

Recovery Satisfaction Construct and Construct-Related: Assessment Measurement from Item Response Theory

Martín Hernani-Merino

mn.hernanim@up.edu.pe

Department of Marketing and International Business
Universidad del Pacífico, Lima, Peru

Giuliana Isabella

Inspere – Institute of Education and Research, São Paulo, Brazil

Enver Gerald Tarazona Vargas

Department of Science
Pontificia Universidad Católica del Perú

José Afonso Mazzon

School of Economics, Management and Accounting
University of São Paulo, São Paulo, Brazil

Abstract

Researchers have explored the antecedents and consequences of recovery satisfaction by creating, improving or using scales. However, scales should be invariant among contexts and cultures. Using item response theory (IRT), a methodological approach that helps measure items, we evaluated some constructs related to it. Results, limitations and future research are discussed.

Keywords: Item response theory, recovery satisfaction.

<https://doi.org/10.21678/jb.2019.1137>

Paper received 23/08/2019 paper accepted 05/12/2019



Hernani-Merino, M., Isabella, G., Tarazona Vargas, E.G. & Mazzon, J.A. (2019) Recovery Satisfaction Construct and Construct-Related: Assessment Measurement from Item Response Theory. Journal of Business, Universidad del Pacífico (Lima, Peru) Vol.11(1): 2-13

Editor in Chief: Prof. Dr. Luis Camilo Ortigueira-Sánchez

Introduction

In failure service situations, companies attempt to recover customer satisfaction. Keeping customers satisfied is necessary to retain existing customers, and it is much less expensive than acquiring new customers (Mittal & Kamakura, 2001). Therefore, many researchers have explored the antecedents and consequences of recovery satisfaction, investigating the relationship with other constructs, such as perception of justice (Isabella & Mazzon, 2014), trust (Zhou, 2013), negative emotions (Kuo & Wu, 2012), switching (Pick & Eisend, 2013) or post-purchase intention (Holloway, Wang, & Parish, 2005). Consumer researchers explore these relationships by creating, improving or using scales. According to Wong, Rindfleisch and Burroughs (2003), marketing researchers are increasingly interested in testing their measurements and theories in different segments. It is common for one scale to be used in different contexts or regions (Steenkamp & Baumgartner, 1998). Thus, scales need to measure constructs independently of the segment, context, culture or nation. Within this paradigm, item response theory (IRT) has emerged, a methodological approach that helps measure items on a scale in an invariant way. Based on that the present research aims to evaluate the measures of the construct related to recovery customer satisfaction using IRT. The use of IRT models for ordinal data type remain rare in the marketing literature; however, researchers have recognized their contributions to measuring latent variables or constructs (Bacon, 2012; De Jong & Steenkamp, 2010). Such research implies that it is also important to consider the dissemination and application of these methods in consumer research.

To reach this goal, this paper used IRT together with the Classical Test Theory (CTT) to assess the proposed measures. The IRT complements the CTT since IRT focuses on the micro aspects of performance, that is, how well individual test items perform in measuring the construct. The main difference between this method and the classic one is that IRT assumes the probability of the correct response depends on the overall ability of the participant in answering each statement, in other words, it measures the “person ability” to answer a questionnaire. In addition, the IRT tests item difficulty, it is more precise and reliable measure compared to the overall measurement scales usually used on the models. An opposite of CTT, IRT considers the item parameter as independents of the sample used to generated parameters.

The IRT enables elaborate shorter scales without loss of reliability, since it permits to evaluate the items from the scale that have more information and permits to estimate (probabilities) on the properties of a scale in non representative samples (Embretson & Reise, 2009).

We opted to test the recovery satisfaction construct and other related constructs (perception of justice, trust, negative emotion, switching and post-purchase intention) during failure service situations because this kind of consumer satisfaction is very important in today's market. To evaluate the scales, we used the telecommunication failure service context. Since, the telecommunication signal is very important for a cell phone to work and, these companies have a high number of customer complaints compared to other segments.

We follow this paper with a theoretical development defining the constructs used in this paper. Then, we describe the analysis from an empirical study and discuss the results. We conclude with relevant managerial and theoretical implications, along with the limitation of the research and suggestion of future directions.

Theoretical Development

The constructs explored in this paper are related to recovery satisfaction. We opted to explore five variables that are related to recovery satisfaction: perception of justice, negative emotions, trust, switching and post intention of repurchase. Note that this paper does not aim to explore the relationships between the constructs.

Recovery satisfaction and its relationships

Satisfaction is widely accepted through the disconfirmation paradigm (Churchill Jr. & Surprenant, 1982; McCollough, Berry, & Yadav, 2000; Oliver, 1993). According to the paradigm, the people's initial expectations are confronted to the outcome and if it did as expected there will be a confirmation of the expectancies. However, if the outcome were worse than expected there will be a negative disconfirmation and if it did better than expected it will drive to a positive disconfirmation. Kuo & Wu (2012) have shown in an online shopping website the significances relations between perceived justice, negative emotions and post-recovery satisfaction. Justice theory and cognitive appraisal theory also suggested that perceived justice affects satisfaction directly and indirectly through emotions (Klaus Schoefer & Ennew, 2005).

In a relationship between customers and businesses, trust is associated with consistency of qualities, competence, honesty, integrity, responsibility and benevolence as well as having a central role in promoting cooperation between the parties (Morgan & Hunt, 1994). Therefore, trust is the expectation of one party with the other party will behave in a predictable manner in a specific situation. The existence of risks and uncertainties is a basic premise for the existence of the trust (Elliott & Yannopoulou, 2007). Confidence includes the predisposition to accept risks based on positive expectations of the intentions, behavior and integrity of another (Klaus Schoefer & Diamantopoulos, 2009). The trust is affected by perceptions of the trustee's ability, integrity, and benevolence. Zhou (2013) explored the three dimension on the context of acquisition cell phones. In this context, the ability is related to the recognition by the customer that the company can fulfill its tasks properly; integrity refers to the company honor its promises and the benevolence means that the provider also benefits from the relationship with the company. Then, exceeding or disappointing clients expectations can strength or weaken the reliability of the service providers. Based on that, we can say that satisfaction is regarded to influence trust directly (DeWitt, Nguyen, & Marshall, 2008; Zhou, 2013). In addition, many researches have empirically shown this relation through service recover situations (Ruyter & Wetzels, 2000).

"Switching is the likelihood of switching, the intent to switch and the actual switching behavior of a buyer to another seller" (Pick & Eisend, 2013, p.187). The most common construct related to Switching is the Switching Barrier or also called Switching Cost (Pick & Eisend, 2013). It consists in the perception of the magnitude of the additional cost required to terminate a relationship and secure an alternative one. The switching barrier involves psychological, emotional, search effort (time) and sometime monetary in the transaction between the customer and the firm (Patterson & Smith, 2003). Jones, Motherbaugh and Beatty (2000) use a more broadly conception, where switching barrier is any factor which make it more difficult or costly for customer to change provider. Some researchers have been using the idea of switching that has the similar concept, but analysis the customer commitment to stay with a service provider independently of how many or which barriers (Bougie, Pieters, & Zeelenberg, 2003). For instance, the perception of quality, customization, personalization or how the brand treats the customer can considered a barrier for a switching. Therefore, in this paper we used the idea of Switching instead of Switching Barrier, having a more broadly view of the concept.

Antecedents' constructs can influence perceived of switching differently according to the number of alternatives or competitor in the market (Pick & Eisend, 2013). In telecommunication services, when customers consider switch their providers they face a number of setup, one of them is the affection related to the company. To keep customers, it is common that telecommunications create special benefits, preferential treatment or rewards. For instance, it is common that providers give to their long-term customers mobile phones or discount for purchase a recent phone. In addition, this marked does not have many players, so it is not difficult that among the companies, customer perceive theirs companies as the best one.

Failure service can make customers experience negative feelings and also modify the level of satisfaction about a service. Based on the successful or unsuccessful service recovery, those levels of satisfaction can be influenced. Smith and Bolton (1998) explain that sometimes the recovery service can be so good that increase the satisfaction level that the customer had previously the failure service. They call this situation as a service recovery paradox. Therefore, a service recovery can influence positively or negatively customer switching (Lang, 2001). On the other hand, it seems clear that when a failed service encounter there will be a positive impact on switching drivers; negative experiences will increase switching (Keaveney, 1995; Pick & Eisend, 2013). In other words, when a failure service that evokes anger and negative emotions happens switching categories can be shaken (Bougie et al., 2003; Isabella & Mazzon, 2014; Xia & Kukar-Kinney, 2013).

Negative feelings from a failure service can modify the level of satisfaction about a service (Pick & Eisend, 2013). Based on the successful or unsuccessful service recovery, those levels of satisfaction can be influenced. However, Smith and Bolton (1998) explain that sometimes the recovery service can be so good that increase the satisfaction level that the customer had previously the failure service. They call this situation as a service recovery paradox. Therefore, a service recovery can influence positively or negatively customer satisfaction and also switching (Lang, 2001).

To conclude, it is not new that satisfaction has a main role on customer's attitude and post intention of repurchase (Holloway et al., 2005; Jones, Mothersbaugh, & Beatty, 2000). After a service failure, customer's intention of repurchase can be largely affected. For instance, through an online shopping experience, it was shown that positive effects on high satisfaction after a recovery on post purchase intention (Kuo & Wu, 2012).

Method

Surveying college students in a metropolitan city we evaluated the construct measurements (items). All the participants had cellphones and had service account with a telecommunication company. Participants answered questions in a computer lab about a telecommunication failure service context. During some sections, we had 265 completed participants' surveys, of which 244 were considered valid. The survey was presented on participant native language at Qualtrics software that allows questions randomization. Students received credit course to participant on the study.

To evaluate the measurement model, we employed existing scales translating to the participants native language (Portuguese) and adapting the wording as necessary to suit the context and the culture. The participants responded to six scales using seven-point rating scales. All the scales were anchored by 1 = "I strongly disagree" and 7 = "I strongly agree". The name of the constructs and the amount of items were: negative emotions (4 items) scales adapted from Schoefer & Ennew (2005) and Schoefer &

Diamantopoulos (2008); trust scale with three dimensions – benevolence, integrity and capacity - adapted from Gefen & Straub (2004) - 12 items; satisfaction scale from Kuo and Wu (2009) - 3 items; an adapted switching scale from Lang (2001) - 3 items and the post-purchase intention scale adapted from Kuo, Wu and Deng (2012) - 3 items. These original scales are described on Table 1.

The IRT of graduated answers selects the most representative's items based on the Likert Scale was proposed by Singh Howell and Rhoads (1990) in marketing context. The measurement model (item analysis) is evaluated in two ways: (1) based on the classical theory (CTT), analyzing mean, standard deviation, estimating Cronbach's alpha coefficient for the construct and the item-total correlation, and (2) based on the Item Response Theory that estimates the parameters of discrimination (a) difficulty (b) through the Graded Response Model (GRM) (Samejima, 1969).

The first task of the IRT is feasible the models to discovery the item parameters. The idea is to evaluate the performance of a subject about an item. In statistic words, the probability of a correct answer [$\pi(\theta)$] depending on: (1) the ability of the subject or theta (θ) and (2) the parameters of items (a_i , b_i and c_i).

The GRM from IRT helps in estimating the parameters and ability (θ) of survey respondents. This model is designed to be applied to instruments with polytomous or ordinal attributes. However, the design is suitable to be used when responses to the item can be characterized as ordinal categorical responses (for three or more categories) as in the case of a Likert scale. An overview of the estimation of the parameters of this model is described in Bazán, Mazzon and Hernani-Merino (2011). Further details of the estimation of the model can be found in Tarazona (2013). Aiming to perform the 2PL-GRM calculations we used the IRTPRO software, which is a recent application for estimation the ability score and the items parameters of IRT.

Analysis

In order to evaluate the scales proposed in this paper, it was applied the Classical Test Theory (CTT), which shows the correlation score of a particular item with the total score and the IRT, in which presents the relationship of an item to the underlying construct reflected in the discrimination parameter.

The table 1 summarizes and consolidates all analyzes made. It reports the statistics of CTT and the parameter estimates from IRT for all items of the six constructs. As CTT measures we present mean item-total correlation and the Alpha if that item is deleted from the construct. As IRT we show the parameters for each construct. The column labeled item-total correlation shows the different weights of the items to each eight construct. Observe that all the items had higher correlations (above 0.50) with their respective constructs. Also note that the items "degree of satisfaction" and "degree of contentment" had a high item-total correlation ($r=0.89$) (in satisfactions) compared with the other items belonging to the remaining seven constructs. The analyses show that all items on the scale are useful in defining their respective constructs. The IRT and the CTT together show that all items on all the scales are useful in defining their respective constructs. The discrimination parameters estimate following the Graded Response Model from Samejima (1969) can be seen at IRT "a". The quality of the discrimination parameter is considered moderated when the values are between 1.0 and 2.0 and high when those are above 2.0 (Hafsteinsson, Donova, & Breland, 2007). Observe in the Table 1 that all items of the model showed high or moderate quality.

Table 1 - Failure service items properties

Construct (Alpha)	Item	CTT Statistics			IRT Model Item Properties						
		Mean (SD)	Item-total correlation	Alpha if item delete	a	b1	b2	b3	b4	b5	b6
Negative Emotion (0.83)	1 The solution that the company gave to me made me angry.	4.20 (2.24)	0.70	0.767	2.81	-0.95	-0.59	-0.39	0.00	0.39	0.87
	2 I got upset with the solution that the company gave to me.	4.26 (2.24)	0.71	0.76	3.91	-0.96	-0.55	-0.35	0.01	0.31	0.70
	3 I had a negative feeling with the treatment that the company offered me.	4.47 (2.25)	0.71	0.761	3.33	-1.04	-0.67	-0.44	-0.14	0.23	0.60
	4 I got unhappy with the solution that the company gave me.	4.21 (2.36)	0.52	0.847	1.56	-1.15	-0.65	-0.48	-0.06	0.41	0.79
Satisfaction (0.94)	1 Degree of contentment (after calling the company)	3.15 (1.61)	0.89	0.912	6.46	-0.69	-0.27	0.04	0.84	1.54	2.28
	2 Degree of satisfaction (after calling the company)	3.12 (1.68)	0.89	0.916	5.96	-0.63	-0.28	0.14	0.82	1.27	2.29
	3 Degree of happiness (after calling the company)	3.33 (1.56)	0.87	0.928	5.74	-0.84	-0.41	-0.10	0.80	1.55	2.30
Post Intention of Purchase (0.87)	1 My intention is to continue using this service company.	4.29 (1.78)	0.80	0.804	4.70	-1.33	-0.87	-0.47	-0.04	0.66	1.31
	2 I would recommend this company service to my friends and relatives.	3.49 (1.77)	0.71	0.841	2.52	-0.91	-0.60	-0.02	0.64	1.34	1.95
	3 I could acquire more services from this company if their services interest me.	4.58 (1.81)	0.63	0.872	1.94	-1.83	-1.26	-0.84	-0.27	0.58	1.31
	4 The probability to continue using the services from this company is high.	4.28 (1.77)	0.76	0.823	3.45	-1.44	-0.91	-0.49	0.01	0.64	1.37
Switching (0.73)	1 I use the services from this company because it is the best choice for me.	4.58 (1.72)	0.56	0.641	1.80	-2.01	-1.44	-0.90	-0.20	0.73	1.46
	2 The service quality this company offers is higher than the service quality of other service providers	3.83 (1.76)	0.60	0.589	2.68	-1.29	-0.78	-0.27	0.50	1.04	1.64
	3 I have grown to like this service provider more than other service providers in this category	3.77 (1.77)	0.51	0.71	1.55	-1.34	-0.88	-0.30	0.49	1.33	2.55
Justice (0.87)	1 The employee seemed to be very interested in my problem	3.71 (1.77)	0.63	0.88	1.67	-1.57	-0.64	-0.21	0.26	1.38	2.67
	2 The company reacted positively when I complained.	3.73 (1.71)	0.73	0.84	2.64	-1.30	-0.61	-0.13	0.29	1.16	2.12
	3 The solution offered by the company was just.	3.56 (1.85)	0.81	0.81	5.48	-0.85	-0.42	-0.07	0.41	1.03	1.52
	4 Overall, the company's complaint handing procedure was fair.	3.65 (1.89)	0.76	0.82	3.98	-0.93	-0.43	-0.04	0.37	0.96	1.52
Trust Integrity (0.80)	1 Promises made by the telecommunication company are likely to be reliable	3.81 (1.68)	0.68	0.739	2.67	-1.30	-0.78	-0.28	0.46	1.31	1.80
	2 I do not doubt the honesty of this company	3.71 (1.65)	0.59	0.781	1.79	-1.49	-0.84	-0.28	0.69	1.49	2.49
	3 I expect that this company will keep promises they make I	3.54 (1.75)	0.71	0.723	3.21	-1.04	-0.51	0.01	0.57	1.21	1.81
	4 I expect that the advice given by this company is their best judgment	3.41 (1.82)	0.55	0.801	1.69	-1.11	-0.49	0.23	0.71	1.51	2.30
Trust Benevolence (0.82)	1 I expect I can count on this company to consider how its actions affect me	3.77 (1.72)	0.63	0.783	2.19	-1.34	-0.71	-0.27	0.36	1.31	2.10
	2 I expect that this company intentions are benevolent	3.95 (1.66)	0.73	0.74	3.22	-1.40	-0.81	-0.41	0.27	1.09	1.88
	3 I expect that this company puts customers' interests before their own	3.07 (1.71)	0.53	0.828	1.62	-0.99	-0.24	0.40	1.13	1.87	2.75
	4 I expect that this company is well meaning	3.85 (1.60)	0.70	0.752	2.63	-1.45	-0.90	-0.35	0.43	1.26	2.09
Trust Competence (0.81)	1 This company is competent	3.63 (1.70)	0.72	0.753	3.51	-1.06	-0.64	-0.17	0.50	1.29	1.83
	2 This company understands the market it works in	3.93 (1.68)	0.60	0.805	1.67	-1.60	-1.09	-0.50	0.52	1.38	2.26
	3 This company knows about signal telecommunication	3.75 (1.74)	0.64	0.787	1.91	-1.50	-0.73	-0.15	0.53	1.33	2.05
	4 This company knows how to provide excellent service	3.22 (1.68)	0.65	0.782	2.59	-0.89	-0.40	0.15	0.89	1.67	2.13

In addition, to discrimination it is important to note the difficulty or the location that the items are presented. Thus, in CTT, mean scores measure the item difficulty; indicating the severity of the item; usually low average denotes a low agreement with the item while a high average suggests a high correlation with the item. For instance in the item "I could acquire more services from this company if their services interest me" (in post intention of repurchase) had the highest average with 4.58 showing that the level of agreement is high.

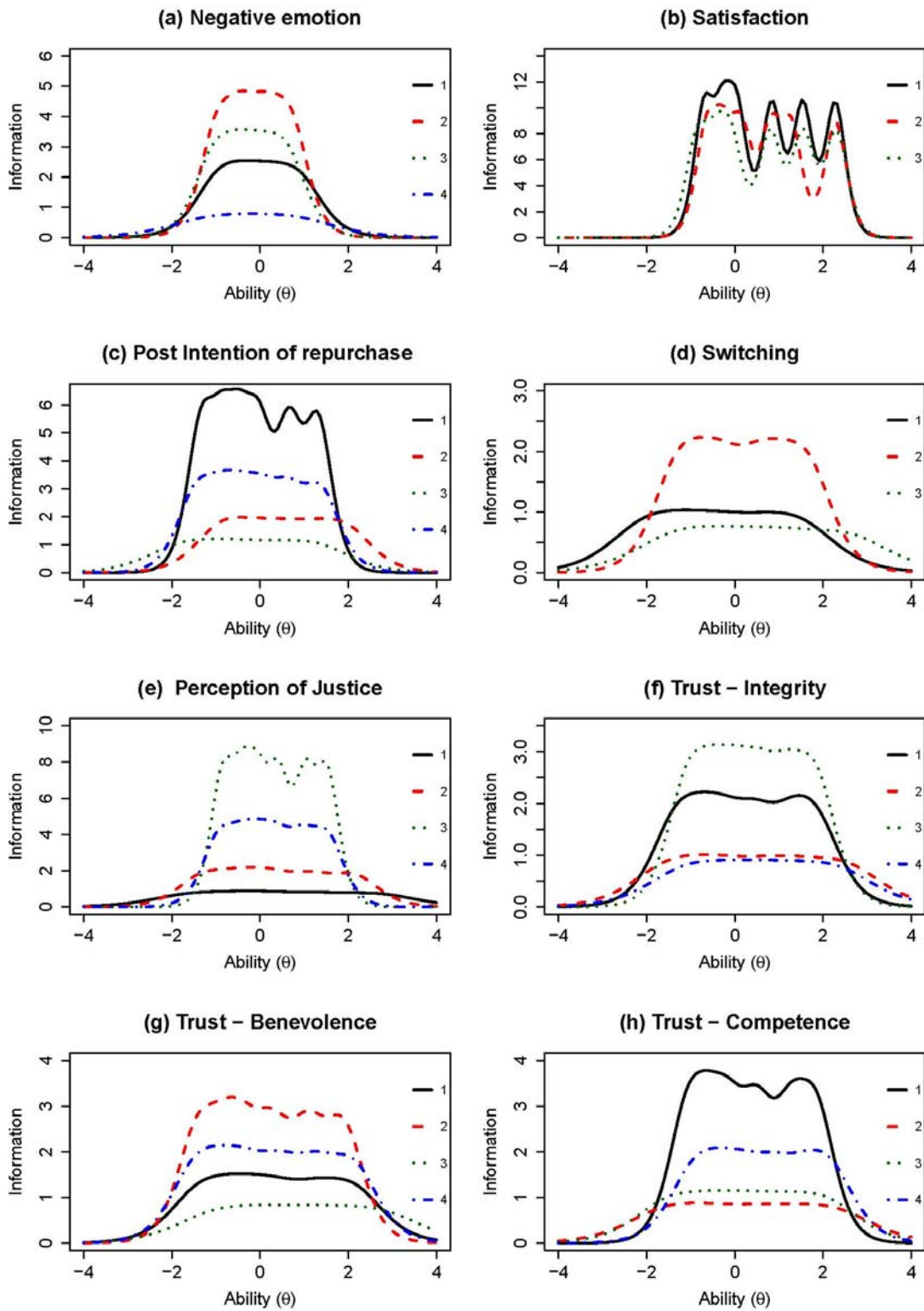
The parameters from the IRT show the difficulty of the items and they are represented in the table with the letter "b". Observe that the items have 7 score points therefore the distance of difficulty between one point to another are 6 (b1 to b6). The item "I expect that this company puts customers' interests before their own" from Trust Benevolence construct (b6 = 2.75) is considered by the participant to be the most difficult to reach the last category (total agreement). On the other hand, the item "I had a negative feeling with the treatment that the company offered me" from Negative Emotion construct is the easier one to reach the last category (b6 = 0.60). The items in accordance find less resistance because it would be more readily accepted by responding in comparison with the other items.

The difference in the difficulty of the items can be used to manage services offered by a company to find potential drivers in different dimensions or constructs under studies (Battisti, Nicolini, & Salini, 2005). For example, in the related recovery satisfaction construct, considering that the participant/customer is satisfied when they score the scale as 6 or 7 (towards total agreement), we can say that the parameter b5 indicates the threshold that must be overcome to achieve ideal levels of satisfaction. However, observe that the items "Degree of happiness (after calling the company)", it the most difficult for customer to evaluate (b5=1.55). Based on the three items of the construct of satisfaction it is the highest score, what means that the association of happiness to satisfaction has a lower quality. In other words, this situation indicates that the item is associated with a lower quality and needs to be by companies to prioritized to achieve an improvement in satisfaction construct. Therefore, the cellular telephone company should pay more attention to this specific item, enabling strategizes more focused on problems specific to bring happiness and consequently satisfaction. Similar interpretations can be obtained in the remaining constructs and considering different levels of scale.

The functions of information indicates that upper range of ability (θ) in which an item or scale could best discriminate between individuals which denotes information more accurately (or reliability). Embretson and Reise (2009) comment that the information functions of the items can identify good or bad item performance. Little information for a particular item means that: it measures something distinct from the other items on the scale; it is poorly written and needs to be rewritten; it is too complex for respondents; or it is taken out of context in the questionnaire.

We also present graphically on Graph 1 the information function of latent trait from eight constructs (a) Negative emotion, (b) Satisfaction, (c) Post Purchase Intention, (d) Switching, (e) Perception of Justice and (f.g.h) Trust - divided in its three dimensions: (f) Integrity, (g) Benevolence and (h) Competence. Note that the items 1,2 and 3 from (a); 1,2 and 3 in (b); 1 and 4 in (c); 2 in (d); 3 and 4 in (e); 1 and 3 in (f); 2 and 4 at (g) and 1 and 4 in (h) have greater amount of information within their respective latent features. Based on this amount of information, these items for each construct could be used to measure the respective constructs in any type of sample. For example, the satisfaction construct, Graphic 1 (b), the item 1, "Degree of contentment (after calling the company)" has more information which would indicate that the item could be stored in a database of items measure and may be used for any type of sample that measures satisfaction.

Graph 1



Graph 1 - Item Response Theory (IRT) analysis – Item information function

Final Consideration

Failure service is a critical issue for companies (McCollough et al., 2000). It is cheaper to keep customer satisfied than reach and acquire a new customer (Mittal & Kamakura, 2001). In this study, we first identified in the literature constructs that were connecting with recovery satisfaction. Using a telecommunication failure service context, we evaluated the measures of each of the constructs from the perspective IRT.

Working with IRT we could obtain more information about the items in the scales. IRT is a great technic to evaluate items from scales, permitting researchers to define short scales, measuring a construct without loss of significant information, or permitting researchers to choose the best item from the construct scales when they cannot use many questions. For instance, in “neuromarketing”, it is common consumers researchers use few questions since it is expensive to run a study inside a fMRI or in PETs (for more information see Isabella, Mazzon, Dimoka (2013)). While the TCC uses Alpha Cronbach to evaluate the reliability of a construct, the IRT uses the information function indicating that the upper range of ability where an item or scale could better discriminate individuals. Therefore, superior information denotes more accurately or reliability of the item in the construct (Embretson & Reise, 2009).

This research contributes to management field since companies can identify specific critical points in the service provided to its consumers. This could be appreciated when we analyzed the difficulty parameter (b) to achieve the highest degree of agreement for items of satisfaction construct. Based on that, companies can develop marketing strategies of services according to the degree of difficulty to match a particular item. For instance, in the satisfaction recovery, it was identified that the item "I am happy with the solution/after calling this company" is the most difficult to achieve a high degree of agreement among the participants. Therefore, companies could provide training to their attendants to improve happiness and satisfaction by call centers.

As all the researches this one has some limitations: (1) we used specific groups - undergraduate students; (2) we run few scales related to recovery satisfaction however many other scales could be used; (3) we also opted for one scale for each construct, different scales could be used to test which one would be the best; (4) following the original scales we opted for a 7 point-Likert scales, other measurement could be tested and (5) we did not explored the relationship between the constructs. Future researches could be made to study the abilities of the respondents in the context of satisfaction in services and many other marketing contexts.

To conclude, the focus of this research was evaluate the measurements of each constructs presented, the IRT possible to evaluate the ability of each of the respondents regarding the information function for each item. This research brings significant contributions to understanding consumer behavior in a service context in the light of a methodological approach little explored in marketing.

References

Bacon, D. R. (2012). Understanding Priorities for Service Attribute Improvement. *Journal of Service Research*, 15(2), 199–214.

Battisti, F. De, Nicolini, G., & Salini, S. (2005). The Rasch Model to Measure Service Quality. *Journal of Services Marketing*, 3(3), 58–80.

Bazán, J. L., Mazzon, J. A., & Hernani-Merino, M. (2011). Classificação de Modelos de resposta ao item Policotômicos com aplicação ao Marketing. *Revista Brasileira de Estatística*, 72(235), 7–39.

Bougie, R., Pieters, R., & Zeelenberg, M. (2003). Angry Customers don't Come Back, They Get Back: The Experience and Behavioral Implications of Anger and Dissatisfaction in Services. *Journal of the Academy of Marketing Science*, 31(4), 377–393.

Churchill Jr., G. A., & Surprenant, C. (1982). An Investigation into the Determinants of Customer Satisfaction. *Journal of Marketing Research*, 19(4), 491–504.

De Jong, M. G., & Steenkamp, J.-B. E. M. (2010). Finite Mixture Multilevel Multidimensional Ordinal IRT Models. *Psychometrika*, 75(1), 3–32.

DeWitt, T., Nguyen, D. T., & Marshall, R. (2008). Exploring Customer Loyalty Following Service Recovery: The Mediating Effects of Trust and Emotions. *Journal of Service Research*, 10(3), 269–281.

Elliott, R., & Yannopoulou, N. (2007). The nature of trust in brands: a psychosocial model. *European Journal of Marketing*, 41(9/10), 988–998.

Embretson, S. E., & Reise, S. P. (2009). *Item Response Theory for Psychologist*. New Jersey: Psychology Press.

Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega*, 32(6), 407–424.

Hafsteinsson, L. G., Donova, J. J., & Breland, B. T. (2007). An item response theory examination of two popular goal orientation measures. *Educational and Psychological Measurement*, 67, 719–739.

Holloway, B. B., Wang, S., & Parish, J. T. (2005). The role of cumulative online purchasing experience in service recovery management. *Journal of Interactive Marketing*, 19(3), 54–66.

Isabella, G., & Mazzon, J. A. (2014). It is not fair to pay more! A study of the perception of justice and price fairness of justice and price fairness in hedonic and utilitarian purchases. In T. Brown & V. Swaminathan (Eds.), *AMA Winter 215*. San Antonio, TX.

Isabella, G., Mazzon, J. A., & Dimoka, A. (2013). Consumer Behavior and Neurophysiology Tools: A Perspective in Latin America. In *CLAV - 6o Congresso Latino-Americano no Varejo: Varejo e Desenvolvimento em Mercados Emergentes*.

Jones, M. a, Mothersbaugh, D. L., & Beatty, S. E. (2000). Switching barriers and repurchase intentions in services. *Journal of Retailing*, 76(2), 259–274.

Keaveney, S. M. (1995). Customer Switching Behavior in Service Industries: An Exploratory Study. *Journal of Marketing*, 59(2), 71–82.

Kuo, Y.-F., & Wu, C.-M. (2012). Satisfaction and post-purchase intentions with service recovery of online shopping websites: Perspectives on perceived justice and emotions. *International Journal of Information Management*, 32(2), 127–138.

Kuo, Y.-F., Wu, C.-M., & Deng, W.-J. (2009). The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services. *Computers in Human Behavior*, 25(4), 887–896.

Lang, M. C. B. (2001). Switching barriers in consumer markets: an investigation of the financial services industry. *Journal of Consumer Marketing*, 18(4), 332–347.

McCollough, M. a., Berry, L. L., & Yadav, M. S. (2000). An Empirical Investigation of Customer Satisfaction after Service Failure and Recovery. *Journal of Service Research*, 3(2), 121–137.

Mittal, V., & Kamakura, W. A. (2001). Satisfaction, Repurchase Intent , and Repurchase Behavior : Investigating the Moderating Effect of Customer Characteristics. *Journal of Marketing Research*, XXXVIII(February), 131–142.

Morgan, R. M., & Hunt, S. D. (1994). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, 58(July), 20–38.

Oliver, R. L. (1993). Cognitive, Affective, and Attribute Bases of the Satisfaction Response. *Journal of Consumer Research*, 20(3), 418–430.

Pick, D., & Eisend, M. (2013). Buyers' perceived switching costs and switching: a meta-analytic assessment of their antecedents. *Journal of the Academy of Marketing Science*, 42(2), 186–204.

Ruyter, K. de, & Wetzels, M. (2000). Customer equity considerations in service recovery: a cross-industry perspective. *International Journal of Service Industry Management*.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrica Monograph*, 17, 1–100.

Schoefer, K., & Diamantopoulos, a. (2008). The Role of Emotions in Translating Perceptions of (In)Justice into Postcomplaint Behavioral Responses. *Journal of Service Research*, 11(1), 91–103.

Schoefer, K., & Diamantopoulos, A. (2009). A Typology of Consumers' Emotional Response Styles during Service Recovery Encounters. *British Journal of Management*, 20(3), 292–308.

Schoefer, K., & Ennew, C. (2005). The impact of perceived justice on consumers' emotional responses to service complaint experiences. *Journal of Services Marketing*, 19(5), 261–270.

Singh, J., Howell, R. D., & Rhoads, G. K. (1990). Adaptive designs for likert-type data: an approach for implementing marketing surveys. *Journal of Marketing Research*, 27(3), 304–321.

Smith, A., & Bolton, R. N. (1998). An experimental investigation of consumer reactions to service failure and recovery encounters: paradox or peril? *Journal of Service Research*, 1(1), 65–81.

Steenkamp, J. B. E. M., & Baumgartner, H. (1998). Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research*, 25(June), 78–90.

Tarazona, E. G. (2013). Modelos alternativos de respuesta graduada con aplicaciones en la calidad de servicios. Pontificia Universidad Católica del Perú. Access at: <http://perustat.com/blog/aplicaciones-de-la-teoria-de-respuesta-al-item-en-la-medicion-de-la-satisfaccion-de-clientes/>

Wong, N., Rindfleisch, A., & Burroughs, J. E. (2003). Do Reverse-Worded Items Confound Measures in Cross-Cultural Consumer Research? The Case of the Material Values Scale. *Journal of Consumer Research*, 30, 72-91.

Xia, L., & Kukar-Kinney, M. (2013). Examining the Penalty Resolution Process: Building Loyalty Through Gratitude and Fairness. *Journal of Service Research*, 16(4), 518–532.

Zhou, T. (2013). An empirical examination of the determinants of mobile purchase. *Personal and Ubiquitous Computing*, 17(1), 187–195.