

# MEASURING *AFFECTIVE RESPONSE TO CONSUMPTION* USING RASCH MODELING

Alexandra Ganglmair, University of Otago  
Rob Lawson, University of Otago

## ABSTRACT

The historic conceptualization of satisfaction has been criticized in recent years as emotions and affect gained an increasingly important position in marketing (Bagozzi, Gopinath and Nyer 1999; Erevelles 1998). This paper describes the development of a scale to measure *Affective Response to Consumption (ARC)* using Rasch Modeling. The concept is an extension to satisfaction that has arisen out of both assessment of literature in satisfaction and attempts to address problems in measuring satisfaction.

Rasch Modeling (Rasch 1960/80) is introduced as a measurement technique that is particularly suited for *ARC*, where items need to differ significantly in the intensity of the concept they represent. The final scale consists of 13 items ranging from *impressed* to *euphoric*. The newly developed scale distinguishes respondents with highly positive experiences, where commonly used scales, such as the Delighted-Terrible scale (Andrews and Withey 1976), fail to discriminate.

## BACKGROUND: CONCEPTUALIZING *ARC*

Historically, satisfaction was conceptualized as a cognitive construct (Westbrook 1989), with Oliver's (1980) expectation-disconfirmation paradigm being regarded as the dominant model (Fournier and Mick 1999; Halstead, Hartman & Schmidt 1994). In the mid 1990s, research had started to not only criticize the overwhelming dominance of this paradigm (Hunt 1993) but also increasingly investigated affective antecedents of satisfaction (e.g. Brockman 1998; Dube-Rioux 1990; Evrard and Aurier 1994; Oliver 1994,1992,1989; Mano and Oliver 1993; Westbrook 1989; Westbrook and Oliver 1991; Wirtz, Mattila and Tan 2000).

In the last few years, a number of studies have been conducted that re-examine established

satisfaction research, especially regarding the difference between satisfaction and other emotional post consumption responses (Gardial, Clemons, Woodruff and Burns 1994; Fournier and Mick 1999; Giese and Cote 2000, Nyer 1998). Growing evidence points towards satisfaction itself being an emotional consumption response, and "not merely a consequence of other emotions" (Nyer, 1998, p.62). Conceptualizing satisfaction as a consumption emotion is not new but was proposed in the early 1980s by Day (1983) and Sirgy (1984) and has recently received renewed attention. Giese and Cote (2000) as well as Gardial et al. (1994) found that consumers swap *satisfaction* for other, highly emotional words when talking about their experiences. Bagozzi, Gopinath and Nyer (1999) make this point very clear when they question the difference between satisfaction and other positive emotions by stating that: "The centrality of satisfaction in marketing studies is perhaps more due to being the first emotion to receive scrutiny in postpurchase behavior research than to constituting a unique, fundamental construct in and of itself (p.201)".

The debate about satisfaction as a cognitive or emotional construct is sometimes confused because when using self-reports like questionnaires, responses to questions about different emotions will always be cognitive. *Emotional feeling* or *affect* occurs when one becomes consciously aware of activities of the emotional system in the brain, which are necessary to process written questions (LeDoux 1996, Cacioppo and Bernstein 1999). Growing evidence, (Gardial et al. 1994; Fournier and Mick 1999; Giese and Cote 2000, Nyer 1998) suggests that satisfaction, happy and pleased etc. belong to the same group of terms commonly referred to as *emotions*. If satisfaction is conceptualized as an emotional response, just like e.g. happy, all terms related to these affective state could be measured on one dimension. As we will measure these emotional states using self-reports we will refer to

them as *emotional responses* or *affect* in order to credit the cognitive process involved.

In line with the above-mentioned research, Ganglmair and Lawson (2002) have conceptualized *Affective Response to Consumption (ARC)*. *ARC* is a construct measuring emotional responses with satisfaction being regarded as one of a large number of possible unfavorable/favorable responses to an experience. This conceptualization enables the inclusion of stronger affective states and shifts the emphasis from the linguistically relatively weak word satisfaction, as derived from the Latin *facere* – to do or make and *satis* – enough (Schumm 1999), to a multitude of positive affects that are found in post-purchase/post-experience situations (Fournier and Mick 1999; Gardial et al. 1994; Giese and Cote 2000). Investigating *ARC* emphasizes the diversity amongst consumers.

*ARC* is concerned with the dimension that relates to unfavorable-favorable consumption experiences (Ganglmair and Lawson, 2003). Although the unidimensionality of emotions (in psychology) and satisfaction/dissatisfaction (in marketing) has been questioned (Larsen, McGraw and Cacioppo 2001; Mackoy and Spreng 1995; Maddox 1981; Swan and Combs 1976), results have been inconclusive. Studies show that mixed emotions are mainly experienced in very complex situations (Larsen et al. 2001) e.g. when moving out of college. The ability to cope with mixed emotions is also limited in Western cultures and likely to be avoided (Festinger 1957; Williams and Aaker 2002). It is further not certain whether respondents think of the same attributes when answering two-dimensional satisfaction/dissatisfaction questions (Mackoy and Spreng 1995) and a number of results presenting support for two-dimensional conceptualizations have been shown to represent statistical artifacts and measurement error (Green, Goldman and Salovey 1993). While it is not denied that mixed emotions can be experienced, this seems to be an exception, rather than the norm (Larsen et al. 2001). Russell and Carroll (1999) come to the conclusion that “for theories about affective feelings, bipolarity is a reasonable assumption” (p.25).

Variations of the Differential Emotions Scale

(Izard 1977) or Pleasure-Arousal-Dominance scale (Mehrabian and Russell 1974) are frequently used to measure emotions in marketing. These two scales were developed to measure the entire range of fundamental emotions and responses to environmental stimuli, respectively, with negative emotions being dominant in both scales. Emotions are further likely to be context specific, with different emotions being of different importance depending on the context in which they are used (Richins 1997). Richins (1997) states “emotions that arise in the context of intimate interpersonal relationships are likely to differ from the emotions experienced when buying a pair of shoes” (p. 129).

The Consumption Emotion Set (CES) was developed in order to provide a set of emotions that cover the entire space of frequently experienced consumption emotions including e.g. worry, shame, envy or peacefulness. *ARC* on the other hand is only concerned with emotions that relate to the unfavorability/favorability of consumption experiences. It started with the term *satisfaction* and includes only items that show more or less favorable expression.

The Delighted-Terrible scale (D-T scale) (Andrews and Withey, 1976) is conceptualized along the same dimension as *ARC* but tries to cover the entire range of experiences from rotten to delighted with seven items. As it frequently produces skewed results, a re-examination and extension of this scale towards more positive terms is warranted.

#### WHY RASCH MODELING SUITS A CONCEPT LIKE *ARC*

Research into scales for measuring satisfaction has been limited, with work done more than 20 years ago by Oliver and Westbrook being the most cited sources for scales (Oliver 1981, 1980; Westbrook 1980, Westbrook and Oliver 1991).

The limited attention to measurement issues in satisfaction research and marketing in general is surprising, as calls encouraging research into these issues have been made in regular intervals (Babin and Griffin 1998; Hunt 1977; Jacoby 1978; LaBarbara 1984). As with a majority of social

sciences, the Classical Test Theory (CTT) is the leading measurement paradigm in marketing (Embretson 1996; Hambleton 1991; Salzberger, Sinkovics and Schlegelmilch 1999), with Churchill's (1979) classical article on *A Paradigm for Developing Better Measures of Marketing Constructs* being the most influential paper for scale development in the marketing discipline.

Currently used scales to measure satisfaction, which are rooted in CTT, regularly show limited discrimination as well as a strong negative skewness (Diener 1984; Diener and Fujita 1995; Peterson and Wilson 1992; Szymanski and Henard 2001). Peterson and Wilson (1992) remark that "virtually all self-reports of customer satisfaction possess a distribution in which a majority of the responses indicate that customers are satisfied ... (and) the modal response to a satisfaction question is typically the most positive response allowed" (p.62). This characteristic of satisfaction scales has been mentioned in passing by several authors e.g. Halstead et. al 1994; Oliver 1981 (cited from Peterson and Wilson, 1994). Attempts to overcome these shortcomings have been limited. ARC extends and complements satisfaction by trying to overcome the inherent skewness in the latter concept in order to provide additional information about respondents traditionally found in the most positive answer category.

The current paper investigates a method for scale development based on an alternative measurement paradigm. Rasch Modeling (Rasch 1960/80) has been widely used in educational measurement and only recently received interest in other social sciences (Embretson and Reise 2000) and marketing in particular (e.g.: Salzberger 2000, Salzberger et al. 1999; Soutar and Cornish-Ward 1997; Soutar and Monroe 2001). Although Georg Rasch, a mathematician and statistician, developed his model in an educational context, he explicitly mentions that measurement problems encountered in medicine, psychology, technology, economics, sociology, linguistics etc. led him towards the development of the Rasch Model (Wright 1980).

Rasch Models belong to the family of latent trait models, which are concerned with measuring an unobservable, latent, variable. One is interested

in the underlying attribute of a person that a measurement score reflects (Ryan, 1983). The model tests whether a single latent trait actually underlies a number of questions that are conceptualized to comprise a unidimensional scale. It further establishes where respondent's are positioned on this latent variable. The underlying attribute of interest can be any latent trait e.g. the level of algebra knowledge a student possesses (in an educational setting) or the amount of Affective Response to Consumption a consumer experienced (in a satisfaction or marketing setting).

Rasch Modeling aims to introduce rigid rules of measurement - similar to physics - into social sciences (Wright 1997). The technique is regarded as a probabilistic alternative to the deterministic Guttman scaling (Andrich 1982; Salzberger et al. 1999; Wright 1997) with the probabilistic qualities of Rasch Models helping to overcome disadvantages of the latter scales, which tend to work quite well for objective information, but produce messy results when the phenomenon of interest is not concrete (DeVellis 1991). Rasch Modeling is based on a mathematical model dealing with the "probabilistic relation between any item's difficulty and any person's ability" (Bond and Fox 2001, p.199]:

$$P_{vi} = \exp(B_v - D_i) / [1 + \exp(B_v - D_i)]$$

where

$P_{vi}$  = probability of person  $v$ , given their level of ability, answering correctly

(e.g. as predicted by the model) to item  $I$

$B_v$  = Location of person  $v$  on Rasch scale and

$D_i$  = Location of  $i$ th item on Rasch Scale.

While extensive discussions of the classic approach and Rasch Modeling can be found elsewhere (e.g. Andrich 1988; Bond and Fox 2001; Embretson 1996; Embretson and Reise 2000; Fischer and Molenaar 1995; Lord, 1980; Lord and Novick 1968; Nunnally and Bernstein 1994; Wright and Stone 1979) the following discussion is limited to aspects of Rasch Modeling and differences between the new technique and CTT that are important when applied in the context of measuring *ARC*.

With Rasch Modeling “A person having greater ability than another should have the greater probability of solving any item of the type in question and similarly, one item being more difficult than another one means that for any person the probability of solving the second item correctly is the greater one” (Rasch 1960/80, p.117). In a marketing context, the term difficulty can be replaced by the amount of a specific content an item stands for e.g. how hard it is to endorse the item, how extreme the item is. Ability can be translated into the characteristic of the person e.g. the person’s innate level of *ARC*.

Linking back to the research traditions of Guttman and Thurstone (Andrich 1988; Engelhard 1990), Rasch Modeling requires differences in the items representing a construct in question (Salzberger et al. 1999; Salzberger 2000; Wright and Stone 1979). The researcher is explicitly asked to generate items covering different intensity levels (Andrich 1988; Salzberger 2000; Salzberger et al. 1999; Wright and Stone 1979) as Rasch Modeling emphasizes that the entire breath of the construct is under investigation.

Rasch Modeling thereby differs from the CTT. The latter relies heavily on the principle of correlation with factor analysis and Coefficient alpha being important reliability indices (Churchill 1979). The use of these quality indices has received considerable critique as it encourages the inclusion of items that tap similar aspects of the construct (Salzberger 2000; Smith 1999; Steinberg and Thissen 1996).

The Rasch Model, as a probabilistic Guttman model, computes item and person fit in relation to the model (Wright 1977), with the items’ observed fit to the model being taken to investigate unidimensionality (Soutar and Monroe 2001). Rasch Software, such as RUMM 2010 (Andrich, Sheridan and Luo 2001), projects items and respondents onto the same dimension and provides indices and visual displays to investigate whether items spread sufficiently along a continuum rather than clumping towards one point of the dimension.

Scales used currently to measure satisfaction show a limited discrimination (Peterson and Wilson 1992) which suggests that only a point on the dimension rather than the entire continuum is

being measured while the overwhelming use of the most positive answer category illustrates that this point is on a moderate point on the continuum of emotions felt towards a consumption experience (Ganglmair and Lawson 2003).

Rasch Modeling represents a simple, yet mathematically elegant approach (Rasch 1960/1980) for scale development. While composite scales might be formed in CTT in order to combine e.g. pleasure and arousal (to measure emotional valence and it’s intensity) Rasch Modeling establishes such a scale in one scale development process. This process might only appear more complex due to its unfamiliarity to a majority of marketing researchers.

#### DEVELOPMENT OF A SCALE TO MEASURE *ARC*

In the current scale development process, it was of particular importance not to rely on terms that are used in the marketing literature, as potentially important areas and intensity levels of satisfaction might not be included in existing research. Terms used in the consumer satisfaction literature were therefore merely a starting point for an extensive item collection process. Taxonomies of emotions from the psychology literature have been used to complement existing terms from the marketing literature.

The Delighted-Terrible scale (D-T Scale) (Andrews and Withey 1976), which has shown favorable results in comparison studies of satisfaction scales (Westbrook 1980) served as a starting point for item generation, as it includes seven different words that are available for expressing different levels of affective response. Only six of these words were investigated in greater detail, as the seventh – *mixed* - was considered too vague. *Contented* and *pleased* have been described as synonyms of satisfaction (Oliver, 1989; Oliver, Rust and Vakri 1997) and *contented* will be an additional starting point for further item development (*pleased* is already included from the D-T scale).

In order to gain a large variety of related words, three differently structured thesauri (Kirkpatrick 1987; Chapman 1992; McCutcheon

1995) were used. Not all thesauri contain the words that have been chosen to seed the investigation and if necessary the terms closest to the ones mentioned above were used. All words within the categories in question were noted for possible inclusion in the scale.

A classic taxonomy of emotions (Shaver et al. 1987) and a categorizations study of emotions (Storm and Storm, 1987) formed the input from the psychology literature. Shaver et al. (1987) paper on Prototypes of Emotion provided words clustered under: *joy*, and under *anger*. In Storm and Storm's (1987) case the items were either part of the category *positive terms without interpersonal references, negative terms related to shame, sadness, and pain*, or *negative terms related to anger, hatred, or disgust*. (For a detailed list on all items please contact the first author.)

After coding and alphabetically sorting the items as well as deleting double entries and different grammatical forms of the same term, 715 words were available for further investigation. Three judges with qualifications in English were presented with an introductory letter and the list of words. The letter briefly outlined the aim of the research and asked judges to mark all items that an average New Zealander would possibly use as response to the question: "*How do you feel about your experience with an excursion train ride*", as

the data collection was going to take place in the last 15 minutes of a 4 hour excursion train ride from Dunedin, New Zealand. The item selection by independent judges was necessary, as the original list contains a large variety of terms, some of which were clearly not suitable as an answer to the above question.

The judges were further asked to rate the chosen items on a five point intensity scale from strongly positive to strongly negative. This rating should provide verification that judges considered the items in a similar way and provided the researchers with an overview of items that have the possibility to address the positive end of the scale.

Twenty-nine words were chosen by all three judges. The indication of direction and intensity of the terms chosen can be seen in Table 1. The three judges rated the items in question very similarly, which reinforces that these terms convey the same meaning, and serves as a first reliability check. The selected list of words can be split into 19 positive and 10 negative terms. This is considered a suitable mix of items, as a large number of positive and very positive items is called for in the item pool for a scale measuring *ARC* in order to provide respondents normally found in the most positive answer category with other options to express their positive feelings.

**Table 1**  
**Items Chosen by Three Judges Including Direction and Strength**

Term	Frequency of rating					Term	Frequency of rating					Term	Frequency of rating					
	++	+	+-	-	--		++	+	+-	-	--		++	+	+-	-	--	
superb	3					fine		3				disappointed					3	
over the moon	3					delighted		3				discontented					3	
in 7 <sup>th</sup> heaven	3					pretty good		3				indifferent					3	
fantastic	3					impressed		3				unhappy					3	
exhilarated	3					happy		3				displeased					2	1
euphoric	3					pleased		3				dissatisfied					2	1
magnificent	2	1				satisfied		3				unsatisfied					2	1
great	2	1				relaxed		2	1			appalled						3
enthralled	2	1				okay		1	2			rotten						3
fabulous	2	1										terrible						3

Of particular interest is the rating of items that are currently used in satisfaction scales. The D-T scale (Andrews and Withey 1976) contains 5 of the terms that can also be found in Table 1. These are *delighted, pleased, satisfied, dissatisfied, unhappy and terrible*. While the judges rated the negative items *unhappy* and *terrible* on two different intensity levels (negative and strongly negative, respectively) all positive terms were rated on the same intensity level, namely positive. This emphasizes two characteristics of currently used scales: It shows that scales developed in CTT contain items of similar intensity (Salzberger 2000; Wright and Stone 1979), and it reinforces that currently used satisfaction scales lack the inclusion of items that reflect feelings reserved for people who have a very positive experience. Existing scales therefore encourage the use of the extreme answer category, as there are no really strong items available that might discriminate between respondents who had a positive, and those who had a very positive experience.

## ANALYSIS AND RESULTS

The data collection took place on an excursion train around Dunedin, New Zealand. The first researcher approached passengers during the last twenty minutes of a four-hour train ride. Acceptance to participate was exceptionally high due to the personal approach, with hardly anyone declining to fill in the questionnaire. Four-hundred-and-twenty-three questionnaires were distributed and collected, 419 of which could be used for further analysis. The questionnaire contained the following question:

*How do you feel about your experience with the Taieri Gorge railway? Please consider every word and tick all those that describe your feelings about the experience.*

Each of the 29 items as chosen by the judges was presented in a box that could be ticked by respondents (binary variables). The demographic profile matched the typical passenger on the train with an average age of 48 years. Almost half of the respondents (48%) were New Zealanders, 27%

Australian, 11 % British and a further 11% were from other countries (mainly USA and the Netherlands).

A preliminary examination using a Guttman pattern showed that only 16 out of 419 respondents ticked any of the negative items. The limited expression of negative affect is likely to be a characteristic of the underlying experience: a holiday excursion train ride that puts people generally in a neutral to positive mood. Negative items had to be eliminated, as the limited number of responses did not justify a quantitative analysis.

The computer software RUMM 2010 (Andrich et al. 2001) was used for data analysis. A first run including the 19 remaining variables revealed that certain items did not fit the unidimensional pattern. These items were gradually removed, similar to a backward stepwise regression (Soutar and Ryan 1999). After removing six items (*okay, pretty good, relaxed, fine, satisfied, pleased*) a point was reached where further elimination of items did not significantly improve the overall characteristics of the scale.

Rasch Modeling enables the projection of the distribution of items and persons onto the same dimension. Figure 1 illustrates the location of items along the dimension of *ARC* as well as the location of respondents. It shows that items and respondents are well spread across the *ARC* scale, with a standard deviation of 1.99 and 1.95 (see Table 2) for items and respondents respectively. This implies that the current scale holds a great deal of information about different experiences respondents had with the excursion train.

The person separation index is an index of internal consistency, similar to Cronbach's alpha (Peck, 2000). The index, ranging from 0-1, stands for "the ratio of true variance to observed variance based on the estimates of a person ability B" (Andrich, 1982, p. 98). The person separation index of the *ARC* scale was 0.844 and the overall powers of the test-of-fit were *good*. The unidimensionality of *ARC* is examined through the items observed fit to that expected by the model (Soutar and Cornish-Ward, 1997) using a  $\chi^2$  test. A significant  $\chi^2$  test (0.000144) was deemed acceptable (Salzberger 2002) as the relatively large sample size leads to over-sensitivity

Figure 1  
Person Item Distribution

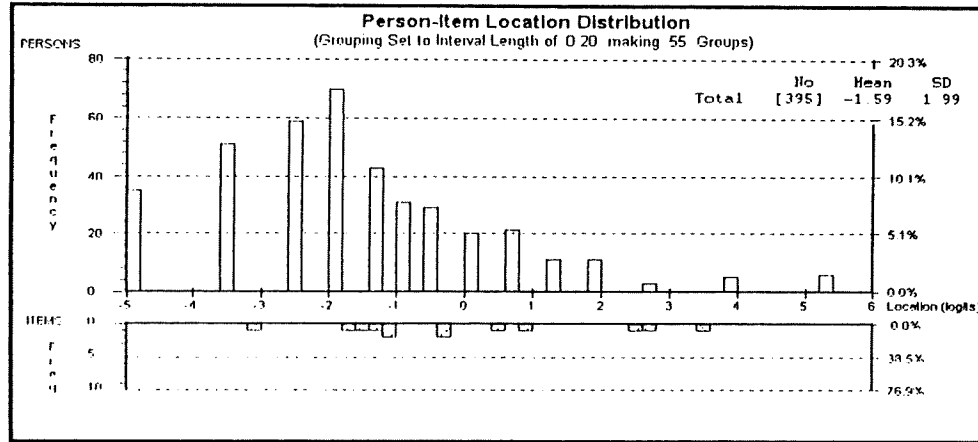


Table 2  
Summary Characteristics of Scale

	Items		Persons	
	Location	Std. Error	Location	Std. Error
Mean	0.00	-0.36	-1.59	-0.24
Std. Deviation	1.95	1.35	1.99	0.51
Item-Trait Interaction: $\chi^2 = 118.932$ (DF = 65, $p < 0.001$ )			Person Separation Index: 0.844	

Table 3  
Individual Item-fit

Item	Location	$\chi^2$ Probability
impressed	-3.02	0.06
happy	-1.66	0.001
great	-1.48	0.20
fabulous	-1.23	0.08
magnificent	-1.09	0.15
delighted	-1.04	0.34
superb	-0.38	0.01
fantastic	-0.34	0.01
enthralled	0.47	0.06
exhilarated	0.93	0.67
in 7th heaven	2.50	0.08
over the moon	2.79	0.95
euphoric	3.54	0.87

of that statistic and the scale still fits to a satisfactory extent (Peck 2000). Table 2 shows summary characteristics of the ARC scale. The mean location of items is commonly fixed to 0. A negative mean location of persons indicates that the respondents found the items included in the final scale relatively hard to endorse (see also Figure 1) which was expected, given the extreme answer categories chosen. A large standard deviation indicates that items and respondents were spread over the entire width of the concept investigated. Small errors are further encouraging.

Table 3 shows the location of the items along the ARC continuum (in logits) and their fit. The items are in order of un-likelihood of endorsement e.g. *impressed* stands for the lowest intensity, while *euphoric* represent the highest intensity of ARC. All items show an acceptable fit. Although

**Table 4**  
**Probabilities for Endorsement**

Group Number	G 14	G 13	G 12	G 11	G 10	G 9	G 8	G 7	G 6	G 5	G 4	G 3	G 2	G 1
Groupsize in %	1	1	1	3	3	5	5	7	8	11	18	15	13	9
<i>impressed</i>	100%	100%	100%	99%	99%	98%	97%	93%	89%	84%	75%	61%	38%	14%
<i>happy</i>	100%	100%	99%	97%	95%	91%	91%	77%	68%	56%	43%	29%	14%	4%
<i>great</i>	100%	100%	99%	97%	94%	90%	89%	74%	64%	52%	39%	25%		
<i>fabulous</i>	100%	99%	98%	96%	92%	87%	87%	69%	58%	46%	33%	21%		
<i>magnificent</i>	100%	99%	98%	95%	91%	85%	85%	66%	54%	42%	30%			
<i>delighted</i>	100%	99%	98%	95%	91%	85%	84%	65%	53%	41%	29%			
<i>superb</i>	100%	99%	96%	91%	83%	74%	73%	49%	37%	27%				
<i>fantastic</i>	100%	99%	96%	91%	83%	73%	73%	48%	36%	26%				
<i>enthralled</i>	99%	97%	91%	81%	68%	55%	54%	29%	20%					
<i>exhilarated</i>	99%	95%	87%	73%	57%	43%	43%	20%						
<i>in 7th heaven</i>	94%	80%	57%	37%	22%									
<i>over the moon</i>	93%	75%	50%	30%										
<i>euphoric</i>	85%	59%	32%											

*(Probabilities > 50% printed black / probabilities > 20% have a grey background)*

*happy* has a  $\chi^2$  probability below the proposed cutoff value of  $p < 0.01$ , removing this variable from the scale does not considerably alter the characteristics of the entire scale regarding unidimensionality and reliability (Salzberger 2002). On the other hand, the variable holds a considerable amount of information as its position on the *ARC* scale is close to the average person location (*happy* -1.66, average person location -1.59 see Table 2 and 3). No significant ( $p < 0.01$ ) main effect was found that would suggest that answers were biased in terms of age groups, gender, education or nationality.

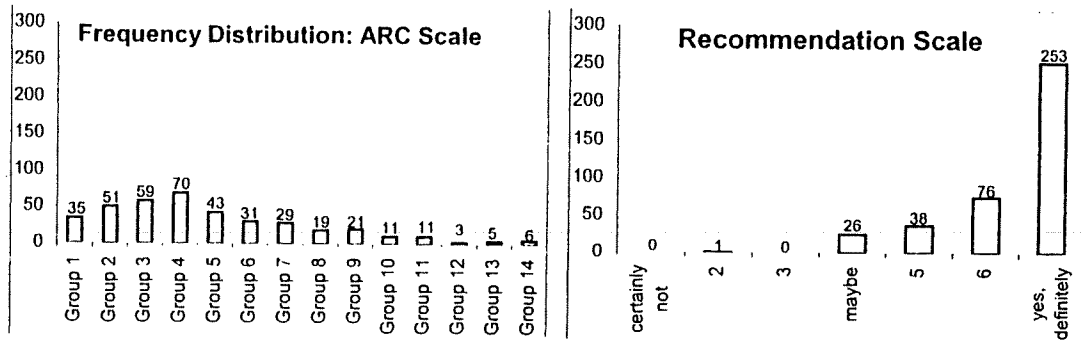
Rasch Modeling uses probabilities to describe the endorsement for each item by groups of respondents. The 14 groups on the *ARC* scale can therefore be described according to the probability of agreeing to ('ticking') each item. As can be seen in Table 4, the members of group 14 (G 14) show the highest intensity of *ARC* and the group consist of respondents with the best experience, while group 1 (G 1) shows the lowest intensity on *ARC* and is least impressed with the experience as measured on the scale.

#### COMPARISON WITH EXISTING SCALES

In order to establish whether the newly developed *ARC* scale discriminates stronger between respondents than currently used satisfaction scales, comparisons of group-membership between the *ARC* scale (14 groups), and ratings on the D-T scale (Andrews and Withey 1976), as well as two seven-point scales measuring bipolar satisfaction (1=completely unsatisfied, 7=completely satisfied) were carried out. An ANOVA was computed, with the group location as derived from RUMM 2010 (Andrich et al. 2001) as the factor variable and answers to the other three satisfaction scales as dependent variables, in order to see whether respondents at different locations in the *ARC* scale differ significantly in their response behavior of conventionally used scales. Although both ANOVA results were significant at a level of  $p < 0.01$  a post-hoc test (Bonferroni) revealed that all differences occur between the 3 lowest rating groups (= Group 1, 2 and 3 in Table 4). The top 11 groups in the *ARC* scale do not differ significantly in their response to any of the two conventional



Figure 2  
Frequency Distribution of ARC and Recommendation Scale



scales and show a mean rating of 6.3 on the two scales. This result clearly shows that the newly developed *ARC* manages to discriminate at the positive end of the experience, where conventional scales fail to detect differences.

*ARC* was further tested against a 7-point scale measuring the likelihood to recommend the trip (1= certainly not, 7= yes, definitely) using the same testes as above. Once again, Anova and Bonferroni produced significant results with all differences in the 3 lowest *ARC* groups. At closer inspection, potential doubts regarding the predictive capabilities of *ARC* can be disregarded. As can be seen in Figure 2 the recommendation scale is highly skewed, with 64% of all respondents choosing the highest available category (7= yes definitely), and fails to discriminate. One possible explanation could be that this recommendation scale does not distinguish between respondents who actively plan to go and recommend the experience and those who will only recommend if they are directly asked – a difference similar to the one between Market Mavens and Opinion Leaders as discussed by Feick and Price (1987). Further research will have to identify a predictive scale or collect longitudinal data to re-examine this phenomenon and establish the behavioral consequences of the *ARC* scale.

## LIMITATIONS

This paper does not claim to present a generalizable scale to measure experiences. Rather it explores the suitability of a specific method of scale development – Rasch Modeling – in a satisfaction/post-consumption emotion context. Further, *ARC* is not intended to replace current satisfaction measurement but expand and complement satisfaction measurement when more information is required about respondents who are using the highest possible answer category of the satisfaction scale.

Future research will apply this scale development approach to different kinds of consumer experiences, as well as in different cultural settings. After multiple applications and replications in various contexts, a selective number of items might be detected that is inherent in all settings and therefore build the core of a generalizable *ARC* scale, while it is fully expected that there will always be different experiences and cultural-specific terms that add to the understanding of the relevant context.

In order to arrive at a scale that shows acceptable unidimensionality characteristics, a number of items that were conceptualized as forming the middle ground of the *ARC* continuum had to be eliminated – particularly noteworthy are *satisfied* and *pleased*, as they have traditionally been key terms. It is suggested that these terms fall

out because they can be regarded in two different ways, e.g. someone might tick *satisfied* and mean *at least satisfied* (positive connotation) while another respondent might tick the same item and mean *just satisfied* (negative connotations). Further research will investigate how variables that cover the middle ground of *ARC* should be treated, in order to add information about respondents who find themselves at this part of the continuum. The removal of negative variables is due to the nature of the experience measured. Additional service experiences or products will have to show if negative affect can be added to the scale.

### CONCLUSION

*ARC* was conceptualized as a scale to measure unfavorable/favorable consumption emotions, especially at the very positive end of the dimension. One of the main tasks the scale to measure *ARC* set out to accomplish was to discriminate between respondents on the most positive end of commonly used satisfaction scales. *ARC* has shown it can achieve a considerable differentiation amongst those people who have traditionally used the highest answer category possible. The very low membership at the top *ARC* groups is expected, as a lot of people are satisfied but not that many experience extremely positive emotions. Further research will have to find if members of the most positive *ARC* group will also be more likely to recommend the experience and/or repurchase/re-experience.

It is reassuring to see that the final *ARC* scale, as derived from Rasch Modeling, orders the items very similarly to the earlier indication the judges made on a five point scale, while adding a lot of additional information at the higher end of the construct. This result shows that items were interpreted by both groups in the same way.

Overall, the *ARC* has shown to be a promising extension to conventional satisfaction scales. It shifts the emphasis from one term – *satisfaction* – to a variety of positive emotions and adds valuable information about respondents at the higher end of the continuum where traditional scales failed to discriminate.

### REFERENCES

- Andrews, Frank M. and Stephen B. Withey (1976), *Social Indicators of Well-Being Americas Perception of Quality of Life*, New York: Plenum Press.
- Andrich, David (1982), "Using Latent Trait Measurement Models to Analyse Attitudinal Data: A Synthesis of Viewpoints," in: *The Improvement of Measurement in Education and Psychology*, D. Spearit, Melbourne: Australian Council for Educational Research, 89–126.
- Andrich, David (1988), *Rasch Models for Measurement*, California: Sage University Paper Series on Quantitative Applications in Social Science.
- Andrich, David, Barry Sheridan and Guanhong Luo (2001), *RUMM2010: Rasch Unidimensional Measurement Models* [Computer Software].
- Babin, Barry J. and Mitch Griffin (1998), "The Nature of Satisfaction: An Updated Examination and Analysis," *Journal of Business Research*, 41, (2), 127–136.
- Bagozzi, Richard P., Mahesh Gopinath and Prashanth Nyer (1999), "The Role of Emotions in Marketing," *Journal of the Academy of Marketing Science*, 27, (2), 184–206.
- Bond, Trevor G. and Christine M. Fox (2001), *Applying the Rasch Model*, Mahwah: Lawrence Erlbaum Associates.
- Brockman, Beverly K. (1998), "The Influence of Affective State on Satisfaction Ratings," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 11, 40–50.
- Cacioppo, John T. and Gary G. Bernstein (1999), "The Affect System: Architecture and Operating Characteristics," *Current Directions in Psychological Science*, 8, (5), 133–137.
- Chapman, Robert L. (ed.) (1992), *Roget's International Thesaurus*, New York: Harper Collins.
- Churchill, Gilbert A (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16, (February), 64–73.
- Day, Ralph L. (1983), "The Next Step: Commonly Accepted Constructs for Satisfaction Research," *International Fare in Consumer Satisfaction and Complaining Behavior*, R. L. Day and H. K. Hunt, (eds.), Bloomington: Indiana University Press, 113–117.
- DeVellis, Robert F. (1991), *Scale Development: theory and application*, Beverly Hills: Sage Publications.
- Diener, Ed (1984), Subjective Well-Being, *Psychological Bulletin*, 95, (3), 542–575.
- Diener, Ed and Frank Fujita (1995), "Methodological Pitfalls and Solutions in Satisfaction Research," in *New Dimensions in Marketing/Quality of Life Research*, J. M. Sirgy and C. A. Samli, Westport: Quorum Books, 27–46.
- Dube-Rioux, Laurette (1990), "The Power of Affective Reports in Predicting Satisfaction Judgements," *Advances in Consumer Research*, 17, 571–576.
- Embretson, Susan E. (1996), "The New Rules of Measurement," *Psychological Assessment* 8, (4), 341–349.

- Embretson, Susan E. and Steven P. Reise (2000), *Item Response Theory for Psychologists*, Mahwah: Lawrence Erlbaum Associates.
- Engelhard, George, J. (1990), "History of Measurement," *Rasch Measurement Transactions*, 4, (4), 126-127, available online at <http://www.rasch.org/rmt/rmt44.htm>.
- Erevelles, Sunil (1998), "The Role of Affect in Marketing," *Journal of Business Research*, 42, (3), 199-215.
- Evrard, Yves and Philippe Aurier (1994), "The Influence of Emotions on Satisfaction with Movie Consumption," *Journal of Satisfaction, Dissatisfaction and Complaining Behavior*, 7, 119-125.
- Feick, Lawrence F. and Linda L. Price (1987), "The Market Maven: A Diffuser of Marketplace Information," *Journal of Marketing*, 51, (January), 83-97.
- Festinger, Leon (1957), *A Theory of Cognitive Dissonance*, New York: Row, Peterson and Company.
- Fischer, Gerhard H. and Ivo Molenaar (1995), *Rasch Models, Recent Developments and Applications*, New York: Springer Verlag.
- Fournier, Susan and David G. Mick (1999), "Rediscovering Satisfaction," *Journal of Marketing*, 63, (October), 5-23.
- Ganglmair, Alexandra and Rob Lawson (2002), "Conceptualizing Affective Satisfaction," in *Proceedings of the 2002 Australia and New Zealand Marketing Academy Conference*, R. N. Shaw, S. Adam S. and H. McDonald, Melbourne: Deakin University, CD ROM.
- Ganglmair, Alexandra and Rob Lawson (2003), "Advantages of Rasch Modeling for the Development of a Scale to Measure Affective Response to Consumption," *Advances in Consumer Research* (Europe), forthcoming.
- Gardial, Sarah F., D. Scott Clemons, Robert B. Woodruff and Mary Jane Burns (1994), "Comparing Consumers' Recall of Prepurchase and Postpurchase Product Evaluation Experiences," *Journal of Consumer Research*, 20, (March), 548-560.
- Giese, Joan L. and Joseph A. Cote (2000), "Defining Consumer Satisfaction," *Academy of Marketing Science Review [Online]* 00(01): Available online: <http://www.vancouver.wsu.edu/amsrev/theory/giese01-00.html>.
- Green, Donald, P., Susan L. Goldman, and Peter Salovey (1993), "Measurement Error Masks Bipolarity in Affect Ratings," *Journal of Personality and Social Psychology* 64, (6), 1029-1041.
- Halstead, Diane, David Hartman and Sandra L. Schmidt (1994), "Multisource Effects on the Satisfaction Formation Process," *Journal of the Academy of Marketing Science*, 22, (2), 114-129.
- Hambleton, Ronald K. (1991), *Fundamentals of Item Response Theory*, Newbury Park California: Sage Publications.
- Hunt, H. Keith (1977), "CS/D - Overview and Future Research Directions," in *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, Hunt H. K., Cambridge, Mass: Marketing Science Institute, 455-488.
- Hunt, H. Keith (1993), "CS/D & CB Research Suggestions and Observations for the 1990's," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 6, 40-42.
- Izard, Carol E. (1977), *Human Emotions*, New York: Plenum.
- Jacoby, Jacob (1978), "Consumer Research: A State of the Art Review," *Journal of Marketing*, 42, (April), 87-96.
- Kirkpatrick, Betty (ed.) (1987), *Roget's Thesaurus of English words and phrases*, Essex: Longman Group.
- LaBarbara, Priscilla and David A. Mazursky (1984), "A Longitudinal Assessment of Consumer Satisfaction/Dissatisfaction: The Dynamic Aspect of the Cognitive Process," *Journal of Marketing Research*, 20, (November), 394-404.
- Larsen, Jeff T., A. Peter McGraw and John T. Cacioppo (2001), "Can People Feel Happy and Sad at the Same Time?" *Journal of Personality and Social Psychology*, 81, (4), 684-696.
- LeDoux, Joseph (1996), *The Emotional Brain: The Mysterious Underpinnings of Emotional Live*, U.S.A.: Touchstone.
- Lord, Frederique M. (1980), *Applications of item response theory to practical testing problems*, Hillside NJ: Erlbaum.
- Lord, Frederique M. and Melvin Novick (1968), *Statistical theories of mental test scores*, Massachusetts: Addison-Wesley, Reading.
- Mackoy, Robert D. and Richard A. Spreng (1995), "The Dimensionality of Consumer Satisfaction/Dissatisfaction: An Empirical Examination," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 8, 54-58.
- Maddox R. Neil (1981), "Two-Factor Theory and Consumer Satisfaction: Replication and Extension," *Journal of Consumer Research*, 8, (June), 97-102.
- Mano, Haim and Richard L. Oliver (1993), "Assessing the Dimensionality and Structure of the Consumption Experience: Evaluation, Feeling, and Satisfaction," *Journal of Consumer Research*, 20, (December), 451-466.
- McCutcheon, Marc (1995), *Roget's Superthesaurus*, Cincinnati: Writer's Digest Books.
- Mehrabian, Albert and James A. Russell (1974), *An Approach to Environmental Psychology*, Cambridge Mass.: MIT Press.
- Nunnally, Jum C. and Ira H. Bernstein (1994), *Psychometric Theory*, New York: McGraw Hill.
- Nyer, Prashanth U. (1998), "The Effects of Satisfaction and Consumption Emotion on Actual Purchasing Behavior: An Exploratory Study," *Journal of Satisfaction, Dissatisfaction and Complaining Behavior*, 11, 62-68.
- Oliver, Richard L. (1980), "A Cognitive Model of Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17, (November), 460-469.

- Oliver, Richard L. (1981), "Measurement and Evaluation of Satisfaction Processes in Retail Settings," *Journal of Retailing* 57, (3), 25-48.
- Oliver, Richard L. (1989), "Processing of the Satisfaction Response in Consumption: A Suggested Framework and Research Propositions," *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, 2, 1-16.
- Oliver, Richard L. (1992), "An Investigation of the Attribute Basis of Emotion and Related Affects in Consumption: Suggestions for a Stage-Specific Satisfaction Framework," *Advances in Consumer Research*, 19, 237-244.
- Oliver, Richard L. (1994), "Conceptual Issues in the Structural Analysis of Consumption Emotion, Satisfaction, and Quality: Evidence in a Service Setting," *Advances in Consumer Research*, 21, 16-22.
- Oliver, Richard L., Roland T. Rust and Sajeew T. Vakri (1997), "Customer Delight: Foundations, Findings, and Managerial Insights," *Journal of Retailing*, 73, (3), 311-336.
- Peck, Bob (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, Perth: Murdoch University.
- Peterson, Robert A. and William R. Wilson (1992), "Measuring Customer Satisfaction: Fact and Artifact," *Journal of the Academy of Marketing Science*, 20, (1), 61-71.
- Rasch, Georg (1960/80), *Probabilistic Models for Some Intelligence and Attainment Tests; With a Foreword and an Afterword by Benjamin D. Wright*, Chicago: MESA Press.
- Richins, Marsha L. (1997), "Measuring Emotions in the Consumption Experience," *Journal of Consumer Research*, 24, (September), 127-146.
- Russell, James A. and James M. Carroll (1999), "On the Bipolarity of Positive and Negative Affect," *Psychological Bulletin*, 125, (1), 3-30.
- Ryan, J.P. (1983), "Introduction to Latent Trait Analysis and Item Response Theory," in *Testing in the Schools*, ed W.E. Hathaway, San Francisco: Jossey-Bass.
- Salzberger, Thomas (2002), personal communication, e-mail.
- Salzberger, Thomas (2000), "An Alternative Way of Establishing Measurement in Marketing Research - Its Implications for Scale Development and Validity," in *Proceedings of the 2000 Australia and New Zealand marketing Academy Conference (ANZMAC)*, ed. A. O'Cass, Gold Coast.
- Salzberger, Thomas, Rudolf R. Sinkovics and Bodo 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.
- Schumm, Walter R. (1999), "Satisfaction," in *Encyclopedia of Human Emotions*, David Levinson, James J. Ponzetti, Jr. and Peter F. Jorgensen, (eds.), New York: Macmillan Reference USA, 583-590.
- Shaver, Phillip, Judith Schwartz, Donald Kirson and Cary O'Connor (1987), "Emotion Knowledge: Further Exploration of a Prototype Approach," *Journal of Personality and Social Psychology*, 52, (6), 1061-1086.
- Sirgy, Joseph M. (1984), "A Social Cognition Model of CS/D: An Experiment," *Psychology and Marketing*, 1, (Summer), 27-44.
- Smith, Anne M. (1999), "Some Problems When Adopting Churchill's Paradigm for the Development of Service Quality Measurement Scales," *Journal of Business Research*, 46, (2), 109-120.
- Soutar, Geoffrey N. and Steven P. Cornish-Ward (1997), "Ownership Patterns for Durable Goods and Financial Assets: a Rasch Analysis," *Applied Economics*, 29, (7), 903-911.
- Soutar, Geoffrey N. and Gary Monroe (2001), "Rasch modeling: An alternative to summated scales," in *ANZAM 2001 - Closing the Divide*, Auckland.
- Soutar, Geoffrey N. and Maria Ryan (1999), "People's Leisure Activities: A Logistic Modeling Approach", *Proceedings of the 1999 Australia and New Zealand Marketing Academy Conference*, ed. Cadeaux, J. and Uncles M., Sydney: UNSW.
- Steinberg, Lynne and David Thissen (1996), "Uses of Item Response Theory and the Testlet Concept in the Measurement of Psychopathology," *Psychological Methods*, 1, (1), 81-97.
- Storm, Christine and Tom Storm (1987), "A Taxonomic Study of the Vocabulary of Emotions," *Journal of Personality and Social Psychology*, 53, (4), 805-816.
- Swan, John E. and Linda J. Combs (1976), "Product Performance and Consumer Satisfaction: A New Concept," *Journal of Marketing*, 40, (2), 25-33.
- Szymanski, David M. and David H. Henard (2001), "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29, (1), 16-35.
- Westbrook, Robert A. (1980), "A Rating Scale for Measuring Product/Service Satisfaction," *Journal of Marketing*, 44, (Fall), 68-72.
- Westbrook, Robert A. (1989), "Product/ Consumption-Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research*, 24, (August), 258-270.
- Westbrook, Robert A. and Richard L. Oliver (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," *Journal of Consumer Research*, 18, (June), 84-91.
- Williams, Patti and Jenifer L. Aaker (2002), "Can Mixed Emotions Peacefully Coexist," *Journal of Consumer Research*, 28, (March), 636-649.
- Wirtz, Jochen, Anna S. Mattila, and Rachel L. Tan (2000), "The Moderating Role of Target-Arousal on the Impact of Affect on Satisfaction: An Examination in the Context of Service Experiences," *Journal of Retailing*, 76, (3), 347-365.

- Wright, Benjamin D. (1977), "Solving Measurement Problems with the Rasch Model," *Journal of Educational Measurement*, 14, (2), 97-116.
- Wright, Benjamin D. (1980), "Foreword," in *Probabilistic Models for Some Intelligence and Attainment Tests; With a Foreword and an Afterword by Benjamin D. Wright*, Georg Rasch, Chicago: MESA Press.
- Wright, Benjamin D. (1997), "Fundamental Measurement for Psychology," *MESA Research Memorandum 64*, available online at <http://www.rasch.org/memo64.htm>.
- Wright, Benjamin D. and Mark H. Stone (1979), *Best Test Design*, Chicago: MESA Press.

**Send correspondence regarding this article to:**

Rob Lawson  
Department of Marketing  
University of Otago  
PO Box 56, Dunedin  
NEW ZEALAND  
fax: + 64(0)3 479 8172  
e-mail: [rlawson@business.otago.ac.nz](mailto:rlawson@business.otago.ac.nz)

---