet, can significantly improve the robustness of qualitative and mixed-method international marketing research.

Design/methodology/approach – The paper describes and evaluates the application of ICT in a recent cross-national enquiry into rapidly internationalising small firms. Online sources were used to gather information on 218 internationalising small firms, in Australia, Canada, Ireland and New Zealand. An e-mail instrument was then administered to verify this data and address information gaps, resulting in 143 usable responses, evenly distributed across locations. Key emerging themes were identified and a representative sub-sample of 53 firms was selected for further in-depth investigation via face-to-face interviews with CEOs.

Findings – The authors contend that such technologies can help to refine sample identification and selection procedures, improve response rates and encourage greater respondent “buy-in” to depth interviews. They also lead to much more targeted lines of enquiry during depth interviews by identifying key research themes and issues, thus enhancing the depth and richness of the insights obtained.

Originality/value – The paper concludes that novel ICT-enabled research approaches as described herein are particularly effective because, compared to conventional survey methods, they are more user friendly and better received by subjects.

Keywords Internet, International marketing, Qualitative methods, Research methods

Paper type Research paper

Introduction

The nature and scope of international marketing enquiry has been widely criticised due to the absence of a tradition of building upon prior research and for a failure to extend its geographic scope in a rigorous and systematic manner (Cavusgil and Das, 1997; Nakata and Huang, 2005). A major contributory factor has been the dominant use of positivist methodologies and associated survey techniques that, it has been argued, have also restricted the theoretical development of the area (Kamath *et al.*, 1986; Daniels, 1991; Inkpen and Beamish, 1994; Sullivan, 1998). Consequently, there is a growing recognition among researchers of the need for greater methodological



pluralism and, in particular, of the value of qualitative approaches in gaining deeper and richer perspectives of the phenomena under investigation. Nevertheless, such approaches also present considerable challenges, not least in identifying and selecting suitable cases in a cross-national setting and in conducting complex qualitative enquiries in a rigorous and resource effective manner.

A focus of much recent international marketing enquiry is on the phenomenon of rapidly internationalising smaller firms (Moen and Servais, 2002; Bell *et al.*, 2003; Kundu and Katz, 2003; Sharma and Blomstermo, 2004, Chetty and Campbell-Hunt, 2004; Knight and Cavusgil, 2004; Rialp *et al.*, 2005) whose behaviour conflicts with the notion of incremental step-wise internationalisation so prevalent in the literature (Leonidou and Katsikeas, 1996). Known variously as “committed internationalists” (Bonaccorsi, 1992) “born-global” firms (McKinsey and Co., 1993) or “international new ventures” (McDougall *et al.*, 1994), these are entrepreneurial firms that have internationalised from inception, or shortly thereafter (Oviatt and McDougall, 1994). Their emergence has been influenced by the globalisation of business activity and the impact of new information and process technologies (Knight and Cavusgil, 1996). In addition, such firms often possess knowledge-based competitive advantages that enable them to offer value-added offerings (McKinsey and Co., 1993).

The overall objective of a recent study undertaken by the authors was to investigate the behaviour of rapidly internationalising small firms in Australia, Canada, Ireland and New Zealand. In adopting an essentially qualitative approach, involving case study methods and in-depth interviews, our objective was to obtain deeper insights into an emerging phenomenon by exploring factors that impacted upon firms’ internationalisation activities and behaviour. However, the main focus of the present paper is not on the results of the study, but rather on international marketing research method, centring on the potential of information communication technologies (ICT), particularly the internet, to enhance qualitative research methodologies.

While employing the internet as a research tool or using qualitative approaches to conduct international research are neither particularly novel ideas in themselves (Craig and Douglas, 2001; Hewson *et al.*, 2002; Martin and Matlay, 2003; Fletcher *et al.*, 2004), we consider that a useful contribution is made to method by combining tools and techniques in a manner not previously developed in the literature. First, we demonstrate how the internet provides a good basis for identifying and constructing multiple cases for further qualitative investigation. Second, we illustrate how an ICT-enabled approach improves both the quality and volume of data obtained. Third, we show how the method provides deeper, and richer perspectives in a broader cross-national context.

The structure of this contribution is as follows: Following a synthesis of the extant literature on international marketing research methodologies, we outline the research process adopted in the present study and describe the various stages involved. Thereafter, we discuss the merits and disadvantages of using ICT-enabled approaches to facilitate international enquiry. We conclude that the approaches described and evaluated can be usefully employed to significantly enhance qualitative international marketing research.

Review and synthesis of the literature

The “ethnocentric” nature of research and the relative absence of comparative enquiry is a recurring theme in the international marketing literature (Cavusgil and Das, 1997).

Aaby and Slater (1989) describe earlier contributions as “a mosaic of autonomous endeavours” which according to Reid (1983), are “too fragmentary and unprogrammable to yield coherent results”. Given these limitations, there have been growing calls for researchers to engage in greater cross-national enquiry and research collaboration that is meaningful in a global context (Sin *et al.*, 1999; Walters, 2001; Farley and Lehmann, 2001).

In terms of theoretical development, Albaumb and Peterson (1984) observe that, with few exceptions, early international marketing research is “fragmentary, generally atheoretic and not sufficiently programmatic to offer anything other than simplistic and incomplete insights into the underlying phenomena of interest”. Others, such as Bradley (1987) and Daniels (1991), articulate similar views. Kamath *et al.* (1986), contend that the dominant use of positivist methodologies has been a major contributory factor. They argue that differences in the context or foci of empirical investigations and the use of diverse definitions, measurement parameters or analytical techniques have resulted in conflicting and contradictory findings that have only served to retard the conceptual and theoretical advancement of the area.

Katsikeas (1991) and Leonidou (1995) argue that the main problems lie in the absence of a tradition of building upon prior research and in inadequate research designs that fail to conduct enquiries in a rigorous and systematic manner. However, other authors (Daniels, 1991; Inkpen and Beamish, 1994; Sullivan, 1998) contend that an (over) emphasis on quantitative methodologies “may unintentionally direct scholars toward undertaking studies that are tightly designed, exactly executed, but of limited relevance and scope” (Nakata and Huang, 2005). Thus, according to Sullivan (1998), while many enquiries have “demonstrated impeccable micro logic” they have, at the same time, avoided “genuine shifts in intellectual direction”. Indeed, more recently, Salzberger and Sinkovics (2006) question whether current paradigms regarding measurement allow to capture fundamental problems of international research altogether.

Clearly, many of these criticisms relate to earlier research in the field. However, a recent meta-analysis by Nakata and Huang (2005) of nearly 600 international marketing contributions between 1990 and 2000 finds that, the situation had not really changed significantly. First, over 70 per cent of all conceptual and empirical studies still involved quantitative research, with surveys having a “methodological monopoly” however, a slight trend towards qualitative approaches was observed between 1996 and 2000. Second, around half of all investigations were single country studies, although this proportion declined (44 per cent) between 1996 and 2000 as the number of comparative and global studies increased. Third, as observed by Walters (2001), the choice of research locations continued to be dominated by North America and Western Europe, notably the USA (64 per cent), and the UK (19 per cent). Consequently, other countries, particularly emerging economies, or other regions, notably Latin America and the Middle East, were badly under-represented. Fourth, that almost 80 per cent of studies involved authors from a single country, which tended to further limit the geographic scope of enquiries, although again, a significant positive trend towards greater cross-national collaboration was evidenced between 1996 and 2000.

Notwithstanding the limited development of the field, what has changed in the ensuing period is a growing acceptance among scholars of the value of qualitative research. This has been influenced by a number of factors, such as the need for a deeper understanding of “how” different markets work and “why” customers in

different countries behave in ways that they do (Ghauri and Gronhaug, 2002), a shift away from studies of large corporations towards enquiries into SMEs (Coviello and McAuley, 1999; Bell *et al.*, 2003; Sinkovics *et al.*, 2005) and advances in information technology (IT) that facilitate the collection of data on a global basis (Craig and Douglas, 2001; Wilson and Laskey, 2003).

In addition, the growing popularity of qualitative methodologies can be attributed to a number of factors (see Sinkovics *et al.* (2005), for a good discussion of these issues). First, information overload coupled with the lack of flexibility of quantitative techniques in attempting to make sense of larger volumes of data. Secondly, increased complexity, due to the fragmentation of markets and emergence of multiplex hyper-reality environments. Thirdly, greater capabilities and confidence of researchers in employing a growing array of qualitative tools and techniques, along with the increased sophistication of such tools due to technological advances. Further evidence of a growing enthusiasm for qualitative international research is the number of recent contributions to method (Marschan-Piekkari and Welch, 2004; Coviello and Jones, 2004; Coviello, 2005).

In this context, Craig and Douglas (2001) contend that 21st century researchers need to incorporate technological advances into research design and methodology to enable them to gather data faster, in a more systematic manner and over a broader geographic scope. They conclude that that as the internet evolves:

It offers the potential to dramatically change the way international marketing research is conducted, both in providing ready access to secondary data, and in providing new means of collecting primary data.

We would further argue that such advances must also be incorporated into the design and execution of qualitative international marketing enquiry.

Turning to the focus of the present research, as an emerging field of enquiry, the extant research on rapidly internationalising small firms is typically, exploratory and qualitative in nature, generally involving case study methods (Coviello and McAuley, 1999). Such approaches reflect the need to develop a solid theoretical base for an under-researched area in order to create theories that provide a greater understanding of the phenomenon, rather than to test them (Jones and Coviello, 2005). Moreover, as such research tends to centre on the behaviour of entrepreneurs and the impact of their personal and/or professional networks on firm strategy (Miesenbock, 1988; Nordstrom, 1990; Welch and Welch, 1996; Coviello and Munro, 1997), it is increasingly argued that qualitative techniques are more “fit” for these purposes (Carson *et al.*, 2000; Hill and McGowan, 2000).

However, a limitation of this emerging field of enquiry is that much of the research to date has been conducted in a single location and involved a relatively small number of cases. Indeed, a meta-analysis by Coviello and McAuley (1999) found that of 16 empirical studies they reviewed, only two, those of Bell (1995) and Gankema *et al.* (1997), were cross-national in nature. They concluded that, although context specific information is valuable, previous research has, perhaps, suffered from an overly myopic perspective and that this narrow focus needed to be urgently addressed.

Nevertheless, the reluctance to undertake larger scale cross-national investigations is understandable, due to the increased logistical complexity of managing the

process in multiple locations and consequent cost and resource implications. These considerations, present particular challenges for qualitative researchers, where an appropriate number of cases need to be identified and developed in each research location. Additionally, there is the added complexity of systematically analysing a larger number of cases and managing greater volumes of data, than is normally required in a single-country context. Then, there is also the need to compare and contrast associations in occurring patterns of response within each location and across different countries. Finally, great care must be exercised in the design and execution of qualitative international research to ensure cross-cultural equivalence (Sinkovics *et al.* (2005), for a comprehensive discussion of these issues).

For researchers engaged in small business enquiry these challenges have been exacerbated in the past by a lack of published information, as many small firms are not publicly listed. There is also a general absence of comprehensive small firm databases from which appropriate sampling frames can be constructed (Carson and Coviello, 1996). Finally, the typical reluctance of small firm decision-makers to respond to conventional survey methodologies means that alternative research approaches must be considered (Carson *et al.*, 2000; Coviello and Jones, 2004).

Given the preceding discussion, we contend that many of the existing criticisms of international marketing enquiry can be addressed by embracing more qualitative methodologies. Furthermore, that ICT, a key driver of rapid internationalisation, can provide researchers with new and innovative approaches to investigate the phenomenon and can be usefully employed to overcome some of the inherent difficulties of conducting cross-national qualitative research. A framework to guide such enquiry, including the present study, is shown in Figure 1.

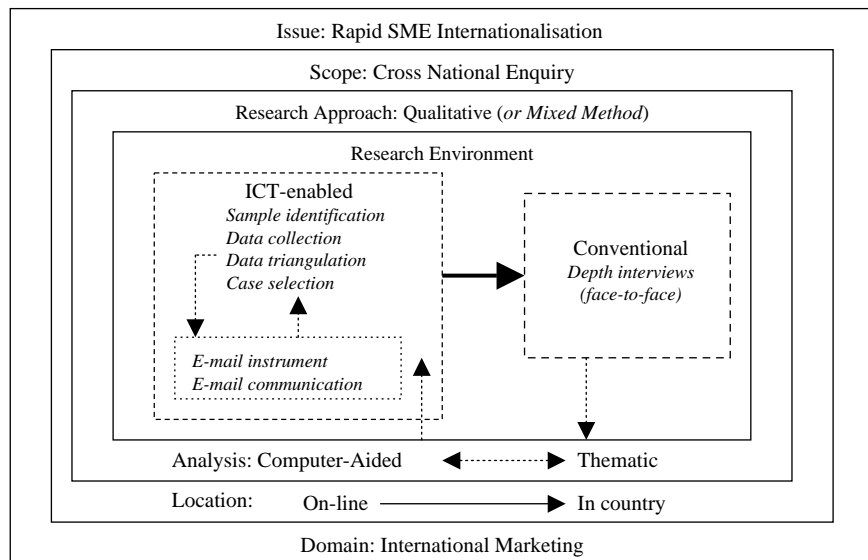


Figure 1.
Research framework

Research design

The present study is used to illustrate the application of this framework and outline the research method. It involved a multi-stage research design and adopted a predominantly qualitative approach. Having reviewed the export marketing and internationalisation literatures, the first stage was to identify a suitable sample of around sixty firms in each of four locations (Australia, Canada, Ireland and New Zealand) and develop cases on them from online sources, as a first step in obtaining sufficient cases for interview. In the second stage, telephone and/or e-mail contact was established with CEOs to address specific information gaps via a tailored e-mail instrument. A pre-designed template containing all available pertinent secondary data on the firm was employed. In the third stage, preliminary analysis of completed instrument was used to identify emerging themes and select a representative sub-sample of around 15 cases in each location for further qualitative investigation via in-depth personal interviews.

The choice of locations was based on the fact that these countries are relatively small open economies (i.e. second or third tier OECD countries) with a high dependence on export-led growth, an internationally active small business sector and comparably high levels of internet penetration. Another factor was the current and previous countries of residence of the co-authors, which provided pre-existing knowledge of the markets and good research networks within them (Buchanan *et al.*, 1988). The rationale for the selected approach was based on experimental research previously conducted among Irish firms, which indicated that:

- A large proportion of the information sought was already in the public domain, on firms' web sites and from other online or published sources.
- The approach yielded positive results in terms of obtaining further information from firms and also in improving response rates.
- It was also very beneficial as subjects were more amenable to being interviewed.

The overall research design is shown in Figure 2 and the procedures adopted are outlined and evaluated thereafter.

Identification and construction of cases

In the first stage of the research, online sources were used to construct a database of some 220 cases (approximately 50-60 in each location). These procedures involved the pre-screening of over 750 firms. Key criteria were that firms should be autonomous, employ fewer than 250 staff, be actively engaged in exporting and have an internet presence. Following McDougall and Oviatt(2003), firms were classified as "rapid internationalisers" if they had entered overseas markets within six years. Non-probability sampling was used to include these firms and exclude those that did not fit other criteria, such as firm size. A cross-sectional sample was sought, although it was anticipated that it would contain a higher proportion of firms from knowledge-based industry sectors.

Sources of information included small business databases, government export promotion organisation web sites, online trade directories, firms' own web sites and those of any identified clients. Following McGuinness and Short (1998) and Hewson *et al.* (2002), internet search engines were also trawled for media sources and profiles

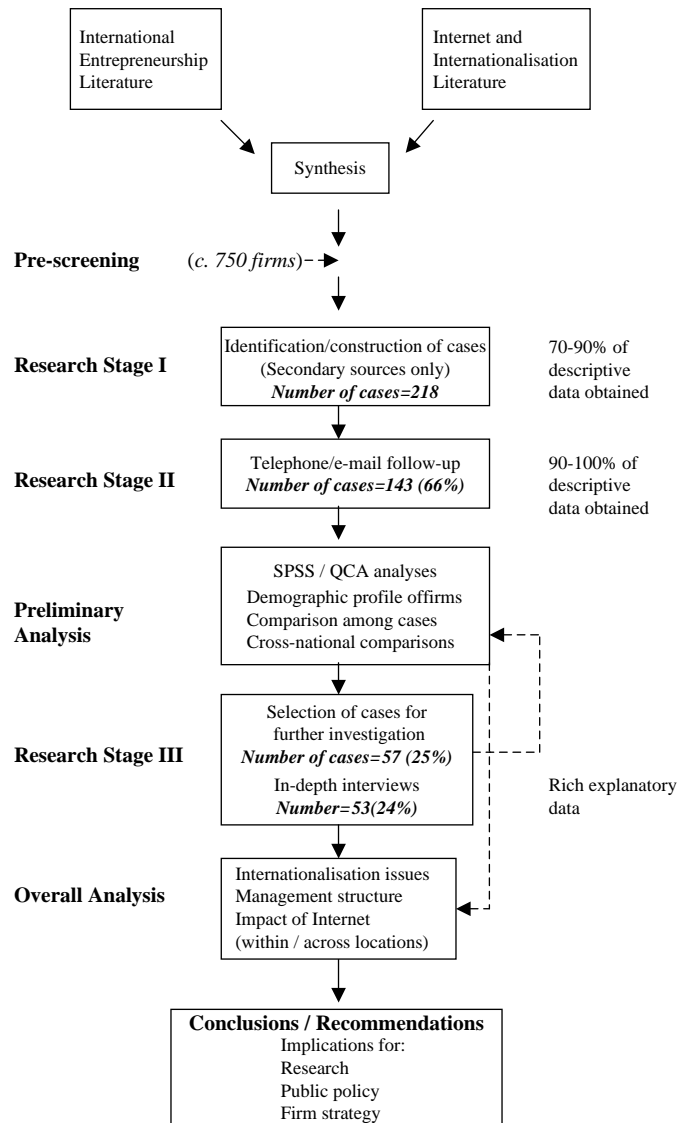


Figure 2.
Research design

of the founder/s or management teams. As the main focus of the enquiry was the rapid internationalisation of internet-enabled firms, the following information was sought[1]:

- firm demographics (age, size, turnover, etc.), history, funding structure;
- main activities, product/service offerings;
- background and prior experience of entrepreneur, or of founding members;
- current export markets/export ratio/market entry modes and pace to markets;

-
- target segments, major clients, partnerships/alliances, suppliers;
 - date first online, and level of internet-enablement; and
 - any other relevant information (e.g. awards received, venture capital (VC) invested).

The objective was to select a heterogeneous sample of rapidly internationalising firms in order to explore key themes from the literature. Given the intention to conduct in-depth interviews, non-proportional, purposive quota sampling (Strauss and Corbin, 1998; Sin *et al.*, 1999; Saunders *et al.*, 2003) was used to ensure that sufficient cases could be identified and clustered in high-density locations, such as Canada's technology triangle, to allow subsequent field research to be conducted efficiently.

All of the information obtained for each firm was summarised onto a pre-designed Company Information Sheet and cases were entered in a SPSS database. During the course of building these cases and prior to any direct contact with firms, various patterns began to emerge on availability of online data and common information gaps across countries. First, details of product offerings, key clients, channel partners or firms' management teams were generally comprehensive. Second, financial and employment data were not always available on firms' web sites, but could often be found elsewhere, for example, from government agency web sites. Notable exceptions were where firms had obtained VC and wanted to promote this fact. Third, while it was often evident from a firm's web site that it had internationalised rapidly, the particular sequence of entry into each market was not always explicit, although inferences could sometimes be drawn.

Nevertheless, 70 per cent or more of the descriptive data sought was readily obtained directly from firms' web sites. Searches of other sites often provided additional useful information and facilitated triangulation to corroborate data obtained from firms' sites. For example, where VC had been obtained, financial information could be found by accessing the VC firm's web sites. Patterns and pace of firm internationalisation could often be deduced by tracking the online business press, which also provided information on agent/distributor agreements, partnerships, contracts obtained or corporate "spin-outs". Sales revenues and employee numbers were sometimes obtained from these sources or government sites. If a firm had received an export or business award, press releases or media interviews provided good background details. For firms with entrepreneurial management teams, the specific capabilities and skills of individual members were often evident from profiles on the firm's web site, or from searches of personal home pages.

Telephone and e-mail instrument follow-up

In the second phase, telephone and e-mail contact was established with CEOs to fill in existing information gaps. The researchers were able to state that they already knew a lot about the firm and wished to ask some very targeted questions. The objectives of the study were explained and permission was sought to forward the Company Information Sheet and a covering letter to an e-mail address specified by the CEO. Respondents were asked to verify the accuracy of the summarised information, address highlighted gaps and add any contextual information pertinent to their internationalisation strategies.

These procedures led to 143 useable responses (66 per cent), fairly evenly distributed across locations (Table I). This return compares very favourably with

those reported by Dillman (2000) and Dillman *et al.* (2002). Another feature was speed of response, with over 30 per cent of firms replying within two weeks, supporting recent literature on the efficacy of e-mail and online surveys (McDonald and Adam, 2003). Following Armstrong and Overton (1977) procedures to estimate non-response bias, no statistically significant differences were found among early and late respondents or between non-responding and responding firms in terms of age, size, export intensity, or other characteristics.

Once the data from the e-mail survey were incorporated into the database, some 90-100 per cent of the descriptive information needs for the study were met. Typically, there was sufficient information to identify the patterns and pace of internationalisation. In some cases it was also possible to deduce, from information provided on channel partners or key clients, factors that had influenced particular market selection and entry decisions.

At this stage, an SPSS package was utilised to profile firms, establish broad patterns of response, develop descriptive statistics (in terms of age, size, export ratios, key markets, pace to markets, etc.) and analyse the relationships between these variables. Qualitative comparative analysis (QCA) techniques were also used to cluster cases with similar characteristics in terms of the firms demographic profiles (age, size, management structures, etc.) and internationalisation behaviour and compare them within and across research locations (for more detailed discussion of QCA methods, see Ragin and Zaret, 1983; Rihoux *et al.*, 2003).

These analyses revealed that over 70 per cent of firms were less than ten years old, had fewer than a hundred employees, internationalised within two years of inception and exported over half their output. In almost 20 per cent of cases, export ratios of 90-100 per cent indicated virtually no domestic sales. Over 60 per cent concentrated on key markets, operating in fewer than five countries and focusing on "lead" markets, such as the USA or Europe. No statistically significant differences were found between the age or size of firms and international performance, in terms of export ratios or speed to markets, nor were any discernible differences by country of origin or industry sector. However, a proportion of cases in each location displayed atypical international expansion strategies, in terms of internationalising in advance of gaining domestic market sales, entering psychologically and/or geographically "distant" markets before targeting countries with more obvious potential, or focusing on particular geographical regions. Motives for these behaviours were not immediately apparent and merited further investigation.

Location	Cases developed (<i>n</i>)	Responses to e-mail instrument (2nd stage) (<i>n</i>)	In-depth interviews conducted (3rd stage) (<i>n</i>)
Ireland	56	39 (70)	13 (23)
Canada	61	44 (72)	16 (26)
New Zealand	51	30 (59)	12 (24)
Australia	50	30 (60)	12 (24)
<i>Total no. of cases</i>	<i>218</i>	<i>143 (66)</i>	<i>53 (24)</i>

Table I.
Sample size, response rates and in-depth interviews

Note: Figures given in parenthesis are percentage

Analyses were also undertaken on managerial issues (such as the background and experience of key decision-maker/s, management structure, etc.). Statistically significant differences were found ($p = 0.05$) between firms that had lone founders and those with management team formations and international performance measures. Typically, firms founded by management teams had higher export ratios and internationalised more rapidly. As management teams, rather than lone entrepreneurs, were involved in the formation of over 70 per cent of cases, this became an important avenue of enquiry. QCA was used to further explore relationships between these management variables and rapid internationalisation, identify other main themes and select a representative sub-sample.

Follow up interviews

Based on these analyses, 57 cases (25 per cent) were chosen for in-depth study. In addition to being rapid internationalisers, as previously defined, key criteria were that export ratios should exceed 50 per cent of total sales and that the sub-sample selected should be representative of the overall sample in terms of firm size, management structure, industry sector and geographic location. For theoretical purposes, cases meeting these criteria and evidencing atypical internationalisation behaviour, or with team formations, were well represented to obtain greater insights into these phenomena and seek richer explanations.

As shown in Table I, 53 face-to-face interviews were conducted over a six-month period in early-mid 2003. It is noteworthy that only four of fifty-seven CEOs approached declined to be interviewed. In all but one case, the reason was that they were not in the country during the interview “window”. As discussed by Saunders *et al.* (2003) and Robson (2002), an important factor in such a high level of co-operation was the incremental rapport developed during the telephone and e-mail exchanges when the researcher’s credibility and the value of the research were established. Indeed, many CEOs commented that they appreciated the approach, as it did not impinge on their time, and were very amenable to being interviewed, with a number actually inviting a visit.

The purpose of these interviews was to seek motives for particular internationalisation behaviours, gain insights into strategy and explore issues relating to management team structures, individual roles and decision-making processes. All of these semi-structured interviews were recorded, transcribed and explored utilising qualitative techniques to undertake thematic analyses. Along with the cases constructed from online secondary, supplementary data obtained from the e-mail instrument, field notes and observation, they provided rich detail the phenomenon investigated. These interviews were especially insightful in terms of the impact of specific events on the internationalisation process.

One of the acknowledged benefits of a qualitative approach is that of the “discovery” of unanticipated phenomenon (Maxwell, 1997; Strauss and Corbin, 1998). As previously indicated, the analyses of the initial data collected from on line sources and the e-mail instrument pointed to the significance of management teams in rapid internationalisation. However, with the exception of the pioneering work of Reuber and Fischer (1997, 2002), this issue has been largely under-researched by international marketing scholars. Thus, it became an important focus of the in-depth interview phase, yielding particularly valuable perspectives on the roles and functions of team members. This avenue of enquiry also identified the specialist skills, contact networks

and market knowledge of individual team members, as well as their ability to leverage external sources of finance, in a manner that could not have been accomplished by means of a quantitative survey. In particular, it revealed that networking was occurring at micro, meso and macro levels, thus supporting the work of Ghauri and Holstius (1996). As a consequence, we contend that the results of our enquiry have the potential to make an important theoretical contribution to the field in terms of the influence, impact and role of management teams.

Overall research approach

Several aspects of the research implementation merit further elaboration. First, although the research stages are shown as sequential in Figure 1, many of the activities, especially in the later stages, were more concurrent than can be easily illustrated. Second, although the approach was essentially qualitative, mixed methods were employed in parallel to gain more holistic perspectives of the issues under investigation. Gaps in secondary information identified in the construction of cases were addressed via the e-mail survey. However, this phase also corroborated or amplified the data gathered from online sources. In turn, both of these phases and the QCA analyses informed and influenced particular lines of enquiry during the in-depth interviews. Third, as the research was “rolled out” perspectives gained from the later stages in one country often provided meaningful explanation for data obtained in the earlier stages in another.

Discussion

A number of observations can be made on issues relating to the methodological approach adopted in the study. As De Meur and Rihoux (2002) observe, a key issue in case study research is developing a comprehensive and intimate knowledge of the cases under investigation. The construction of cases from online sources provides very detailed information on both the focal firms and the CEO or management team. This obviates the need to survey firms for information that is already in the public domain and any gaps can be addressed via a tightly focused e-mail instrument. In the present enquiry, the Company Information Sheet developed for this purpose demonstrated that the researcher had already gathered a large amount of data on the firm, thereby showing serious intent. It clearly identified specific information gaps that respondents were asked to address. It also encouraged respondents to verify or amend the information already obtained and enabled the researcher to triangulate data from various secondary and primary sources.

Second, greater prior knowledge of cases enables qualitative researchers to identify an appropriate sub-sample more effectively and to pursue much more targeted lines of enquiry when conducting in-depth interviews. As Wass and Wells (1994) contend, greater prior knowledge allows researchers to probe for more accurate factual information and seek opinions more effectively during in-depth interviews. Moreover, the approach can be used to select cases either on a representative basis, or in order to construct a “theoretical” sample.

Third, the overall research design follows Jones (1985, p. 45) who observes that the in-depth interview provides:

... the means to understand why persons act as they do, and to understand the meaning and significance they give to their actions, in such a way that they can tell interviewer in their own terms.

Thus, we support Tull and Hawkins' (1990) conclusion that the in-depth interview is the most appropriate method to probe in any detail, attitudes and behaviours of individual owners/managers. This is particularly important in small business research, given the widely recognised influence of the key decision maker/s on the firm's strategic direction.

Fourth, and in relation to the use of the internet and e-mail instruments, we strongly concur with Craig and Douglas (2001) on the need to incorporate new technologies into research design and methodology. In the present enquiry, the use of internet sources to construct cases and the e-mail instrument employed to address information gaps and/or corroborate online information were both highly effective, as were e-mail interactions with CEOs. We would further argue that, in light of the resource constraints facing many researchers and the cost of conducting face-to-face interviews in multiple locations, even greater use could be made of allied ICT, for example, via the use of video-links or CUCMe cameras, to conduct in-depth interviews. Furthermore, as new ICT emerge, their potential to facilitate international marketing research must be recognised.

The evolution of the internet also provides greater opportunities to track firms over time and will open up exciting new possibilities for longitudinal research. We would consider this to be particularly important, for in addition to the failure to build on previous research identified by Katsikeas (1991) and Leonidou (1995), there is a marked absence of longitudinal international marketing enquiry. In the present study, a database containing comprehensive information on over 140 firms has been established. In addition fairly detailed, but as yet incomplete data is also available on a further 75 cases. By accessing firms' web sites and other online sources on a regular basis, this database is being continuously expanded and updated to enable researchers to monitor developments that impact on the firms' international behaviour and performance over time. It is also useful for tracking significant changes, such as when firms are acquired by third parties or cease trading, or when management team members exit or new ones join.

Finally, as already noted, evidence that CEOs appreciate this less time-consuming approach and that it is highly conducive to developing incremental rapport between researcher and respondent means that relationships established with key decision-makers offer ample opportunities for further enquiry on other international marketing issues via e-mail or telephone. It is interesting to report that some of the firms in the study now include the researchers on their distribution lists and send regular updates on recent developments and that some CEOs remain in regular e-mail contact.

Up to this point, discussion has focused on the benefits derived from such approaches (Table II). However, in the interest of providing an objective assessment, some of the limitations should also be noted. First, developing cases from secondary sources requires greater up-front investment in terms of research time and effort compared to conventional survey techniques. The latter often rely on established databases and assume that they are sufficiently accurate to allow the researcher to obtain a large enough valid response to manipulate the data.

Second, the method described precludes the inclusion of firms that do not have an online presence. Depending on the nature of the enquiry, this may lead to an unrepresentative sample if, for example, particular industry sectors have lower internet uptake than others. However, as previously noted, a web presence was a key

Table II.
Benefits and limitations
of the research approach

Research phase	Benefits	Limitations	Comments
Construction of cases from secondary sources	<p>Finely targeted sample</p> <p>Develops good pre-knowledge of firms and sample</p> <p>Large volume of information can be obtained</p>	<p>Limited to firms with web sites</p> <p>Web site language issues</p> <p>Information gaps on firm web sites</p> <p>Accuracy of web site information</p> <p>Greater up-front investment in research time and effort required</p> <p>More complex to manage than a conventional survey</p> <p>Cost of international phone calls (although this is falling)</p> <p>Working across time zones has advantages and disadvantages</p> <p>Reliability of information provided by source</p>	<p>May not be appropriate for all forms of IM enquiry</p> <p>The use of other online sources to address gaps is important</p> <p>Data from various sources can be triangulated</p> <p>More systematic than using conventional survey</p> <p>“User friendly” approach, appreciated by respondents, demonstrates “serious intent” on part of researcher/s</p> <p>Much faster response than mail survey</p> <p>Erroneous information can be amended/updated</p> <p>Higher respondent “buy-in” opportunities for further e-mail or telephone discussions</p> <p>Facilitates securing appointments for in-depth interviews</p> <p>Target high-density locations in countries</p> <p>Use of new ICT (e.g. web cameras) to conduct in-depth interviews will reduce research costs</p> <p>Offers great potential for longitudinal research and tracking of cases over time</p> <p>Offers opportunities for more “holistic” research perspectives</p>
Telephone/e-mail follow-up	<p>Very focused lines of enquiry address specific information gaps</p> <p>Higher response rates</p> <p>Speed of response</p> <p>Corroborates data already gathered</p> <p>Develops rapport with respondent</p> <p>Potential for on-going dialogue</p>		
In-depth interviews	<p>Preliminary analyses lead to more targeted lines of enquiry and provide richer and deeper insights into the phenomena under investigation</p>	<p>High costs of travel</p> <p>High investment of research time and effort</p>	
Overall research approach	<p>Enables collection of quantitative and qualitative data</p> <p>Helps to identify key themes/issues</p> <p>Facilitates cross-national enquiry</p>	<p>Systematic qualitative analysis is required, which takes greater time than quantitative analyses</p>	

selection criterion for cases in the present study. This problem is likely to diminish as more and more firms establish an internet presence, but may remain an issue due to variable internet uptake in different countries, or due to cultural factors that influence the extent to which firms are prepared to disclose information on their web sites (Fletcher *et al.*, 2004). It is also worth noting that the approach is currently more suited to business research as opposed to consumer studies, although the rapid growth of personal web sites may such enquiries more feasible in the future.

Third, the issue of the language used on firms' web sites is also a potential limitation. Given the countries selected, it did not present any difficulties in the present study, but could be problematic in enquiries focusing on other locations. Recent, as yet unpublished, exploratory studies by the authors, in collaboration with researchers from other countries (Brazil, Chile, Czech Republic and Poland) suggest that the approach described in this paper is "fit for the purpose" yielding similar outcomes in terms of the quantity and quality of online data available to construct cases, high response rates to the e-mail instrument and subject buy in for interviews. Thus, any limitations are likely to stem from researchers weak language capabilities, rather than from the research approach *per se*. In addition, we would anticipate that any difficulties of this nature are likely to decline in the future as a growing number of firms that wish to operate internationally develop multi-language web sites.

Finally, some of the data required for selection and classification purposes may be hard to obtain from firm web sites. Nevertheless, a great deal of this information can be found by comprehensive searches of other online sources. In this respect, it should be observed that published data for smaller firms that are not publicly listed is equally difficult to obtain by other means and that in any event such firms tend to be reluctant to divulge sensitive information in conventional surveys.

From the preceding discussion, it is clear that researchers need to take cognisance of the limitations of internet-enabled research approaches, regardless of whether their enquiries adopt quantitative, qualitative or pluralistic methodologies. However, we would contend that the advantages of employing ICT to enhance qualitative international marketing enquiry are significant, not just as a means of facilitating the research process, but also in terms of enabling researchers to gain deeper, richer insights on new phenomena in a multiplex and rapidly changing global environment. It is only by gaining such broader perspectives that scholars can aspire to influence "genuine shifts in intellectual direction".

Conclusion

This contribution outlines the qualitative research approach adopted in a recent enquiry into the rapid internationalisation of small firms. We conclude that the use of ICT to construct cases for further in-depth analysis offers significant advantages over more conventional research methods. These include improving sample identification and selection procedures, higher response rates and greater respondent "buy-in" better identification of key research themes leading to more targeted lines of in-depth enquiry and; enhancing the quality and richness of the qualitative data obtained.

While the focus of our own research is rapidly internationalising small firms, we would also contend that the approach could be usefully applied in other areas of international marketing enquiry and has wider application in international management research in general. Finally, we conclude that the emergence of new ICT and the growing volume of information available on the internet presents interesting possibilities to undertake greater cross-national enquiry than has been the case to date and to enhance

such studies, both in terms of data collection methods and comparative analyses. Such technologies can also facilitate the development of longitudinal studies that will, over time, offer richer insights into new and interesting phenomena.

Note

1. An interview guideline can be provided by the authors on request.

References

- Aaby, N. and Slater, S.F. (1989), "Management influences on export performance: a review of empirical literature 1978-88", *International Marketing Review*, Vol. 6 No. 4, pp. 7-26.
- Albaum, G. and Peterson, R.A. (1984), "Empirical research in international marketing: 1976-1982", *Journal of International Business Studies*, Vol. 15, pp. 161-73.
- Armstrong, J.S. and Overton, T.S. (1977), "Estimating nonresponse bias in mail surveys", *Journal of Marketing Research*, Vol. 14 No. 3, pp. 396-402.
- Bell, J. (1995), "The internationalisation of small computer software firms: a further challenge to 'stage' theories", *European Journal of Marketing*, Vol. 29 No. 8, pp. 59-74.
- Bell, J., McNaughton, R., Young, S. and Crick, D. (2003), "Towards an integrative model of small firm internationalisation", *Journal of International Entrepreneurship*, Vol. 1 No. 4, pp. 339-62.
- Bradley, M.F. (1987), "Nature and significance of international marketing: a review", *Journal of Business Research*, Vol. 15, pp. 205-19.
- Bonaccorsi, A. (1992), "On the relationship between firm size and export intensity", *Journal of International Business Studies*, Vol. 23 No. 4, pp. 605-35.
- Buchanan, D., Boddy, D. and McCalman, J. (1988), "Getting in, getting on, getting out, and getting back – the art of the possible", in Bryman, A. (Ed.), *Doing Research in Organizations*, Routledge, London, pp. 53-67.
- Carson, D. and Coviello, N.E. (1996), "Qualitative issues at the marketing/entrepreneurship interface", *Marketing Intelligence & Planning*, Vol. 14 No. 6, pp. 51-8.
- Carson, D., Gilmore, A., Perry, C. and Gronhaug, K. (2000), *Qualitative Marketing Research*, Sage, London.
- Cavusgil, S.T. and Das, A. (1997), "Methodological issues in empirical cross-cultural research: a survey of the management literature and a framework", *Management International Review*, Vol. 37, pp. 71-96.
- Chetty, S. and Campbell-Hunt, C. (2004), "A strategic approach to internationalization: a traditional versus a 'born-global' approach", *Journal of International Marketing*, Vol. 12 No. 1, pp. 57-81.
- Coviello, N. (2005), "Integrating qualitative and quantitative techniques in network analysis", *Qualitative Market Research An International Journal*, Vol. 8 No. 1, pp. 39-60.
- Coviello, N.E. and Jones, M.V. (2004), "Methodological issues in international entrepreneurship research", *Journal of Business Venturing*, Vol. 19, pp. 485-508.
- Coviello, N.E. and McAuley, A. (1999), "Internationalisation and the smaller firm: a review of contemporary empirical research", *Management International Review*, Vol. 39 No. 2, pp. 223-57.
- Coviello, N.E. and Munro, H. (1997), "Network relationships and the internationalisation process of small software firms", *International Business Review*, Vol. 6 No. 4, pp. 361-86.
- Craig, C.S. and Douglas, S.P. (2001), "Conducting international marketing research in the twenty-first century", *International Marketing Review*, Vol. 18 No. 1, pp. 80-90.

-
- Daniels, J.D. (1991), "Relevance in international business research, a need for more linkages", *Journal of International Business Studies*, Vol. 22, pp. 77-86.
- De Meur, G. and Rihoux, B. (2002), *L'Analyse Quali-Quantitative Comparee (AQQC-QCA). Approche, Techniques et Applications en Sciences Humaines*, Academia-Bruylant, Louvain-la-neuve, p. 35.
- Dillman, D.A. (2000), *Mail and Internet Surveys: The Tailored Design Method*, 2nd ed., Wiley, New York, NY.
- Dillman, D.A., Eltringe, J.L., Grove, J.L. and Little, R.J.A. (Eds) (2002), *Survey Non Response*, Wiley, New York, NY.
- Farley, J.U. and Lehmann, D.R. (2001), "The important role of meta-analysis in international research in marketing", *International Marketing Review*, Vol. 18 No. 1, pp. 70-9.
- Fletcher, R., Bell, J. and McNaughton, R. (2004), *International e-Marketing*, Thompson International Business Press, London.
- Gankema, H.G., Snuit, H.R. and van Dijken, K.A. (1997), "The internationalisation process of small and medium sized enterprises: an evaluation of the stage theory", in Donkels, R. and Miettinen, A. (Eds), *Entrepreneurship and SME Research: On Its Way to the Next Millennium*, Ashgate Publishing, Aldershot, pp. 185-97.
- Ghauri, P.N. and Gronhaug, K. (2002), *Research Methods in Business Studies: A Practical Guide*, 2nd ed., Financial Times Prentice-Hall, London.
- Ghauri, P.N. and Holstius, K. (1996), "The role of matching in the foreign market entry process in the Baltic states", *European Journal of Marketing*, Vol. 30 No. 2, pp. 75-88.
- Hewson, C., Yule, P., Laurent, D. and Vogel, K. (2002), *Internet Research Methods: A Practical Guide for the Social and Behavioural Sciences*, Sage, London.
- Hill, J. and McGowan, P. (2000), "Small business and enterprise development: questions about research methodology", *International Journal of Entrepreneurial Behaviour and Research*, Vol. 5 No. 1, pp. 5-18.
- Inkpen, A. and Beamish, P. (1994), "An analysis of twenty-five years of research in the JIBS", *Journal of International Business Studies*, Vol. 25, pp. 703-15.
- Jones, S. (1985), "Depth interviewing", *Applied Qualitative Research*, Gower, Aldershot, pp. 45-55.
- Jones, M.V. and Coviello, N.E. (2005), "Internationalization: conceptualizing an entrepreneurial process of behaviour in time", *Journal of International Business Studies*, Vol. 36 No. 3, pp. 284-303.
- Kamath, S., Rosson, P.J., Patton, D. and Brooks, M. (1986), "Research on success in exporting: past present and future, in rosson", in Rosson, P.J. and Reid, S.D. (Eds), *Managing Export Entry and Expansion*, Praeger, New York, NY, pp. 398-421.
- Katsikeas, C. (1991), "A review of the literature on export marketing problems", in Piercy, N. et al. (Eds), *Preparing Marketing for the New Millennium*, pp. 601-25.
- Knight, G. and Cavusgil, S.T. (1996), "The born global firm: a challenge to traditional internationalization theory", *Advances in International Marketing*, JAI Press, Greenwich, CT, pp. 11-26.
- Knight, G. and Cavusgil, S.T. (2004), "Innovation, organizational capabilities, and the born-global firm", *Journal of International Business Studies*, Vol. 35 No. 2, pp. 124-41.
- Kundu, S.K. and Katz, J. (2003), "Born-international SMEs: bi-level impacts of resources and intentions", *Small Business Economics*, Vol. 20, pp. 25-47.
- Leonidou, L.C. (1995), "Export barriers: non-exporters' perceptions", *International Marketing Review*, Vol. 12 No. 1, pp. 4-25.

- Leonidou, L.C. and Katsikeas, C.S. (1996), "The export development process: an integrative review of empirical models", *Journal of International Business Studies*, Vol. 27 No. 3, pp. 517-51.
- McDonald, H. and Adam, S. (2003), "A comparison of online and postal data collection methods in marketing research", *Marketing Intelligence & Planning*, Vol. 21 No. 2, pp. 85-95.
- McDougall, P.P. and Oviatt, B.M. (2003), "Some fundamental issues in international entrepreneurship", available at: www.usasbe.org/knowledge/whitepapers (accessed 30 June 2005).
- McDougall, P.P., Shane, S. and Oviatt, B.M. (1994), "Explaining the formation of international new ventures: the limits of theories from international business research", *Journal of Business Venturing*, Vol. 9, pp. 469-87.
- McGuinness, K. and Short, T. (1998), *Research on the Net*, Old Bailey Press, London.
- McKinsey and Co. (1993), *Emerging Exporters: Australia's High Value-Added Manufacturing Exporters*, Australian Manufacturing Council, Melbourne.
- Marschan-Piekkari, R. and Welch, C. (2004), "Qualitative research methods in international business: the state of the art", in Marschan-Piekkari, R. and Welch, R. (Eds), *Handbook of Qualitative Research Methods for International Business*, Edward Elgar, Cheltenham, pp. 5-24.
- Martin, L.M. and Matlay, H. (2003), "Innovative use of the internet in established small firms: the impact of knowledge management and organisational learning in accessing new opportunities", *Qualitative Market Research*, Vol. 6 No. 1, pp. 18-26.
- Maxwell, J.A. (1997), "Designing a qualitative study", in Bickman, L. and Rog, D.J. (Eds), *Handbook of Applied Social Research Methods*, Sage, London, pp. 69-100.
- Miesenbock, K.J. (1988), "Small business and exporting: a literature review", *International Small Business Journal*, Vol. 6 No. 2, pp. 42-61.
- Moen, O. and Servais, P. (2002), "Born global or gradual global? Examining the export behavior of small and medium-sized enterprises", *Journal of International Marketing*, Vol. 10 No. 3, pp. 49-72.
- Nakata, C. and Huang, Y. (2005), "Progress and promise: the last decade of international marketing research", *Journal of Business Research*, Vol. 58, pp. 611-8.
- Nordstrom, K.A. (1990), *The Internationalization Process of the Firm – Searching for New Patterns and Explanations*, Stockholm School of Economics, Stockholm.
- Oviatt, B.M. and McDougall, P.P. (1994), "Toward a theory of international new ventures", *Journal of International Business Studies*, Vol. 25 No. 1, pp. 45-64.
- Ragin, C. and Zaret, D. (1983), "Theory and method in comparative research: two strategies", *Social Forces*, Vol. 61, pp. 731-54.
- Reid, S.D. (1983), "Managerial and firm influences on export behavior", *Journal of the Academy of Marketing Science*, Vol. 11 No. 3, pp. 323-32.
- Reuber, A.R. and Fischer, E. (1997), "The influence of the management team's international experience on the internationalization behaviours of SMEs", *Journal of International Business Studies*, Vol. 28 No. 4, pp. 807-25.
- Reuber, A.R. and Fischer, E. (2002), "Foreign sales and small firm growth: the moderating role of the management team", *Entrepreneurship Theory & Practice*, Vol. 27, pp. 29-46.
- Rialp, A., Rialp, J. and Knight, G.A. (2005), "The phenomenon of early internationalizing firms: what do we know after a decade (1993-2003) of scientific inquiry?", *International Business Review*, Vol. 14 No. 2, pp. 147-66.
- Rihoux, B., De Meur, G., Yamasaki, S. and Ronsse, S. (2005), "An inventory of good QCA practice", available at: <http://smalln.spri.ucl.ac.be/Didactics1GB.htm#1bis> (accessed 12 September 2005).

-
- Robson, C. (2002), *Real World Research*, 2nd ed., Blackwell, Oxford.
- Saunders, M.N.K., Lewis, P. and Thornhill, A. (2003), *Research Methods for Business Students*, 3rd ed., Financial Times Prentice Hall, Harlow.
- Salzberger, T. and Sinkovics, R.R. (2006), "Reconsidering the problem of data equivalence in international marketing research – contrasting approaches based on CFA and the Rasch model for measurement", *International Marketing Review*, Vol. 23 No. 4, pp. 390-417.
- Sharma, D.D. and Blomstermo, A. (2003), "The internationalization process of born globals: a network view", *International Business Review*, Vol. 12 No. 6, pp. 739-53.
- Sin, L.Y.M., Cheung, G.W.H. and Lee, R. (1999), "Methodology in cross-cultural consumer research: a review and critical assessment", *Journal of International Consumer Marketing*, Vol. 11, pp. 75-96.
- Sinkovics, R.R., Penz, E. and Ghauri, P.N. (2005), "Analysing textual data in international marketing research", *Qualitative Marketing Research: An International Journal*, Vol. 8 No. 1, pp. 9-32.
- Strauss, A.L. and Corbin, J.M. (1998), *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Sage, Thousand Oaks, CA.
- Sullivan, D.P. (1998), "Cognitive tendencies in international business research: implications to the matter of a narrow vision", *Journal of International Business Studies*, Vol. 29, pp. 837-63.
- Tull, D.S. and Hawkins, D. (1990), *Marketing Research, Measurement and Method*, 5th ed., Macmillan, New York, NY.
- Walters, P.G. (2001), "Research at the 'margin' challenges for scholars working outside the 'American-European' domain", *International Marketing Review*, Vol. 18 No. 5, pp. 468-73.
- Wass, V.J. and Wells, P.E. (1994), "Research methods in action", in Wass, V.J. and Wells, P.E. (Eds), *Principles and Practices in Business Management Research*, Dartmouth, Aldershot, pp. 1-34.
- Welch, D.E. and Welch, L.S. (1996), "The internationalization process and networks: a strategic management perspective", *Journal of International Marketing*, Vol. 4 No. 3, pp. 11-28.
- Wilson, A. and Laskey, N. (2003), "Internet based marketing research: a serious alternative to traditional research methods", *Marketing Intelligence & Planning*, Vol. 21 No. 2, pp. 79-84.

Further reading

- Blomstermo, A., Eriksson, K. and Sharma, D.D. (2004), "Domestic activity and knowledge development in the internationalization process of firms", *Journal of International Entrepreneurship*, Vol. 2 No. 3, pp. 239-58.
- Gabrielsson, M. and Kirpalani, V.H. (2004), "Born globals: how to reach new business space rapidly", *International Business Review*, Vol. 13 No. 5, pp. 555-71.
- Hashai, N. and Almor, T. (2004), "Gradually internationalizing 'born global' firms: an oxymoron?", *International Business Review*, Vol. 13 No. 4, pp. 465-83.

Corresponding author

Sharon Loane can be contacted at: sp.loane@ulster.ac.uk