International market segmentation: issues and perspectives

Jan-Benedict E.M. Steenkamp\textsuperscript{a, *}, Frenkel Ter Hofstede\textsuperscript{b}

\textsuperscript{a}Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
\textsuperscript{b}Carnegie Mellon University, Graduate School of Industrial Administration, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Abstract

With the increasing globalization of the business world, international segmentation becomes an ever more important concept in marketing. The globalization forces now at work push many companies to extend or reorganize their marketing strategies across borders and target international segments of consumers. It is the purpose of this paper to review the international market segmentation literature and to identify its future prospects and threats. We critically assess the current status of international market segmentation research and provide a systematic overview of 25 previous empirical studies with respect to the samples used for segmentation, segmentation bases and methods, geographic configuration of segments, and validation efforts. We discuss a number of conceptual and methodological issues that deserve more attention if international market segmentation is to fulfill its high potential. The conceptual issues include construct equivalence of the segmentation basis used, level of aggregation in the segmentation process, and choice of the segmentation basis. The methodological issues include measure equivalence and sample equivalence of the segmentation basis, segmentation methods employed, and whether national sample sizes should be proportional to population sizes. We describe a case study to illustrate and integrate the various issues and conclude with suggestions for future research to stimulate further advances in the area.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: International market segmentation; International marketing; Segmentation

1. Introduction

International market segmentation has become an important issue in developing, positioning, and selling products across national borders. It helps companies to target potential customers at the international-segment level and to obtain an appropriate positioning across borders. A key challenge for companies is to effectively deal with the structure of heterogeneity in consumer needs and wants across borders and to target segments of consumers in different countries. These segments reflect geographic groupings or groups of individuals and consist of potential consumers who are likely to exhibit similar responses to marketing efforts.

A natural form of international segmentation is to adopt a \textit{multi-domestic} strategy where each country represents a separate segment (Jeannet & Hennessey, 1998). A multi-domestic strategy amounts to selection of countries on the basis of their local advantages. Traditionally, multinational companies implemented such multi-domestic strategies by tailoring national brands to the needs shared by groups of consumers...
in the same country. In such an approach no coordination is required between countries, products are produced locally and are tailored to satisfy local needs. Distinct advertising, distribution, and pricing strategies are developed for targeting consumers in each country, and competition is managed at a national level. Competitive moves are conducted on a country-to-country basis and do not take the developments in other countries into account. Thus, in segmenting their markets, firms operating according to a multi-domestic approach can suffice with the standard segmentation techniques that are developed for domestic markets (Jeannet & Hennessey, 1998).

International segmentation becomes a particularly challenging issue when companies adopt a global or pan-regional strategy, that is, a strategy integrated across national borders. In many industries, national borders are becoming less and less important as an organizing principle for international activities, rendering multi-domestic strategies less relevant (Yip, 1995). Developments accelerating this trend include regional unification, shifts to open economies, global investment, manufacturing, and production strategies, expansion of world travel, rapid increase in education, literacy levels, and urbanization among developing countries, convergence of purchasing power, lifestyles and tastes, advances in information and communication technologies, the emergence of global media, and the increasing flow of information, labor, money, and technology across borders (Gielens & Dekimpe, 2001; Hassan & Katsanis, 1994; Hassan & Kaynak, 1994; Parker & Tavassoli, 2000; Yip, 1995). Many global companies such as Coca-Cola, McDonald’s, Sony, British Airways, Ikea, Toyota, and Levi-Strauss have successfully integrated their international strategies. The forces that are now at work drive many companies to extend their operations abroad and target international market segments. By globalizing their strategies, such companies benefit from several advantages, including cost reductions through economies of scale, improved quality of products, and increased bargaining and competitive power (Levitt, 1983; Yip, 1995).

Still, companies cannot serve the entire heterogeneous population of (a region of) the world with fully standardized marketing strategies. Many companies recognize that groups of consumers in different countries often have more in common with one another than with other consumers in the same country. Hence, they choose to serve segments that transcend national borders (Hassan & Katsanis, 1994). International segmentation aids the firm in structuring the heterogeneity that exists among consumers and nations and helps to identify segments that can be targeted in an effective and efficient way. As argued by Walters (1997, pp. 165–166):

When significant heterogeneity characterizes the international market context, tools are needed which can assist in the identification of underlying patterns of similarity which can provide a platform for global integration at the strategic and operational levels. The segmentation construct offers great promise in this respect. ... Segmentation is therefore particularly important in enterprises that wish to develop and implement successful global marketing strategies.

International segmentation offers a solution to the standardization versus adaptation debate in that it creates the conceptual framework for offering products and/or marketing programs that are standardized across countries by targeting the same consumer segment(s) in different countries (Verhage, Dahringer, & Cundiff, 1989). When using a similar marketing strategy in multiple countries, economies of scale will lead to a reduction in the average costs of production, advertising, and distribution. If, at the same time, consumers in the targeted segments share the same needs, such strategies can also be highly effective (Yip, 1995). Hence, international segmentation combines the benefits of standardization (e.g., lower costs, better quality) with the benefits of adaptation (e.g., close to needs of consumers).

Despite the obvious importance of international market segmentation for marketing as a discipline in general and international marketing in particular, it has received relatively little attention in the literature. In a review of about 900 articles on international marketing, it was found that just over 1% (11 papers) dealt directly with international market segmentation (Aulakh & Kotabe, 1993). A similar observation applies to international versus domestic market segmentation: “Segmentation is a central issue in domestic marketing strategy. Yet, in international markets, it has received little attention” (Douglas & Craig, 1992, p. 312). One
exception is the review by Walters (1997), which focuses especially on the process and international business aspects of international segmentation. We expand upon this work by providing an in-depth treatment of the key conceptual and methodological issues one should address when doing international segmentation research. Our point of departure is a systematic overview of previous empirical international segmentation studies. These studies are subsequently related to the various issues raised in international market segmentation. These conceptual and methodological issues deserve more attention if international market segmentation is to fulfill its potential. We bring these issues together and apply them in an illustrative case study involving an international segmentation challenge faced by one of the largest global consumer packaged goods companies. We conclude with suggestions for future research on international market segmentation.

2. Previous empirical research in international segmentation

Table 1 provides an overview of the key features of 25 international market segmentation studies. For each study, we provide information on the sample, segmentation basis, segmentation method, results, and validation, and briefly comment on some notable features of the study. In Fig. 1 the same studies are classified according to (1) their geographic coverage (whether non-Triad countries were included in the study or not), (2) the level of the segmentation basis used (country, region, or consumer), and (3) the nature of the study (exploratory versus model-based). Although Table 1 and Fig. 1 are more or less self-explanatory, several comments are in order.

First, whereas the total number of studies on international market segmentation has been limited up to 1990, it is encouraging to see that the interest in international segmentation increases over time, as evidenced by the relative large number of studies that were published after 1990 (14 out of 25 studies).

Second, international market segmentation research is truly international in that many different countries from all (inhabited) continents are included. Although there is a bias toward Triad countries (see Fig. 1), it is encouraging to see that research has also been conducted in other parts of the world. Hofstede’s (1980) seminal study may have served as exemplar.

Third, the studies have used a wide variety of segmentation bases and segmentation methods. Some studies used information on countries (or regions within countries), aggregated across consumers to the country (or region) level or information pertaining to the countries (or regions) themselves (e.g., climate, legal regime). Other studies used disaggregate, individual-specific information of consumers. The country-level segmentation bases typically included a combination of economic, political, geographic and demographic information (e.g., Helsen, Jedidi, & DeSarbo, 1993; Huszagh, Fox, & Day, 1986) or cultural variables (e.g., Hofstede, 1980; Sirota & Greenwood, 1971; Steenkamp, 2001). The set of individual-level segmentation bases is more heterogeneous, ranging from domain-specific characteristics such as attribute evaluations (Moskowitz & Rabino, 1994), attitudes (Verhage et al., 1989), and risk and brand loyalty ratings (Yavas, Verhage, & Green, 1992) to intermediate constructs such as means-end chains (Ter Hofstede, Steenkamp, & Wedel, 1999) to general characteristics such as values (Kamakura, Novak, Steenkamp, & Verhallen, 1993). The diversity of segmentation bases attests to the versatility of international market segmentation, but there is an overrepresentation of studies using characteristics of countries. Only eight studies used responses from individual consumers as compared to fifteen using country-level bases. Information on countries often is acquired relatively easily through published secondary data sources (e.g., publications of Euromonitor, the United Nations, and the World Bank), but the results of such country clusters do not always provide relevant information for managerial decisions (Helsen et al., 1993; Nachum, 1994). Presumably, the lack of available cross-national consumer data and the difficulty and cost involved in its collection, relative to the avail-

---

1 The segmentation basis is the set of criteria used to group consumers into international segments.

2 Following Ohmae (1985), a distinction is made between Triad and non-Triad countries. Collectively, the Triad regions produce the overwhelming share of the world’s GDP, and participation in the Triad regions is generally recommended (Jeannet & Hennessy, 1998). This renders market segmentation focusing on these regions especially relevant.
Table 1
Overview of previous empirical international market segmentation studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Segmentation basis</th>
<th>Method</th>
<th>Results</th>
<th>Validation</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Askegaard and Madsen (1998)</td>
<td>20,000 respondents from 79 regions from 15 European countries</td>
<td>average response per region to 138 food-related items</td>
<td>(i) factor analysis on region-level responses (ii) cluster analysis on region scores on 41 factors extracted</td>
<td>12 geographically contiguous segments following language borders</td>
<td>several factor analytic and cluster methods yielded largely the same results</td>
</tr>
<tr>
<td>2</td>
<td>Boote (1983)</td>
<td>899 respondents from UK, Germany, France</td>
<td>29 psychographic statements</td>
<td>Q factor analysis on individual responses</td>
<td>4 segments of which 2 have substantial membership in all 3 countries</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>Dawar and Parker (1994)</td>
<td>35 countries covering all continents</td>
<td>% of labor force engaged in retail sector</td>
<td>cluster analysis</td>
<td>2 segments, high versus low engagement</td>
<td>no</td>
</tr>
<tr>
<td>4</td>
<td>Day et al. (1988)</td>
<td>96 countries covering all continents</td>
<td>18 country characteristics reflecting economic development</td>
<td>(i) factor analysis on country characteristics (ii) cluster analysis on country scores on three factors extracted</td>
<td>6 geographically dispersed segments</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Helsen et al. (1993)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>12 Triad countries</td>
<td>23 country characteristics reflecting economic development</td>
<td>(i) factor analysis on country characteristics</td>
<td>2 and 3 segment solutions examined; both solutions contained a large segment of most West European countries</td>
<td>little similarity with product-specific segmentation (see below)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Triad countries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>annual unit sales data per country for TV, VCR, and CD-players for 14 years</td>
<td>latent class estimation of the Bass model per product</td>
<td>3 segments (TV, VCR) and two segments (CD-player), geographically dispersed</td>
<td>little similarity with sociodemographic segmentation (see above)</td>
<td>large differences among the 3 segmentation solutions indicate limitations of country-based segmentation using product-specific bases</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>315 managers from 14 countries covering four continents</td>
<td>12 scales for personal and interpersonal values</td>
<td>Q factor analysis on country average scores on the scales</td>
<td>5 segments, of which two were contiguous</td>
<td>substantial agreement in within-segment value profiles with student data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>French (German)-speaking Switzerland was assigned to Latin (Germanic) segment, providing evidence for within-country cultural value heterogeneity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>88,000 employees from 40 countries covering all continents</td>
<td>varying sets of items related to work goals</td>
<td>(i) theoretical reasoning and factor analysis on country average item scores</td>
<td>8 segments, most of which tended to be geographically close</td>
<td>results partially replicated with SSA</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(ii) cluster analysis on country scores on the four dimensions retained followed by judgmental regroupings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Segmentation basis</th>
<th>Method</th>
<th>Results</th>
<th>Validation</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 Huszagh et al. (1986)</td>
<td>21 Triad countries</td>
<td>9 country characteristics reflecting economic development</td>
<td>cluster analysis on country scores</td>
<td>5 segments of which one was geographically contiguous</td>
<td>largely same results were obtained for different clustering algorithms and for slightly varying sets of characteristics</td>
<td>high variation in product acceptance rates across 27 product categories for 5 European countries assigned to same segments sheds doubt on usefulness of segmentation basis only one segment the same as in Hofstede (1980); results also very different from other related studies (Ronen &amp; Shenkar, 1985)</td>
</tr>
<tr>
<td>10 Kale (1995)</td>
<td>17 Western European countries</td>
<td>Hofstede's (1980) national-cultural dimensions</td>
<td>cluster analysis on country ratings on the Hofstede dimensions</td>
<td>3 segments, none of them completely contiguous</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>11 Kamakura et al. (1993)</td>
<td>1573 consumers from Italy, West Germany, UK</td>
<td>Rokeach instrumental and terminal values</td>
<td>clusterwise rank logit on individual-level rankings of the top-nine instrumental and terminal values</td>
<td>5 segments, all showing a strong national orientation</td>
<td>including segment information over country information leads to significant improvement in $R^2$ for 40 out of 59 psychographic and benefit items</td>
<td>incomplete rankings reduced discriminability of segments</td>
</tr>
<tr>
<td>12 Kumar et al. (1998)</td>
<td>14 Western European countries</td>
<td>annual unit sales data per country for 5 consumer durables for 1970s – 1990</td>
<td>cluster analysis on the Bass coefficients of innovation and imitation per product</td>
<td>3 segments per product; segment composition not consistent across countries</td>
<td>no correspondence with sociodemographic segmentation of Helsen et al. (1993)</td>
<td>cluster solutions indicate that countries tend to group based on time of introduction, geographical proximity, and cultural/economic similarity</td>
</tr>
<tr>
<td>13 Kumar et al. (1994)</td>
<td>11 Western European countries</td>
<td>market potential of each country on 6 criteria in 8 different machinery industries</td>
<td>interactive multicriteria approach for solving problems of multobjective decision models using Augmented Weighted Tchebycheff procedure (i) stepwise regression of national innovativeness on 10 country characteristics (ii) cluster analysis on country scores on four significant characteristics</td>
<td>different segments of countries industries are found dependent on tradeoffs between criteria</td>
<td>multicriteria approach superior to personal evaluation by managers</td>
<td>the only study incorporating international market segmentation in decision support systems</td>
</tr>
<tr>
<td>14 Lee (1990)</td>
<td>70 countries covering all continents</td>
<td>4 country characteristics reflecting economic development</td>
<td>(i) stepwise regression of national innovativeness on 10 country characteristics (ii) cluster analysis on country scores on four significant characteristics</td>
<td>5 segments each containing countries from at least 2 continents</td>
<td>no</td>
<td>segments interpreted using Rogers' (1995) scheme but no scores on national innovativeness provided</td>
</tr>
<tr>
<td>Study</td>
<td>Methodology</td>
<td>Sample Size</td>
<td>Findings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moskowitz and Rabino (1994)</td>
<td>(i) factor analysis on respondents’ ideal points on sensory attributes (ii) cluster analysis on respondents’ factor scores</td>
<td>808 consumers from 4 (unidentified) countries 10 sensory attributes of flavored soda</td>
<td>3 cross-national segments no generate larger variation in optimal product formulation than countries choice of segments based on a priori notions and not strongly supported by actual results</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ronen and Kraut (1977)</td>
<td>(i) factor analysis on respondents’ factor scores smallest space analysis (SSA) on rank order correlations across the average scores on the 22 items for each pair of countries and visual clustering on SSA locations</td>
<td>4000 technicians from 14 mostly European countries 22 items related to work goals</td>
<td>4 segments, of which 2 are contiguous no</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sethi (1971)</td>
<td>(i) factor analysis on country characteristics reflecting economic development (ii) cluster analysis on country scores on four factors extracted</td>
<td>91 countries covering all continents 29 country characteristics reflecting economic development</td>
<td>7 geographically dispersed segments no segments varied much in number of countries; USA was outlier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sirota and Greenwood (1971)</td>
<td>Q factor analysis on country average scores on the items</td>
<td>about 13,000 employees from 25 countries covering all continents 14 items related to work goals</td>
<td>6 segments, of which 4 are contiguous Ronen and Kraut (1977) reanalyzed the data with SSA and visual clustering and many differences were found no segment of 6 ‘independents’ seems rest category indicating that about 25% of countries could not be properly assigned only study using Schwartz’s theory-derived framework of national culture</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steenkamp (2001)</td>
<td>(i) factor analysis on country scores on 11 dimensions (ii) two-stage cluster analysis on country scores on 4 factors extracted</td>
<td>24 countries covering 5 continents Hofstede’s (1980) 4 and Schwartz’s (1994) 7 national-cultural dimensions</td>
<td>7 segments, 4 of which geographically largely contiguous no</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Ter Hofstede et al. (1999)           | mixture model incorporating within- and between-country heterogeneity in response behavior using country membership as concomitant variables | 2961 consumers from 11 EU countries means–end chain data on yogurt | 4 cross-national segments, including one pan-European segment assessment of predictive validity; comparison with K-means clustering high predictive validity; model outperforms cluster analysis; significant differences in response tendencies found within and between countries
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Segmentation basis</th>
<th>Method</th>
<th>Results</th>
<th>Validation</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>21 Ter Hofstede et al. (2002)</td>
<td>1966 consumers from 120 regions in 7 EU countries</td>
<td>importances of 6 image attributes for outlets selling meat within each region</td>
<td>constrained hierarchical Bayes model based on similarity in inferred image importances of consumers in different regions identifying contiguous geographic segments</td>
<td>5 contiguous segments of which 3 are cross-national and 2 largely country-specific</td>
<td>comparison with unrestricted segmentation and country segmentation indicated strong performance of model</td>
<td>model accommodates a broad set of constraints on segments, which allows managerial objectives to be included in the segmentation process</td>
</tr>
<tr>
<td>22 Vandermerwe and L’Huillier (1989)</td>
<td>173 regions from 18 European countries</td>
<td>5 geographic and sociodemographic region-level characteristics</td>
<td>cluster analysis on region scores</td>
<td>6 geographically contiguous segments transcending national borders</td>
<td>no</td>
<td>choice of region-level variables biases results to contiguous segments; variables range from nominal to ratio level creating problems in cluster analysis screen test indicates 2 segments: segments 2 and 3 not well separated on attitudes</td>
</tr>
<tr>
<td>23 Verhage et al. (1989)</td>
<td>571 consumers from France, Netherlands, Norway, USA</td>
<td>14 attitudes toward energy conservation</td>
<td>cluster analysis on individual-level ratings</td>
<td>3 cross-national segments</td>
<td>$Q$ factor analysis and cluster analysis indicated same number of clusters; significant differences between segments for 9 out of 18 conservation behaviors</td>
<td></td>
</tr>
<tr>
<td>24 Wedel et al. (1998)</td>
<td>3900 consumers from 6 EU countries</td>
<td>rank order of 9 LOV values</td>
<td>mixture model which accounts for different sampling designs, including stratification by country cluster analysis on individual-level ratings</td>
<td>5 cross-national segments including 2 pan-European and one country-specific segment</td>
<td>comparison with mixture model which ignores complex sampling design</td>
<td>accounting for complex sampling design leads to different and better interpretable results</td>
</tr>
<tr>
<td>25 Yavas et al. (1992)</td>
<td>781 consumers from 6 countries covering three continents</td>
<td>risk and brand loyalty ratings for bath soap and toothpaste</td>
<td>cluster analysis on individual-level ratings</td>
<td>4 cross-national segments</td>
<td>no</td>
<td>conceptual relation between the products used for segmentation not clear</td>
</tr>
</tbody>
</table>
ability of information about countries, has discouraged researchers to study international consumer segments.

Fourth, it should be noted that most studies are of a rather exploratory nature, especially the studies that segmented countries. As shown in Fig. 1, few studies test specific hypotheses or base their work on a specific conceptual model (e.g., Helsen et al., 1993; Kumar, Ganesh, & Echambadi, 1998; Ter Hofstede et al., 1999). The exploratory nature of the studies limits the generalizability of specific findings. In Section 4.3, we will elaborate on methodological developments relevant for international segmentation.

Fifth, the studies involving responses from individual consumers have been conducted mainly in Triad countries. One exception is the study of Yavas et al. (1992), which also included consumers from Thailand and Saudi Arabia. A reason for the focus on Triad countries is their economic importance. In addition, the market research agency infrastructure in Triad countries is better developed and respondents may be more familiar with the instruments (e.g., rating scales) often used in such studies. Advanced marketing research is feasible, however, among less favored populations provided appropriate safeguards are taken such as adaptation of the measurement instrument (Steenkamp & Burgess, 2002).

Sixth, despite the growing importance of marketing decision support systems (Van Bruggen & Wierenga, 2000), and the important role market segmentation can play in such systems, only one study has integrated
international market segmentation in a decision support system (Kumar, Stam, & Joachimsthaler, 1994). They combined information on product-country markets with decision-makers’ objectives, constraints, and strategies, to identify the most attractive product-country segments through an interactive process.

Seventh, little attention is given to the validation of the solution. Eleven studies conducted no validation effort while the validation in most other studies was largely limited to an informal comparison of results obtained by different estimation methods (e.g., Askegaard & Madsen, 1998; Huszagh et al., 1986). One of the exceptions is Kamakura et al. (1993). They explicitly tested the statistical significance of the increase in explained variance in 59 psychographic and benefit items when international segment information was added to country information. If the international segments add insights over and above country member-

---

**Fig. 2. Issues in International Market Segmentation.**

### Conceptual issues

<table>
<thead>
<tr>
<th>Level of aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• individuals</td>
</tr>
<tr>
<td>• regions</td>
</tr>
<tr>
<td>• countries</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Choice of segmentation basis</th>
</tr>
</thead>
<tbody>
<tr>
<td>• identifiability</td>
</tr>
<tr>
<td>• substantiability</td>
</tr>
<tr>
<td>• accessibility</td>
</tr>
<tr>
<td>• stability</td>
</tr>
<tr>
<td>• actionability</td>
</tr>
<tr>
<td>• responsiveness</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Construct equivalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>• functional equivalence</td>
</tr>
<tr>
<td>• conceptual equivalence</td>
</tr>
<tr>
<td>• category equivalence</td>
</tr>
</tbody>
</table>

### Methodological issues

<table>
<thead>
<tr>
<th>Measure equivalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>• calibration equivalence</td>
</tr>
<tr>
<td>• translation equivalence</td>
</tr>
<tr>
<td>• score equivalence -individual items</td>
</tr>
<tr>
<td>• -across all items</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sampling equivalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>• choice of respondent</td>
</tr>
<tr>
<td>• sampling frame</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Segmentation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>• heuristic</td>
</tr>
<tr>
<td>• model based</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>• equal across countries</td>
</tr>
<tr>
<td>• proportional to population sizes</td>
</tr>
</tbody>
</table>

### Toward better international market segmentation

- two-stage segmentation
  - country clustering
  - cross-national consumer segmentation
- model-based segmentation
  - theory and prior knowledge
  - managerial considerations
- correction for response styles
- sample reweighting
ship—which may be considered as an a priori geographic segmentation scheme—the increase in $R^2$ should be significant. This was the case for 40 out of 59 variables, which is far better than chance ($p < 0.001$). Another noteworthy exception is Kumar et al. (1994) who validated their interactive, multicriteria approach to identify high-potential international country segments. They compared the segmentation solution to qualitative, managerial decision making and found poorer results for the latter approach. Managers arrived at solutions that were strictly dominated by other, more attractive solutions, implying a poor decision was reached, or at solutions where only one of the criteria was emphasized.

Eighth, some of the studies found spatial patterns of the segmentation solution in that the segments were geographically largely contiguous (e.g., Askegaard & Madsen, 1998; Hofstede, 1980; Ter Hofstede, Wedel, & Steenkamp, 2002). This is an important finding, since it renders targeting such segments more feasible. After all, if cross-national consumer segments are geographically widely dispersed, the increased costs of logistics would reduce the benefits of a global strategy.

Finally, most studies identified segments that span multiple countries. This is not particularly surprising when segmentation was based on country-level information, but cross-national segments were also found in studies using data from regions rather than countries (e.g., Vandermerwe & L’Huillier, 1989) or individual consumers from different countries (e.g., Moskowitz & Rabino, 1994; Ter Hofstede et al., 1999; Yavas et al., 1992; Wedel, Ter Hofstede, & Steenkamp, 1998). This provides support for the contention that groups of consumers in different countries are often more similar to each other than to other consumers in their own country (Hassan & Katsanis, 1994). As such, the combined evidence of these studies supports a basic condition for global marketing strategies, viz., the existence of cross-national consumer segments.

Table 1 indicates that we are witnessing an emerging body of international market segmentation research. Nevertheless, there are still a number of conceptual and methodological issues that need to be addressed if this area is to fulfill its high potential for marketing theory and practice. We will discuss conceptual issues (construct equivalence, level of aggregation, and choice of segmentation basis) in Section 3 and methodological issues (measure equivalence, sample equivalence, segmentation method, and sampling issues) in Section 4. Fig. 2 provides an overview of the various issues and the flow of the discussion.

3. Conceptual issues

3.1. Level of aggregation

The level of aggregation used in international segmentation ranges from no aggregation to country-level aggregation (Fig. 1). No aggregation implies that data on individual consumers from different countries constitute the basic information for segmentation. Segments formed in this way consist of consumers from different countries that are similar on the segmentation basis employed. The common way to aggregate in international segmentation is geographically. Country-level aggregation pools all the information across consumers within each country and/or uses information pertaining to the countries themselves (e.g., legal regime) and forms groupings of countries based on their similarities of the segmentation basis in question. A less “extreme” form is aggregation of people within regions (e.g., Askegaard & Madsen, 1998; Vandermerwe & L’Huillier, 1989).

Geographic (country or region) segmentation plays an important role when firms pursue a geographic differentiation strategy. For example, Philips Electronics basically treated the geographic segments of northern and southern Europe differently when marketing its personal care products. A rationale for geographic segmentation, especially when based on country aggregates, is that it results in segments that are accessible and cost effective through centralization of activities such as production, sales force management, service support, and logistics (Amine & Cavusgil, 1986; Ter Hofstede et al., 2002). On the other hand, geographic segments may overlook the differences that exist between consumers in these countries, affecting the responsiveness of segments. Moreover, the segmentation basis employed often is neither theoretically motivated nor is the managerial relevance of the segmentation variables established (Nachum, 1994). Whereas the discussion concentrates on consumers, the formulation generalizes to decision-makers including industrial customers.
1994). It is frequently unclear why particular variables are included: “some subjectivity is required in selecting the type and number of variables used to cluster the countries in groups” (Day, Fox, & Huszagh, 1988, p. 15).

Disaggregate international consumer segmentation provides a better way to identify segments that are similar in terms of needs. Still, targeting a consumer segment that exists in many countries is not always cost efficient from a logistics perspective. Especially in industries where distribution costs constitute a large part of the total costs, such as in retailing and in industries dealing with perishable products, geographically dispersed consumer segments will often not allow profitable entry strategies to be pursued.

3.2. International segmentation bases

3.2.1. Choice of the segmentation basis

The choice of the segmentation basis is one of the most crucial factors in international segmentation. The segmentation basis is a set of characteristics that is used to assign consumers to segments. As stated by Day et al. (1988, p. 14): “That a segmentation approach is essential in ... international markets is no longer questioned. Rather, the bases for segmentation become the focus...” [emphasis in the original].

A key distinction can be made between general and domain-specific segmentation bases (cf., Wedel & Kamakura, 1998). General bases are independent of the domain in question and can be further divided into observable and unobservable bases. Examples of general observable bases include geographic locations (regions, countries), economic indicators, political characteristics, and demographics. Two key instances of general unobservable bases are consumer values and life-styles. Domain-specific bases such as brand penetration rates, attitudes, benefit importances or domain-specific attitudes, depend on the particular domain/product.

Six criteria commonly used to evaluate segmentation bases (Wedel & Kamakura, 1998) are: identifiability (extent to which distinct segments can be identified), substantiality (related to segment size), accessibility (degree to which segments can be reached with promotional and distributional efforts), stability (temporal dynamics of segments), actionability (extent to which the segments provide a basis for the formulation of effective marketing strategies), and responsiveness (whether segments respond uniquely to marketing efforts targeted at them). Responsiveness is closely related to the issue of standardization of elements of the marketing mix (Farley & Lehmann, 1994). Responsive international segments provide a way to target consumers with standardized marketing mixes across borders. Table 2 provides an assessment of general observable, general unobservable, and domain-specific segmentation bases for international market segmentation.4

General observable bases are easy to identify and typically result in substantial international segments. The type of data involved in these segmentation bases is available from published sources for a large range of countries. This even applies to culture (Hofstede, 1980; Schwartz, 1994). Their accessibility will be relatively high, although typically lower than in domestic segmentation since complete and comparable international media profiles often are not available, and unless the segments are geographically contiguous, geographically widely dispersed segments can be obtained (Ter Hofstede et al., 2002). Since these bases themselves are highly stable, they tend to yield stable international segments. On the negative side, the actionability is low since these bases provide little guidance for developing international marketing strategies. Responsiveness is low, as the link between these bases and differentiated responses to marketing actions typically is weak.

Two general unobservable bases enjoy some popularity in international segmentation research—values and lifestyle.5 Since the performance differs considerably between these two bases, they will be discussed separately. Value-based segmentation leads to rather well-identified and substantial international segments. Distinct value segments of substantial size have been identified in international segmentation research (e.g., Kamakura et al., 1993; Wedel et al., 1998). Validated measurement systems of short to manageable length

---

4 Although some differences in evaluation exist with respect to specific bases belonging to a particular broad category, there is significant commonality in their performance on the various criteria.  
5 Lifestyle segmentation is more popular in applied than in academic research. Table 1 reports only one study but a number of proprietary international lifestyle segmentation schemes exist, including VALS-2, Eurostyles, and Roper Starch Worldwide (Kotabe & Helsen, 2000).
exist, including the nine-item List of Values (Kahle, 1986), the 36-item Rokeach Value Survey (Rokeach, 1973), and the 44-item Schwartz Value Survey (Schwartz, 1992). Especially the Schwartz Value Survey has been shown to exhibit high construct equivalence across scores of countries covering all continents (Schwartz, 1992; Schwartz & Sagiv, 1995). Given that value priorities are central to people’s self-concept, and hence highly stable (Burgess, 1992), this basis should yield highly stable international segments. On the other hand, little is known about the sociodemographic and media profile of international value segments, leading to low accessibility. Since values are remote from the product context and relations between values and specific behaviors are typically weak (Pitts & Woodside, 1984), actionability of international value segments is rather low. One can expect differential responsiveness with respect to advertising strategies (Pollay, 1984) but the responsiveness to other marketing mix elements is less clear. The actionability and responsiveness of international value segmentation is increased if values are linked to product attributes and benefits in a means–end framework (Ter Hofstede et al., 1999).

Lifestyle segments are difficult to identify. The lifestyle construct pertains to how consumers express their values and personalities in their daily lives through various activities, interests, and opinions (AIOs). There are few if any generally accepted and validated cross-cultural lists of AIOs publicly available and the identification of international lifestyle segments is problematic (Kotabe & Helsen, 2000). International lifestyle segmentation can yield substantial segments. For example, Kotabe and Helsen (2000) describe a (proprietary) Roper global lifestyle segmentation scheme consisting of five segments with segment sizes varying between 10% and 23%. Since media profiles are typically part of an international lifestyle segmentation study, accessibility of the segments through promotional activities should be reasonable. Accessibility in terms of distribution is low, however, since segments can be widely dispersed geographically and retail sectors tend to differ across countries. The lifestyle construct is not rigorously grounded in theory and is farther removed from the self, so the stability of international lifestyle segments will be low. Actionability and responsiveness are expected to be modest. Previous domestic research indicates that lifestyles are more closely related to behavior than values are (Vinson, Scott, & Lamont, 1977). However, we expect that this does not apply to international lifestyle segmentation as apparently the same AIO items may mean different things in different countries and thus lack construct equivalence (see below).

Domain-specific segments should be rather easy to identify given that they are typically based on a limited number of well-understood and well-researched variables, such as attribute evaluations (Moskowitz & Rabino, 1994) or risk and loyalty ratings (Yavas et al., 1992). On the other hand, they reflect a more direct response to the socio-cultural environment, which differs between countries, and this should hamper identifiability of international segments (cf. Kamakura et al., 1993). Domain-specific segments can be substantial—e.g., Moskowitz and Rabino (1994) identified three international segments, with segment sizes varying between 26% and 39%. Accessibility will be low since such segments typically do not exhibit strong relations with media profiles and may be geographically dispersed across countries (but see Ter Hofstede et al., 2002). They are expected to be less stable than segments based on values or general observable bases since they are more directly influenced by changes in the socio-cultural environment.

Table 2
Evaluation of international segmentation bases

<table>
<thead>
<tr>
<th>Bases</th>
<th>Criteria</th>
<th>Identifiability</th>
<th>Substantiality</th>
<th>Accessibility</th>
<th>Stability</th>
<th>Actionability</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>General observable</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>–</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>General unobservable</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>Values</td>
<td>–</td>
<td>+</td>
<td>±</td>
<td>–</td>
<td>±</td>
<td>–</td>
<td>±</td>
</tr>
<tr>
<td>Lifestyle</td>
<td>±</td>
<td>+</td>
<td>–</td>
<td>±</td>
<td>–</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>Domain-specific</td>
<td>±</td>
<td>+</td>
<td>–</td>
<td>±</td>
<td>–</td>
<td>±</td>
<td>±</td>
</tr>
</tbody>
</table>

++: very good; +: good; ±: moderate; –: poor.

and are far removed from the self. Domain-specific bases typically deal with variables that are directly linked to marketing decisions of the firm (a good example is Moskowitz & Rabino, 1994). Hence, we may expect high actionability and responsiveness.

3.2.2. Construct equivalence of the segmentation basis

Any rigorous international segmentation study requires that its segmentation basis exhibits construct equivalence. Construct equivalence refers to whether the basis used for segmentation is equivalent across countries. Three types of construct equivalence can be distinguished: (1) functional equivalence—whether the basis serves the same purpose or function across countries; (2) conceptual equivalence—whether the basis occurs in different countries and whether it is expressed in the same way; (3) category equivalence—whether the basis belongs to the same specific class of objects or activities across countries (Craig & Douglas, 2000; Kumar, 2000).

If construct equivalence of the segmentation basis has not been met, international segments are likely to be masked by different meanings of the basis in different countries. If consumers in different countries share the same basis but cross-national differences exist with respect to the meaning of the segmentation basis or if the segmentation basis has no meaning at all in some of the countries, segments are likely to be based on these differences instead of the desired similarities in the segmentation basis. The various types of construct equivalence typically are less easy to achieve for domain-specific bases such as attribute importances or purchase incidence as they reflect a more direct response to the consumer’s socio-cultural environment.

For most general bases, the different types of construct equivalence are easier to achieve (although the score on those variables may still carry different meanings in different cultures; see Section 4.1.3). A notable exception is segmentation based on lifestyle. Lifestyle is closely tied to the socio-cultural environment (Kamakura et al., 1993) and its theoretical content has not been identified rigorously.\(^6\) Consequently, lifestyle may lack construct equivalence across countries—and especially conceptual and category equivalence—giving rise to limited applicability of lifestyle in international segmentation (e.g., Beatty, Homer, & Kahle, 1987). Consider the case of the well-known VALS segmentation scheme (Mitchell, 1983). The key VALS-2 dimensions in the USA (Japan) are self-orientation and resources (life orientation and attitudes toward social change), indicating lack of conceptual equivalence.\(^7\) Many of the VALS-2 items are also highly domestic (US) in content and relevance, and thus lack broad conceptual equivalence. One example is: “The Federal Government should encourage prayers in public schools.”

Most studies reviewed in Table 1 do not provide detailed information on construct equivalence, but some observations can still be made. Several studies segment countries on characteristics reflecting economic development (Day et al., 1988; Helsen et al., 1993; Huszagh et al., 1986; Lee, 1990; Sethi, 1971). Most of these variables may be expected to possess construct equivalence, but there are exceptions. Urbanization is commonly defined as the percentage of people living in urban areas, but the categorization of areas as urban or rural need not be the same across countries. The whole notion of urbanization takes on a different meaning in a densely populated country like the Netherlands (where every village is within 30 km of a larger city) versus a country characterized by large metropolitan areas with relatively empty space in between (e.g., USA, France, Australia). Lee (1990) used the number of scientists and engineers per 1000 people as a segmentation variable, but occupational categories do not always have the same meaning and status across countries. Dawar and Parker (1994) used the percentage of the labor force engaged in the retail sector as segmentation basis, but one may wonder whether definitions of retail sector are comparable across countries, especially given the enormous variety of companies active in various forms of reselling.

Hofstede (1980) and Kale (1995) segmented countries based on Hofstede’s national—cultural dimensions. However, even Hofstede’s seminal framework may exhibit only partial construct equivalence. In a

---

\(^6\) In contrast, the theoretical structure of two somewhat related international segmentation bases, viz. values and personality is much more rigorously defined (Church & Burke, 1994; Schwartz, 1992).

\(^7\) More information can be obtained on the VALS website: http://future.sri.com/VALS/valsindex.shtml.
replication study (Hofstede, 1994), a fifth dimension was identified: long-term versus short-term orientation, while uncertainty avoidance was not replicated. Steenkamp (2001) combined Hofstede’s (1980) and Schwartz’s (1994) national–cultural frameworks and identified one dimension, egalitarianism versus hierarchy, which was not included in the Hofstede framework.

Ter Hofstede et al. (1999) segmented the European market, based on means–end chain data on yogurt while Moskowitz and Rabino (1994) segmented the market for flavored soda. It is not clear whether product class definitions are the same across countries. Yogurt may belong to different categories in different countries and what is considered a soft drink also differs between countries (Craig & Douglas, 2000).

3.2.3. Two-stage international segmentation

Table 2 reveals a basic tradeoff between construct equivalence—leading to well-identified international segments—accessibility, and stability on the one hand versus actionability and responsiveness on the other hand. In particular, general observable (domain-specific) bases are high (low) on identifiability, accessibility, and stability, and low (high) on responsiveness and actionability. On most criteria, the general unobservable bases are in between. Thus, those bases that exhibit higher construct equivalence and yield more accessible and stable segments tend to be poor on marketing relevance. On the other hand, those bases that yield actionable and responsive segments produce solutions that are not always well accessible or stable, and the segmentation solution may be contaminated by cross-national differences in the meaning of the basis per se. This dilemma is one of the key conceptual problems in international market segmentation. It can be addressed if we recognize that construct equivalence is a matter of degree. The more alike the socio-cultural environment is among countries, the more equivalent domain-specific constructs as well as segment stability within a cluster are likely to be higher than that for the unsegmented solution.

In the second (disaggregate) phase, data can be collected on domain-specific bases among consumers in the different countries belonging to the target geographic segment. These disaggregate data form the basis for the identification of cross-national segments of consumers within the geographic segment selected. These segments will be more responsive to marketing efforts and, especially if the segments are geographically contiguous, their accessibility will also be higher.

4. Methodological issues

4.1. Measure equivalence

Measure equivalence refers to whether the measures used to operationalize the segmentation basis are comparable across countries. A distinction can be made between calibration equivalence, translation equivalence, and score equivalence (Craig & Douglas, 2000; Kumar, 2000).

4.1.1. Calibration equivalence

Calibration equivalence refers to equivalence in monetary units, measures of weight, distance, and volume, and perceptual cues such as color and shape. This type of equivalence is straightforward to achieve when dealing with physical and monetary measurement units. However, it may present problems with other variables, including a number of sociodemographics. Consider a “simple” variable like level of education. A common scale used in the USA is: grade 8 or less; grades 9–11; high school; 1–3 years of college; college (4 years); attended or completed graduate
In Europe, the educational system is quite different, and the term “college” is not appropriate. In fact, in countries like Germany, the Netherlands, and Belgium, there are several levels of post-high-school training, which sometimes take the same number of years to complete but represent quite different levels of educational attainment. Calibration equivalence is problematic, unless one recodes the responses into very broad categories (e.g., lower versus higher educated). Note that such recoding practices will improve equivalence but lead to a loss of information and thus negatively affect the stability of segments.

The scale for household income is difficult to calibrate internationally due to differences in tax regime and purchasing power. One solution is to ask for net income after taxes including all allowances. However, a net income of US$15,000 is below the poverty line for a couple with two children in the USA while it puts you into the wealthy segment of the population in India. Calibration equivalence is also an issue when dealing with retail and household scanning data. Market definitions, definitions of (sub)types of categories and brands, of what constitutes a private label, etc., differ widely, even within the European Union. While we see great progress in domestic segmentation modeling, due to the growing availability of these data (Leeflang & Wittink, 2000; Wedel, Kamakura, & Böckenholt, 2000), the lack of calibration equivalence is a serious barrier to their use in international segmentation.

The studies reported in Table 1 do not provide much information on calibration equivalence. In general, it does not seem to be a problem since the response scales used in disaggregate research are fairly standard, and the units of most secondary variables are also straightforward, internationally. Several studies, however, include education, GNP/capita, and/or income (e.g., Day et al., 1988; Helsen et al., 1993; Sethi, 1971; Vandermerwe & L'Huillier, 1989) as segmentation variables, which suffer from the problems noted above. Nevertheless, given that all these studies also included many other variables for which calibration equivalence is not an issue, it can be concluded that lack of calibration equivalence is not a serious biasing factor in these studies.

For example, this scale is used on the VALS website: http://future.sri.com/VALS/valsindex.shtml.

### 4.1.2. Translation equivalence

Translation equivalence indicates that the measurement instrument is interpreted similarly by respondents in different countries. Translation equivalence is much broader than verbatim translation of the items. As emphasized by Kumar (2000), the focus should not be on verbatim translation but on constructing items in the local language that convey the intended meaning. This requires (1) careful back-translation and (2) extensive pretesting of translations in different countries. Achieving translation equivalence is a very complicated process, especially since equivalent words may not be available in other countries. For example, there is no obvious cultural equivalent for the Dutch word *gezelligheid* (sitting together with a number of people and feeling good because of a purported unity in feelings and emotions) in other languages. In the Netherlands, quite some brands are positioned on *gezelligheid*, which complicates international benefit segmentation studies for these brands.

International market segmentation studies are not always careful in their linguistic translation. Only about half the studies reviewed in Table 1 that collected primary data explicitly noted that back-translation procedures were followed (e.g., Askegaard & Madsen, 1998; Sirotta & Greenwood, 1971; Ter Hofstede et al., 1999). It was even less clear whether the various items conveyed the intended meaning. For example, Askegaard and Madsen (1998) questioned whether their translated items had similar meanings in the different countries.

The various studies provide little, if any, information on international pretesting of the translated questionnaires. Exceptions include Sirotta and Greenwood (1971) and Ter Hofstede et al. (1999). The former study reported “exhaustive pretesting” (p. 54). Ter Hofstede et al. (1999) developed the questionnaire, partially based on qualitative information from a bilingual country (Belgium), and pretested it (after back-translation) in three successive waves involving six other countries, before the main study was carried out in 11 EU countries. Some studies (Dawar & Parker, 1994; Hofstede, 1976) circumvented the translation issue by having all respondents respond to an English questionnaire. Dawar and Parker (1994, p. 85) argued that this “helps to overcome potential language-based response biases,” but ignored that differ-
ential language fluency and interpretation of English can create even more serious biases.

4.1.3. Score equivalence

Score equivalence refers to the equivalence of the observed scores on the measures. Lack of score equivalence can be due to cross-national differences in response to (1) individual measures and/or to (2) the complete set of measures (Baumgartner & Steenkamp, 2001; Steenkamp & Baumgartner, 1998). In both cases, observed scores are not comparable across countries. This will lead to biased international segmentation solutions, in which true cross-national similarities and differences are confounded by bias in observed scores.

Steenkamp and Baumgartner (1998) outlined a procedure to test for score equivalence at the level of individual measures, using confirmatory factor analytic procedures for repeated measurements of constructs. It yields individual-level scores corrected for lack of score equivalence and these scores can be used in subsequent analyses, including international segmentation (see Cadogan, Paul, Salminen, Puumalainen, & Sundqvist, 2001; Deshpande, Farley, & Webster, 2000; Steenkamp, Ter Hofstede, & Wedel, 1999 for applications). Baumgartner and Steenkamp (1998) extended this approach to combined emic–etic scales (Berry, 1989). They showed how the confirmatory factor analysis model can be modified to yield cross-nationally comparable scores on such scales which consist of a combination of country-specific and cross-nationally standardized measures. This procedure is useful if a segmentation basis exhibits only partial construct equivalence.

The issue of measure-specific score equivalence not only applies to primary data but also to secondary data. Note that several studies have segmented the international market based on secondary data (see Table 1). However, these statistics are not often comparable across countries. For example, employment data are sometimes manipulated for political purposes. Even the wealth of countries (e.g., GNP or GNP/capita) is difficult to compare due to different tax morals. Secondary data from transitional economies and emerging consumer markets are especially hard to compare (Batra, 1999). Notwithstanding these considerations, score equivalence has been typically ignored in international segmentation studies using secondary data. The very fact that the data have been published earlier does not assure the need for evaluating score equivalence, e.g., by comparing estimates from different sources and by constructing composite indices. The procedure described by Steenkamp and Baumgartner (1998) can also be used on secondary data provided the number of observations is sufficiently large and multiple indicators/estimates per construct are available.

There can also be lack of score equivalence in the complete set of measures due to cross-national differences in response styles (Baumgartner & Steenkamp, 2001). The term response style refers to a person’s tendency to systematically respond to questionnaire items on some other criterion that what the items were supposed to measure (Paulhus, 1991). Stylistic responding tends to affect all measures of a given segmentation basis and as such is not easily detected by the procedure outlined by Steenkamp and Baumgartner (1998), which is based on differential item functioning. The best-known response styles are acquiescence bias, extreme responding, and social desirability responding (Paulhus, 1991). There is evidence that countries may differ systematically on these response styles (Chen, Lee, & Stevenson, 1995; Stening & Everett, 1984). Baumgartner and Steenkamp (2001) described a procedure to purge observed scores from stylistic responding by regressing the raw summed scale scores on response bias indices.

Rossi, Gilula, and Allenby (2001) developed a Bayesian hierarchical method to correct for response styles that arise in discrete rating scales. Their method treats observed ratings as discrete realizations of the true (latent) score on a continuous scale and some individual-level response-style parameters. The re-

---

9 A further distinction can be made between metric and scalar equivalence. Metric equivalence refers to the equivalence in measurement units while scalar equivalence exists if measure intercepts are equal across countries. See Steenkamp and Baumgartner (1998) for details.

10 For example, an article in the Wall Street Journal Europe (1999) stated that “The government and international institutions explicitly recognize that a large part of the Italian economy goes unreported, and add a certain percentage to GDP to make up for these leakages when measuring its size.” Needless to say, such mark ups are estimates, suffering from reliability problems.

11 Kumar (2000) referred to this as analysis equivalence.
Response-style parameters represent the overall mean and variance across all items and thus capture common patterns such as extreme responding and the tendency to use the midpoint of the scale. Predictions of the continuous (unbiased) scores on items can be obtained through Bayesian shrinkage estimation, which provides individual-level estimates that are more stable and reliable (Blattberg & George, 1991). Although their method was developed for domestic research, it holds considerable promise for international segmentation.

Cross-national differences in response styles have been most commonly studied for rating scale data. Recently, Ter Hofstede et al. (1999) showed that these biasing effects may also occur for scales of multiple dichotomous items, i.e., pick-any/K data collection methods. Building upon item response theory, they developed an international segmentation model, which distinguishes idiosyncratic response styles from the actual strength of the person’s belief. If the strength exceeded the person’s threshold (response style), the person indicates a belief, otherwise not. The smaller the threshold, the higher the propensity is to indicate a belief. Thresholds were allowed to vary within and across countries. In a study involving 11 EU countries, they showed that the null hypothesis of equal threshold means across countries was clearly rejected \( p < 0.0001 \), indicating that there are differences in response tendencies across countries on dichotomous scales as well.

Whereas in international segmentation research, rating scales and dichotomous scales may lack score equivalence, it is less likely to occur for constant sum scales or ranking. Constant sum scales force respondents to make tradeoffs. Although extreme responding can still occur (by giving either many or very few points to a given item), it avoids yea-saying. On the negative side, constant sum scales are cognitively more burdensome than rating or dichotomous scales and cannot be used easily in less developed countries where mathematics education is often less advanced.

Ranking is easier than constant sum, and it does not suffer from response tendencies in that in some countries, consumers rate more extreme, have a greater tendency to agree with items, or tick off a category more often. It forces respondents to use each category in a scale once and only once (assuming that no ties are permitted which is the usual procedure and guards best against response styles). However, a ranking task is not appropriate when absolute scores on items are required (e.g., purchase intentions). In addition, it may suffer from order effects and since it does not allow ties, it forces people to make distinctions they may not perceive. But to the extent that such effects do not differ systematically between countries and absolute scores are not required, ranking is an attractive scoring method in international segmentation research.

It is worrisome to note that score equivalence has not received much attention in international segmentation research (cf., Table 1). None of the studies examined score equivalence at the level of individual items, probably because the analytical procedure is recent. A mere two studies reported in Table 1 used ranking data (Kamakura et al., 1993; Wedel et al., 1998), while Hofstede (1976) used constant sum scales. Ter Hofstede et al. (1999) is the only study to explicitly model and correct for response styles. We believe that lack of attention to score equivalence is one of the reasons why international segmentation studies often report a heavy country influence.

### 4.2. Sampling equivalence

A final type of equivalence that should be fulfilled is sampling equivalence. This concept has frequently led to confusion because it is often (erroneously) equated with obtaining a sample with the same socio-demographic characteristics in different countries. Kumar (2000, p. 236) makes the following important observation: “The stress is not on the equivalence of the method used or the profile from which the sample has been drawn, but rather on the equivalence of the information collected from the sample.” This should focus on (1) the choice of the target population, and (2) the sampling frame. What the target relevant population is, depends on the purpose of the study. It will often comprise all primary decision makers. Due to cross-cultural differences, a sample of primary decision makers may very well differ in background characteristics (Kumar, 2000).

The second issue concerns the extent to which the sample is representative of the relevant target population. In developed countries, sampling frames are typically easily available (an exception is Japan; Kumar, 2000). This is not the case in most developing or emerging countries. Also, questionnaires often
require basic amount of schooling, which excludes a substantial part of the population in developing countries. While in developed countries, representativity of any given sample is easily checked against census data, reliable census data are unavailable in some other countries. Moreover, census data often are not the appropriate comparison standard as one strives for representativity with respect to the relevant target population, not the population at large. A second best option, if sample representativity is difficult to obtain, is to select matched samples in the different countries (Sekaran, 1983). This could be achieved by collecting data among a particular group of people (e.g., students) or in one of the country’s large metropolitan areas (e.g., Lagos, Sao Paulo, Paris). Of course, the matching process itself is ambiguous. The definition of “students” and the social classes they come from, differs across countries, and the socio-cultural position of Lagos in Nigeria may not necessarily be similar to the position of Paris in France or of Sao Paulo in Brazil.

All international segmentation studies collecting data among respondents appear to have sampling equivalence. Askegaard and Madsen (1998), Boote (1983), Kamakura et al. (1993), Moskowitz and Rabino (1994), Ter Hofstede et al. (1999, 2002), and Wedel et al. (1998) have all taken efforts to collect a representative sample of respondents, although none of the studies provided concrete evidence. Matched samples of managers were used by Hofstede (1976, 1980), Ronen and Kraut (1977), and Sirotka and Greenwood (1971). Yavas et al. (1992) have matched samples of middle income women residing in major urban centers. The samples collected by Verhage et al. (1989) were matched on income, household expenditures on fuel and heating, and the number of frost days per year. Sample equivalence in Dawar and Parker (1994) seems the most ambiguous. They used samples of people studying at INSEAD. It is questionable though, whether students from, e.g., Zimbabwe or Syria that can study at INSEAD come from the same socio-economic background compared to the country at large as students from, say, France or Germany.

4.3. International segmentation methods

Several methods have been used to identify international market segments. Roughly, a distinction can be made between heuristic and model-based methods. **Heuristic** methods rely on Q- or R-factor analysis and cluster analysis. Q-factor analysis has been demonstrated not to be an appropriate technique for segmentation (Dillon & Goldstein, 1984; Stewart, 1981). Although Q-factor analysis was used in some early international segmentation studies (Boote, 1983; Hofstede, 1976; Sirotka & Greenwood, 1971), it is encouraging to observe that it has not been employed in recent studies (see Table 1).

Cluster analysis is more appropriate than Q-factor analysis and is the most frequently used international segmentation technique (see Table 1). Despite its popularity, cluster analysis has a number of limitations. It generates deterministic classifications often based on subjective optimization criteria and performs less well when variables are highly correlated. It does not fit within the framework of standard statistical theory and does not provide reliability judgments of the results. A two-stage procedure involving R-factor analysis and cluster analysis is also popular in international market segmentation (e.g., Askegaard & Madsen, 1998; Helsen et al., 1993; Moskowitz & Rabino, 1994; Steenkamp, 2001). An initial set of items is reduced to a smaller subset of factors and the scores of the objects on those factors are subjected to cluster analysis. It eliminates the problem of correlated variables but it introduces the issue of the selection of the number of factors and most of the problems associated with cluster analysis remain. Further, the combination of factor analysis and cluster analysis is not recommended in the cluster analysis literature because it does not make full use of the available information in the data. Arabie and Hubert (1994) demonstrated that it is better to apply factor analysis after cluster analyzing the full set of variables.

**Model-based international segmentation** is based on a particular representation of reality. It allows researchers to identify international segments using cross-national theories of consumer behavior. For example, Helsen et al. (1993) and Kumar et al. (1998) used epidemiological theory (Bass, 1969) to model diffusion of innovation in countries. Ter Hofstede et al. (1999) developed a mathematical model of psychological links between different aspects of consumer product-related knowledge based on means–end chain theory (Gutman, 1982). They also showed that, in terms of statistical and managerial criteria, an exploratory technique, i.e., K-means clustering, led to a
less attractive segmentation solution than their model-based approach. Another advantage of a model-based approach is that it allows one to make normative inferences based on managerial considerations that are relevant for international firms.

A powerful statistical approach to model-based international market segmentation is the finite mixture model, which identifies international segments and at the same time estimates unknown parameters that characterize such segments. Such models have been employed in the international market segmentation studies of Helsen et al. (1993), Kamakura et al. (1993), Ter Hofstede et al. (1999), and Wedel et al. (1998). Finite mixture models overcome the limitations of cluster analysis. They allow hypothesis testing and estimation according to basic statistical principles. Limitations of the finite mixture approach to international segmentation are that it is not easily implemented and may suffer from local optima in estimation.

More recently, Bayesian sampling-based approaches to finite mixture models have been proposed (Gilks, Richardson, & Spiegelhalter, 1995). They have a number of desirable properties for international segmentation. First, these sampling-based approaches extend to model specifications of within-segment heterogeneity. The standard finite mixture approach to segmentation has recently received some criticism since it assumes that segments of homogenous consumers underlie the data-generating process. One could question whether international segments of consumers exist or that they are mere constructions yielding an artificial partition of an underlying continuous distribution. Some authors claim that models of continuous heterogeneity would be more useful (Allenby & Rossi, 1999) since such models do not impose restrictive assumptions of within-segment homogeneity. On the other hand, models of continuous heterogeneity rely on the correct specification of probability distributions and such distributional assumptions are likely to be violated as well. Sampling-based approaches allow researchers to specify flexible models that capture heterogeneity within international segments.

The direct evidence on whether international segments of consumers really exist is limited. As noted in Section 2, previous work segmenting consumers from multiple countries has consistently found cross-national segments. Our review indicates that quite some of these studies did not correct for response styles, that construct and translation equivalence are not always unambiguously established, etc., all of which work against finding cross-national segments. The results of these studies support the premise of the existence of international marketing segments, although, with the exception of Ter Hofstede et al. (2002), between- versus within-segment heterogeneity is not directly investigated. Ter Hofstede et al. (2002) compared the amount of within- versus between-segment heterogeneity in store image attribute importances. In their contiguous international segmentation model, between-segment heterogeneity was found to be more than twice as large as within-segment heterogeneity. When the contiguity constraint was relaxed (as it does not bear on the existence of international segments, only on their accessibility), between-segment heterogeneity was 3.3 times as large as within-segment heterogeneity. Thus, they found support for the existence of international segments. Nevertheless, the within-segment heterogeneity was non-negligible. Therefore we recommend accommodating within-segment heterogeneity in future international segmentation studies to assess to what extent international segments are truly separated.

A second desirable property of a Bayesian sampling-based approach to international market segmentation is that it allows researchers to take a modular approach to building international segmentation models, accommodating the many methodological complexities and conceptual problems one encounters in international segmentation. Comprehensive models can be constructed by adding constituent model components to a core international segmentation model. The unknown parameters of this comprehensive model can then be estimated simultaneously using conditional distributions of the model parameters, which generally leads to estimable models. For example, the measurement model of Rossi et al. (2001), which captures differences in response styles, can be added as one of several constituent components to a comprehensive international segmentation model. As another example, the two-stage segmentation approach proposed in Section 3.2.3 can be integrated in one model, by combining a model of country segmentation with one of international segmentation of consumers.

Third, if prior knowledge is available about certain aspects of the international segmentation structure,
they can be incorporated in the prior distribution of
the model. Ter Hofstede et al. (2002) developed a
Bayesian segmentation method, which imposes a
spatial prior on segments. This improves the identi-
fication of spatially concentrated international seg-
ments by incorporating prior knowledge obtained
from previous studies in the model. The method was
used to identify international segments in the EU
based on store image data and was shown to provide
responsive segments, despite the spatial prior imposed
on the geographic configuration of the segments.

Finally, Bayesian sampling-based techniques facil-
itate normative modeling for designing optimal
international marketing strategies. If objective functions
of international segmentation strategies can be measured
and formulated, Bayesian decision theory can be
applied by generating (posterior) draws from the
objective function, thereby integrating out the uncer-
tainty in the data. Such an approach may be optimal
from a decision point of view, but it assumes that the
quality of decisions can be elicited from managers and
can be characterized by a mathematical specification.
None of the studies in Table 2 addressed the nature of
cost structures of international segmentation strat-
gegies.12 The measurement and specification of cost
structures and utility structures of managers deserves
more attention in order for such an approach to be
successful.

Fig. 1 shows that in the last decade there is a
distinct trend toward more model-based research in
international segmentation. This is highly encourag-
ing, but also necessary if international market seg-
mentation is to gain the place and research effort
 accorded to domestic market segmentation. Although
model-based approaches require more statistical
expertise of researchers, we believe the future in
international segmentation belongs to model-based
approaches, integrating consumer behavior and mar-
keting theories with sound statistical methods and
further exploring the possibilities for normative mod-
els of optimal international segmentation decisions.
Bayesian sampling-based techniques offer much
potential for future improvements in this area.

4.4. Sample size versus population size

In international market segmentation on disaggre-
gate, consumer-level data, the issue emerges whether
country sample sizes should be proportional to their
population sizes, i.e., should the ratio of sample to
population sizes be constant across countries or not? If
international samples are not proportional to popula-
tion sizes, the pooled sample is not representative of
the pooled population and some kind of reweighting
should be applied in the estimation stage. Thus, there
are two approaches leading to proper international
segmentation results. Either (1) reweighting is applied
to samples with sizes disproportional to population
sizes, or (2) proportional samples are drawn, which
renders reweighting obsolete.

In international marketing research, samples are
often stratified by country, with approximately the
same sample size in each country. This results in
samples that are not proportional to population
sizes. When data are averaged across respondents
within countries (e.g., Hofstede, 1976, 1980), this
does not pose a serious problem. The key issue is to
obtain samples that are sufficiently large for reliable
estimation of country means, which constitute the
input for the segmentation process.

All international segmentation studies using indi-
vidual-level data are based on national sample sizes
that are not proportional to population sizes and most
of these studies have not applied any form of
reweighting (e.g., Boote, 1983; Moskowitz and
Rabino, 1994; Verhage et al., 1989; Yavas et al.,
1992). These results will be biased because each
subject in the sample is assigned the same weight,
while the chances of a subject being included in the
sample is not uniform across countries. For example,
when sample sizes are equal across countries, a person
from Ireland would have a probability of being chosen
that is 20 times larger than a person from Germany
(the population of Germany and Ireland is about 81
million and 4 million, respectively). The net result is
that the solution is biased toward respondents from
smaller countries. Obviously, this is more problematic
when country populations are very different.

More specifically, when a specific segment exists in
a populous country, this segment should be relatively
large. However, if the sample sizes are equal across
countries, traditional estimation procedures that do not

---

12 However, Kumar et al. (1994) included contribution margin
in their model and used simulated data on this variable in their
empirical application.
take the sampling design into account are very likely to overlook such a segment and may give more weight to smaller segments in less populous countries. Verhage et al. (1989), for example, pooled data from the USA, France, the Netherlands, and Norway without reweighting despite the dramatically different ratios of sample to population sizes (e.g., the population of the USA is over 60 times larger than the population of Norway). In such a design, the USA and France are vastly underweighted, which biases the results. Indirectly, this can be observed in their segment solution. The second and third segments are dominated by Norway and the Netherlands, while more than two-thirds of the US and French respondents belong to the first segment. Too much weight is given to less populated countries while heterogeneity in populous countries is not adequately captured.

In cluster analysis, reweighting procedures can be employed by weighting the cluster solution by the selection probabilities (see Wedel & Kamakura, 1998). Wedel et al. (1998) recently proposed a pseudo-maximum likelihood estimation approach to accommodate sampling designs in mixture models. The method provides consistent estimates of the model parameters when data arise from disproportionate samples. They applied their approach to value rankings from six EU countries with approximately equal sample sizes from each country while population sizes varied between 10 million and 81 million. They found that standard maximum likelihood led to misleading results, including an overestimation of the number of segments and biased estimates of segment-level parameters. Pseudo-maximum likelihood yielded unbiased and well interpretable results. The pseudo-maximum likelihood procedure was also used by Ter Hofstede et al. (1999) in a study involving disproportionate samples from 11 EU countries. Straightforward methods for accommodating sampling designs for Bayesian sampling-based estimation of mixture models have not been proposed. Future research could shed more light on this issue.

An alternative to reweighting is to draw samples proportional to population sizes. This sampling design is ‘self-weighting’ and renders reweighting obsolete. None of the studies examined in Table 1 had proportional samples. Proportional sampling may lead to sample sizes being too small for smaller countries to allow reliable estimation of country-specific parameters and parameters pertaining to segments with large proportions of consumers in these countries. This creates problems as even in global marketing strategies, the country-specific element still plays an important role, due to, e.g., differences in legislation, language, and retail structure (Jeannet & Hennessey, 1998). Another self-weighting sampling design is complete random sampling, where consumers are drawn directly from the total population instead of national populations. This procedure, however, introduces additional complexities with respect to the construction of the international sampling frame.

In sum, analyzing pooled data is appropriate when one is looking for cross-national segments but requires either reweighting or proportional samples. We believe that reweighting of disproportionate samples is the best solution to the sample size versus the population size issue. Thus far, it has only been applied in two studies reviewed in Table 1 (Ter Hofstede et al., 1999; Wedel et al., 1998), but deserves more attention in future research.

5. Integrating the various issues: a case study

The problem faced by one of the largest consumer packaged good companies in the world may serve to illustrate and integrate the various issues and guidelines described in this review (Fig. 2). Note that the purpose of the illustration is not to define best practices. This company is highly dependent on new product introductions to ensure future profitability and to keep store brands at bay. The company traditionally adopted an unsegmented strategy, in which new products are introduced at a pan-European level. It was aware that not all consumers are equally receptive to new products and that the size of the segment of “best” prospects might differ between countries. An international segmentation strategy was called for.

One of the authors worked with the company to address this problem. A two-stage approach was adopted. In the first (aggregate) phase, a qualitative clustering on EU countries was conducted, using economic and company criteria (e.g., size of the market, presence of the company in those markets).

---

13 In fact, in this study, sample sizes were neither equal across countries nor proportional to country populations: Norway: 59, USA: 90, France: 100, Netherlands: 268.
This resulted in two segments: the five largest EU countries (Germany, UK, France, Italy, and Spain, representing 80% of the EU population) and the other countries. It was decided to focus on the former segment. In the second (disaggregate) stage, data were collected on consumer innovativeness. Consumer innovativeness is the predisposition to buy new products and brands rather than to remain with previous choices and consumption patterns. It is a secondary disposition in that it is operant only in a specific domain of contexts and roles, namely, the consumer setting (Steenkamp et al., 1999). It serves the same purpose or function in different countries (functional equivalence), i.e., purchasing of new products and brands, is interpreted similarly in different countries (conceptual equivalence), and is applicable to a wide variety of categories, including durables and nondurables (category equivalence) (see Steenkamp et al., 1999). Thus, it possesses construct equivalence. Given the consumer-oriented focus of the company and the purpose of the study, the level of aggregation was the individual consumer. The segmentation basis was a domain-specific basis.

A measurement instrument was developed, consisting of a number of items, inspired by previous academic and applied research. The items were developed in English and pretested in another European country, after back-translation. After the pretest, the items were back-translated in German, French, Italian, and Spanish and administered to representative samples in each country. Items were scored on Likert scales. Calibration equivalence was not an issue with the items per se and the response scale was also well known in all countries. However, we encountered calibration problems with background variables (e.g., education) used for segment accessibility purposes, which necessitated recoding of these data to broad categories. Translation equivalence was assured (to the extent possible) using back-translation and input from local branches of the pan-European market research agency.

Score equivalence was ascertained as follows. First, the measurement instrument was balanced, with about the same number of positively and negatively worded items. This guards against yea-saying (Baumgartner & Steenkamp, 2001). Second, we added a few unrelated items that were shown in the pretest to generate extreme responses. Extreme response style as computed on these items was not correlated with the summed score on the innovativeness instrument.

Third, we analyzed item-specific score equivalence using the procedure outlined by Steenkamp and Baumgartner (1998). No problems emerged. Sample equivalence was achieved by the random sampling procedures and the existence of comprehensive sampling frames in each country.

We used a model-based segmentation approach, using Rogers’ (1995) innovation theory of consumer behavior, and managerial considerations. According to Rogers’ theory, 2.5% of the consumers are innovators, 13.5% early adopters, 34% early majority, 34% late majority, and 16% laggards. Managerial considerations were incorporated by combining the segments of innovators and early majority. The company deemed a segment of 2.5% of consumers to be not sufficiently substantial to warrant separate consideration. Rogers’ theory was put to use by applying cutoff scores to the distribution of summed scores across the total pooled sample (n = 29,830). The cutoff scores were determined to reflect Rogers’ distribution of segments as closely as possible at the pooled level. That is, the 16% (approximately) respondents with the highest summed score were assigned to the early adopters segment, the next 34% (approximately) to the segment of early majority, etc. The size of the four cross-national segments—early adopters, early majority, late majority, and laggards—differed dramatically between countries, the size of the segment of early adopters being nearly three times as large in the most innovative country compared to the least innovative country. Thus, our model was methodologically straightforward but firmly based in an established theoretical model of consumer behavior.

We had to reweight the country samples before constructing the distribution of pooled scores since there was no relation between sample sizes and population sizes. Otherwise, especially the UK would have been vastly overweighted, which would have led to seriously biased conclusions. The cross-national segments were validated against actual purchase behavior and profiled with sociodemographic and other characteristics. Apart from core sociodemographics, the available data (e.g., media) differed vastly across countries.

The resulting international segmentation scheme is currently used in a variety of product categories to develop new product introduction and marketing
strategies. Compared to the previous unsegmented strategy, it was estimated that the effectiveness of new product introductions using the new international segmentation scheme could be increased by (on average) about 50%. Our results also indicated that a test market for a new product (in terms of the percentage of people purchasing the new product in the first months after introduction) yields substantially better results if it is carried out in the most innovative country compared to less innovative countries, due to dramatic differences in size of the early adopters segment. Our conclusions were consistent with the company’s actual test market experiences. As it happened, for a variety of historical and other reasons, this company conducted its test markets for Europe most often in the most innovative country, and subsequently observed that sales goals based in this test market were frequently not met in other countries. The new international segmentation scheme helped to put domestic test market results in a European perspective and to develop more realistic marketing plans.

In summary, this case study shows that it is possible to integrate the specific conceptual and methodological issues discussed in this paper, and to adopt the guidelines for better international market segmentation. It is interesting to note that the very fact that the study was based on an established theoretical model increased managerial confidence in the results considerably.

6. Conclusions and future research

6.1. Conclusions

The body of international segmentation research is growing. Previous international segmentation studies have applied a wide range of segmentation bases and have provided preliminary evidence of consumer segments that exist in different countries. The latter finding is important for the feasibility and further development of standardized international marketing strategies. Still, the review revealed a number of serious issues that need attention in order for international segmentation to fulfill its potential. Fig. 2 summarizes the most critical conceptual and methodological issues for international segmentation research. The case study illustrated these issues in a practical setting.

More specifically, construct equivalence does not appear to be generally examined in a rigorous way in international market segmentation studies. There is also a preponderance of country-level international segmentation studies. Such segments are not easy to link to marketing strategies as the actionability of such results per se is low (except when used in a two-stage procedure) and the within-country heterogeneity is ignored.

A key conceptual problem in international segmentation is that general (domain-specific) segmentation bases tend to be more (less) construct equivalent and yield more (less) accessible and stable segments, but tend to be less (more) actionable and responsive. The two-step approach may alleviate this problem. First, country clusters are identified using general bases, and subsequently, consumer segments are identified within a target cluster, based on domain-specific bases.

Although calibration equivalence does not pose great problems in published research, it is worrisome to note that translation equivalence is much less certain. International segmentation studies are not always careful in their linguistic translation and few studies provide any evidence on international pretesting of the measures. Score equivalence has not received much attention although methods for testing and correcting for response styles are available. On the other hand, sampling equivalence has attracted much attention in published research, and is well established.

There is a healthy trend toward model-based international segmentation research, based on specific models of conceptual and statistical models of consumer behavior. Such work increases the generalizability of the results over and above the specific countries studied, and makes it much easier to build on previous work, e.g., by refining the underlying theory, by relaxing statistical assumptions, and by improving the testing procedures. Finally, the issue of the sample size versus the population size has hardly received any attention in the literature. This can lead to seriously biased conclusions.

Thus, when we assess published international segmentation research on the different issues discussed in the paper, a mixed picture emerges. On some issues, much progress has been made, while on a number of other aspects, much remains to be done. As argued in this review, international segmentation can be improved by adopting and developing two-step, model-
based approaches, taking into account prior knowledge (e.g., contiguity of segments), managerial objectives, and theories of consumer behavior. Such studies should correct for response tendencies to ensure score equivalence and take implicit stratified sampling designs into account. Bayesian sampling-based techniques seem to offer much potential for the development of comprehensive international segmentation models that integrate conceptual and measurement issues in one unified framework.

6.2. Future research

We will conclude with a number of suggestions for future research. First, among the criteria for effective segmentation described in Section 3.2.1, we mentioned the temporal stability of segments. Over time, the number of segments, segment sizes, and structural properties of international segments may change. To our knowledge, this issue has not received rigorous attention. Some recent studies have modeled the dynamic properties of domestic segments though. Kamakura, Kim, and Lee (1996) extended a mixture model of logits to accommodate changes in segment sizes and structural properties of segments over time. Ramaswamy (1997) developed a latent Markov model to study the temporal shifts of preference segments. In addition to changes in size and structural properties of segments, this model allows for the formation of completely new segments. Other extensions have been proposed by Wedel and Kamakura (1998, pp. 159–179), where changes of segment sizes and segment-specific parameters (i.e., the segmentation basis) are captured by a latent curve across time. The integration of such approaches and advances in latent curve modeling (Meredith & Tisak, 1990; Steenkamp & Baumgartner, 2000) and time series modeling (Dekimpe & Hanssens, 2000) may provide powerful tools for studying temporal changes in international segments. Other extensions have been proposed by Wedel and Kamakura (1998, pp. 159–179), where changes of segment sizes and segment-specific parameters (i.e., the segmentation basis) are captured by a latent curve across time. The integration of such approaches and advances in latent curve modeling (Meredith & Tisak, 1990; Steenkamp & Baumgartner, 2000) and time series modeling (Dekimpe & Hanssens, 2000) may provide powerful tools for studying temporal changes in international segments. This allows researchers to move from static international segmentation solutions to dynamic solutions, to integrate changes in the marketplace into the firm’s marketing strategy, and to test substantive hypotheses concerning changes in the marketplace. For example, one could examine the often suggested, but never tested, hypothesis of global convergence of consumer needs and wants (Levitt, 1983; Yip, 1995).

Second, most international segmentation studies are of an exploratory nature and the segmentation bases are selected on ad hoc criteria. The segmentation results of such studies help marketers develop marketing programs, but they do not provide explicit guidelines for the development of optimal marketing programs, nor do they provide insights into the effectiveness of marketing mix instruments. In the domestic segmentation literature, mixture models of brand choice have been proposed that identify segments of consumers sharing similar price and sales promotion elasticities (e.g., Kamakura & Russell, 1989; Gupta & Chintagunta, 1994). Applying these models in an international context is not straightforward because they rely on scanner panel data of brand choices. For most product categories, the availability of particular brands will differ considerably across countries. In addition, currencies and average price levels differ among countries. This poses conceptual difficulties in interpreting and comparing brand parameters and elasticities across markets and suggests that more research are warranted to resolve this issue.

Third, there is a need for models that integrate international segmentation in managerial decision-making. The segmentation outcome should also be based on managerial objectives. Kumar et al. (1994) made an important step in this direction. It is the only study in which international market segmentation is explicitly integrated in a marketing decision support system. Bayesian decision models may also be fruitfully employed, combining statistical formulations of reality with objective functions to be maximized. More research is needed in this area, especially on international cost and profit structures, to develop formal policy-oriented models that provide allocations of marketing efforts that are optimal in terms of profits for international firms.

Fourth, with the advent of database marketing and e-commerce, attention has shifted to the identification of individual-level response parameters for the implementation of direct marketing strategies (e.g., Lee-flang & Wittink, 2000; Mahajan & Venkatesh, 2000). Companies like Amazon.com provide individualized recommendations, based on search and purchase behavior, to individuals around the world. On the other hand, although the Internet has reduced the costs of targeting individual consumers around the world, those costs are not negligible and will not become so
in the near future. Companies like Procter and Gamble, Clorox, and General Motors have pared down product assortments to target larger segments with a more limited variety of products (Business Week, 1996). Unilever is shedding 1200 of its 1600 brands to focus more on brands that appeal to global consumer segments. For some products like books or music, individual-level models and marketing on an international scale is feasible but for many other products, if not most, this is not the case. International segmentation still offers significant global scale advantages in production, distribution, and advertising, which are not shared by individual-level approaches and thus international segmentation strategies are called for. Future research could shed more light on the conditions that make international segmentation more profitable than targeting customers individually, and vice versa.

Fifth, our review of previous empirical studies demonstrated that segments of consumers exist in different countries. This was found to be true for various segmentation bases. Still, the studies upon which this result is based were limited to Triad countries, and most studies were conducted in a single continent. The existence of worldwide segments is a prerequisite for the successful implementation of truly global marketing strategies. Additional research is needed to assess the structure of international consumer segments worldwide. Some preliminary results are provided by a proprietary worldwide study, conducted for a large multinational under guidance of one of the authors. The methodology developed by Ter Hofstede et al. (1999) was applied to a sample collected in different continents, including both Triad and non-Triad countries. The results did not demonstrate any true global segments, but four out of five segments included a substantial amount of consumers from many countries in different continents. One segment was specific to China. The analyses revealed that failure to account for differences in population sizes would have masked the existence of a China-specific segment consisting of over 100 million consumers. More research is needed, involving other segmentation bases and rigorous segmentation methods applied to worldwide datasets.

Sixth, for the further development of international segmentation, it is important that score equivalence of segmentation bases are established. As discussed in Section 4.1.3, lack of score equivalence may mask similarities across borders and hamper the identification of cross-national segments. A two-step approach can be taken. In the first step, methods of Baumgartner and Steenkamp (1998, 2001), Rossi et al. (2001), and Steenkamp and Baumgartner (1998) can be used to correct responses for measurement bias. The corrected data can be used as input for international segmentation techniques in a second step. Ideally, the two steps are integrated in a simultaneous model for bias correction and international market segmentation. Bayesian sampling-based methods seem very promising for this purpose since they allow combining different submodels (e.g., measurement model and segmentation model) into one comprehensive international segmentation model.

Finally, an important issue that we did not discuss in this review is the effect that competition may have on international segmentation strategies. When targeting cross-national segments, the presence and nature of potential competitors will have implications for optimal competitive positioning and profitability of segments. Up to now, the international segmentation literature generally has ignored this issue. Not only is the structure of consumer heterogeneity more complex in international markets as compared to domestic markets, also is competition more heterogeneous. After all, the set of competitors operating in a particular cross-national segment is likely to include country-oriented companies pursuing domestic and multidomestic strategies. Recent model-based approaches for (domestic) market segmentation and competitive positioning (e.g., Chintagunta, 1994; Gruca, Sudharshan, & Kumar, 2001; Wedel & DeSarbo, 1996) should be extended to the competitive arena of international firms. Future research should develop models that simultaneously segment the market and identify an optimal positioning in a cross-national segment while taking into account varying competitive positions in different countries. Such models might allow for country-specific positioning, but, given the emergence of pan-regional and global strategies, the real challenge will be to estimate optimal global positioning, which may entail a less-than-optimal positioning in particular countries. Such modeling efforts would further increase the managerial relevance of international market segmentation research.

In summary, a lot has been achieved in three decades of international marketing research. However,
the time of exploratory research should be over. Future research will benefit from taking a more rigorous, theory-grounded, model-based approach. There are a number of opportunities for and threats to the further development of international market segmentation research. We have offered suggestions how to deal with them and which issues require more research effort. It is our hope that this paper will stimulate further advances in this important area of marketing.

Acknowledgements

The authors are indebted to Michel Wedel, Wagner Kamakura, and the anonymous reviewers for valuable comments on an earlier version of the paper.

References


