DIFFERENT ENGLISHES? INVESTIGATING EQUIVALENCY OF THE AFFECTIVE-RESPONSE-TO CONSUMPTION SCALE AMONGST GEOGRAPHICALLY DISPARATE GROUPS OF ENGLISH SPEAKERS

Alexandra Ganglmair-Wooliscroft, University of Otago, Dunedin, New Zealand
Ben Wooliscroft, University of Otago

ABSTRACT

Satisfaction surveys conducted in global industries frequently have respondents fill in questionnaires in English, even though respondents are a mix of native and non-native English speakers. Using Kachru's (1985) concentric circles representing the cultural embeddedness of the English language, this study investigates whether such a mix of respondents consistently interpret emotional terms used to describe evaluations of their consumption experience. Affective-Response -to -Consumption (ARC) includes a variety of terms that can be used to describe feelings about the favorable/unfavorable evaluation of a consumption experience. The concept is applied to examine whether different groups of respondents, ranging from native English speakers of different origin (e.g. U.S.A. or UK) to those from countries where English is a foreign language, assign consistent meaning to emotional terms included in the ARC. Rasch Modeling, particularly useful and convenient when investigating sub-groups within a dataset, is used to examine whether results from subjects with different English backgrounds can be meaningfully integrated. The investigation finds that there are differences in the meaning of emotional terms between different English speaking groups, emphasizing the importance of investigating data equivalence when administering surveys in an international setting.

INTRODUCTION

Consumer satisfaction, long regarded as one of the cornerstones of marketing (Babin and Griffin 1998; Churchill and Suprenant 1982, Peterson and Wilson 1992, Taylor 2008, Yi 1990), has been the topic of numerous studies with many investigating antecedents or consequences of the concept (e.g. Anderson and Fornell 1994, Fornell 1992, Halstead, Hartman and Schmidt 1994, Halstead, Jones, and Cox 2007, Nyer 1998, Parker and Mathews 2001). However, with a few exceptions (e.g. Oliver 1980; 1981, Westbrook 1980, Westbrook and Oliver 1991), the measurement of the concept has received comparably less attention (Babin and Griffin 1998) although calls for its advancement have been made at regular intervals (Hunt 1977, Peterson and Wilson 1992, Diener and Fujita 1995). Affective -Response -to -Consumption (ARC) (Ganglmair-Wooliscroft 2007, Ganglmair and Lawson 2003a, 2003b) has been developed in light of these calls for further research.

ARC builds on a stream of research acknowledging the role of emotions in consumption research and views satisfaction as one of many possible responses to an experience (Bagozzi, Gopinath, and Nyer 1999; Fournier and Mick 1999; Giese and Cote 2000, Hicks, et al. 2005, O'Shaughnessy and O'Shaughnessy 2003, Soderlund and Rosengren 2004). The concept includes a multitude of favorable/unfavorable emotional

terms covering different intensity levels. ARC extends satisfaction measurement towards a more inclusive view of post-consumption emotion measurement). The concept captures the entire continuum of unfavorable/favorable emotional responses to consumption experiences ranging from highly negative emotional responses like "terrible," to extremely positive consumption responses like "fabulous" and "overjoyed;" covering a range of nuances between these extremes. It thereby offsets the common problem of highly skewed results in conventional satisfaction measurement (Diener and Fujita 1995, Fornell et al. 1996, Peterson and Wilson 1992).

In global industries or within multinational companies, measures of satisfaction must deal with another challenge, as respondents frequently possess different cultural backgrounds. Previous research has investigated the suitability of different measurement approaches when investigating data equivalence in multi-lingual datasets using Structural Equation Modeling (e.g. Steenkamp and Baumgartner 1998) or Rasch Modeling (Ewing, Salzberger, and Sinkovics 2005; Salzberger and Sinkovics 2006). However, data equivalence is not only an issue in multi-lingual surveys but also in surveys conducted in one language, often English, with different groups of native speakers who use their own form of colloquial English as well as non-native speakers.

English native speakers from distinctive geographic regions (e.g. the U.S.A. and the UK) might use different emotional terms to express the same emotional state; or the same emotional term might be used by culturally distinct groups of English speakers to express different emotional intensities. Additionally, in survey research conducted in a multinational setting (e.g. in multinational companies or in global industries like tourism), non-native speakers are frequently asked to fill in questionnaires in English, bringing in their own culturally influenced interpretation of words expressing emotional states. This poses a psychometrical as well as a managerial challenge in post-consumption emotion research.

Usiing Rasch Modeling (Rasch 1960/80), a measurement approach that is particularly suitable when investigating equivalence between different groups in a dataset (Salzberger 2009, Salzberger and Sinkovics 2006), this study seeks to establish to what extent survey results from different English speaking groups can be meaningfully combined. It examines the equivalence of meaning and intensity of ARC's favorable/unfavorable consumption emotions for different English speakers using English language data collected from U.S. Americans, British, Indians, Northern Europeans and other European citizens.

Previously, the concept has been successfully applied when investigating experiences with an excursion train (Ganglmair and Lawson 2003a, 2003b) and cell-phone U.S.A.ge (Ganglmair-Wooliscroft 2007). Extending ARC to another context; experiences with long-haul flights, is a secondary goal of this study.

Affective -Response -to -Consumption

Affective -Response -to -Consumption (ARC) emphasizes the role of affect and emotional responses to consumption (Bagozzi, et al. 1999; Fournier and Mick 1999; Gardial, et al. 1994; Giese and Cote 2000; Hicks, et al. 2005, O'Shaughnessy and O'Shaughnessy 2003). Starting from the term *satisfaction* the concept includes emotional terms that are more or less favorable responses to an experience. ARC thereby does not consider these emotional terms as predecessors of satisfaction as advocated by an important research stream in marketing (Mano and Oliver 1993; Oliver 1997; Westbrook and Oliver 1991), but regards satisfaction as one of a large number of emotional responses to an experience. This is in line with studies that struggled to discriminate between satisfaction and other

closely related constructs (Gardial, et al. 1994; Giese and Cote 2000, Sonderlund and Rosengren 2004) which lead Bagozzi and his colleagues (1999, p.201) to conclude that ".. it is unclear whether satisfaction is phenomenologically distinct from many other positive emotions" and to question whether it is a "unique, fundamental construct in and of itself."

Usiing ARC, the emphasis shifts from one term - *satisfaction* - towards a multitude of terms, many of them highly positive. The inclusion of these terms helps to overcome the high positive skewness and limited discrimination observed in many satisfaction studies (Fornell, et al. 1996, Peterson and Wilson 1992).

Viewing satisfaction as a consumption emotion is also consistent with the psychology literature where satisfaction, although not a basic (Ortony and Turner 1990) or higher order emotion (Shaver, et al. 1987; Storm and Storm 1987), is described as an affective concept[1], very similar to pleasure, joy and happiness (Lewis, Haviland-Jones, and Barrett 2008) and is included in seminal classification studies of emotions by Shaver, et al. (1987) and Storm and Storm (1987).

In addition, the majority of psychology (and marketing) research conducted relies on questionnaires and self-reports of emotions, where it is inevitable that respondents use analytical thought which is necessary for introspection as well as for formulating the required written or verbal answer (Buck and Georgson 1997). Neuro-psychological findings further question a strict distinction between cognitive and emotional/affective constructs as they consider emotional feelings and cognitive judgments to be heavily intertwined and interdependent processes (Damasio 1994; Forgas 1995, Goleman 1996; LeDoux 1996, Pham, Cohen, Pracejus, and Hughes 2001, Slovic, Finuncane, Peters and MacGregor 2007, Zajonc 1980).

The dimensionality of emotions is another ongoing discussion in psychology (Russell and Carroll 1999) and marketing (Babin and Griffin 1998, Maddox 1981, Westbrook and Oliver 1991). While there is evidence that satisfaction and dissatisfaction or happiness and sadness are independent emotions that can be experienced simultaneously (Dube, Belanger and Tudeau 1996, Herzberg, Mausner and Synderman 1959, Maddox 1981; Swan and Combs 1976), some studies suggest that the experience of complex simultaneous emotions is the exception rather than the norm (Green, Goldman and Salovey 1993, Kahneman 1999, Russell and Carroll 1999) and that people are "either happy *or* sad a lot of the time" (Larsen, McGraw, and Cacioppo 2001, p.684, emphasis added). Russell and Carroll (1999, p.25) therefore conclude that "for theories about affective feelings, bipolarity is a reasonable assumption".

Compared to other scales used to investigate affect like the Differential Emotions Scale (DES) (Izard 1977) or Pleasure Arousal Dominance (PAD) Scale (Mehrabian and Russel 1974), ARC (Ganglmair and Lawson 2003a, Ganglmair-Wooliscroft 2007) specifically includes items relevant for investigating consumption experiences rather than the entire range of human emotions (DES) or responses to environmental stimuli (PAD) and it has more positive emotional terms.

In order to provide a full set of emotions that are experienced in a consumption situation, Richins (1997) developed the Consumption Emotion Set (CES). CES covers the entire spectrum of emotions that "serve as antecedents, consequences, or moderating variables" in a consumption context (Bagozzi, et al. 1999, p.190), including items like loneliness and guilt that emphasize the state of the individual rather than the evaluation of the consumption experience itself.

[1] While brain activities during an emotion generally remain outside consciousness, the conscious experience or awareness of emotions is frequently referred to as *emotional feeling* or *affect* (Cacioppo and Bernston 1999, LeDoux 1996)

Acknowledging the historic importance of satisfaction in marketing research, ARC builds on and extends satisfaction research by providing detailed information about the favorable/unfavorable outcome of an experience. ARC thereby shifts the focus from one, (arguably a) mildly positive term, satisfaction, towards the inclusion of many often highly favorable assessments, reducing the positive skewness frequently found in satisfaction studies (Fornell et al. 1996, Peterson and Wilson 1992).

English Language Cross-National Studies

Cross-national surveys generally include sub-samples containing respondents from different countries with nationality often taken as a not ideal, but acceptable approximation of cultural affiliation (Salzberger and Sinkovics 2006). Multi-lingual datasets have to be investigated in terms of qualitative and quantitative data equivalence in order to draw meaningful comparisons and conclusions from the results (Salzberger, Sinkovics and Schlegelmilch 1999; Singh 1995; van de Vijver and Tanzer 2004).

Since the mid-1950s, English has become a truly global language (Crystal 1999). It is spoken by more people than any other language and has been referred to as the most desirable *lingua franca* of our time (Crystal 1998; Qiong 2004). Although figures are vague, experts estimate that more than 400 million people speak it as their native language, out of which about 320 million live in the U.S.A. and 56 million in Great Britain (Qiong 2004). However, more people speak English as a foreign or second language than those speaking it as their native language (Mondiano 1999).

In the 1980s Kachru (e.g. 1985, 1986) proposed a model of three concentric circles representing the historical and political spread of the English language. The *Inner Circle* contains countries where English is the mother tongue (e.g. Australia, Ireland, New Zealand UK, and U.S.A.). Within countries of the *Inner Circle*, the English language reflects the country's history and culture and although similarities are overwhelming, differences in pronunciation, style, grammar or choice of vocabulary do exist (Ilson 1985). The *Outer Circle* refers to countries that were typically British or U.S. colonies and retained English as a second, institutionalized or official language (e.g. India, Kenya, Singapore). Finally, the *Extended Circle* consists of countries where English is spoken as a foreign language (Bruthiaux 2003, Kachru 1985, Qiong 2004). The distinction between *Outer* and *Extended Circle* is sometimes seen as problematic (Bruthiaux 2003, Lowenberg 2002); e.g. in Scandinavia, according to Kachru's original model a country of the *Extended Circle,* English is spoken fluently by a large portion of the population. Some researchers are therefore suggesting tha these countries be classified as between the *Outer* and *Extended Circle* (Bruthiaux 2003; Qiong 2004).

Kachru's (1985) model was also criticized for relying too much on historical, political and geographical boundaries (Bruthiaux 2003; Lowenberg 2002). However, in the opinion of many it is still a reasonable starting point for investigating the spread of the English language (Bruthiaux 2003, Mondiano 1999) as it emphasizes the broad cultural background of English speakers all over the world (Kirkpatrick 2007).

Survey research in global industries or within multinational companies frequently encounters respondents that are a mix of native and non-native English speakers. Many of these customers are happy to fill in a questionnaire in English, although their cultural and/or linguistic background is very different as they might belong to any of Kachru's (1985) circles. It is important for academics, interested in advancing the quality of global satisfaction measures, and for practitioners, who might use results from

these surveys as benchmarks or as dependent variables in subsequent model building, to establish whether for example U.S. Americans and Europeans, for whom English is a second language, describe the same underlying emotional intensity when they claim to be *happy*, *delighted* or *dissatisfied*.

The main aim of this study is to investigate to what extent combining English survey responses, gained from diverse nationalities with different English language backgrounds, provide psychometrically sound and managerially useful results. A pre-requisite of the main aim of this study, and a secondary goal, is the extension of ARC towards a new context: investigating experiences on long-haul flights. Testing ARC in another setting represents a step in working towards developing a generalizable base of ARC items that are applicable in many different contexts.

METHODOLOGY

Satisfied passengers are considered a key competitive advantage in the tourism industry and many airlines assess the concept on a regular basis (Chen, 2008; Oyewole 2001). Due to the characteristics of this industry, a multi-national customer mix is frequently found. The current study is based on 578 questionnaires collected by a European airline on long-haul flights between Europe and the United States and Europe and India.

This particular airline collects questionnaires on approximately ten pre-determined seats per aircraft to allow an efficient and non-interruptive collection process of the questionnaires. In the current sample, this resulted in 317 responses from U.S. citizens and 261 questionnaires collected from other nationalities, a sample split that, according to the airline, is typical of the proportion of passenger responses generally received on their English surveys.

The U.S.A.. is part of Kachru's (1985) *Inner Circle*, but due to the size of the group, U.S. citizens were retained as a separate group. Amongst non-U.S. citizens, 22 respondents come from the United Kingdom, Australia, and New Zealand, countries of Kachru's (1985) *Inner Circle*, and 87 respondents come from India, a country of the *Outer Circle* and the third largest English speaking nation after the U.S. and the UK (Kachru 1986). Thirty questionnaires were filled in by Scandinavians, classified between the *Outer* and *Extended Circle* in order to acknowledge their high English proficiency (Bruthiaux 2003). The remaining 122 passengers, labeled *Combined English Speakers*, are citizens of an array of mostly European countries with various linguistic and cultural backgrounds, representing a very heterogeneous customer group.

For this study the original ARC items, previously applied to an excursion train (Ganglmair and Lawson 2003a, 2003b) and adapted to investigate ARC with cell-phones (Ganglmair-Wooliscroft 2007), were further modified and refined to suit the current investigation of ARC with long-haul flights. In order to acknowledge the different linguistic backgrounds of airline passengers, three experts were called upon (Rossiter 2002), including a U.S. native speaker and two Europeans who have lived in a country of the *Inner Circle* for several years and have an interest in emotion research. In a collaborative, iterative process (Douglas and Craig 2007), the team went over the original ARC item list and evaluated the suitability of items when investigating passengers' emotional responses about a long-haul flight. Although it is acknowledged that the final ARC list includes some terms (e.g. *over the moon*) that might not be commonly used in particular English speaking regions, the experts agreed that a majority of people would still understand the general meaning of those items. Rasch Modeling, the statistical process used to investigate the equivalence of items across different English speakers in this study, is also able to detect if the meaning of an item is inconsistent between groups of English speakers. The statistical technique will be discussed in the following paragraphs.

Table 1 shows the ARC question and the 26 dichotomous emotional terms that are included in the scale for long-haul flights. The 17 positive and 9 negative emotional terms were presented to passengers in randomized order.

ARC for long-haul flights was investigated using Rasch Modeling (Rasch 1960/80) an alternative measurement approach that has been used extensively in several social science disciplines, particularly education and psychology (Bond and Fox 2007) but has only gained recent attention in marketing (Ewing, et al. 2005; Salzberger 2009; Salzberger and Sinkovics 2006; Singh 2004; Soutar and Cornish-Ward 1997). The Rasch Model is a mathematically elegant model that is believed to represent an ideal form of measurement (Bond and Fox 2007; Fischer and Molenaar 1995; Linacre 1992; Rost 2001).

TABLE 1

Items Included in ARC on Long –Haul Flights

How do you feel about this flight? Please consider every word and tick all those that describe your feelings about this flight.			
average	dreadful	great	over the moon
contented	ecstatic	happy	pleased
delighted	euphoric	horrible	satisfied
disappointed	excellent	in seventh heaven	superb
discontented	fabuloU.S.	magnificent	terrible
displeased	fantastic	overjoyed	unhappy
dissatisfied	good		

Note: Items are presented to each respondent in randomized order.

Rasch Modeling has been referred to as a probabilistic alternative to the deterministic Guttman Scaling (Wright 1997) and belongs to the family of logit models. The Rasch Measurement Model provides a mathematical model for measurement, as follows:

$$P_{ni}(x_{ni}=1)=\frac{e^{\beta_n-\delta_i}}{\left(1+e^{\beta_n-\delta_i}\right)}$$

In the original binary Rasch Model depicted here, the probability of a positive response (agreeing to an item, in comparison with disagreeing) depends on the endorsability of that item, referred to as item characteristic and represented by the item location parameter δ_i, and the person characteristic, operationalized by the person parameter β_n (Andrich 1988; Bond and Fox 2007; Wright and Stone 2004).

The model is theory centered and derived a-priori to define measurement (Wright 1992). Empirical data will always diverge from this ideal form to a certain degree (Salzberger, et al. 1999; Wright 1997) and researchers are encouraged to learn from ill-fitting data (Wright 1997) and work

towards a better fit of the data to this ideal mathematical measurement model (Bond and Fox 2007; Linacre 1999).

In other words, and contrary to Classical Test Theory where additional parameters can be added so a model explains the data better, by applying the Rasch Model, a theoretically driven measurement approach is taken where the model is considered an ideal form of measurement and the characteristic and quality of the data need to be examined in order to determine if measurement is actually achieved. If the Rasch Model does not provide satisfactory results, it is the researcher's task to investigate the quality of the data, rather than to adapt the model to better suit the current purpose.

When investigating the fit of the data to the Rasch Model, several fit statistics and visual displays need to be investigated. Contrary to approaches based in Classical Test Theory that refer to aggregate statistics like variances, co-variances and means, the Rasch Model emphasizes the fit of individual items (Salzberger and Sinkovics 2006) examined using Chi-square tests that compare the actual scores based on proportions in a certain number of groups to the expected scores based on probabilities (*a non-significant result is preferable*) (Andrich, Sheridan, and Lou 2003a). In addition, as these tests uncover any divergence from the model, a special fit residual is provided, revealing items that systematically under- or over-discriminate when compared to the theoretical model (*fit residuals smaller than ± 2.5 are acceptable*) (Andrich, et al. 2003a; Salzberger 2009)).

An overall fit statistic is provided by summing the result of individual Chi-square statistics over all items (Andrich, et al. 2003a). Additionally, a Person-Separation-Index, interpreted similarly to classical reliability statistics (Salzberger 2009), and an overall Power-of-Test-of-Fit (Andrich et al. 2003a) is provided. The software program, RUMM2020 (Andrich et al. 2003b), further offers a number of graphical displays to visually inspect the fit of the data to the model.

When dealing with a broad concept like ARC, Rasch Modeling provides advantages over conventional approaches based in Classical Test Theory as the model works best when items differ in intensity and are spread over the dimension under investigation (Embretson and Reise 1999, Singh 2004). As ARC sets out to capture the entire favorable–unfavorable evaluation of an experience, we are of the opinion that Rasch Modeling is particularly suitable.

Additionally, the model can be applied to all possible scale formats and deals readily with dichotomous data. The use of dichotomous (yes/no) answer categories is considered beneficial as it reduces the impact of culturally and/or linguistically influenced interpretation of answer categories and reduces the cognitive demand on respondents, particularly with groups of respondents in this study having different English language proficiency.

Salzberger and his colleagues provide extensive investigations of the Rasch Model and other, more conventional approaches based in Classical Test Theory (CTT) most notably Steenkamp and Baumgartner's (1998) Multi-Group Confirmatory Factor Analysis (MG-CFA) approach when dealing with cross-national/cultural datasets. They come to the conclusion that the Rasch Model is particularly useful and convenient when investigating data equivalence in cross-cultural research (Ewing, et al. 2005, Salzberger and Sinkovics 2006, Salzberger 2009).

The most important aspect of the Rasch Model is *Specific Objectivity* (Rasch 1977). It says that if the data fits the model, item parameters and person parameter are independent from one another (Fischer and Molenaar 1995) and that the model has to be "invariant against all possible groupings of respondents" (Salzberger and Sinkovics 2006, p.395). *Specific Objectivity* explains why the Rasch Model is particularly useful and efficient when investigating cross-cultural

datasets: *Testing whether items are interpreted in the same way by different groups within the dataset, e.g. airline passengers with different English-speaking backgrounds, and therefore testing whether the model is invariant for different national/cultural groupings in the dataset is merely a special case for testing if the data fits the model* (Salzberger and Sinkovics 2006).

Non-invariance is indicated by Differential Item Functioning (DIF). DIF is found if an item has different meaning and therefore takes a different position for respondents from different groups, or when an item does not fit the model at all for one or more groups in the dataset. DIF is tested using a two-way ANOVA (Andrich, et al. 2003a, 2003b) and can also be investigated visually. If an item functions differently for groups in the data set (DIF occurs), the item is split and investigated separately for each respective sub-group, a feature conveniently enabled in the software program (Andrich, Sheridan, and Lou 2003b). This investigation will show whether the item is used inconsistently between sub-groups, but consistently within the group, in which case the item is split and retained as a separate item for each group. If it is used inconsistently within the group, the item is eliminated for the respective sub-group in the dataset. If an item is not used consistently within any sub-group, it is removed completely from further analysis.

The following analysis section is split into two parts. In the first part, the psychometric properties of ARC are examined and the concept's general suitability when measuring emotional responses in the context of long-haul flights is confirmed. Having established the suitability of ARC to measure subjects' experiences with long-haul flights, the second part will explore the perceived meaning and intensity of favorable/unfavorable ARC consumption emotions for different English speakers.

Investigation of Psychometric Properties of ARC in the Context of Long-Haul Flights

For the analysis, all 578 cases were entered into the Rasch Modeling software RUMM2020 (Andrich, et al. 2003b). A Rasch Analysis consists of an iterative process where at each step the individual items are investigated using fit statistics and graphics provided and the overall fit statistics are examined. Taking all fit indices as well as visual displays into account, items that do not fit the model to a satisfactory extent are removed one at a time.

When all 578 cases were included in the analysis, initial Rasch Model results were poor with individual item and summary statistics showing unacceptable results. As required by the Rasch Model and discussed above, the quality and suitability of the data to measure ARC had to be further investigated. DIF analysis suggested that the inadequate results were driven by the *Combined English Speakers,* a highly heterogeneous sub-group representing Kachru's (1985) *Extended Circle.* The group consists of respondents from a variety of European countries who speak English as a second language. As expected, these respondents incorporate their various culturally influenced interpretations of emotional terms included in ARC, leading to inconsistent results within that group. This subgroup of 122 respondents was subsequently removed from further analysis (see discussion in the next section).

The reduced dataset fit the Rasch Model much better. During the following iterative process investigating the fit of individual items to the Rasch Model, a number of ARC items were highlighted that did not fit to a satisfactory extent. Each item was investigated following the process outlined above examining overall numerical and graphical fit indicators provided by the software program and checking for DIF. If DIF occurred the item was split and re-investigated. Only after the described process suggested that the item was used

inconsistently within all subgroups it was removed from further analysis.

This led to the elimination of *good, satisfied, contented, excellent, fantastic, disappointed, discontented* and *dissatisfied.* All remaining items fit the Rasch Model to a satisfactory extent, showing fit residuals within ±2.5 and non-significant Chi-square probabilities after Bonferroni adjustment (Andrich, et al. 2003a). At the end of this process, every item was also checked for item bias (Differential Item Functioning) between different English speaking groups (Salzberger 2009, Salzberger and Sinkovics 2006) with *happy* showing some Differential Item Functioning (DIF) between groups and the item was split into two groups for U.S. versus other English speakers (to be discussed in detail in the next section).

The final ARC scale for long-haul flights, with all fit statistics at an acceptable level (Andrich, et al. 2003a), shows results largely consistent with previous applications of ARC evaluating the unfavorability/favorability of an experience with an excursion train (Ganglmair and Lawson 2003a, 2003b) or with a cell-phone (Ganglmair-Wooliscroft 2007). **Table 2** shows the overall fit statistics of the ARC Scale for long-haul flights for respondents with different English backgrounds. The overall Chi square value is not significant, which is preferred, and the overall Test of Fit based on the Person Separation Index is *Good* (categories range from *Too Low/Low/Reasonable/Good* to *Excellent*).

Table 2

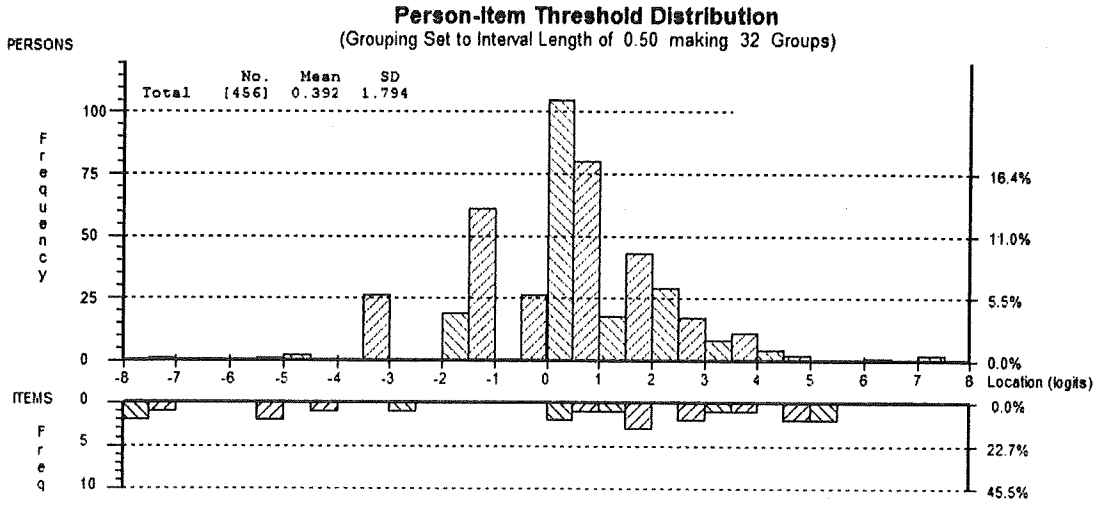
Summary Statistics for ARC Long-haul Flight

Item Trait Interaction		Reliability Indices	
Total Item Chi Square	183.9	Separation Index	0.69
Total Deg of Freedom	166	Power of Test-of-Fit	GOOD
Total Chi Square Prob.	0.16		

The order of items, and implied intensity of emotional terms included in ARC, also agrees with previous applications of the concept (Ganglmair and Lawson 2003a, Ganglmair-Wooliscroft 2007). **Figure 1** shows the distribution of items (bottom half of the graph) and the distribution of persons (top half of the graph) on the ARC continuum for long-haul flights. As can be seen, items spread across the entire ARC dimension and respondents are normally distributed along the continuum. ARC therefore provides more detailed information about consumers' evaluations of long-haul flights and stronger discrimination among respondents than traditional satisfaction scales that are frequently heavily skewed (Peterson and Wilson 1992). A discussion of specific items in the ARC continuum follows.

The statistics confirm the applicability of ARC when investigating experiences with long-haul flights and indicate that the items fit the theoretical Rasch Model well. The following section will examine the interpretation and meaning of ARC items for different English speaking groups in greater detail.

Figure 1
Item/Person Location for ARC with Long-haul Flights



PERCEIVED INTENSITY OF FAVORABLE/UNFAVORABLE ARC ITEMS FOR CONSUMERS WITH DIFFERENT ENGLISH LANGUAGE SKILLS

As discussed, the dataset contained five groups of English speaking respondents who differ regarding their geographical origin and their level of English proficiency: a group of respondents from the U.S.A. (part of Kachru's *Inner Circle*); respondents from other countries with English as their first language (UK, Australia and New Zealand – also part of Kachru's *Inner Circle*); respondents from India, a country where English is considered an official second language (Kachru's *Outer Cirlce*); respondents from Scandinavia, countries who are well known for their high English proficiency (placed in between *Outer* and *Extended Circle* (Bruthiaux 2003, Lowenberg 2002) and other respondents of various nationalities (Kachru's *Extended Circle*) who were nevertheless prepared to fill in a questionnaire in English, called *Combined English Speakers*.

The fifth group of *Combined English Speakers* consisted of mainly European respondents (e.g. Albanians, Lithuanians, and Spanish) with no more than seven respondents from each country, making additional groupings impossible due to insufficient sizes for analysis. Airline passengers from these countries were happy to fill in a questionnaire in English but interpreted the emotional terms included in the ARC scale inconsistently and ascribed different meaning to the emotional items included. As a consequence, English speakers from countries in the *Extended Circle* – where English is truly a foreign language – cannot be included in the same analysis with other English speaking groups and had to be removed from the following analysis[2]. Exploring emotional experiences on a detailed level therefore provides a challenge for quantitative survey research on long-haul flights that will need further investigation.

[2] As previously discussed, the Rasch Model provides an ideal form of measurement. If data does not fit the model, it is the quality of the data that has to be questioned – in this case the inconsistent interpretation of the meaning of an emotional term – rather than the model that needs to be adjusted by adding another parameter, etc.

After deletion of that sub-group, the remaining analysis consisted of U.S. citizens, the group of the *Inner Circle* (UK, Australia, and New Zealand), *Outer Circle* (India) and Scandinavia – countries situated between the *Outer* and *Extended Circle*. **Table 3** shows the item location for the final ARC (long-haul flights) for different English speaking groups. The location is measured in logits and can be interpreted as an interval scale (Peck 2000, Soutar and Cornish-Ward 1997).

Items representing the feelings attached to a most negative experience are *terrible, horrible* and *dreadful*, followed by *unhappy, displeased, disappointed* and *average*, with the last term still indicating anegative evaluation of the experience. This evaluation is followed by *happy* in its variations (to be discussed in more detail in the following section), *pleased, great* and *delighted.* This weaker positive area of the ARC continuum contains terms that generally form the most positive emotional states incorporated in traditional satisfaction measurement (Oliver 1997) including items included in the frequently used Delighted–Terrible Scale (Westbrook 1980, Westbrook and Oliver 1991, Danaher and Haddrell 1996). Additionally, ARC provides the opportunity for respondents to express truly outstanding experiences, described by feelings such as *superb, fabulous., overjoyed* or *euphoric.*

Table 3
Item Location for Different English-Speaking Groups: ARC with Long-Haul Flights

Item	Location
Horrible	-7.79
Terrible	-7.79
Dreadful	-7.08
Unhappy	-5.34
Displeased	-5.31
Disappointed	-4.44
Average	-2.98
Happy: India *(Outer Circle)*	**0.20**
Pleased	0.50
Happy: UK, Australia, New Zealand *(Inner Circle)*	**0.79**
Happy: U.S.A.	**1.25**
Happy *Scandinavia (between Outer & Extended Circle)*	**1.67**
Delighted	1.68
Great	1.71
Superb	2.56
FabuloU.S.	2.84
Magnificent	3.42
Overjoyed	3.97
In 7th heaven	4.75
Ecstatic	4.90
Over the moon	5.14
Euphoric	5.36

As has been found in previous applications (Ganglmair-Wooliscroft 2007, Ganglmair and Lawson 2003a), items like *satisfied, good,* or *contented* do not fit the Rasch Model leading to an item-gap at the weakly positive area of ARC. This is probably due to respondents unsystematically using items like *satisfied* interpreting it as either positive – 'I am at least satisfied' – or mildly negative – 'I am just satisfied'. Further study will be required to explain this phenomenon.

During the scale development process, every item was checked for DIF between different English speaking groups as suggested by Salzberger (2009) and Salzberger and Sinkovics (2006). The uneven group size between U.S. citizens and other English speakers has thereby been taken into account when interpreting results with a conservative interpretation of the Two-Way Anova and an emphasis on the graphical display provided by the software (Andrich, et al. 2003b).

The term *happy* showed some DIF between the included English speaking groups – indicating that the term is assigned different emotional intensity depending on the geographic and cultural background of the linguistic group. *Happy* was therefore first split into two groups: U.S. Americans and other English speakers. This resulted in *happy* fitting well for the U.S. sample, but not well for the remaining respondents as can be seen in **Figure 2 and 3** with the group means following the Item Characteristic Curve (ICC) closely for the U.S. sample (Figure 2) while they diverge considerably from the ideal curve in **Figure 3**. An ICC curve represents the ideal position of an item determined by the Rasch Model; group means are represented by dots.

When the item was split further to account for *Inner, Outer,* and *Extended Circle* the item-fit was improved considerably with all fit statistics within acceptable limits. The term *happy* is therefore interpreted substantially different between English speaking groups, while its assigned meaning is consistent within these groups. Respondents from India interpret it as the least positive emotional evaluation of an experience, followed by respondents from the U.S.A. who interpret the term *happy* slightly more favorable than UK, Australian and New Zealand citizens. Scandinavians ascribe the most positive emotional intensity to the term *happy.*

Figure 2
***U.S.* sample: ICC and Observed Group Means for Happy**

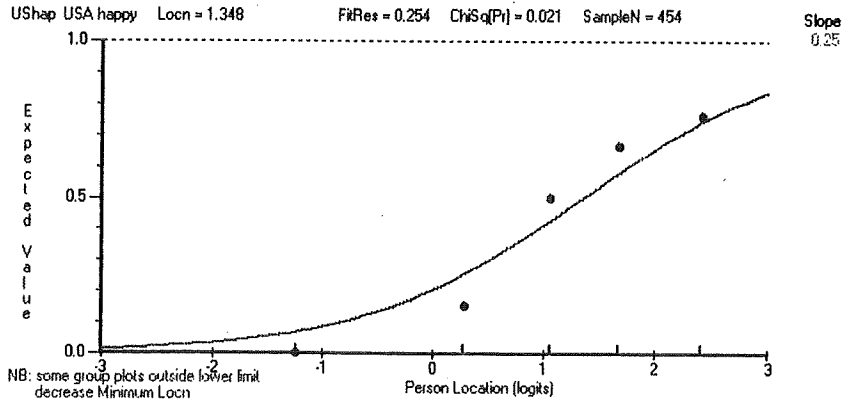
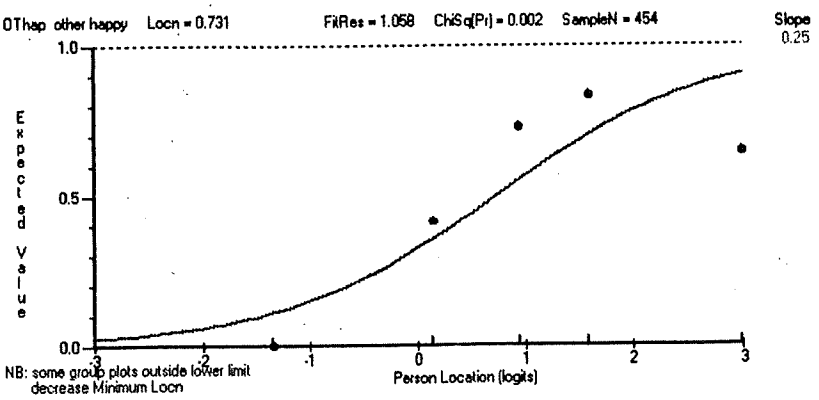


Figure 3

***Other English Speakers* sample: ICC and Observed Group Means for Happy**



All other emotional terms included in the ARC scale for long-haul flights are used consistently within and between the investigated English speaking groups, showing that the terms carry the same meaning and emotional intensity.

DISCUSSION AND CONCLUSION

This study investigated to what extent English survey responses completed by diverse nationalities with different English language backgrounds provide psychometrically sound and managerially useful results and, as a pre-requisite of that main aim, to examine whether Affective Response to Consumption (ARC) can be applied in the context of measuring experiences with long-haul flights.

ARC extends global satisfaction measures by including a range of emotional terms that can be used to describe an experience. The high number of positive and highly positive emotional terms reduces the skewed answer pattern frequently found in traditional satisfaction measures. The airline traditionally includes an overall satisfaction question in its questionnaires asking respondents to indicate their agreement to *Overall I feel at ease on this flight today* on a 5-point scale. Compared to that traditional question, ARC provides substantially less skewed results (skewness ARC: -0.23; Overall I feel at ease: 1.17). Previous studies, applying ARC when investigating experiences with an excursion train (Ganglmair and Lawson 2003) and with cell-phones (Ganglmair-Wooliscroft 2007) included the Delighted-Terrible Scale (Andrews and Withey 1976) and came to very similar results. Based on this series of findings, we believe that ARC provides a psychometrically superior alternative to traditional overall satisfaction measures.

Extending ARC to investigate experiences with long-haul flights confirms the applicability and suitability of the outlined scale development process to measure Affective Response to Consumption in different industries. It is expected that after further extension of ARC to measure favorable/unfavorable consumption experiences in a variety of other industries, a set of items will emerge that builds the transferable base for measurement of the ARC concept, while the outlined scale development process enables the addition of industry specific items providing context specific richness and detail.

To arrive at a psychometrically sound ARC scale capturing experiences with long haul flights, the terms *good, satisfied, contented, excellent, fantastic, disappointed,*

discontented and *dissatisfied* had to be removed as they did not fit the Rasch Model to an acceptable extent. The elimination of weakly positive items like *good* and *satisfied* has also been seen in a previous application of ARC with excursion trains (Ganglmair and Lawson, 2003b). The Rasch Model highlights that these items are used inconsistently across and within groups. As is common in quantitative approaches, while it is possible to detect that items do not perform as expected, it is not possible to conclude why. It is speculated that people might variously consider these items as mildly positive ('at least satisfied') or mildly negative ('just satisfied') but further qualitative analysis has to be carried out to establish less tentative explanations. The elimination of the term *satisfied* in ARC applications is in line with previous research that questions the extraordinary role of satisfaction in marketing as customers generally do not use the term to describe an experience (Fournier and Mick 1999; Giese and Cote 2000).

A reviewer pointed out that the eliminated items *contented, satisfied, discontented* and *dissatisfied* might be an indication of satisfaction being an empirically different construct from ARC. Indicators regarding residual statistics provided by the Rasch software do not support this conclusion. Additionally, neither a separate Rasch Analysis carried out on the eliminated items, nor the inclusion of various super-items (consisting of combinations of *contented, satisfied, discontented,* and *dissatisfied*) provided acceptable results. Additional research will have to investigate if satisfaction and ARC are indeed empirically separate dimensions.

The main aim of this study is to investigate to what extent survey results from different English speaking groups provide psychometrically sound and subsequently managerially useful results. Overall, this study shows that exploring potential differences in answer behavior of sub-groups is important when English surveys are filled in by respondents with different cultural backgrounds. The results suggest that native English speakers or respondents from countries that have an established, strong relationship with the English language (e.g. Northern Europe) use emotional terms associated with the favorability/ unfavorability of an experience consistently.

However, respondents from Kachru's (1985) Extended Circle, who have very different cultural and linguistic backgrounds, prescribe inconsistent meaning to emotional terms included in ARC. This unpredictable interpretation and assignment of different emotional intensity to individual words provides a challenge for quantitative survey research in an international context. Even if respondents are happy to fill in a questionnaire in their second-language, the results may not be comparable with native English speakers' responses.

The results suggest that the inclusion of non-native English speakers in survey results should be treated with caution and other options might have to be considered. For example, if a majority of non-native English speakers share another common native language (e.g. Spanish) a translated version of the questionnaire could be tested. If respondents speak a variety of native languages, it might be necessary to revert to simple questions or alternative answer formats. However, any alternative approach has to establish whether equivalence of responses is achieved before combined results provide psychometrically and managerially useful information.

When sub-groups of English native speakers or groups from countries with strong cultural ties to the English language are compared, the interpretation of emotional terms included in ARC is very similar. Crystal (1999) suggests that English native speakers adjust the local version of their language when acting in an international context, which seems to be the case here. *Happy* is the only term included in the ARC

continuum that signifies different emotional intensity depending on membership in these English speaking groups with Indian respondents more readily describing an experience as *happy* followed by the three groups; the *Inner Circle* (UK, Australia, and New Zealand), U.S. Americans and Scandinavians. The observed difference in intensity assigned to *happy* in various language areas is a reflection of cultural influences on language and appears to fit the stereotypical characteristics of these groups of people. Indians, for example, often pictured as vibrant, outgoing and joyful, interpret happy as an emotional state of relatively less positive intensity and Indians would describe feelings they gain from an experience relatively more easily as *happy*.

Respondents from the U.S.A. generally interpret the term *happy* as slightly more positive than respondents from other countries within Kachru's (1985) *Inner Circle.* This result might be somewhat surprising given the stereotypical characteristic of the reserved British, who make up the majority of this group. However, the generally similar meaning *happy* has in the U.S. and other countries within the *Inner Circle* reflects the historical connection between these countries but also the strong ties in music, film and other areas of popular culture. Scandinavians are frequently described as reserved and not showing emotions easily, a result that is also reflected in the ARC scale, where they assign the strongest positive emotional meaning to the word *happy*.

Once the different interpretations of *happy* is taken into account, native speakers and respondents from countries with strong links to the English language can be meaningfully combined. When the position of a respondent from the U.S.A. and an Indian citizen on the ARC scale reflects the same level of emotional state experienced, in spite of their identical inherent emotional state, compared to the Indian respondent, the American respondent will find it harder to endorse the item *happy.* Further research can therefore use these ARC results to investigate different antecedents or consequences of ARC.

The Rasch Model (Rasch 1960/80) has been shown to provide an effective and efficient method for examining data equivalency between English speaking groups. The model conveniently enables the comparison of subgroups that might present different or inconsistent answer patterns. Marketing academics and managers use global satisfaction measures such as ARC as benchmarks or as dependent variables in models explaining antecedents of the construct or subsequent behavior. Being aware of potential inconsistencies in groups' answer behaviors is essential to ensure the quality of measurement instruments in cross-national satisfaction and post-consumption emotion studies.

REFERENCES

Anderson, E. W. and C. Fornell. (1994), "A CU.S.tomer Satisfaction Research ProspectU.S.," in R. T. Rust and R. L. Oliver (eds.), *Service Quality: New Directions in Theory and Practice*, ThoU.S.A.nd Oaks: Sage, 241–268.

Andrich, D. (1988), *Rasch Models for Measurement*, California: Sage University Paper Series on Quantitative Applications in Social Science.

Andrich, D., B. S. Sheridan and G. Luo. (2003a), "Displaying the Rumm 2020 Analysis," working paper, RUMM Laboratory, Perth.

_____ (2003b), *Rumm2020: Rasch Unidimensional Measurement Models*, Perth: RUMM Laboratory.

Andrews, F. M. and S. B. Withey. (1976), *Social Indicators of Well-Being Americas Perception of Quality of Life*. New York: Plenum Press.

Babin, B. J. and M. Griffin. (1998), "The Nature of Satisfaction: An Updated Examination and Analysis," *Journal of Business Research,* 41, (2), 127–136.

Bagozzi, R. P., M. Gopinath, and P. U. Nyer. (1999), "The Role of Emotions in Marketing," *Journal of the Academy of Marketing Science*, 27, (2), 184-206.

Bond, T. G. and C. M. Fox. (2007), *Applying the Rasch Model (2nd edition).* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Bruthiaux, P. (2003), "Squaring the Circles: Issues in Modeling English Worldwide," *International Journal of Applied Linguistics*, 13, (2), 159-78.

Buck, R. and M. Georgson. (1997), "Consumer Emotion Space: An Investigation of Semantic Space and Context Effects in Self-Reported Emotion Elicitation," *Advances in Consumer Research* 24, 431–437.

Cacioppo, J. T. and G. G. Bernston. (1999), "The Affect System: Architecture and Operating Characteristics," *Current Directions in Psychological Science,* 8, (5), 133–137.

Chen, C. (2008), "Investigating structural relationships between service quality, perceived value, satisfaction, and behavioral intentions for air passengers: Evidence from Taiwan", Transportation Research A – Policy, 42, (4), 709-717.

Churchill, G. A. and C. Suprenant. (1982), "An Investigation into the Determinants of Customer Satisfaction," *Journal of Marketing Research* 19 (November), 451–504.

Crystal, D. (1998), "The Future of English: An Anecdotal Introduction," in D. Lynch and A. Pilbeam (eds.), *Heritage and Progress, Proceedings of the SIETAR European Congress 1998*, 6-16. Bath.

_____ (1999), "The Future of Englishes," *English Today*, 58, (2), 10-20.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*, New York: Harper Collins.

Danaher, P. J. and V. Haddrell. (1996). "A Comparison of Question Scales Used for Measuring Customer Satisfaction," *International Journal of Service Industry Management*, 7, (4), 4-26.

Diener, E. and F. Fujita. (1995), "Methodological Pitfalls and Solutions in Satisfaction Research," in J. M. Sirgy and C. A. Samli (eds.), *New Dimensions in Marketing/Quality of Life Research*, Westport: Quorum Books, 27–46.

Douglas, S. P. and C. S. Craig. (2007), "Collaborative and Iterative Translation," *Journal of International Marketing*, 15, (1), 30-43.

Dube, L., M.-C. Belanger, and E. Tudeau. (1996), "The Role of Emotions in Health Care Satisfaction," *Journal of Health Care Marketing,* 16, (2), 45–51.

Embretson, S. E. and S. P. Reise. (2000), *Item Response Theory for Psychologists.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ewing, M. T., T. Salzberger, and R. R. Sinkovics. (2005), "An Alternate Approach to Assessing Cross-Cultural Measurement Equivalence in Advertising Research," *Journal of Advertising*, 34, (1), 17-36.

Fischer, G. H. and I. W. Molenaar. (1995), *Rasch Models, Recent Developments and Applications.* New York: Springer Verlag.

Forgas J. P. (1995), "Mood and Judgment: The Affect Infusion Model (AIM)," *Psychological Bulletin*, 117, (1), 39-66.

Fornell, C. (1992), "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing* 56 (January), 6–21.

Fornell, C., M. D. Johnson, E. W. Anderson, J. Cha, and B. E. Bryant. (1996), "The American Customer Satisfaction Index: Nature, Purpose, and Findings," *Journal of Marketing,* 60 (October), 7–18.

Fournier, S. and D. G. Mick. (1999), "Rediscovering Satisfaction," *Journal of Marketing*, 63 (October), 5-23.

Ganglmair-Wooliscroft, A. (2007), "A Comparison of Affective Response to Consumption in Two Contexts" *Der Markt; Zeitschrift für Absatzwirtschaft und Marketing* (article in English), 46, (1/2), 36-49.

Ganglmair, A. and R. Lawson. (2003a), "Advantages of Rasch Modeling for the Development of a Scale to Measure Affective Response to Consumption," *European Advances in Consumer Research*, 6, 162-68.

_____ (2003b), "Measuring Affective Response to Consumption Using Rasch Modeling," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 16, 198-21.

Gardial, S. F., D. S. Clemons, R. B. Woodruff, and M. J. Burns. (1994), "Comparing Consumers' Recall of Prepurchase and Postpurchase Product Evaluation Experiences," *Journal of Consumer Research*, 20 (March), 548-56.

Giese, J. L. and J. A. Cote. (2000), "Defining Consumer Satisfaction," *Academy of Marketing Science Review* online at http://www.vancouver.wsu.edu/amsrev/theory/giese01-0.html.

Goleman, D. (1996). *Emotional Intelligence*, London: Bloomsbury.

Green, D. P., S. L. Goldman, and P. Salovey. (1993), "Measurement Error Masks Bipolarity in Affect Ratings," *Journal of Personality and Social Psychology,* 64, (6), 1029–1041.

Halstead, D., D. Hartman, and S. L. Schmidt. (1944), "Multisource Effects on the Satisfaction Formation Process," *Journal of the Academy of Marketing* Science 22, (2), 114–129.

Halstead D. and M. A. Jones, and A. N. Cox. (2007), "Satisfaction Theory and the Disadvantaged Consumer," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 20, 15-35.

Herzberg, F., B. MAusner, and B. B. Synderman. (1959). *The Motivation to Work*, New York : John Wiley & Sons.

Hicks, J. M., T. J. Page Jr., B. K. Behe, J. H. Dennis, and R.T. Fernandez. (2005), "Delighted Consumers Buy Again," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 18, 94-104.

Hunt, H. K. (1977), "CS/D - Overview and Future Research Directions," in H. K. Hunt (ed.), *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, 455–488. Cambridge, Mass: Marketing Science Institute.

Ilson R. (1985), "Diversity in Unity: American and British English," *English Today*, 4 (Oct), 7-11.

Izard, C. E. (1977), *Human Emotions.* New York: Plenum.

Kachru, B. B. (1985), "Standards, Codification, and Sociolinguistic Realm: the English Language in the Outer Circle," in R. Quirk and H.G. Widdowson (eds.), *English in the World,* 11-30. Cambridge: Cambridge University Press.

_____ (1986), "The Indianization of English," *English Today*, 6 (April), 31-33.

Kahneman, D. (1999), "Objective Happiness," in D. Kahneman, E. Diener and N. Schwarz (eds.), *Well-Being: The Foundations of Hedonic Psychology*, New York: RU.S.sell Sage Foundation, 3–25.

Kirkpatrick, A. (2007), *World Englishes: Implications for International Communication and English Language Teaching.* Cambridge: Cambridge University Press

Larsen, J. T., A. P. McGraw, and J. T. Cacioppo. (2001), "Can People Feel Happy and Sad at the Same Time?," *Journal of Personality and Social Psychology,* 81, (4), 684–696.

LeDoux, J. (1996), *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York: Touchstone.

Lewis M., J. M. Haviland-Jones, and Barrett L. Feldman. (2008), *Handbook of Emotions 3rd edition.* New York: The Guilford Press.

Linacre, J. M. (1992), "Why Fuss about Statistical Sufficiency?," *Rasch Measurement Transactions* 6 (3), 230. http://www.rasch.org/rmt/rmt63c.htm.

_____ (1999). Understanding (or Misunderstanding?) the Rasch Model, *Rasch Measurement Transactions* 13, (3), 706. online at http://www.rasch.org/rmt/rmt133i.htm.

Lowenberg, P. H. (2002), "Assessing English proficiency in the Expanding Circle," *World Englishes,* 21, (3), 431-35.

Maddox, R. N. (1981), "Two-factor Theory and Consumer Satisfaction: Replication and Extension," *Journal of Consumer Research,* 8, (1), 97-102

Mano, H. and R. L. Oliver. (1993), "Assessing the Dimensionality and Structure of the Consumption Experience: Evaluation, Feeling, and Satisfaction," *Journal of Consumer Research,* 20 (December), 451- 66.

Mehrabian, A. and J. A. Russell. (1974), *An Approach to Environmental Psychology.* Cambridge, Mass.: MIT Press.

Mondiano, M. (1999), "International English in the Global Village," *English Today 58*, 15, (2), 22-8.

Nyer, P. U. (1998), "The Effects of Satisfaction and Consumption Emotion on Actual Purchasing Behavior: An Exploratory Study," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 11, 62–68.

Oliver, R. L. (1980), "A Cognitive Model of Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research,* 17 (November), 460–469.

_____ (1981), "Measurement and Evaluation of Satisfaction Processes in Retail Settings," *Journal of Retailing,* 57, (3), 25–48.

_____ (1997). *Satisfaction; A Behavioral Perspective on the Consumer.* Boston: Irwin McGraw-Hill.

Ortony, A. and T. J. Turner. (1990), "What's Basic about Basic Emotions," P*sychological Review,* 97, (3), 315–31.

O'Shaughnessy, J. and N. J. O'Shaughnessy. (2003), *The Marketing Power of Emotion.* New York: Oxford University Press.

Oyewole, P. (2001), "Consumer's Socio-Demographic Characteristics and Satisfaction with Services in the Airline Industry", *Services Marketing Quarterly*, 23, (2), 61–80.

Parker, C. and B. P. Mathews. (2001), "Customer Satisfaction: Contrasting Academic and Consumers' Interpretations," *Marketing Intelligence and Planning,* 19, (1), 38–44.

Peck, B. (2000). *Ethnicity: A Study of Participation, Aspirations and Achievement in Education,* A thesis submitted in part completion of the requirements of the degree of Doctor of Education, Murdoch University, Perth.

Peterson, R. A. and W. R. Wilson. (1992), "Measuring Customer Satisfaction: Fact and Artifact," *Journal of the Academy of Marketing Science,* 20, (1), 61–71.

Pham M. T., J. B. Cohen, J. W. Pracejus, and G. D. Hughes. (2001) "Affect Monitoring and the Primacy of Feelings in Judgment," *Journal of Consumer Research,* 28 (September), 167-188.

Qiong H. X. (2004), "Why China English Should Stand Alongside British, American, and the Other 'World Englishes'," *English Today,* 78, (2), 26-32

Rasch, G. (1960/80), *Probabilistic Models for Some Intelligence and Attainment Tests; With a Foreword and an Afterword by Benjamin D. Wright.* Chicago: MESA Press.

_____ (1977), "On Specific Objectivity: An Attempt at Formalizing the Request for Generality and Validity of Scientific Statements," *Danish Yearbook of Philosophy* 14, 58–93. http://209.238.26.90/memo18.htm.Richins 1997.

Richins, M. L. (1997), "Measuring Emotions in the Consumption Experience," *Journal of Consumer Research,* 24 (September), 127–46.

Rossiter, J. R. (2002), "The C-OAR-SE procedure for scale development in marketing," *International Journal of Research in Marketing,* 19, (4), 305–335.

Rost, J. (2001), "The Growing Family of Rasch Models," in A. Boomsma, M. A. van Duijn and T. A. Snijders (eds.), *Lecture Notes in Statistics: Essays on Item Response Theory,* New York: Springer Verlag, 3-24.

Russell, J. A. and J. M. Carroll. (1999), "On the Bipolarity of Positive and Negative Affect," *Psychological Bulletin,* 125, (1), 3–30.

Salzberger T. (2009) *Measurement in Marketing Research: An Alternative Framework,* Cheltenham, UK: Edward Elgar.

Salzberger T. and R. Sinkovics. (2006), "Reconsidering the Problem of Data Equivalence in International Marketing Research – Contrasting Approaches Based on CFA and the Rasch Model for Measurement," *International Marketing Review*, 23, (4), 390-417.

Salzberger, T., R. R. Sinkovics, and B. B. Schlegelmilch. (1999), "Data Equivalence in Cross-Cultural Research: A Comparison of Classical Test Theory and Latent Trait Theory Based Approaches," *Australasian Marketing Journal,* 7, (2), 23–38.

Shaver, P., J. Schwartz, D. Kirson, and C. O'Connor. (1987), "Emotion Knowledge: Further Exploration of a Prototype Approach," *Journal of Personality and Social Psychology,* 52, (6), 1061–86.

Singh, J. (1995), "Measurement Issues in Cross-National Research," *Journal of International Business Studies*, 26, (3), 597-619.

_____ (2004), "Tackling Measurement Problems with Item Response Theory," *Journal of Business Research*, 57, (2), 184-208.

Slovic, P., M. L. Finuncane, E. Peters, and D. G. MacGregor. (2007), "The Affect Heuristic," *European Journal of Operational Research,* 177, 1333-1352.

Soderlund, M. and S. Rosengren. (2004), "Dismantling 'Positive Affect' and its Effects on Customer Satisfaction: An Empirical Examination of Customer Joy in a Service Encounter," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 17, 27-41.

Soutar, G. N. and S. P. Cornish-Ward. (1997), "Ownership Patterns for Durable Goods and Financial Assets: a Rasch Analysis," *Applied Economics,* 29, (7), 903–11.

Steenkamp, J-B. and H. Baumgartner. (1998), "Assessing Measurement Invariance in Cross-National Consumer Research," *Journal of Consumer Research*, 25, (1), 78-90.

Storm, C. and T. Storm. (1987), "A Taxonomic Study of the Vocabulary of Emotions," *Journal of Personality and Social Psychology,* 53, (4), 805–16.

Swan, J. E. and L. J. Combs. (1976), "Product Performance and Consumer Satisfaction: A New Concept," *Journal of Marketing,* 40, (2), 25–33.

Taylor, S.A. (2008), "Reconciling Satisfaction, Emotions, Attitudes, and Ambivalence within Consumer Models of Judgment and Decision Making: A Cautionary Tale," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 21, 41-65

van de Vijver, F. and N. K. Tanzer. (2004), "Bias and Equivalence in Cross-Cultural Assessment: An Overview," *Revue Europeenne de Psychologie Appliqué*, 54, (2), 119-35.

Westbrook, R. A. (1980), "A Rating Scale for Measuring Product/Service Satisfaction," *Journal of Marketing*, 44(4), 68-72.

Westbrook, R. A. and R. L. Oliver. (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," *Journal of Consumer Research,* 18 (June), 84–91.

Wright, B. D. (1992). IRT in the 1990s: Which Models Work Best, Ben Wright's opening remarks in his invited debate with Ron Hambleton, Session 11.05, AERA Annual Meeting 1992, *Rasch Measurement Transactions* 6, (1), 196–200. available online at http://www.rasch.org/rmt/rmt61a.htm.

_____ (1997), "Fundamental Measurement for Psychology," *MESA Research Memorandum* 64. http://www.rasch.org/memo64.htm.

Wright, B. D. and M. H. Stone. (2004), *Making Measures.* Chicago: The Phaneron Press.

Yi, Y. (1990), "A Critical Review of Consumer Satisfaction," in V. A. Zeithmal (ed.), *Review of Marketing*, Chicago: American Marketing Association, 68–123.

Zajonc, R. (1980), "Thinking and Feeling," *American Psychologist*, 35(2y), 151–175.

Acknowledgements: The authors would like to thank the editor and three anonymous reviewers for their helpful and constructive suggestions. A special thank-you goes to Walter Feldkircher at Austrian Airlines for supporting this research and collecting the data.

Send correspondence regarding this article to:

Alexandra Ganglmair-Wooliscroft
Department of Marketing
University of Otago
PO Box 56
Dunedin, New Zealand
Tel: +64 (0)3 4798167
Fax: +64 (0)3 4798172
e-mail: Alexandra.Ganglmair@otago.ac.nz