ISSN (PRINT): 1994 - 7658 ISSN (ONLINE): 1314 - 0817

VARNA UNIVERSITY OF MANAGEMENT

European Journal of Tourism Research

http://ejtr.vumk.eu

Adapting and Validating Scale of Customer Engagement in Online Travel Communities

Peter J. Mkumbo [1], Dandison C. Ukpabi [2*] and Heikki Karjaluoto [3]

[1] Department of Parks, Recreation and Tourism Management, Clemson University, USA. E-mail: pmkumbo@g.clemson.edu

[2] Marketing Department at University of Jyväskylä, Finland. E-mail: dandison.c.ukpabi@jyu.fi

[3] Marketing Department at University of Jyväskylä, Finland. E-mail: heikki.karjaluoto@jyu.fi

* Corresponding author

Abstract

Increasing attention towards customer engagement has caused its measurement to remain a highly debated issue among scholars. This study adapts and validates measurement scales for customer engagement in Online Travel Communities. It builds on previous studies on scale development for customer engagement. Data were collected from 450 members of Online Travel Communities in eight countries: Australia, Canada, Hong Kong (Chinese territory), New Zealand, Singapore, South Africa, the United Kingdom and the United States of America. The process of adapting and validating the scale involved exploratory factor analysis, testing for differential item functioning, examination of item response theory, confirmatory factor analysis, testing for invariance and criterion validity. The results found that three dimensions (affection, absorption, and interaction) suitably and adequately measure customer engagement in Online Travel Communities.

Keywords: Customer engagement; Online Travel Communities; Psychometric scale development and validation

Citation: Mkumbo, P., C., Ukpabi, D. and Karjaluoto, H. (2020). Adapting and Validating Scale of Customer Engagement in Online Travel Communities. *European Journal of Tourism Research* 25, 2501

Introduction

The literature on consumer engagement (CE) is growing. Since its adoption into the consumer research literature, several studies have applied it in different contexts, such as brand community (Algesheimer, Dholakia, & Herrmann, 2005), Facebook (Cheung, Lee, & Jin, 2011), automobiles (Sarkar & Sreejesh, 2014) and mobile phones (Dwivedi, 2015), among others. The growing interest is due to scholarly evidence that highlights that CE is a driver of customer trust, value, affective commitment, satisfaction and loyalty (Bowden, 2009; Brodie, Hollebeek, Juric, & Ilic, 2011; Vivek, Beatty, Dalela, & Morgan, 2014). Additionally, other critical indicators of brand performance, such as profit, sales growth and return on investment, have been linked to CE (Harrigan, Evers, Miles, and Daly, 2017; Hollebeek, 2011).

While some of these studies are contextualised offline (Moreau, 2011; Sarkar and Sreejesh, 2014), social media has been the context for the majority (Cheung *et al.*, 2011; Dessart, Veloutsou, and Morgan-thomas, 2016; Hollebeek, Glynn, and Brodie, 2014). This is mainly due to social media's fostering of CE through relationship creation and sustenance (Sashi, 2012), interaction and value co-creation and customer experience management (Brodie, Ilic, Juric, & Hollebeek, 2013). Similarly, social media aggregates brand enthusiasts into communities where they connect, share experiences of hospitality and travel services and offer common programmes that influence and advance a brand (Dessart, Veloutsou and Morgan-Thomas, 2016). Thus, customer engagement defined as '. . . the level of an individual customer's motivational, brand-related and context-dependent state of mind characterised by specific levels of cognitive, emotional and behavioural activity in direct brand interactions' (Hollebeek, 2011, p. 790) has continued to attract both scholarly and practitioner attention in recent times.

Although some attempts have been made to measure CE, there is still no scholarly consensus on the most appropriate dimensions for which populations (Table 1). Furthermore, within the hospitality and tourism domain, So, King, and Sparks (2014) and Harrigan *et al.* (2017) have differed on their dimensions and scale for CE. While So *et al.* (2014) proposed a 25-item CE scale with 5 dimensions, Harrigan *et al.* (2017) contended that 3 dimensions with 11 items are sufficient. To this end, Harrigan *et al.* (2017, p. 607) recommended that 'future research should validate the CE scale and model using random samples in countries with varying cultures.' Consequently, by building on these two studies, the aim of this study is to contribute to this debate by adapting and validating the CE scale that suits OTCs with samples drawn from eight countries: Australia, Canada, Hong Kong (Chinese territory), New Zealand, Singapore, South Africa, the United Kingdom (UK) and the United States of America (USA). Thus, our study's key contributions to the hospitality and tourism literature is that we adapt, examine and validate the CE scale with tourism brands as proposed by So *et al.* (2014) and applied by Harrigan *et al.* (2017) in online travel communities (OTCs), which constitute critical engagement platforms between hospitality and tourism service providers and customers. To this end, consumers' interest in travel sites, such as TripAdvisor, Booking.com and Expedia, among others, have continued to grow, and they remain helpful in travel decisions (Xiang & Gretzel, 2010).

This paper proceeds as follows: in the next section, we present the literature review; section three discusses the methodology; section four presents the results; and section five offers the discussion, implications and limitations.

Literature review

CE has enjoyed increasing attention in practitioner and consumer behaviour literature (Harrigan *et al.*, 2017) because of its enduring benefits to firms in relation to other customer-centric activities, such

as advertising and loyalty programmes (So *et al.*, 2014). Since the extension of engagement into consumer behaviour research began, it has been applied in different research domains (Table 1). The table (Table 1) also contains how different studies have measured CE including the dimensionality, number of items used and sample diversity. This is in a bid to demonstrate that there is no uniform measurement of CE in extant research. However, different customer touchpoints across a wide spectrum of a firm's activities, such as advertising, product or service offerings or even an event, engender engagement platforms (Vivek, Beatty and Morgan, 2012), with brand communities accentuating critical engagement platforms in which closer interaction between the firm and customers occurs. Interestingly, social media has enjoyed the most attention in the CE body of knowledge because it enhances real-life and two-way communication between the firm and customers (Dessart, Veloutsou and Morgan-Thomas, 2016).

Extant research has documented antecedents as well as consequences of CE. We refer to Van Doorn *et al.* (2010) for explicit discussion of antecedents and consequences of customer engagement behaviour. towards both the firm and the customer. Hospitality and tourism services are experience-based; hence, they are personal and memorable and often regarded as high-involvement consumption contexts (Hur *et al.* (2017). Similarly, participation in OTCs builds relationships with fellow customers and with the brand (Casaló, Flavián and Guinalíu, 2010a).

CE measurement scale and its dimensions

Despite the growth and increasing attention towards CE, its measurement remains one of the most debated topics in the general service literature (Hollebeek *et al.*, 2014; Vivek *et al.*, 2014) and tourism research domain (Harrigan *et al.*, 2017; So *et al.*, 2014). With contextual differences, these studies have either measured CE unidimensionally or multidimensionally (Table 1). However, most subsequent studies dwelled extensively on the affective, cognitive and behavioural dimensions (Dessart *et al.*, 2016). As an extension of affective commitment, Bowden, (2009) argued that the application of affective CE implies an emotional state, such as enjoyment, passion and enthusiasm, towards the focal firm and/or brand. Additionally, cognitive CE embodies a customer's attention, absorption and sustained active interests in the firm and/or its brand (Brodie *et al.*, 2011; Hollebeek, 2013). Finally, behavioural CE has been conceptualised as the vigour and energy that encompass a customer's interaction with a brand (Brodie *et al.*, 2011). Interestingly, CE measurement within hospitality and tourism research has included five dimensions: enthusiasm, identification, attention, absorption and interaction (Harrigan *et al.*, 2017; So *et al.*, 2014).

Enthusiasm

A customer's positive perception of a service or product leads to greater interest and provides fertile ground for CE (van Doorn *et al.*, 2010), with further interaction leading to enthusiasm. Enthusiasm reflects one of the ways in which customers emotionally connect to a brand (Hollebeek, 2013), and it implies a 'strong level of excitement' regarding the firm or focal brand (So *et al.*, 2014, p. 308). Consumers who join brand communities and/or recommend a brand to others are driven by enthusiasm for that brand. Although the literature is unclear regarding whether enthusiasm differs from passion (Hollebeek, 2011, 2013), the two are often regarded as critical parts of CE's emotional component (Hollebeek, 2011). In the context of CE measurement with hospitality and tourism brands, So *et al.* (2014) maintained that there are five dimensions, including enthusiasm, while Harrigan *et al.* (2017) contended that enthusiasm can be merged with absorption.

Table 1. *Dimensions and measurement of CE in extant research*

Author(s)	Context	Dimensionality		Number of items	Sample diversity (Number of countries)
		Uni-dimensional	Multidimensional		
Algesheimer *et al.* (2005)	Brand community	Community engagement	-	4	1
Cheung *et al.* (2011)	Facebook users	-	Vigour, absorption, dedication	18	1
Enginkaya and Esen, (2014)	Online	-	Trust, dedication, reputation	14	1
Hollebeek *et al.* (2014)	Social media users	-	Cognitive, refers to "a consumer's degree of positive brand-related affect, activation	10	1
So *et al.* (2014)	Travellers	-	Identification, enthusiasm, attention, absorption, interaction	25	1
Sarkar and Sreejesh, (2014)	Car owners	Active engagement	-	4	1
Vivek *et al.* (2014)	Students	-	Conscious attention, enthused participation, social connection	10	1
Dwivedi, (2015)	Mobile phone users	-	Vigour, dedication, absorption	17	1
Dessart, Veloutsou, and Morgan-Thomas, (2016)	Facebook users	-	Affective (enthusiasm, enjoyment); cognitive (attention, absorption); behavioural (sharing, learning, endorsing)	22	3
Harrigan *et al.* (2017)	Amazon Mechanical Turk	-	Identification, absorption, interaction	11	1

Identification

Identification originated in social identity theory (So *et al*., 2014), which holds that an individual's self-concept represents his or her personal identity (the 'I') and the group with which he/she associates (Tajfel, 1982). In the marketing context, a consumer's sense of identification with a firm and/or brand is based on the ability of the firm and/or brand to satisfy his or her self-definitional goal (So *et al*., 2014), which can occur through either cognitive or affective processes (Stokburger-Sauer, Ratneshwar, & Sen, 2012). The cognitive process comprises the consumer's perception of his or her own personality traits in relation to the brand while the affective process comprises the memorable experiences pertaining to the brand (Stokburger-Sauer *et al*., 2012). Interestingly, both So *et al*. (2014) and Harrigan *et al*. (2017) found support for identification as a valid factor for measurement of CE in hospitality and tourism brands.

Attention

From the organizational behaviour literature, attention partly constitutes the cognitive component of CE (Hollebeek, 2013; Vivek *et al*., 2012). It is the customer's focus and conscious participation in issues relating to a brand and/or firm (So *et al*., 2014; Vivek *et al*., 2012). Customers who are engaged with a brand learn more and spend more time thinking about the brand (So *et al*., 2014). The use of attention as a factor of CE has been measured through various methods in the general marketing literature. For instance, Vivek *et al.'s* (2014) measurement of conscious attention is similar to Hollebeek's (2011) dimensions of immersion and activation, supporting attention as a valid factor of CE. However, in the measurement of CE in hospitality and tourism brands, So *et al*. (2014) considered attention a key dimension, whereas a study by Harrigan *et al*. (2017) found no support for its inclusion.

Absorption

Absorption is an extension of flow theory (So *et al*., 2014), which describes the state of total immersion that an individual encounters when engrossed in a given activity (Nakamura & Csikszentmihalyi, 2014). In the context of CE, absorption is a state wherein the customer experiences intrinsic enjoyment and is fully concentrated, happy with and engrossed in the brand (So *et al*., 2014; Vivek *et al*., 2012). There is no scholarly consensus on how absorption is measured as a dimension of CE. For instance, while Cheung *et al*. (2011) and Dwivedi, (2015) have acknowledged absorption as a distinct dimension of CE, Dessart *et al*. (2016) consider it a component of the cognitive dimension. Notably, the two studies that examined scale measurement and validation of CE in hospitality and tourism brands found support for its inclusion (Harrigan *et al*., 2017; So *et al*., 2014).

Interaction

As a dimension of CE, interaction is critical because it constitutes the window through which engagement takes place, and it is practical because it involves communicating one's feelings about the brand (So *et al*., 2014). Brand communities constitute an important forum for brand enthusiasts to demonstrate their connection with the brand (Merrilees, 2016). However, time and distance pose critical challenges to this platform (Ukpabi & Karjaluoto, 2017). By contrast, social media has liberalised CE; through many platforms, engaged customers can post, write reviews, blog and share content (images, videos) of their favourite brands and experiences. While there is a scarcity of scholarly evidence that recognises interaction as a distinct factor in measuring CE, it has been recognised as a valid measurement of CE in hospitality and tourism brands (Harrigan *et al*., 2017; So *et al*., 2014).

Continuance participation

The introduction of a piece of information system is usually accompanied with studies on users' attitudes and intentions to using it (Boateng, Adam, Okoe & Anning-Dorson (2016; Israel, Tscheulin & Zerres, 2019). Within the information system literature, the technology acceptance model (TAM) has been prominently used to examine users' attitude at the pre-adoption stage (Ooi & Tan, 2016). However, exposure to the system can influence can influence their attitude to continue or discontinue its usage. Thus, continuance usage intention, which is underpinned by users' satisfaction, is critical to the survival of a piece of an information system (Bhattacherjee, 2001). As many online communities are facing critical challenges of retaining customers (Zhou *et al.* 2012), thus, understanding the interrelationships between customer engagement and continuance participation would be important for theory and practice. Extant studies have linked continuance intention to customer loyalty as they related to post adoption behaviours (Zhou *et al.* 2012) and the consequence of customer satisfaction (Anastassova, 2011; Cao *et al*., 2013; Moise, Gil-Saura & Ruiz-Molina, 2018).

Methodology

Survey design and data collection

As mentioned previously, this study builds on So *et al.*, (2014) customer engagement in tourism brands by adapting the scale in online travel communities. To test, validate, and adapt the scale that measures CE in the context of online travel communities, a questionnaire was designed and was distributed, using online panel company, to panels in eight countries, including Australia, Canada, Hong Kong (Chinese territory), New Zealand, South Africa, Singapore, the United Kingdom and the United States of America. The English-speaking countries were chosen due to limited resources to translate the questionnaire into other non-English languages. The designed questionnaire targeted only members of OTCs who reside in the eight countries. An individual was considered a member of an OTC if they had either one or multiple user accounts with an online travel website, such as TripAdvisor or LonelyPlanet, or had liked a social media page related to tourism and travel and are regularly following posts on those pages, websites or blogs.

Measures

Measurement scales for CE were adapted from the initial list of 28 items in So *et al.* (2014). The scale was composed of five dimensions: (1) identification, (2) enthusiasm, (3) attention, (4) absorption, and (5) interaction (Table 3). To check for content validity this list of items was emailed to 17 researchers who are professors or have deep expertise in tourism marketing, tourism management, and quantitative research methods especially, development of psychometric measures. In addition to these researchers, five tourism marketing managers were also consulted. These researchers and marketing managers were asked to assess the content validity of each dimension and the overall domain validity of the scale. They were provided with operational definition of customer engagement as behavioural manifestations that have a brand or firm focus, beyond purchase, resulting from motivational drivers (Verhoef, Reinartz, & Krafft, 2010); and online travel community as a group of individuals with shared travel, tourism or hospitality interests brought together by a travel, tourism or hospitality related brand using an online platform. Of the 17 researchers that were emailed, 14 responded with mixed responses. Some commented on the list of items that it is too long and, in many contexts, it is impractical to ask all those questions or there will be a serious concern with common method bias (Baumgartner & Weijters, 2012; Podsakoff, Mackenzie, & Podsakoff, 2012). Some researchers commented on the wording of the items so as to properly adapt the scale and make it suitable for customer engagement in online travel communities. A few other researchers expressed their doubts that five dimensions are too much, there are a lot of redundant items and that some need to be dropped out. All researchers seemed to agree that customer engagement in online travel communities

is a sub-scale/sub-component of the general customer engagement in tourism brands thus, in part, supporting the conceptualization and justification put forward in earlier sections of this study. All marketing managers responded and their comments, as among researchers above, were a little mixed. Critical comments were on the number of items being too long; the wording of items needs to be general but focusing on specific dimension and remain relevant on online travel communities, and item statements need to be shorter to reduce ambiguity among non-native English speakers.

The authors in this study decided to retain all 28 items, made sure that the wording of items sufficiently and fully capture the nomological radii of the dimensions and also use of plain language to reduce ambiguity. Table 2 shows a list of all items and dimensions.

Continuance participation (CPA) was used as a criterion variable for examining criterion validity and the scale for CPA was adopted from Hur, Terry, Karatepe, & Lee, (2017). The scale contains seven items (Table 3).

Pre-test, a priori power analysis, and data collection
A pre-test of the questionnaire was conducted in which a total of 50 respondents filled out the questionnaire. The distribution to panels was done under the help of an online panel company. Preliminary analyses (descriptive and correlation statistics) were conducted to gain a picture of the study, how variables are related so as to inform improvements if there were any in the final questionnaire. Power analysis was of most importance.

A two-stage *a priori* power analysis was conducted; stage i) to determine a minimum sample size to detect model misfit (Kline, 2016) and stage ii) to estimate minimum sample size for sufficient power to detect an effect when testing for construct predictive validity and thus avoiding type II error (Farrokhyar, Reddy, Poolman, & Bhandari, 2013). Power analysis for stage (i) was conducted using semPower() package in R (Moshagen, 2018; Moshagen & Erdfelder, 2016) and the code argument contained the following minimum specifications for fit indices at multiple phases; power = 0.80, AGFI = 0.95, CFI = 0.95, GFI = 0.95, RMSEA = 0.05, alpha = 0.05, df = 994, and p = 47. Power analysis at this stage suggested a minimum sample size to detect model misfit is 110. Stage (ii) of power analysis was conducted using G*Power (Faul, Erdfelder, Buchner, & Lang, 2019), an online tool useful for *a priori* power analysis. In conducting power analysis using G*Power, a minimum correlation (*r*) between variables of .2 (based on pre-test results) and the highest correlation of .8 were used. The result suggested a sample size bigger than 53 will be needed. Since this project had other objectives beyond those mentioned in this study it was decided to target a minimum of 450 participants so as to allow for advanced and sophisticated analyses at subgroup levels. Minimum quotas for each country were abruptly suggested based on the country's population as follows: Australia (50), Canada (50), Hong Kong (30), New Zealand (30), South Africa (30), Singapore (30), the UK (50) and the USA (80).

Filters were set in the questionnaire to exclude untargeted participants and those who did not qualify for this study. Questionnaire validation was set 'force response' for questions related to measurement scales and 'request response' for general demographic questions. Algorithms were programmed to discard all incomplete responses. Since pre-test results showed that the median time to complete the questionnaire was 7.5 minutes; an extra filter was set to discard all questionnaires that were completed below two-thirds of the median time (i.e. below five minutes). In the final version of the questionnaire, items were measured on a seven-point Likert-type scale strongly disagree (1) to strongly agree (7). The questionnaire was distributed to online panels using Qualtrics algorithms. The analysis of data went through numerous rigorous stages namely; exploratory factor analysis (EFA), a

two-phase differential item functioning (DIF) testing, item response theory (IRT) testing, confirmatory factor analysis (CFA) for assessing measurement models, test for invariance (configural and factorial) and finally CFA for testing criterion validity of the scale in the structural model. The analysis procedure is presented in the following sections.

Table 2. *List of items and dimensions for CE adapted from So et al., (2014)*

Dimension	Code	Item statement
Identification (CEID)	ID1	When someone criticizes my online travel community, it feels like a personal insult.
	ID2	I am very interested in what others think about my online travel community.
	ID3	When I talk about my online travel community, I usually say WE rather than THEY.
	ID4	This online travel community's successes are my successes.
	ID5	When someone praises this online travel community, it feels like a personal compliment.
Enthusiasm (CEEN)	EN1	I spend a lot of my discretionary time thinking about this online travel community.
	EN2	I am heavily into this online travel community.
	EN3	I am passionate about this online travel community.
	EN4	My days would not be the same without this online travel community.
	EN5	I am enthusiastic about this online travel community.
	EN6	I feel excited about this online travel community.
Attention (CEAT)	AT1	I would like to learn more about this online travel community.
	AT2	I pay a lot of attention to anything about this online travel community.
	AT3	Anything related to this online travel community grabs my attention.
	AT4	I concentrate a lot on this online travel community.
	AT5	I spend a lot of time thinking about this online travel community.
	AT6	I focus a great deal of attention on this online travel community.
Absorption (CEAB)	AB1	When I am interacting with this online travel community, I forget everything else around me.
	AB2	Time flies when I am interacting with this online travel community.
	AB3	I get carried away when I am interacting with this online travel community.
	AB4	It is difficult to detach myself when I am interacting with this online travel community.
	AB5	I become immersed when I am interacting with this online travel community.
	AB6	I feel happy when I am interacting intensely with this online travel community.
Interaction	IT1	In general, I like to get involved with this online travel

(CEIT)		community's discussions.
	IT2	I am someone who likes to actively participate in this online travel community.
	IT3	I am someone who enjoys interacting with like-minded people in this online travel community.
	IT4	In general, I thoroughly enjoy exchanging ideas with other people in this online travel community.
	IT5	I often participate in the activities of this online travel community.

Table 3. *List of CPA items adopted from Hur, Terry, Karatepe, & Lee, (2017)*

Code	Item statement
CPA1	I intend to continue participating in this online community
CPA2	My intentions are to continue being a member of this online travel community
CPA3	If I could, I would like to continue participating in this online travel community
CPA4	I will continue being an active member of this online travel community
CPA5	I will continue reading posts in this online travel community
CPA6	I intend to continue sharing my travel opinions in this online travel community
CPA7	I intend to continue sharing my travel experiences in this online travel community

Analysis and results

A total of 450 questionnaires were completed. Of them, one was disregarded as it was found (using Mahalanobis Distance a multivariate outlier analysis in SPSS version 25) to be an outlier. The data were checked for normality assumptions using Q-Q-plot, skewness, and kurtosis (Tabachnick & Fidell, 2012) in SPSS; and was found to be multivariate normal.

Study sample characteristics

Table 3 shows the demographics of the study sample. Table 4 also shows the forms of engagement in online travel communities among the individuals who were involved in this study. The majority (n = 204) participate by reading others' opinions and recommendations, while those who read others' posts and post their own travel experiences equalled 192. The smallest group (n = 54) included individuals who mainly post their experiences and travel opinions.

Table 4. *Participants' demographics*

Demographic variables		**Frequency**	**Percentage**
Country	Australia	58	12.9
	Canada	73	16.2
	Hong Kong	41	9.1
	New Zealand	31	6.9
	Singapore	37	8.2
	South Africa	46	10.2
	UK	82	18.3
	USA	82	18.3
Gender	Female	262	58.1

	Male	185	41.0
	LGBTQ*	3	0.9
Marital status	Single	172	38.1
	Married	221	49.0
	Divorced	25	5.5
	Widowed	8	2.0
	Co-habiting	24	5.3
Age range (years)	18–28	105	23.3
	29–38	145	32.2
	39–48	81	18.0
	49–58	54	12.0
	59–68	41	9.1
	69+	24	5.5
Education	High school	178	39.5
	Bachelor's degree	196	43.5
	Master's degree	59	13.1
	PhD	11	2.4
	No formal education	6	1.6
Nature of participation	By posting my travel experiences and opinions	54	12.0
	By reading others' opinions and recommendations	204	45.2
	By both reading others' posts and posting my own travel experiences and opinions	192	42.6

Notes: * LGBTQ – lesbian, gay, bisexual, transgender and queer

Participants in this study were also asked to mention and rank up to three OTCs that they engage with the most. A sum of 290 online travel communities was mentioned in a total of 1,353 frequencies. Table 5 shows the list of OTCs, with a total of 20 frequencies and above, in which TripAdvisor, Expedia, and Trivago were the top three,

Table 5. *Top online travel sites mentioned*

S/N	OTC/Site/Platform	Frequency
1	TripAdvisor	188
2	Expedia	117
3	Trivago	102
4	Booking.com	54
5	Facebook	52
6	Hotels.com	37
7	Airbnb	32
8	Travelocity	26
9	Agoda	22
10	Kayak	22
11	Me Want Travel	22

Exploratory factor analysis

To reduce dimensions and identify those that stand out, exploratory factor analysis (EFA) was conducted using SPSS version 25 under maximum likelihood (ML) method of estimation and promax an oblique rotation. Factors were extracted using the eigenvalue cut-off of 1 and above. The results show KMO measure of sampling adequacy = .976 and Bartlett's test of sphericity $\chi^2_{(378)} = 15421.21(p < .001)$ suggesting that model matrix significantly differ from the identity matrix (Dziuban & Shirkey, 1974; Field, Miles, & Field, 2012; Kaiser, 1974) and thus the data characteristics had met the necessary conditions for factor analysis (Conway & Huffcutt, 2003; Floyd & Widaman, 1995; Yong & Pearce, 2013). EFA results suggested three factors to be extracted. The three factors combined accounted for 77.1% of the total variance (Table 6). EFA was re-run under the same settings as previously with an addition of a specification of three factors to be extracted. An item was retained for further analysis if a factor it belongs to explains at least or near 50% of its (item) variance (Janssens, Wijnen, Pelsmacker, & Kenhove, 2008). From the three factors, six items (AT1, AT2, AT3, EN5, EN6, and ID2) were found to have weak loadings and potentially cross-loading onto multiple dimensions, they were thus dropped out (Little, Lindenberger, & Nesselroade, 1999). The three factors with thirteen items that are retained accounted for 87% of the total variance.

Table 6. *EFA pattern Matrix*

Item	Factor		
	1	2	3
EN1	0.799	-0.101	0.196
ID1	0.761	-0.179	0.085
ID4	0.752	0.185	-0.105
ID5	0.713	0.149	-0.024
EN4	0.710	-0.009	0.232
AT5	0.694	-0.029	0.282
AT6	0.685	0.034	0.232
ID3	0.678	0.276	-0.122
EN2	0.676	0.109	0.148
EN3	0.655	0.237	0.052
AT4	0.620	0.126	0.202
ID2	0.589	0.390	-0.146
AT3	0.422	0.299	0.127
EN5	0.430	0.371	0.054
AT2	0.390	0.299	0.127
EN6	0.420	0.375	0.124
IT4	-0.154	0.975	0.071
IT3	0.014	0.916	-0.034
IT2	0.140	0.778	-0.003
IT1	0.099	0.772	0.043
IT5	0.196	0.573	0.175
AT1	0.394	0.545	-0.014
AB5	-0.025	0.038	0.911
AB3	0.037	0.077	0.811
AB4	0.210	-0.097	0.803
AB2	-0.033	0.203	0.714
AB1	0.311	-0.116	0.667
AB6	-0.156	0.498	0.576

Testing for differential item functioning (DIF)

Differential item functioning happens when item parameters differ across groups in the same population i.e. the way the item is being interpreted differ across groups within the population of interest (Nunnally & Bernstein, 1994). DIF helps to identify items that are biased towards one group of the population, it is essentially a test for item bias. DIF along with item response theory (IRT) forms an integral part of item analysis in developing psychometric measures. An ideal scale should not have items that function significantly different across groups within the same population (Boateng, Neilands, Frongillo, Melgar-Quiñonez, & Young, 2018; Dima, 2018). In this study retained items, under each of the three extracted factors, were subjected into a series of tests for differential item functioning in SPSS version 25. Factor scores were used as the conditioning variables since they are latent variables and are computed using both psychometric and measurement theories, thus are less contaminated with measurement errors (Slocum, Gelin, & Zumbo, 2004). Two grouping variables, gender, and country of residence were used to test DIF. Specifically each item under each extracted factor was tested to see if it functions not significantly different across gender and also across residents in different geographical locations in six continents, eight countries to make sure that the items and the overall scale function the same way regardless of the cultural background of the participant. A *p*-value of 0.01 was used as a cut-off point (Slocum *et al.*, 2004). After a series of DIF tests, nine more items (AT4, AT5, AT6, ID3, ID4, ID5, AB1, AB6, and EN3) were found to be significantly functioning differently across either gender or country of residence, uniform or non-uniform DIF (Hanson, 1998; Woods & Grimm, 2011), they were therefore, dropped out. At this point thirteen items were retained in the three factors; one factor contained five items while the two other factors contained four items each (Table 7). The dimension that hosts items ID1, EN1, EN2, and EN4 was renamed to affection in line with earlier study by Hollebeek *et al.*, (2014) and similar findings by van Tonder & Petzer (2018).

Table 7. Extracted factors and retained items

Item	Factor		
	1	2	3
IT4	**0.952**	0.100	-0.173
IT3	**0.940**	-0.031	-0.018
IT2	**0.848**	-0.028	0.102
IT1	**0.796**	-0.022	0.150
IT5	**0.689**	0.157	0.151
AB3	0.046	**0.877**	0.004
AB5	-0.010	**0.860**	0.071
AB4	-0.074	**0.747**	0.256
AB2	0.187	**0.736**	-0.044
EN1	-0.015	-0.001	**0.948**
EN2	0.203	0.035	**0.713**
ID1	-0.080	0.080	**0.690**
EN4	0.117	0.170	**0.699**

Looking at Tables 6 and 7 one would see a key observation from exploratory factor analysis that the majority of items in the dimensions of identification (CEID) and attention (CEAT) loaded, even if not strongly, onto the same factor while item ID1 loads onto enthusiasm dimension hence the dimension was renamed to affection. These results partly echoed those of Harrigan *et al.* (2017) and van Tonder & Petzer (2018). The dimension of absorption (CEAB) retained four items while the dimension of interaction (CEIT) retained its all five items (Table 7).

Examining item characteristics

Item response theory (IRT) was used to examine item characteristics. IRT also commonly referred as latent train theory (LTT) (Linden, 2016a) evolved due to the limitations of classical test theory (CTT) (Crocker & Algina, 2006; Gulliksen, 1950; Novick, 1966; Woodbury, 1963) that it does not routinely position in the core of its concerns how individuals at different levels of the construct being studied perform on the items, of a scale aimed at measuring the underlying latent dimension (Finch, Immekus, & French, 2016; Raykov & Marcoulides, 2011). To test for IRT in this study, two libraries in the R programming environment, mirt() (Chalmers, 2012) and ltm()(Rizopoulos, 2006) were used. Graded response model (GRM) fitted the data better (AIC = 15495.42; BIC = 15963.62; LogLik = -7633.711) than generalized partial credit model (GPCM) (AIC = 15697.39; BIC = 16165.59; LogLik = -7734.697). Therefore, GRM was a model of preference over GPCM. Additionally, a two-parameter IRT model (2PL) was found to be more relevant and informative to use over one-parameter Rasch model (1PL). Attention was paid to five main statistical outputs; item fits, item difficulty coefficient *b*, item discrimination coefficient *a*, and endorsement of alternative responses on a 7-point Likert type scale of an item (Linden, 2016b; Nering & Ostini, 2010).

All items in all dimensions were found to be fit as assessed by RMSEA (values range .000 to .042) and all Chi-square values under given degrees of freedom were non-significant. In terms of item difficulty, all items were found to be measuring difficulty on a good range below-average, average and above-average. All items appeared to discriminate well responses on a 7-point Likert type scale as evaluated using *a* coefficients being bigger than 1 (Linden, 2016; Linden, 2016a). A 7-point Likert type scale appears to be the right scale to use as each possible responses on the scale received distinct endorsement from participants in all items except item ID1 (under EEN dimension) in which response 3 (somewhat disagree) did not receive distinct endorsement, however, it was retained as it manifested suitable item difficulty and discrimination coefficients (Baker & Kim, 2017; Linden, 2016; Nering & Ostini, 2010; Thissen, Pommerich, Billeaud, & Williams, 1995). So, the 13 items in three dimensions were all retained for further analyses. The data file was then randomly split into two files; calibration sample (n = 225) and validation sample (n = 224). The majority of subsequent analyses are first conducted using the calibration sample and then the findings are corroborated using validation sample as the standard procedure in scale development (Linacre, 1994).

Measurement models

A series of confirmatory factor analysis (CFA) was conducted using the covariance-based structural equation modeling (CB-SEM) approach. EQS software version 6.4 was used to analyse the data where maximum likelihood (ML) and robust estimations were used in all stages of analysis of measurement and structural models, additionally robust results (including Satorra-Bentler scaled Chi-square S-Bχ2) are presented as Mardia's coefficient was larger than 5 (Byrne, 2006). Models were assessed at both; dimensional and higher-order levels. Results showed that measurement models for both at dimensional and higher-order are well specified for both; calibration sample (S-Bχ2 = 104.79, df = 62, p < 0.001, SRMR = 0.039, CFI = 0.98, RMSEA = 0.056) and validation sample (S-Bχ2 = 109.9, df = 62, p < 0.001, SRMR = 0.037, CFI = 0.98, RMSEA = 0.059) as well as for continuance participation (S-Bχ2 =

17.6, df = 14, p = 0.09, CFI = 0.99, SRMR = 0.02, RMSEA = 0.04) which will be used later for predictive validity (Table 8). In addition, Table 6 presents means (M) and standard deviations (SD) of items. All three dimensions in calibration and validation samples passed discriminant validity assessment (Table 9) in which square roots of AVEs were observed to be larger than correlations between factors (Fornell & Larcker, 1981).

Table 8. *Calibration sample: Measurement model*

Dimension	Code	M	SD	Loading (λ)	AVE	α	CR (ρ)	2nd Order (λ)	SQRT-AVE	Fit Indices
Enthusiasm (CEEN)	ID1	3.72	1.74	0.68						
	EN1	3.48	1.81	0.90						
	EN2	3.69	1.78	0.89						
	EN4	3.53	1.80	0.91	0.72	0.91	0.91	0.95	0.85	S-Bχ^2 = 11.79, df = 1, p < 0.01, SRMR = .028, CFI = 0.98, RMSEA = 0.148 90%CI = [.074, .234]
Absorption (CEAB)	AB2	4.04	1.75	0.87						
	AB3	3.81	1.78	0.91						
	AB4	3.47	1.80	0.91						
	AB5	3.83	1.79	0.92	0.81	0.94	0.94	0.88	0.90	S-Bχ^2 = 4.27, df = 1, p = 0.12, SRMR = .011, CFI = 1, RMSEA = .071 90%CI = [.000, .166]
Interaction (CEIT)	IT1	4.38	1.66	0.89						
	IT2	4.22	1.62	0.90						
	IT3	4.42	1.66	0.91						
	IT4	4.50	1.63	0.89	0.80	0.95	0.95	0.86	0.90	S-Bχ^2 = 7.15, df = 5, p = 0.21, SRMR = 0.014, CFI = 1, RMSEA =0 .044 90%CI = [.000, .109]
	IT5	4.16	1.73	0.89						

Higher order GOF: S-Bχ2 = 104.79, df = 62, p < 0.001, SRMR = 0.039, CFI = 0.98, RMSEA = 0.056, 90%CI = [.036, .073] n = 225

Validation sample: Measurement model

Dimension	Code	M	SD	Loading (λ)	AVE	α	CR (ρ)	2nd Order (λ)	SQRT-AVE	Fit Indices
Enthusiasm (CEEN)	ID1	3.69	1.65	0.67						
	EN1	3.54	1.73	0.92						
	EN2	3.67	1.74	0.92						
	EN4	3.62	1.87	0.90	0.74	0.91	0.92	0.96	0.86	S-Bχ^2 = 4.02, df = 1, p = .13, SRMR = 0.013, CFI = 0.99, RMSEA = 0.067 90%CI = [.000, .163]
Absorption (CEAB)	AB2	4.12	1.71	0.81						
	AB3	3.70	1.66	0.91	0.77	0.93	0.93	0.92	0.88	S-Bχ^2 = 5.97, df = 1, p = 0.51,

	AB4	3.39	1.70	0.89						SRMR = 0.014, CFI = 0.99, RMSEA = 0.094 90%CI = [.000, .185]
	AB5	3.79	1.62	0.89						
Interaction (CEIT)	IT1	4.33	1.49	0.91	0.79	0.95	0.96	0.85	0.89	S-Bχ^2 = 15.48, df = 5, p < 0.01, SRMR = 0.02, CFI = 0.98, RMSEA 0.079, 90%CI = [.044, .153]
	IT2	4.18	1.52	0.92						
	IT3	4.50	1.51	0.87						
	IT4	4.49	1.47	0.86						
	IT5	4.09	1.60	0.87						

Higher order GOF: S-Bχ2 = 109.9, df = 62, p < 0.001, SRMR = 0.037, CFI = 0.98, RMSEA = 0.059, 90%CI = [.040, .076] n = 224

Continuance participation (CPA) (N = 449)

Construct	**Code**	**M**	**SD**	**Loading (λ)**	**AVE**	**α**	**CR (ρ)**	**2nd Order (λ)**	**SQRT-AVE**	**Fit Indices**
Continuance participation (CPA)	CPA1	5.01	1.21	0.83	0.66	0.94	0.94	NA	0.81	S-Bχ2 = 17.6, df = 14, p = 0.09, CFI = 0.99, SRMR = 0.02, RMSEA = 0.04, 90%CI = [.000, .061]
	CPA2	4.97	1.28	0.87						
	CPA3	4.97	1.32	0.89						
	CPA4	4.91	1.31	0.87						
	CPA5	5.36	1.16	0.72						
	CPA6	4.83	1.45	0.74						
	CPA7	4.84	1.47	0.76						

Table 9. *Factor correlation matrices*

Calibration sample					**Validation sample**			
S/N	**Dimension**	**1**	**2**	**3**	**Dimension**	**1**	**2**	**3**
1	CEEN	(0.85)			CEEN	(0.86)		
2	CEAB	0.84	(0.90)		CEAB	0.85	(0.88)	
3	CEIT	0.82	0.76	(0.90)	CEIT	0.81	0.78	(0.89)

Diagonal values in brackets are the square roots of AVEs

Testing for invariance

In addition to DIF, testing for invariance assesses whether the scale measures the same construct across different samples drawn from the same population (Meredith, 1964; Putnick, Diane & Bornstein, Mark, 2016). While DIF assessed how individual items functions across demographical groups in the same population, invariance in this study assessed both items and dimensions across different samples from the same population. Testing for invariance is an essential step to make sure that differences, if there are any, between samples are not due to scale's psychometric properties (Lee, 2018; Meredith, 1993). Two assessments were conducted to test for invariance between calibration and validation samples; configural and metric invariances (Cheung & Rensvold, 2002; Lee, 2018; Timmons, 2010). On the one hand, testing for configural invariance assesses if the general construct structure is consistent across different samples; calibration and validation in this study. On the other hand, metric

invariance examines if pattern coefficients (loadings) do not vary significantly across calibration and validation samples (Byrne, 2012; Kline, 2016). In both assessment procedures, EQS version 6.4 software was used. The results for configural invariance suggested that the proposed scale as examined across two samples, calibration (S-Bχ2 = 104.79, df = 62, p < 0.001, SRMR = 0.039, CFI = 0.98, RMSEA = 0.056) and validation (S-Bχ2 = 109.9, df = 62, p < 0.001, SRMR = .0037, CFI = 0.98, RMSEA = 0.059) is configural invariant (S-Bχ2 = 214.52, df = 124, $\Delta\chi$2 = 0.17, SRMR = 0.038, CFI = 0.98, RMSEA = 0.057, 90%CI = [.044, .070]) meaning that the construct structure is consistent across different samples from the same population. In examining metric invariance, constraints were placed in each pair of loading path in calibration and validation samples. The results for metric invariance also suggested that item loadings are invariant (S-Bχ2 = 228.76, df = 137, $\Delta\chi$2 = 14.24, Δdf = 13, p = 0.36, SRMR = .054, CFI = .98, RMSEA = .055, 90%CI = [.042, .067]) across the two samples (Byrne, 2006). In addition to the fit indices of overall metric invariance model, all applied loading constraints across two samples were found to be statistically not significant suggesting that the loadings do not vary significantly across samples of online travel communities. After having confirmed the proposed scale is invariant, the two samples; calibration and validation were re-combined into one and used as a single sample in the subsequent analyses and examinations.

Testing for criterion validity

Criterion validity could simply be defined as the extent to which there is a significant relationship between a given construct (which the scale is being developed for) and performance on another construct of particular relevance, commonly referred to as criterion variable (Cronbach & Meehl, 1955; DeVellis, 2017; Raykov & Marcoulides, 2011). Criterion validity exists in two main forms concurrent validity and predictive validity (Cronbach & Meehl, 1955; Dima, 2018).

Concurrent criterion validity is the degree to which the construct (which a scale is developed for) has a stronger relationship with a theoretically-supported relevant criterion variable made at the time of data collection or shortly afterward (Cronbach & Meehl, 1955; DeVellis, 2017). Predictive criterion validity is the extent to which a construct (which the scale is developed to measure) predicts the response to another criterion variable which it is expected to relate with (Boateng *et al*., 2018; Chan, 2014) and therefore the scale should be able to predict expected behavioural intention or actual behaviour in the future.

In this study, continuance participation (CPA) is used as a criterion construct. An increase in customer engagement with the online travel community should sustain participation of that particular customer on the OTC platform, thus CPA is a suitable criterion construct to examine predictive criterion validity. The scale for CPA is adopted from Hur, Terry, Karatepe, & Lee, (2017). The fact that data for both proposed construct and criterion construct were collected at the same time, makes it (CPA) also a relevant criterion construct for evaluating concurrent criterion validity. Prior to examining criterion validity, a Harman's single factor test was conducted to examine for common method bias in which was found that a single factor accounts for 39% of the total variance. Even though a single factor accounted for less than 50% of the total variance (Harman, 1976) a common latent factor (CLF) was added into the structural model to control for common method bias by accounting for common variance (Craighead, Ketchen, Dunn, & Hult, 2011; Eichhorn, 2014; Podsakoff *et al*., 2012; Podsakoff & Organ, 1986). EQS version 6.4 software was used. The structural model was adequately specified (S-Bχ2 = 284.86, df = 130, p <.001SRMR = 0.060, CFI = 0.98, RMSEA = 0.052, 90%CI = [.043, .060]) and as expected customer engagement (CE) in online travel communities (OTC) significantly predicts continuance participation (CPA) (Figure 1, Table 8) a standard deviation increase in customer engagement in OTC increases continuance participation by 0.832 standard

deviation. In a plain language, a unit increase in CE on a 7-point Likert scale increases CPA by 0.832 units on a 7-point Likert scale. Customer engagement in OTC explains 69% of the variance in continuance participation (Figure 1). The results supported that the scale is criterion valid.

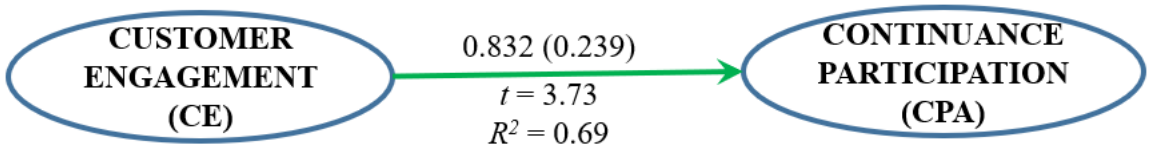


Figure 1. *Examining criterion validity*

Table 10. *Examining criterion validity*

Criterion variable	Predictive variable	B	Beta	S.E.	*t*-statistic	*p* - value	R^2	Result
CPA	CE	0.239	0.832	0.064	3.734	0.00011	0.69	Supported

Key: B = Unstandardized path coefficient; Beta = Standardized path coefficient; CE = Customer engagement; S.E. = Standard Error; R^2 = Coefficient of determination; CPA = Continuance participation.

Discussion and conclusion

While So *et al.* (2014) suggested five factors for general customer engagement with tourism brands, this study suggests three dimensions to be used when measuring customer engagement in online travel communities. Findings of this study partly support those of Harrigan *et al.*, (2017) (who used only Facebook fans) but the factor structures of the construct differ. The items in three extracted factor were further examined through DIF and IRT to make sure that each item is consistent in measuring what it is supposed to measure regardless of demographic differences among participants in samples drawn from the same populations (Baker & Kim, 2017; Hanson, 1998). Additionally, the scale invariance was examined for configural and metric and was found to be invariant. Finally, the scale was found to be criterion valid. A total of 15 items with poor psychometric properties were dropped out.

The key objective of this study was to adapt a general customer engagement scale with tourism brands into online travel communities (OTCs). Contrary to So *et al.* (2014), this study suggests that customer engagement in online travel communities should be measured using three dimensions of affection, absorption and interaction

Furthermore, these findings suggest that i) the construct domain of CE in OTC is smaller than the general customer engagement in tourism brands; ii) the nomological radius of CE construct in OTC is sufficiently covered by the dimensions of affection, absorption, and interaction, while attention and identification are redundant; they bring nothing unique to the higher-order construct of CE in OTC.

Theoretically, this study contributes to the existing literature by adapting and validating the CE scale to OTCs. The findings of this study suggest that CE in OTC could be operationalized as a three-dimensional construct. The three dimensions in their over-identification forms, thus providing an opportunity to assess model fit properly from the dimensional level. In covariance-based SEM, the proper assessment of model fitness, especially at the level of the measurement model, is vital before

one can look at how constructs relate structurally (Tabachnick & Fidell, 2012; Hair, Celsi, Money, Samouel, & Page, 2015; Kline, 2015). This is because covariance-based SEM models are founded on how the estimated model can best reproduce the covariance and variance matrices of the observed data; the only way of knowing that is by assessing both the measurement and the structural models' fitness using established fit indices (absolute, incremental and parsimony) (MacCallum, 1986; Schermelleh-Engel, Moosbrugger, & Müller, 2003; Kline, 2015). While this scale could be broadly used in different customer engagement contexts, it is most appropriate in measuring customer engagement in online travel communities on a 7-point Likert or semantic scale.

Managerially, this study presents two main implications: Managers can confidently measure CE using the three dimensions of affection, absorption, and interaction. Correctly measuring CE in OTCs is important because OTCs have become competitive marketing channels through which tourism and travel-related firms can attract, interact with, convert and retain customers. As such, management of the activities needed to engender CE is critical for continuous patronage.

Managers of tourism and travel-related businesses should understand that enhancing individual dimensions of CE is important for driving overall CE in their business' OTCs. In addition to educating members and promoting different offerings and prices, managers should consciously create opportunities to enhance CE among existing and potential customers. Members can be encouraged to interact with one another by sharing travel-related experiences, reviewing destinations/attractions and recommending destinations to visit, where possible. CE in OTCs could also be fostered when members of the OTC are considered first for offers, incentives, rebates, and discounts before they are made available to the general public via commercial media.

Limitations and future research

While the sample of this study was composed of individuals who are residents of eight countries from five continents, it included only English-speaking economies. This might limit the generalization of the findings within English speaking economies.

The debate regarding CE in OTCs is yet to mature. This study presents with confident the findings especially the adaption and validation of CE in OTCs. Due to limited space, this publication is unable to examine nomological validity of the scale; instead it is presented in a separate publication. Other authors are encouraged to validate this study in other non-English speaking destinations and markets and to explore how the construct of customer engagement in OTCs fits in the wider nomological network (Kock, Josiassen, & Assaf, 2018) by including other variables like the credibility of OTCs platforms and other forms of behavioural intentions.

Acknowledgement

This study benefitted from the grant to the second author by the Jenny and Antti Wihuri Foundation, Finland (Grant Number – 00180408)

References

Algesheimer, R., Dholakia, U. M., & Herrmann, A. (2005). The Social Influence of Brand Community: Evidence from European Car Clubs. *Journal of Marketing*, 69(3), 19–34.

Anastassova, L. (2011). Tourist loyalty and destination brand image perception: the case of Sunny beach resort, Bulgaria. *European Journal of Tourism Research*, 4(2), 191-204.

Baker, F. B., & Kim, S.-H. (2017). *The Basics of Item Response Theory Using R*. Springer. https://doi.org/10.1007/978-3-319-54205-8

Baumgartner, H., & Weijters, B. (2012). Commentary on "Common Method Bias in Marketing: Causes, Mechanisms, and Procedural Remedies." *Journal of Retailing*, 88(4), 563–566. https://doi.org/10.1016/j.jretai.2012.10.003

Bhattacharya, C. B., & Sen, S. (2003). Consumer-Company Identification: A Framework for Understanding Consumers' Relationships with Companies. *Journal of Marketing*, 67(2), 76–88.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioural Research: A Primer. *Frontiers in Public Health*, 6(June), 1–18. https://doi.org/10.3389/fpubh.2018.00149

Bowden, J. L.-H. (2009). The Process of Customer Engagement: A Conceptual Framework. *Journal of Marketing Theory and Practice*, 17(1), 63–74. https://doi.org/10.2753/MTP1069-6679170105

Brochado, A., Rita, P., & Gameiro, C. (2015). Exploring backpackers' perceptions of the hostel service quality. *International Journal of Contemporary Hospitality Management*, 27(8), 1839–1855. https://doi.org/10.1108/IJCHM-03-2014-0145

Brodie, R. J., Hollebeek, L. D., Juric, B., & Ilic, A. (2011). Customer Engagement: Conceptual Domain, Fundamental Propositions, and Implications for Research. *Journal of Service Research*, 14(3), 252–271. https://doi.org/10.1177/1094670511411703

Brodie, R. J., Ilic, A., Juric, B., & Hollebeek, L. (2013). Consumer engagement in a virtual brand community: An exploratory analysis. *Journal of Business Research*, 66(2013), 105–114. https://doi.org/10.1016/j.jbusres.2011.07.029

Byrne, Barbara M. (2012). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*. New York, USA: Routledge.

Byrne, Byrne M. (2006). *Structural Equation Modeling with EQS: Basic Concepts, Applications, and Programming* (2nd Editio). Mahwah, NJ: Routledge.

Casaló, L. V., Flavián, C., & Guinalíu, M. (2010). Antecedents and Consequences of Consumer Participation in On-Line Communities: The Case of the Travel Sector. *International Journal of Electronic Commerce*, 15(2), 137–167. https://doi.org/10.2753/JEC1086-4415150205

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, 48(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Chan, E. K. H. (2014). Standards and Guidelines for Validation Practices: Development and Evaluation of Measurement Instruments. In B. D. Zumbo & E. K. H. Chan (Eds.), *Validity and Validation in Social, Behavioural, and Health Sciences* (pp. 9–24). Springer International Publishing.

Cheung, C. M. K., Lee, M. K. O., & Jin, X.-L. (2011). Customer Engagement in an Online Social Platform: A Conceptual Model and Scale Development. In Thirty Second International Conference on Information Systems (pp. 1–8). Shanghai, China.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling*, 9(2), 233–255.

Conway, J. M., & Huffcutt, A. I. (2003). A Review and Evaluation of Exploratory Factor Analysis Practices in Organizational Research. *Organizational Research Methods*, 6(2), 147–168. https://doi.org/10.1177/1094428103251541

Craighead, C. W., Ketchen, D. J., Dunn, K. S., & Hult, G. T. M. (2011). Addressing common method variance: Guidelines for Survey Research on Information Technology, Operations, and Supply Chain Management. *IEEE Transactions on Engineering Management*, 58(3), 578–588. https://doi.org/10.1109/TEM.2011.2136437

Crocker, L., & Algina, J. (2006). *Introduction to Classical and Modern Test Theory*. Wadsworth Pub. Co.

Cronbach, L. J., & Meehl, P. E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin*, 52(4), 281–302.

Dessart, L., Veloutsou, C., & Morgan-thomas, A. (2015). Consumer engagement in online brand communities: a social media perspective. *Journal of Product & Brand Management*, 24(1), 28–42. https://doi.org/10.1108/JPBM-06-2014-0635

Dessart, L., Veloutsou, C., & Morgan-Thomas, A. (2016). Capturing consumer engagement: duality, dimensionality and measurement. *Journal of Marketing Management*, 32(5–6), 399–426. https://doi.org/10.1080/0267257X.2015.1130738

DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th Editio). Thousand Oaks, USA: SAGE Publications, Inc.

Dima, A. L. (2018). Scale Validation in Applied Health Research: Tutorial for a 6-step R-based Psychometrics Protocol. *Health Psychology and Behavioural Medicine*, 6(1), 136–161. https://doi.org/10.1080/21642850.2018.1472602

Dwivedi, A. (2015). A higher-order model of consumer brand engagement and its impact on loyalty intentions. *Journal of Retailing and Consumer Services*, 24(2015), 100–109. https://doi.org/10.1016/j.jretconser.2015.02.007

Dziuban, C. D., & Shirkey, E. C. (1974). When is a Correlation Matrix Appropriate for Factor Analysis? *Psychological Bulletin*, 81(6), 358–361.

Eichhorn, B. R. (2014). *Common Method Variance Techniques* (No. AA11-2014). Cleveland.

Enginkaya, E., & Esen, E. (2014). Dimensions of Online Customer Engagement. *Journal of Business, Economics and Finance*, 3(1), 106–114.

Farrokhyar, F., Reddy, D., Poolman, R. W., & Bhandari, M. (2013). Why Perform a priori sample size calculation? *Canadian Journal of Surgical*, 56(3), 207–213. https://doi.org/10.1503/cjs.018012

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2019). G*Power 3.1.9.4 Manual. Germany.

Faul, Franz, Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioural, and biomedical sciences. *Behaviour Research Methods*, 39(2), 175–191.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. SAGE Publications, Inc.

Finch, W. H., Immekus, J. C., & French, B. F. (2016). *Applied Psychometrics Using SPSS and AMOS*. Charlotte, NC - USA: Information Age Publishing Inc.

Floyd, F. J., & Widaman, K. F. (1995). Factor Analysis in the Development and Refinement of Clinical Assessment Instruments. *Psychological Assessment*, 7(3), 286–299. https://doi.org/10.1037/1040-3590.7.3.286

Fornell, C., & Larcker, D. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39–50. https://doi.org/10.2307/3151312

Gallaugher, J., & Ransbotham, S. (2010). Social Media and Customer Dialog Management at Starbucks. *MIS Qarterly Executive*, 9(4), 197–212.

Gulliksen, H. (1950). *Theory of Mental Tests*. New York: Wiley.

Hair, J. F., Celsi, M., Money, A., Samouel, P., & Page, M. (2015). *The Essentials of Business Research Methods* (3rd Editio). New York: Routledge.

Hanson, B. A. (1998). Uniform DIF and DIF Defined by Differences in Item Response Functions. *Journal of Educational and Behavioural Statistics*, 23(3), 244–253. Retrieved from https://www.jstor.org/stable/1165247

Harman, H. H. (1976). *Modern Factor Analysis* (3rd Revise). Chicago, Illinois, USA: University of Chicago Press.

Harrigan, P., Evers, U., Miles, M., & Daly, T. (2017). Customer engagement with tourism social media brands. *Tourism Management*, 59(2017), 597–609. https://doi.org/10.1016/j.tourman.2016.09.015

Hollebeek, L. (2011). Exploring customer brand engagement: definition and themes. *Journal of Strategic Marketing*, 19(7), 555–573. https://doi.org/10.1080/0965254X.2011.599493

Hollebeek, L. D. (2011). Demystifying customer brand engagement: Exploring the loyalty nexus. J*ournal of Marketing Management*, 27(7–8), 785–807. https://doi.org/10.1080/0267257X.2010.500132

Hollebeek, L. D. (2013). The customer engagement/value interface: An exploratory investigation. *Australasian Marketing Journal,* 21(2013), 17–24. https://doi.org/10.1016/j.ausmj.2012.08.006

Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer Brand Engagement in Social Media: Conceptualization, Scale Development and Validation. *Journal of Interactive Marketing*, 28(2014), 149–165. https://doi.org/10.1016/j.intmar.2013.12.002

Hur, K., Terry, T. K., Karatepe, O. M., & Lee, G. (2017). An exploration of the factors influencing social media continuance usage and information sharing intentions among Korean travellers. *Tourism Management*, 63(2017), 170–178. https://doi.org/10.1016/j.tourman.2017.06.013

Janssens, W., Wijnen, K., Pelsmacker, P. De, & Kenhove, P. Van. (2008). *Marketing Research with SPSS* (1st Edition). Pearson Education.

Kaiser, H. F. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34, 111–117.

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling (Methodology in the Social Sciences)* (Fourth Edi). New York: The Guilford Press.

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (Fourth). New York, USA: The Guilford Press.

Kock, F., Josiassen, A., & Assaf, A. G. (2018). Scale Development in Tourism Research: Advocating for a New Paradigm. *Journal of Travel Research*, 58(7), 1227–1229. https://doi.org/10.1177/0047287518800391

Ku, E. C. S. (2014). Distributed Fascinating Knowledge Over an Online Travel Community. *International Journal of Tourism Research*, 16, 33–43. https://doi.org/10.1002/jtr

Kumar, V., Aksoy, L., Donkers, B., Venkatesan, R., Wiesel, T., & Tillmanns, S. (2010). Undervalued or Overvalued Customers: Capturing Total Customer Engagement Value. *Journal of Service Research*, 13(3), 297–310. https://doi.org/10.1177/1094670510375602

Lee, S. T. H. (2018). Testing for Measurement Invariance: Does Your Measure Mean the Same Thing for Different Participants? Retrieved from https://www.psychologicalscience.org/observer/testing-for-measurement-invariance

Linacre, J. M. (1994). Sample Size and Item Calibration or Person Measure Stability. *Rasch Measurement Transactions*, 7(4), 328. Retrieved from https://www.rasch.org/rmt/rmt74m.htm

Linden, W.J. (Ed.). (2016). *Handbook of Item Response Theory: Applications (Volume 3:)*. Boca Raton, FL- USA: A Chapman & Hall/CRC.

Linden, Wim J. (Ed.). (2016a). *Handbook of Item Response Theory: Models (Volume 1:)*. Boca Raton, FL-USA: Chapman & Hall/CRC.

Linden, Wim J. (Ed.). (2016b). *Handbook of Item Response Theory: Statistical Tools (Volume 2:)*. Boca Raton, FL- USA: A Chapman & Hall/CRC.

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On Selecting Indicators for Multivariate Measurement and Modeling with Latent Variables: When "Good" Indicators Are Bad and "Bad" Indicators Are Good. *Psychological Methods*, 4(2), 192–211.

MacCallum, R. C. (1986). Specification Searches in Covariance Structure Modeling. P*sychological Bulletin*, 100(1), 107–120. https://doi.org/10.1037/0033-2909.100.1.107

Meredith, W. (1964). Notes on Factorial Invariance. *Psychometrika*, 29(2), 177–185.

Meredith, W. (1993). Measurement Invariance, Factor Analysis and Factorial Invariance. *Psychometrika*, 58(4), 525–543.

Merrilees, B. (2016). Interactive brand experience pathways to customer-brand engagement and value co-creation. *Journal of Product & Brand Management*, 25(5), 402–408. https://doi.org/10.1108/JPBM-04-2016-1151

Moreau, C. P. (2011). Inviting the Amateurs into the Studio: Understanding How Consumer Engagement in Product Design Creates Value. *Journal of Product Innovation Management*, 28, 409–410.

Moreno, H. (2017). Three Rules of Customer Engagement. Retrieved October 5, 2018, from https://www.forbes.com/sites/forbesinsights/2017/11/03/3-rules-of-customer-engagement/#392e48e85609

Moshagen, M. (2018). Package 'semPower' 1.0.0. Germany. https://doi.org/10.1080/10705511.2014.950896>.License

Moshagen, M., & Erdfelder, E. (2016). A New Strategy for Testing Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal,* 23(1), 54–60. https://doi.org/10.1080/10705511.2014.950896

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In M. Csikszentmihalyi (Ed.), *Flow and the Foundations of Positive Psychology* (pp. 239–263). Springer, Nertherlands. https://doi.org/10.1007/978-94-017-9088-8

Nering, M. L., & Ostini, R. (Eds.). (2010). *Handbook of Polytomous Item Response Theory Models.* Routledge.

Novick, M. R. (1966). The Axioms and Principal Results of Classical Test Theory. *Journal of Mathematical Psychology*, 3(1), 1–18.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill, Inc.

Podsakoff, P. M., Mackenzie, S. B., & Podsakoff, N. P. (2012). Sources of Method Bias in Social Science Research and Recommendations on How to Control It. *Annual Review of Psychology*, 63, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Podsakoff, P. M., & Organ, D. W. (1986). Self-Reports in Organizational Research: Problems and Prospects. *Journal of Management*, 12(4), 531–544. https://doi.org/10.1177/014920638601200408

Putnick, Diane, L., & Bornstein, Mark, H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review*, 41, 71–90. https://doi.org/10.1016/j.dr.2016.06.004.Measurement

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to Psychometric Theory*. Routledge, New York, NY 10017, USA.

Rizopoulos, D. (2006). ltm: An R Package for Latent Variable Modelling and Item Response Theory Analyses. *Journal of Statistical Software*, 17(5), 1–25. https://doi.org/http://www.jstatsoft.org/v17/i05/

Sarkar, A., & Sreejesh, S. (2014). Examination of the roles played by brand love and jealousy in shaping customer engagement. *Journal of Product & Brand Management*, 23(1), 24–32. https://doi.org/10.1108/JPBM-05-2013-0315

Sashi, C. M. (2012). Customer engagement, buyer-seller relationships, and social media. *Management Decision*, 50(2), 253–272. https://doi.org/10.1108/00251741211203551

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online*, 8(2), 23–74. https://doi.org/10.1002/0470010940

Slocum, S. L., Gelin, M. N., & Zumbo, B. D. (2004). *Statistical and Graphical Modeling to Investigate Differential Item Functioning for Rating Scale and Likert Item Formats*. Vancouer: Edgeworth Laboratory, University of British Columbia.

So, K. K. F., King, C., & Sparks, B. (2014). Customer Engagement with Tourism Brands: Scale Development and Validation. *Journal of Hospitality and Tourism Research*, 38(3), 304–329. https://doi.org/10.1177/1096348012451456

Stokburger-Sauer, N., Ratneshwar, S., & Sen, S. (2012). Drivers of consumer – brand identification. *International Journal of Research in Marketing*, 29(2012), 406–418. https://doi.org/10.1016/j.ijresmar.2012.06.001

Tabachnick, B. G., & Fidell, L. S. (2012). Using Multivariate Statistics (6th Editio). United States of America: Pearson Education Inc.

Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, 33, 1–39.

Thakur, R. (2016). Understanding Customer Engagement and Loyalty: A Case of Mobile Devices for Shopping. *Journal of Retailing and Consumer Services*, 2016, 151–163. https://doi.org/10.1016/j.jretconser.2016.06.004

Thissen, D., Pommerich, M., Billeaud, K., & Williams, V. S. L. (1995). Item Response Theory for Scores on Tests Including Polytomous Items with Ordered Responses. *Applied Psychological Measurement*, 19, 39–49.

Timmons, A. C. (2010). Establishing Factorial Invariance for Multiple-group Confirmatory Factor Analysis. Timmons, A. C. Manhatan, Kansas. Retrieved from http://crmda.dept.ku.edu/resources/kuantguides/22.Factorial_Invariance_Guide.pdf.

Ukpabi, D., & Karjaluoto, H. (2017). Influence of Social Media on Corporate Heritage Tourism Brand. In R. Schegg & B. Stangl (Eds.), *Information and Communication Technologies in Tourism* 2017 (pp. 697–709). Cham: Springer. https://doi.org/https://doi.org/10.1007/978-3-319-51168-9_50

van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P. C. (2010). Customer Engagement Behaviour: Theoretical Foundations and Research Directions. *Journal of Service Research*, 13(3), 253–266. https://doi.org/10.1177/1094670510375599

van Tonder, E., & Petzer, D. J. (2018). The Interrelationships Between Relationship Marketing Constructs and Customer Engagement Dimensions. *Service Industries Journal*, 38(13–14), 948–973. https://doi.org/10.1080/02642069.2018.1425398

Verhoef, P. C., Reinartz, W. J., & Krafft, M. (2010). Customer Engagement as a New Perspective in Customer Management. *Journal of Service Research*, 13(3), 247–252. https://doi.org/10.1177/1094670510375461

Vivek, S. D., Beatty, S. E., Dalela, V., & Morgan, R. M. (2014). A Generalized Multidimensional Scale for Measuring Customer Engagement. *Journal of Marketing Theory and Practice*, 22(4), 401–420. https://doi.org/10.2753/MTP1069-6679220404

Vivek, S. D., Beatty, S. E., & Morgan, R. M. (2012). Customer Engagement: Exploring Customer Relationships Beyond Purchase. *Journal of Marketing Theory and Practice*, 20(2), 122–146. https://doi.org/10.2753/MTP1069-6679200201

Woodbury, M. A. (1963). The Stochastic Model of Mental Testing and an Application. *Psychometrika*, 28, 391–394.

Woods, C. M., & Grimm, K. J. (2011). Testing for Nonuniform Differential Item Functioning with Multiple Indicator Multiple Cause Models. *Applied Psychological Measurement*, 35(5), 339–361. https://doi.org/10.1177/0146621611405984

Xiang, Z., & Gretzel, U. (2010). Role of Social Media in online travel information search. *Tourism Management*, 31(2010), 179–188. https://doi.org/10.1016/j.tourman.2009.02.016

Yong, A. G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079

Received: 22/07/2019
Accepted: 20/10/2019
Coordinating editor: Stanislav Ivanov

Appendix 1: Test for differential item functioning (DIF)

Item	Grouping variable	Significance (*p*-value)	
		Uniform DIF	Non uniform DIF
EN1	Gender	0.095	0.236
	Country of residence	0.645	0.229
ID1	Gender	0.389	0.290
	Country of residence	0.460	0.061
ID4	Gender	0.002*	0.031
	Country of residence	0.061	0.501
ID5	Gender	0.004*	0.001*
	Country of residence	0.037	0.781
EN4	Gender	0.552	0.598
	Country of residence	0.604	0.981
AT5	Gender	0.074	0.981
	Country of residence	0.005*	0.103
AT6	Gender	0.003*	0.080
	Country of residence	0.019	0.081
ID3	Gender	0.049	0.103
	Country of residence	0.004*	0.027
EN2	Gender	0.895	0.941
	Country of residence	0.921	0.016
EN3	Gender	0.002*	0.045
	Country of residence	0.613	0.301
AT4	Gender	0.006*	0.001*
	Country of residence	0.041	0.482
IT4	Gender	0.763	0.282
	Country of residence	0.615	0.766
IT3	Gender	0.106	0.944
	Country of residence	0.797	0.101
IT2	Gender	0.049	0.823
	Country of residence	0.391	0.823
IT1	Gender	0.243	0.985
	Country of residence	0.568	0.329
IT5	Gender	0.068	0.836
	Country of residence	0.811	0.328
AB4	Gender	0.091	0.799
	Country of residence	0.430	0.846
AB5	Gender	0.795	0.829
	Country of residence	0.462	0.473
AB3	Gender	0.072	0.592
	Country of residence	0.039	0.482
AB2	Gender	0.051	0.377
	Country of residence	0.536	0.921
AB1	Gender	0.006*	0.218
	Country of residence	0.341	0.801

AB6	Gender	0.291	0.001*
	Country of residence	0.393	0.825

*Significant DIF ($p < .01$)

Appendix 2: Examining Item Response Theory (IRT) - Item fit indices

Absorption dimension item fit indices

Item	Chi Square	df	RMSEA	*p* - value
AB2	22.209	26	0.000	0.677
AB3	14.079	20	0.000	0.826
AB4	33.024	21	0.036	0.056
AB5	23.676	19	0.023	0.209

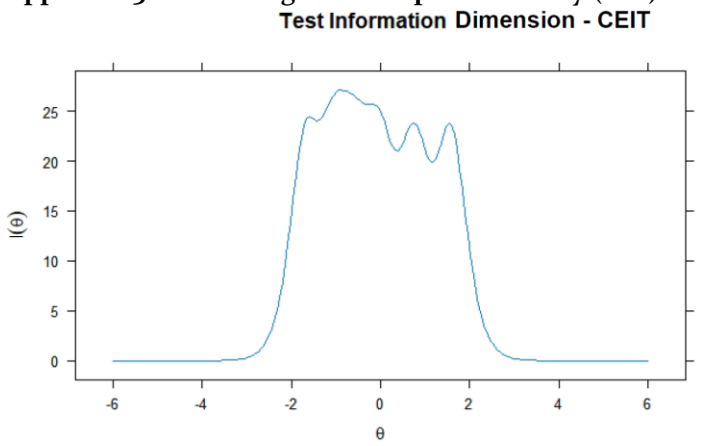
Affection dimension item fit indices

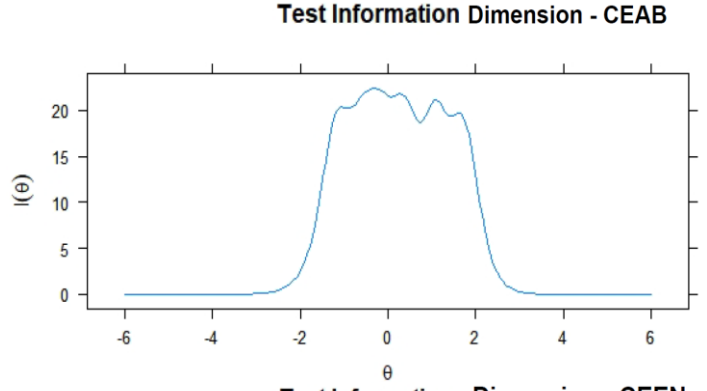
Item	Chi Square	df	RMSEA	*p* - value
ID1	52.487	23	0.029	0.059
EN1	31.824	23	0.029	0.104
EN2	35.069	23	0.034	0.051
EN4	32.770	23	0.022	0.205

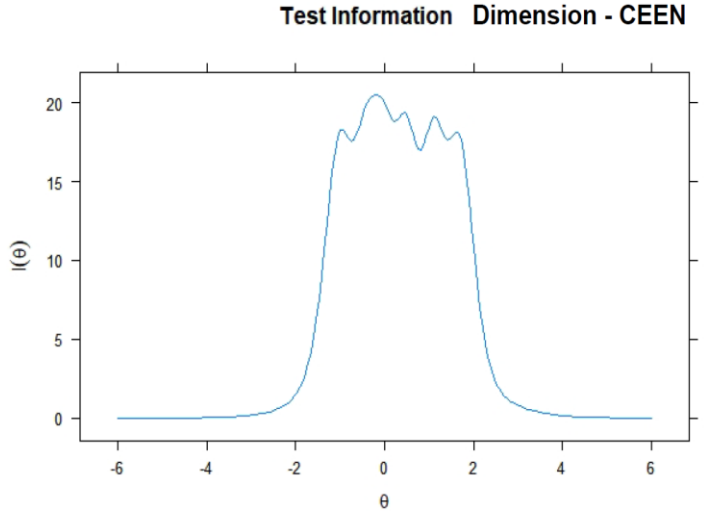
Interaction dimension item fit indices

Item	Chi Square	df	RMSEA	*p*-value
IT1	35.127	25	0.030	0.086
IT2	29.615	25	0.020	0.239
IT3	45.195	25	0.042	0.079
IT4	43.441	25	0.033	0.061
IT5	37.765	25	0.020	0.223

Appendix 3: Examining Item Response Theory (IRT) - Item test information







Appendix 4: Examining Item Response Theory (IRT) - Item *a* & *b* coefficients
Affection dimension-Item coefficients

Item	Item discrimination (*a*)	*b1* (2PL 2-1)	*b2* (2PL 3-2)	*b3* (2PL 4-3)	*b4* (2PL 5-4)	*b5* (2PL 6-5)	*b6* (2PL 7-6)
ID1	2.048	-0.834	-0.477	0.014	1.12	2.262	3.583
EN1	3.619	-0.482	-0.049	0.225	0.875	1.724	2.303
EN2	3.315	-0.522	-0.019	0.54	1.396	2.217	2.761
EN4	3.493	-0.301	0.081	0.492	1.111	2.169	2.650

Affection

Absorption dimension-Item coefficients

Item	*a*	*b1*	*b2*	*b3*	*b4*	*b5*	*b6*
AB2	3.258	-1.427	-0.826	-0.391	0.206	0.935	1.567
AB3	4.753	-1.167	-0.618	-0.163	0.393	1.105	1.611
AB4	4.532	-1.022	-0.360	0.122	0.520	1.117	1.851
AB5	4.860	-1.240	-0.601	-0.198	0.336	1.041	1.716

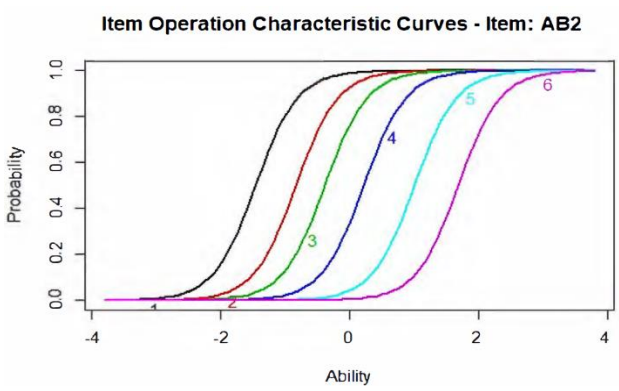
CEAB =Absorption dimension

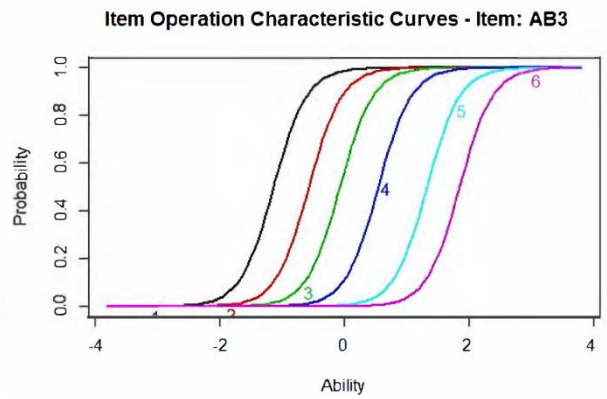
Interaction dimension-Item coefficients

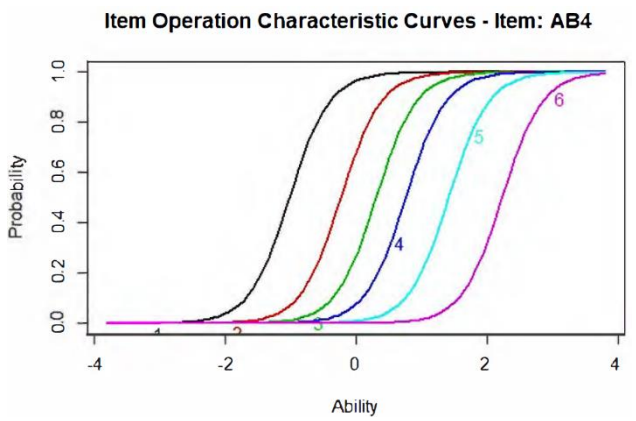
Item	*a*	*b1*	*b2*	*b3*	*b4*	*b5*	*b6*
IT1	3.685	-2.065	-1.247	-0.611	0.167	1.030	2.120
IT2	3.100	-1.767	-1.104	-0.328	0.342	1.306	2.390
IT3	3.589	-1.757	-1.029	-0.436	0.071	1.053	2.253
IT4	3.319	-1.887	-1.234	-0.678	0.047	0.938	1.983
IT5	2.997	-1.384	-0.990	-0.288	0.352	1.276	2.401

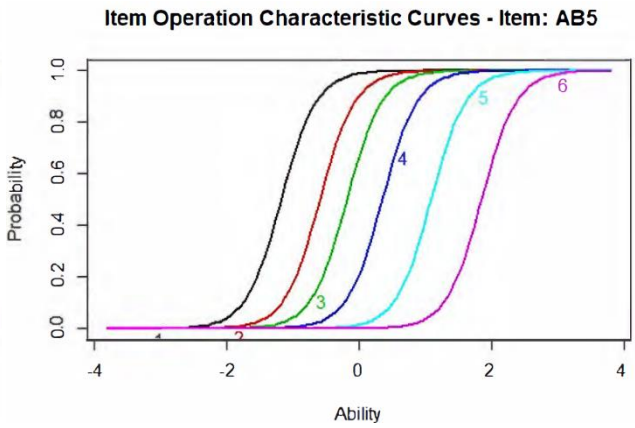
CEIT = Interaction dimension

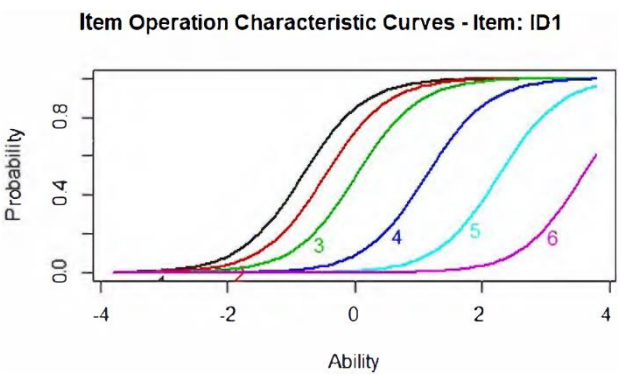
Appendix 5: Examining Item Response Theory (IRT) -Item operation characteristic curves

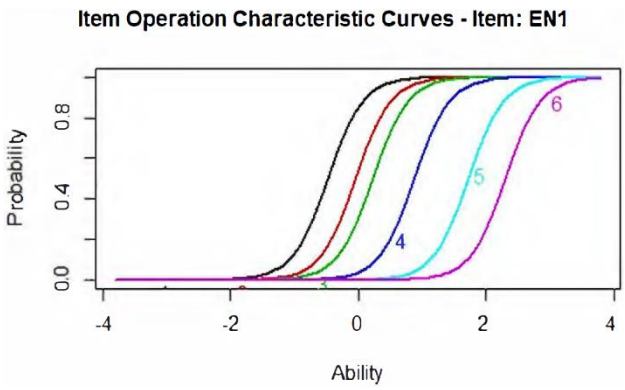




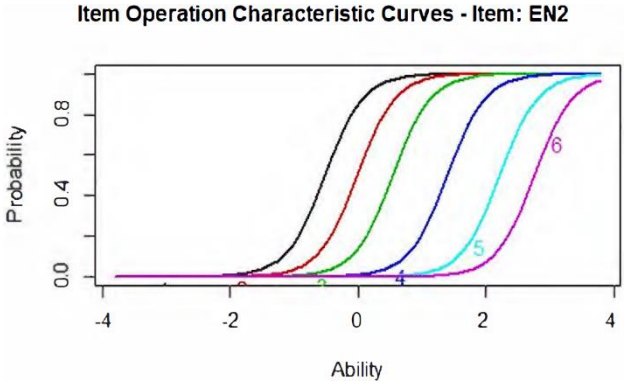


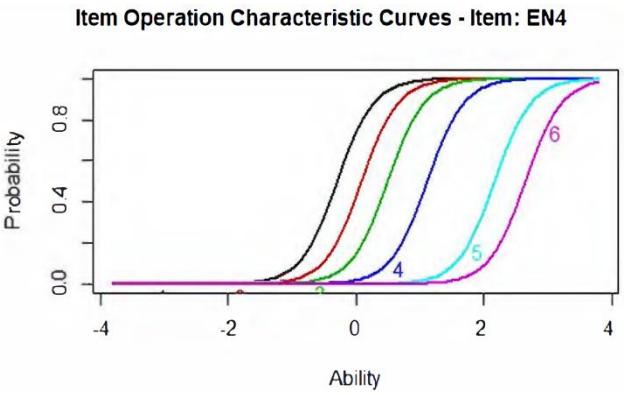


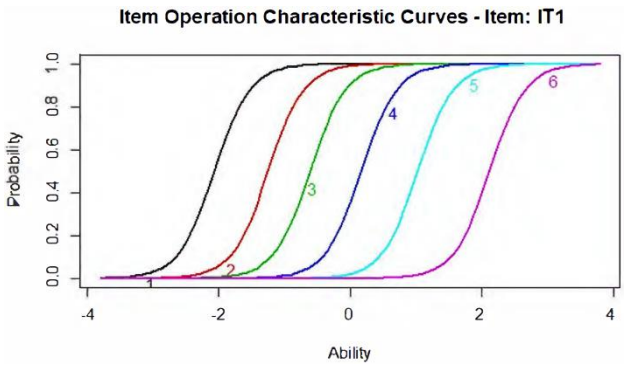


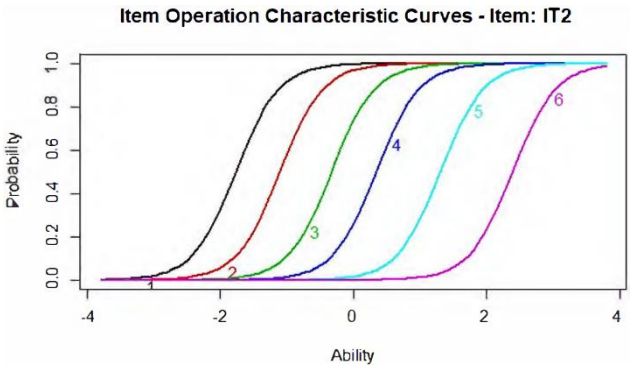


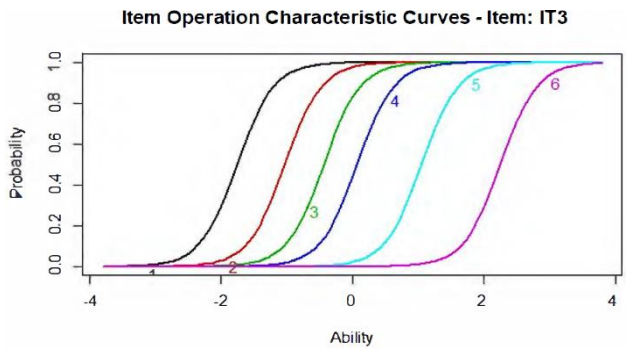
Appendix 5: Examining IRT - Item operation characteristic curves continue

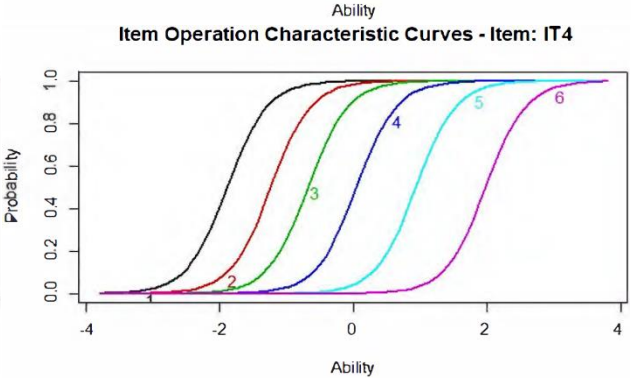


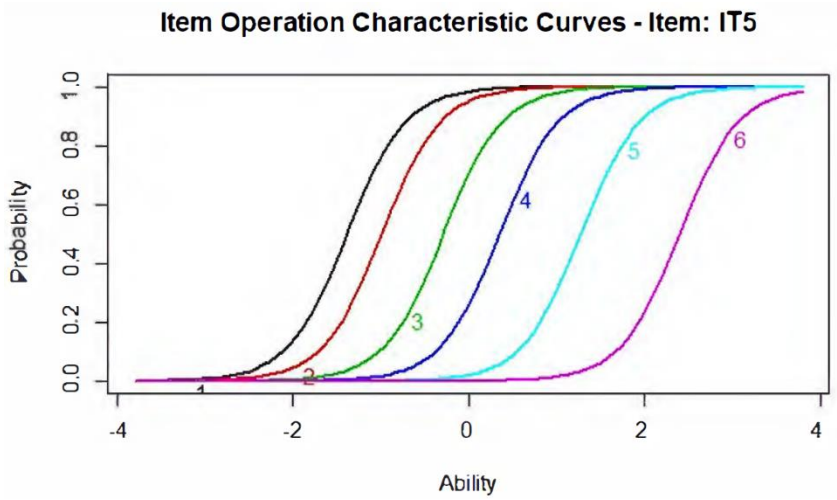




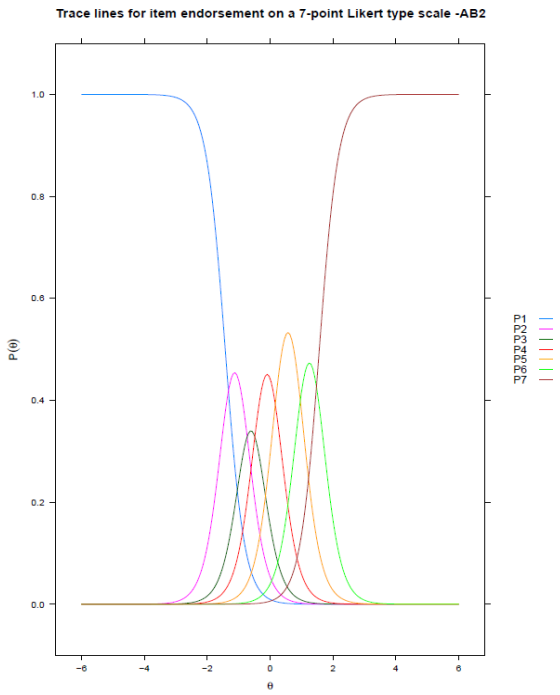


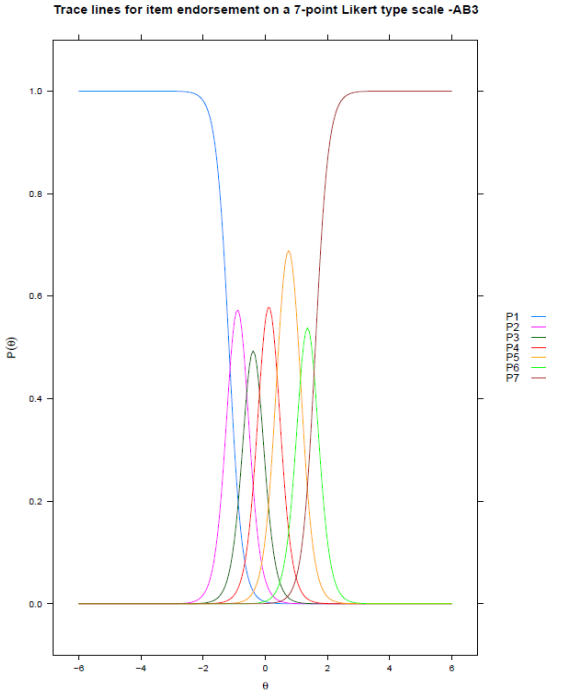


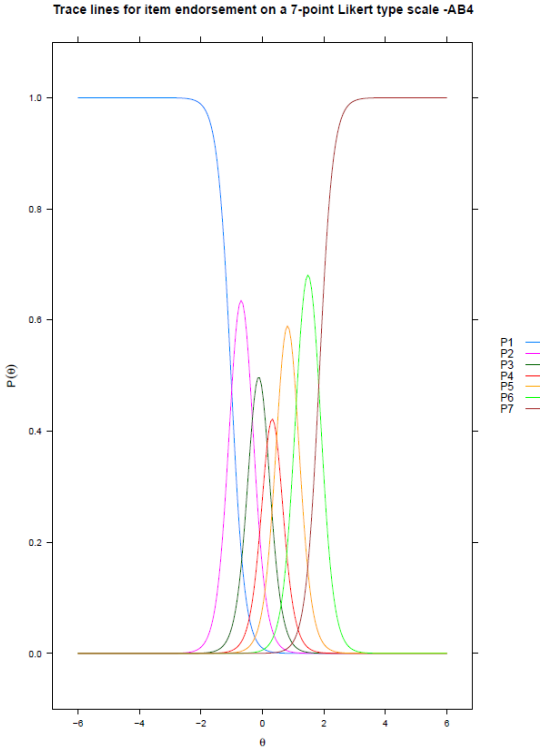


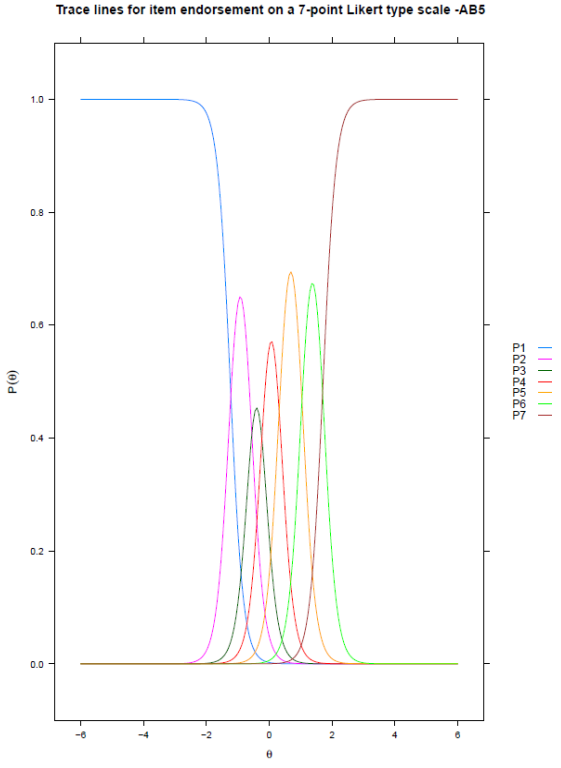


Appendix 6: Examining IRT - Item endorsement on a 7-point Likert type scale

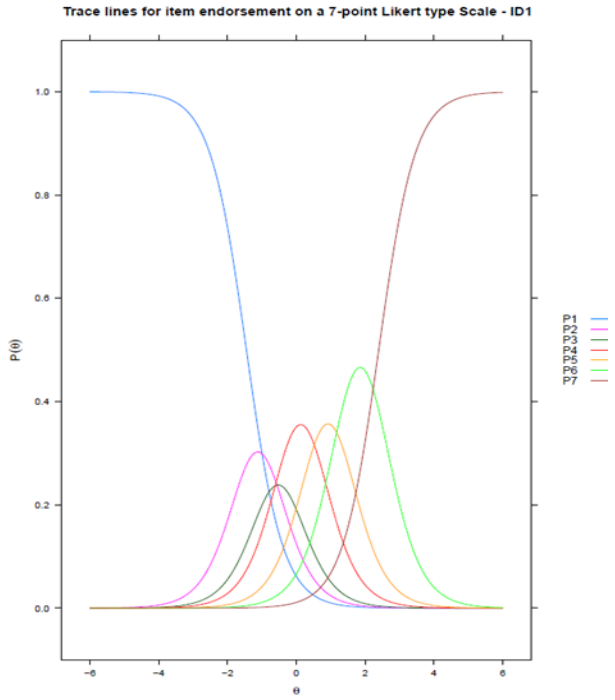


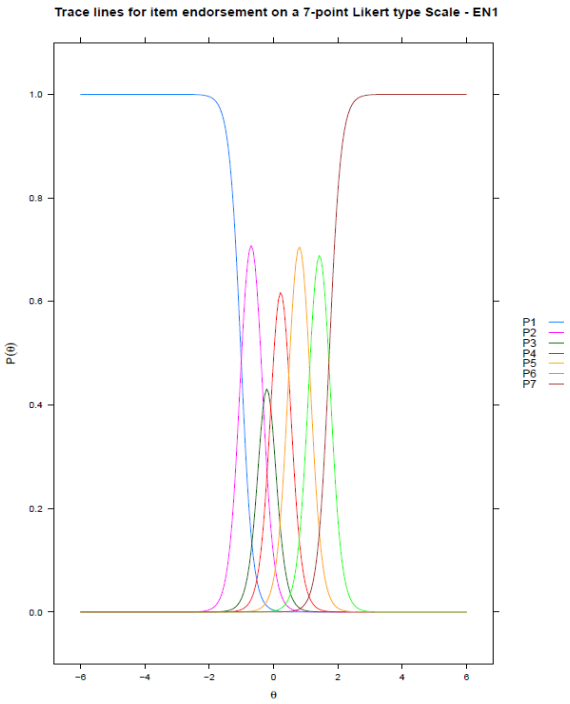


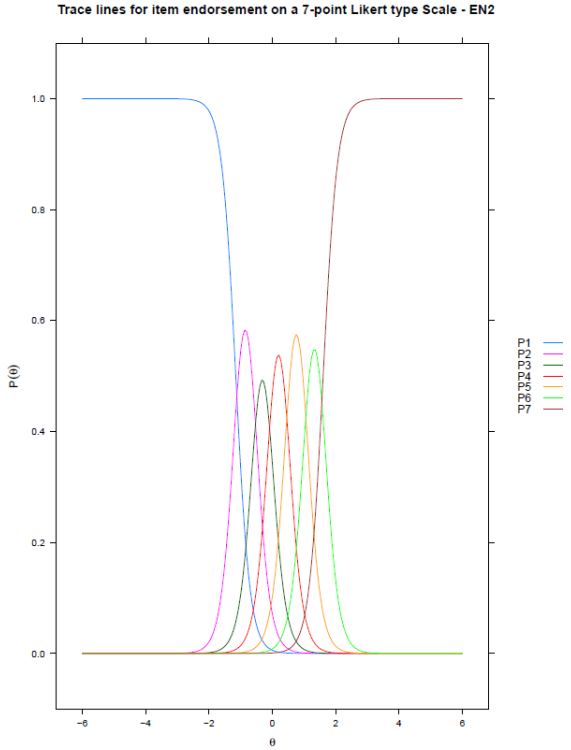


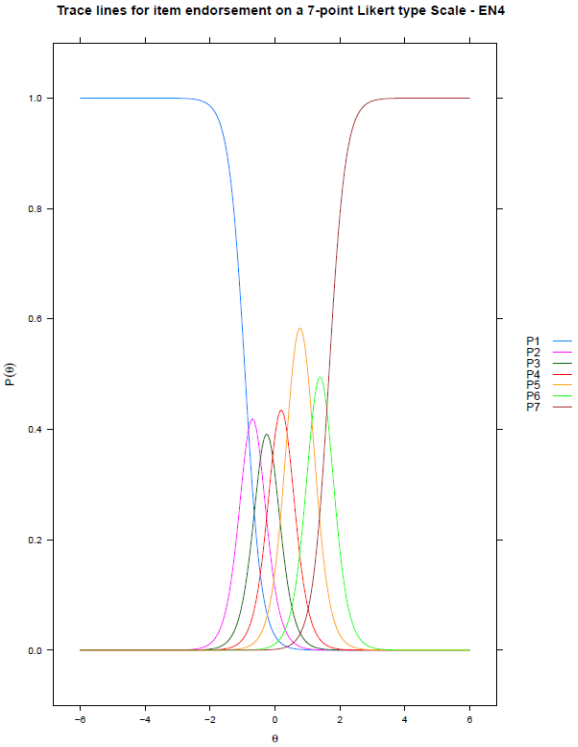


Appendix 6: Examining IRT - Item endorsement on a 7-point Likert type scale continue

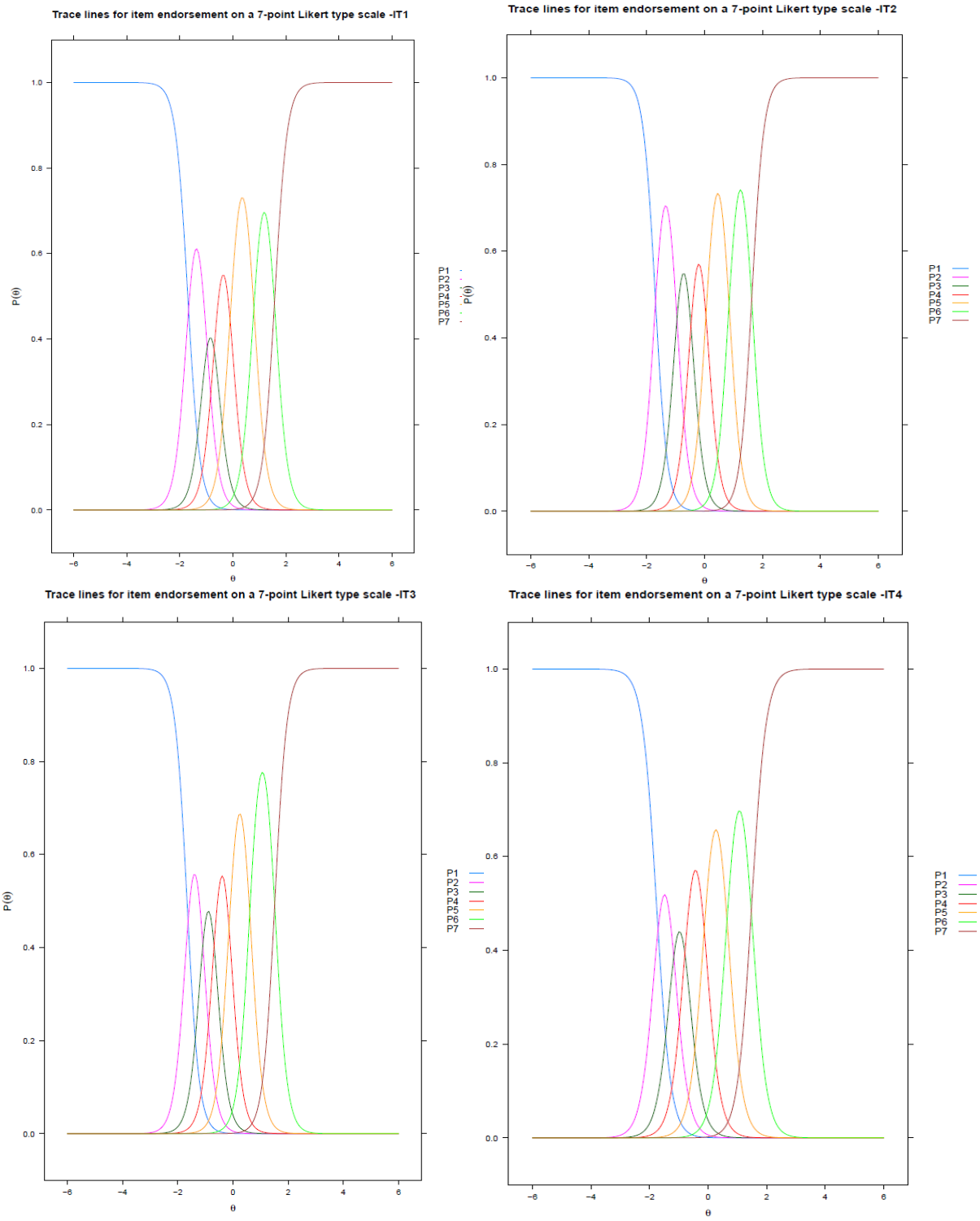








Appendix 6: Examining IRT - Item endorsement on a 7-point Likert type scale continue



Appendix 6: Examining IRT - Item endorsement on a 7-point Likert type scale continue