IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 29, No. 3, June, 2024, pp. 1051-1064
Journal homepage: http://iieta.org/journals/isi

Analyzing Customer Engagement with Gamification Approach in the Banking Sector Using the Rasch Model

Putri Taqwa Prasetyaningrum[1,2*], Purwanto Purwanto[1,3], Adian Fatchur Rochim[4]

[1] Doctoral Program of Information Systems, Universitas Diponegoro, Semarang 50241, Indonesia
[2] Faculty of Information Technology, Universitas Mercu Buana Yogyakarta, Yogyakarta 55283, Indonesia
[3] Department of Chemical Engineering, Faculty of Engineering, Universitas Diponegoro, Semarang 50275, Indonesia
[4] Faculty of Engineering, Universitas Diponegoro, Semarang 50275, Indonesia

Corresponding Author Email: putritaqwa@students.undip.ac.id

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/isi.290323

Received: 9 October 2023
Revised: 11 March 2024
Accepted: 25 March 2024
Available online: 20 June 2024

Keywords:
customer engagement, gamification, mobile banking applications, Rasch models, analysis instruments

ABSTRACT

This study investigates the instruments used to measure consumer engagement in mobile banking applications that employ gamification strategies. Customer engagement plays a crucial role in the effectiveness of e-marketing strategies, particularly for relationships, products, services, and brands. Gamification has emerged as a promising approach for promoting content marketing. Primary data from active users of mobile banking applications were collected from various banks and analyzed using the Rasch model to assess the instrument's efficiency. The study involved 451 participants and considered demographics, customer involvement, psychological aspects, game elements, and 26 dimensions. The results indicate that the average Outfit Mean Square (MNSQ) scores for individuals and items obtained from analyzing 49 questions with the Rasch model fall within the desirable range of 0.5-1.5. The mean values of INFIT and OUTFIT MNSQ, as well as INFIT and OUTFIT ZSTD, closely approximate the desired level. Additionally, the instrument demonstrates strong reliability with a coefficient alpha of 0.94. In conclusion, the findings suggest that the respondents understood the instruments employed in this study well and effectively facilitated data collection.

1. INTRODUCTION

The rapid advancement and widespread adoption of digital technology, particularly through mobile applications [1], have revolutionized numerous sectors, including banking field [2]. Mobile banking applications epitomize this digital transformation, offering unprecedented convenience by enabling transactions without physical cash or paperwork, thereby reshaping the business landscape and influencing social dynamics [3]. In recent times, mobile banking usage has experienced a remarkable and impressive surge, as evidenced by the substantial escalation in global mobile application downloads. The meteoric rise in mobile application downloads, from 140.7 billion in 2016 to 204 billion in 2019, underscores the global transition towards digital banking solutions. However, this shift also unveils a significant challenge: sustaining user engagement. With data revealing that only 32% of mobile application users engage with an app more than ten times, and 25% abandon it after a single use, the banking sector faces a pressing need to reimagine strategies for retaining customer interest and participation [4]. However, this shift also unveils a significant challenge: sustaining user engagement. With data revealing that only 32% of mobile application users engage with an app more than ten times, and 25% abandon it after a single use, the banking sector faces a pressing need to reimagine strategies for retaining customer interest and participation [5].

These statistics clearly indicate the limited level of user engagement with mobile apps. Consequently, businesses operating these apps encounter a significant challenge in sustaining user engagement [6]. Gamification, the integration of game elements into non-gaming contexts [7], emerges as a potent strategy to enhance user engagement [8] by leveraging psychological drivers to foster deeper interaction with mobile banking applications [9]. Despite the promising potential of gamification, empirical research exploring its impact on user engagement within mobile banking remains sparse [6].

A deeper understanding of app engagement is provided by recent study by Ho and Chung [10]. In addition, previous literature was limited because many studies discussed gamification only as a context [11] and did not establish a connection to any theories already in use to explain how gamification elements drive motivational processes [12]. Recognizing the distinctive qualities of every individual, gamification systems can help provide personalized solutions for individual users and can increase their engagement and satisfaction [13-15].

This gap is particularly pronounced in understanding how gamification can transform user behaviour and elevate the banking experience. Our study addresses this lacuna by focusing on the effectiveness of gamification in boosting customer engagement in the banking sector, an area of critical

importance given the competitive and fast-evolving nature of digital finance.

Studies examining the use of gamification in banking have found that it significantly improves system performance and consumer engagement [16, 17]. Furthermore, the mean values of INFIT and OUTFIT MNSQ, along with INFIT and OUTFIT ZSTD, exhibit a close approximation to the desired threshold. Moreover, the instrument exhibits a strong level of reliability, as evidenced by a coefficient alpha of 0.94 [18]. Rodrigues' (2016) research on gamification in e-banking, measures the absolute fit index by determining how well the model is adapted to the sample data using SEM (Structural Equation Modelling) modeling. As a result of the findings, e-banking use is influenced by both simplicity of use and enjoyment [19]. Conducting assessments of convergent validity plays a pivotal role in the measurement model by evaluating the extent to which the indicators align with a specific variable. The evaluation of discriminant validity entails the utilization of the square root of the Average Variance Extracted (AVE) [17].

The findings indicate reliable discriminant validity, supported by correlations among each dimension that are below the square root of the Average Variance Extracted (AVE) [20]. Another study was conducted by Aydınlıyurt et al. [21], using Cronbach's alpha and AVE to measure individual behavioral influence instruments in gamification mobile applications. Hamidi and Safareeyeh's [22] study utilized SPSS and Cronbach's alpha to assess how employing CRM systems in m-banking adoption affected customer happiness and engagement, which is said to be the most crucial element in the banking industry's performance. Dzandu et al., (2022) According to his research, gamification of mobile money payments can increase consumer value by having a positive social impact. Measurement tools were used in Dzandu's study together with the SPSS method and SEM-PLS (Structural Equation Modeling-Partial Least Squares). The research demonstrates a strong and favorable connection between customers, social impact theory, and gamified mobile money payments (GMMP) [23]. In his study, Nasirzadeh and Fathian [18] used SPSS statistics to assess data on how gamification may be customized for people with various demographic and psychological features, particularly in the banking industry.

Among the various measurement options available, the Rasch measurement is a valuable approach. Its primary advantage lies in its ability to assess whether a scale measures independently of a particular sample. Moreover, it provides valuable insights at both the scale and item levels, enabling the adaptation of existing scales to specific research contexts [24]. The Rasch model, a mathematical framework, utilizes probability estimation to evaluate the measurement characteristics of a rating scale [25]. The Rasch model, being a latent logistic model, prioritizes precise measurement. It posits that an individual's likelihood of providing a correct response to a test item is influenced by their ability claims and the item's level of difficulty, both derived from empirically validated data [26]. Rasch analysis aims to enhance the precision of evaluations for both individuals and items, thereby making valuable contributions to various aspects of validity and accuracy [27]. Rasch's model characterizes respondent proficiency (people ability) and difficulty as latent variables [28-30]. It presupposes that the item's slope or discrimination parameter would be constant across all items (i.e., how well the item can identify variations in one's talents) [25]. One advantage of the Rasch model, as opposed to the Item Response Theory (IRT) model, is its capability to facilitate comparisons between individuals' abilities and the difficulty of items, irrespective of the specific sample or set of items used [31]. Furthermore, the Rasch measurement model provides more comprehensive insights into products, models, and the fit with human attributes. By enabling prompt identification of certain measurement issues, the Rasch measurement model acts as a valuable complement to traditional test theory and approaches based on Item Response Theory (IRT) [32]. Rasch analysis can be utilized for evaluating the psychometric properties of existing instruments and developing new ones [33]. Rasch analysis, which is mostly used for instrument creation, shows benefits in terms of item reduction [30]. Research conducted by Brush and Soutar [34] investigated property scale measurements and invariance using Rasch analysis in tourism and recreation studies.

Researchers recommend investigating using Rasch whether the measurement invariant assumption holds, before conducting substantive testing [34]. According to Stolt et al. [35], the application of Rasch analysis in the nursing profession serves as a tool that lends support to the validation of tools created in nursing science. It is definitely advisable to read Rasch's study of nursing science. In the study by Goswami et al. [36], Rasch analysis was used to evaluate the dimensions, model data fit, item difficulty, individual stigma, distribution of items and individuals across items-people, and assessment scale functions for each HPASS subscale (Stereotypes, Discrimination, and Prejudice). According to Hergesell [24], Rasch's measurement is to overcome several weaknesses inherent in the Classical Test Theory (CTT) in research in the field of tourism. Opportunities to improve data compliance with Rasch models and fine-tune scale are provided. In research conducted by Soeharto and Csapó [37], the results of Rasch's analysis showed that students adapted inductive reasoning tests met the criteria of validity and reliability based on Rasch's parameters. Rasch analysis is the best testing method since it provides so much data regarding individual items, their dependability, and the suitability of response formats [38].

This paper examines several variables encompassing demographics, customer engagement, psychological factors, and game elements. Prior studies have identified a range of game elements employed in the banking sector, comprising a total of 21 essential elements. Elements include announcements, points, awards, ratings, badges, scores, tasks, feedback, leaderboards, bid hunt, timers, levels, shares, social interactions, penalties, avatars, lotteries, virtual prizes, epic meanings, informative content, and random prizes. These game elements have been recognized for their potential to increase customer engagement and promote desired behavior in a banking context [39]. The 26 dimensions consist of demographics, customer engagement, self-efficacy, accountability, belongingness, announcements, points, awards, ratings, badges, scores, tasks, feedback, leaderboards, bid hunt, timer, levels, share, social interaction, penalties, avatars, lotteries, virtual prizes, epic meanings, informative content and random prizes.

Furthermore, this research pioneers the application of the Rasch model to evaluate the measurement tools used to assess customer engagement in gamified banking environments. The Rasch model, with its robust mathematical foundation, offers a unique lens through which to examine the alignment between user engagement levels and gamification elements, ensuring

precise, scalable, and interpretable measurements. By employing the Rasch model, we aim to provide a nuanced understanding of customer engagement dynamics in mobile banking, contributing valuable insights to both academic discourse and practical applications in the banking industry.

In summary, this study not only explores an innovative approach to enhancing customer engagement through gamification but also introduces a sophisticated analytical framework for measuring engagement levels. By bridging these two pivotal areas, our research endeavors to enrich the literature on digital banking and gamification while offering actionable strategies for banking institutions to captivate and retain customers in the digital age.

2. RELATED WORK

2.1 Overview of the Rasch analysis

The Rasch model, a mathematical framework, utilizes probability estimates to evaluate the reliability and precision of rating scales, offering a robust methodology for assessment [25]. Considering the nonlinearity of raw scores, it is not reasonable to assume that the difference between two consecutive raw scores represents a consistent interval. This is one of the reasons why the Rasch measurement technique is employed. Some of these concerns are addressed by Wright [40] in their discussion. In an assessment of social skills using the Likert scale, let's consider the raw scores obtained by three individuals. Jim obtains a raw score of 20, Sue obtains a raw score of 22, and Jen obtains a raw score of 32. Based on the coding of raw score points, where Strongly Agree is assigned a code of 4, Agree as 3, Disagree as 2, and Strongly Disagree as 1, we can deduce that Jen demonstrates superior levels of social skills in comparison to the other alternatives. This inference is based on the fact that Jen's raw score corresponds to the highest raw score category, indicating stronger social skills. Rasch analysis employs a logit scale to rank the difficulty of items and the abilities of respondents, where higher scores indicate higher levels of ability or difficulty, while lower scores indicate the opposite. Before performing a Rasch analysis, there are a few preliminary steps that must be taken into account, including: item size, person size, Wright Map, match statistics, Rasch reliability index, and rating scale analysis. To demonstrate the practical application of Rasch's analytical methodologies in school psychology, this study illustrates their implementation by creating and piloting a self-report rating scale [33]. A logit scale was used in Rasch's study to rate participant ability and item difficulty, with higher scores denoting greater ability or difficulty and vice versa [41].

2.2 Gamification

Gamification is the strategic integration of games into non-gaming contexts or activities, aimed at enhancing customer engagement and fostering long-term retention [42]. The majority of studies investigating the use of gamification in banking have come to the conclusion that it significantly improves system performance and increases consumer engagement. The use of gamification to a variety of industries, including banking, will provide outstanding outcomes [16, 17]. This is because there is a clear link between gamification and the desire to utilize mobile banking services, which is supported by their widespread use. By employing a gamification strategy and creating suitable financial services, designers have the opportunity to enhance the appeal of banking operations. This, in turn, can lead to increased customer satisfaction and engagement. By comprehending essential concepts and effectively implementing a balanced selection of gamification techniques, it becomes possible to attain widespread consumer acceptance and value [16]. In a study by Rodrigues et al. [17], they developed a conceptual model to examine the effects of gamification on bank websites, and they demonstrated that it significantly affects web page design elements, user friendliness, and intent to use e-banking, among other things. Rahi et al. [43] study offered a solution for gamified internet banking and noted that by using gamification to make internet banking websites enjoyable and profit from the reward system, users' intentions to adopt the system and recommend it to others will improve social networking. According to related studies, gamification significantly enhanced customer interactions, use intention, and interest in social connections. For predicting behavioral intent in gamified commercial apps and linked enterprises, the researchers put out a conceptual model [44].

2.3 Demographic information

Previous research employed gender and age when using demographic data for various people in gamification systems. Orji [45], says that gender influences behavior, and men and women view behavior modification tactics differently. Lastly, Orji proposed gender as a reliable strategy and recommended the adoption of a gender-specific approach for design purposes. In a correlated investigation, Orji et al. [46] demonstrated that the techniques they employed were more effective in persuading women.

Oyibo et al. [47] investigated the role that culture plays in the influence of age and gender on the social influence of persuasive technology and demonstrated that there are significant differences between women and men and between younger and older people in collectivist and individualist cultures. Additionally, they furnished recommendations for tailoring personalization strategies based on cultural factors such as age and gender. PBL (points, badges, and leaderboards) is an interesting concept. According to some academics who looked into elderly people's gaming experiences, the majority of the currently available gamification systems are designed for younger audiences, have little relevance to parents, and cause them great stress when used [48]. Additionally, Denden et al. [49] have demonstrated that as children get older, their preferences, play motives, and play styles change from being performance-focused to being resolution, choice, and enjoyment-focused. They advise designers to concentrate more on objectives that favor setting and emotional connection as people age.

2.4 Gamification elements in banking

Recent advancements in the use of gamification systems have enhanced the requirement for diversified understanding between many people in numerous domains [15, 50, 51] Gamification is frequently linked to game elements, business applications, information systems, and entertaining tools. Fun-focused gamification endeavors to facilitate customer self-realization and foster sustained utilization of the system over the long term.

Table 1. Gamification elements studied in research

No	Reference	Gamification Element	Description
1	[50]	Announcements	User announcements, game start menus and game-first hints.
2	[9, 15, 52-58]	Point	Points that can be traded in for money or specifics. Calculating points for customers based on their success or performance in the banking system.
3		Reward	Users who complete challenges are rewarded.
4		Rank	User rating.
5		Badge	Awarding superior customers with badges or medals as a symbol of their accomplishments and a sign that they belong to a particular group.
6		Score	Note the results for each section's score. This is done in an effort to keep consumers motivated to raise their total grade.
7	[59-61]	Task	Task completion menu consisting of banking tasks.
8		Feedback	Feedback items can be entered for both the user and the bank, making it possible to distinguish between one customer's wants and another.
9		Leaderboard	Performance by every consumer to foster a spirit of cooperation and competitiveness.
10		Hunt for offers	Users search for deals that can be redeemed for discounts or other forms of promotions.
11	[50]	Timer	There is a timer for each game.
12	[50]	Levevl	Delivering personalised presents to consumers at different levels based on their performance
13	[56]	Share	Users should be encouraged to share. This encourages the sharing of information in society. to boost user motivation.
14	[15, 52, 60-62]	Social Interactive	Possibility for customers to socialize with others through technology.
15	[60, 62]	Penalty	Restricting service provision in response to negative customer behavior in the banking system.
16	[15, 60, 62]	Avatar	Use of a customer icon (avatar) as a thumbnail in a private profile is conceivable.
17	[15, 62]	Lottery	Conduct customer lotteries and award the winner.
18	[15, 60, 62]	Virtual reward	Giving customers virtual presents with the option to collect them as collections and using virtual currency with uses in the financial system.
19	[52, 62]	Epic meaning	Fostering in customers a sense of belonging to something greater than themselves and working for a greater good, even if it may not necessarily directly benefit them.
20	[15, 62]	Informing	Provide them with more banking benefits service.
21	[15, 62, 63]	Random reward	Giving consumers impromptu and unusual presents

The review found 21 different gamification elements (see Table 1). Social networks and social interactions inside a system, for example, may fall under more than one part; however, for the sake of this review, each aspect is categorized under a different heading.

Table 1 categorizes various gamification elements used in banking to engage customers, such as announcements for user interaction, points based on performance that can be exchanged, rewards for completing challenges, rankings, badges for achievements, scores to motivate improvement, task completion menus, feedback mechanisms, leaderboards, promotional hunts, timed, games, levels offering personalized rewards, sharing features, social interaction opportunities, penalties for negative behaviors, avatars for personalization, lotteries, virtual rewards, epic meanings to foster belonging, informational services, and random rewards. Each element is designed to enhance user experience, foster community, and incentivize participation in the banking system.

2.5 Customer engagement

Extensive scrutiny of the concept of engagement has been conducted across multiple disciplines, including marketing, management, information systems, and information technology management [64]. Information systems and information technology experts view user engagement as a part of a more significant sedentary flow or condition that denotes a person's level of involvement and enjoyment in an activity [65].

Engaging customers strengthens long-term ties with them, which boosts corporate competitiveness [66]. Bowden states participation is a "psychological process" involving emotional and cognitive components [67]. According to Vivek et al. [68], customer engagement is the level of involvement and interaction with a company's goods and initiatives that customers or organizations initiate. Businesses create environments to enhance user involvement to significantly increase their chances of success Field [69].

Previous studies have looked into how customers assess the features of Islamic banks; Ahmad et al. [69] found that customers have a more favorable opinion of the quality of service in Islamic banks than in conventional banks. Customer preferences were consistently reflected in the aspects of tangibility and empathy that were used to evaluate the quality of the services. According to Estiri et al. [70], cost, product offerings, and other service qualities and value propositions are important factors in customer satisfaction but do not always translate into loyalty. Echchabi and Nafiu Olaniyi [71] observed that, among other things, branch location and service quality influenced Malaysian clients' preferences for Islamic banking. In a similar vein, Amin et al. [72] connected client loyalty with the bank image. The role of digital technology in transformation of regional models of household financial behavior analyzed by Zhavoronok et al. [73] found that innovation processes play an important role in national economies, which is largely due to the process of digitalization of financial relations in all developed countries.

3. METHOD

This study used five stages, as illustrated in Figure 1. These stages include (1) Formulation of a Scale, (2) Data collection, (3) Data analysis, (4) Test validity and reliability, (5) Development of an instrument rating scale. A detailed description of each stage is given below.

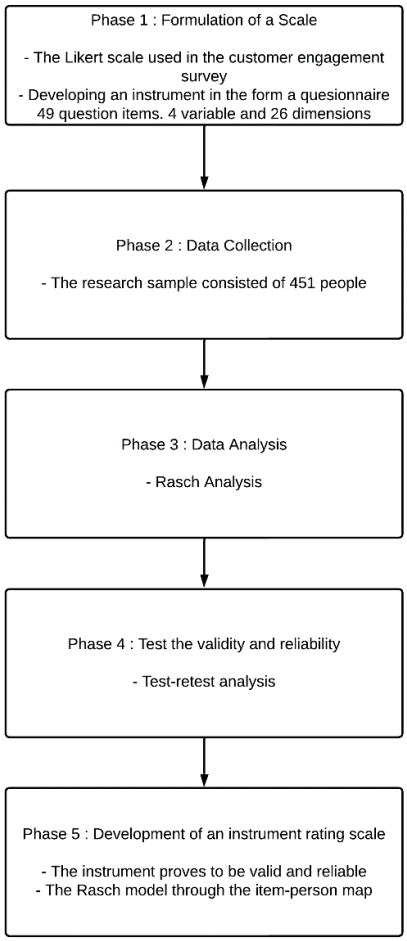

Figure 1. Research stage

3.1 Formulation of the scale

The customer engagement survey employed a Likert scale with the following response options: 1: Strongly Disagree, 2: Disagree, 3: Uncertain or Neutral, 4: Agree, and 5: Completely Agree. The Rasch model converts the ordinal item scores from the Likert rating scale into an interval scale called "logarithmic unit odds" (logit). In practice, most logit values fall within the 5.00 and 5.00 [74]. Item and person fitness statistics reveal the degree to which the results acquired are appropriate, dependable, and consistent with the underlying measure and offer details about the measurement's accuracy. After formulating a Likert scale, the next step is developing an instrument in the form of a questionnaire that is tailored to the needs of customer involvement with a gamification approach in the banking sector consisting of 48 statements. The instrument uses a Likert scale (1-5) which produces ordinal data. The assessment criteria for analyzing the results are derived from the standards established by Fisher Jr [75]. Table 2 below provides a guide, as outlined by Fisher Jr., for assessing the quality of the instrument.

3.2 Data collection

Data collected through measurement plays a crucial role in establishing the connection between empirical observations and quantitative mathematical representations. In this study, primary data from user questionnaires in the field of mobile banking were utilized, encompassing multiple banks across Indonesia. The research sample comprised 451 individuals who were active users of mobile banking applications. The data was collected among mobile banking users in Yogyakarta, Central Java, and North Sumatra, Indonesia.

Table 2. Instrument quality assessment standards [73]

Criterion	Poor	Fair	Good	Very Good	Excellent
Targeting*	>2 errors	1-2 errors	<1 error	< .5 error	< .25 error
ItemModel Fit Mean Square Range Extremes	< .33 ->3.0	.34-2.9	.5-2.0	.71-1.4	.77-1.3
Person and item Measurement Reliability	< .67	.67-80	.81- .90	.91- .94	> .94
Person and Item Stata Separated	2 or less	2-3	3-4	4-5	>5
Celling effect % maximum extreme scores	>5%	2-5%	1-2%	.5-1%	<5%
Floor effect % minimum extreme scores	>5%	2-5%	1-2 %	.5-1%	<5%
Variance in data explained by measures	<50%	50-60%	60-70%	70-80%	>80%
Unexplained variance in contrasts 1-5 of PCA of residuals	>15%	10-15	5-10%	3-5%	<3%

Data Collection Process: This study employed a comprehensive data collection process aimed at capturing a wide array of perceptions and interactions users have with gamification elements in banking apps. The research targeted a demographically diverse group of participants to ensure the inclusivity of the data. Participants were recruited through online forums, social media platforms, and email invitations sent to bank customers, with the selection criteria focusing on individuals who have used mobile banking applications within the last six months.

The survey instrument was carefully designed to gather quantitative and qualitative data on users' experiences with gamification. Questions were developed to assess the frequency of app usage, specific gamification features encountered, and the perceived impact of these features on engagement levels. To enhance the reliability and validity of the survey, items were pre-tested in a small pilot study, leading to refinements based on feedback.

The Rasch model was chosen for its robust capability to transform ordinal survey responses into interval measures, which are essential for accurate comparisons and assessments of user engagement. This model allows for the evaluation of both item difficulty (in this context, the engagement level required to interact with a gamification feature) and person ability (the user's likelihood to engage with the app due to gamification), providing a nuanced analysis of gamification's impact The application process involved several key steps:

(1) Data Preparation: Responses were coded and entered into the Rasch analysis software, ensuring that the data met the model's assumptions.

(2) Model Fitting: The Rasch model was applied to the dataset, with iterative adjustments made based on fit statistics and item-person interaction analyses. This process helped identify and address any anomalies or misfitting items, ensuring the model's accuracy.

(3) Interpretation of Results: The final output included person measures (indicating levels of user engagement) and item measures (reflecting the engagement potential of each

gamification feature). These metrics were then interpreted in relation to our research questions, providing insights into how different gamification elements influence user engagement in mobile banking apps.

3.3 Data analysis

Using Rasch analysis, the collected questionnaire data was assessed, enabling the conversion of ordinal data into interval data. According to Bond and Fox (2015), the Rasch model is widely recognized as the most suitable approach for fundamental analysis in human sciences, particularly when questionnaires are employed, resulting in ordinal measurements [26]. According to Cavanagh and Waugh [76], the probabilistic nature of the Rasch model enables accurate prediction of individuals' responses to all items that conform to the measurement model. This prediction can be achieved by utilizing the person parameter as the measure of individuals' abilities and the item parameter as the measure of item difficulty, both on the same scale.

This analysis uses Rach's mathematical model to estimate an individual's proficiency level based on their answer patterns to instrument items. The mathematical formulation of the Rasch One-Parameter Logistics (1PL) Model is as follows [77]:

$$P(Xi = 1) = c + (1 - c)/(1 + \exp(b - d))$$

In the above formula, P(Xi=1) is the probability that respondent i answered correctly in item i. c is a constant parameter that describes the probability of the respondent answering correctly on a very easy item. b is the item difficulty parameter which describes the level of difficulty of the i-th item. d is the difference between the skill level of respondent i and the level of difficulty of item i.

3.4 Test the validity and reliability

Upon analyzing the data, the validity and reliability tests of the developed instruments that were conducted affirm their accuracy in measuring customer engagement within the banking sector, specifically utilizing a gamification approach.

3.5 Development of an instrument rating scale

Once the instrument has been established as valid and reliable, the next step is to create a final scale that effectively assesses individual skills in customer engagement within the banking sector using a gamification approach. In order to measure the instrument, this study employs a Rasch analysis methodology, which represents a contemporary test analysis technique capable of surpassing the limitations of classical test theory. By utilizing the Rasch model, it becomes possible to ascertain the precision between the developed tests and the individuals being evaluated. This model further facilitates the examination of whether the developed test accommodates various levels of competency among the individuals being assessed through its item-person map feature. This map integrates two crucial pieces of information: the arrangement of items according to their difficulty levels and the arrangement of individuals based on their measured abilities. More specifically, the item maps offer valuable insights into the difficulty levels of the items within the assessment, identifying the most challenging and easiest items in the test.

4. RESULT AND DISCUSSION

4.1 Result

4.1.1 Formulation of the scale

Of the 26 dimensions of aspects related to customer involvement with a gamification approach in the banking sector using a Likert scale at intervals of 1-5 it shows that the scale most chosen by respondents is 5, which means that respondents' ratings tend to give a very agreeing assessment of items that are given. The following Figure 2 presents the results of responses from 451 respondents to the Likert scale items used in this study.

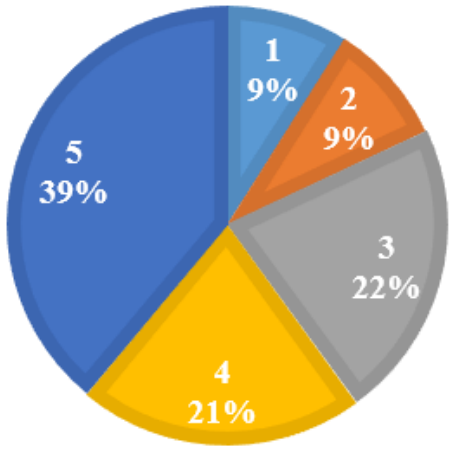

Figure 2. Respondent items

Based on the analysis of the formulation of the scale from 451 respondents showed that item 3 (undecided/neutral) and item 5 (strongly agree) respectively 22% and 21%. In items 1 and 2 regarding customer involvement, respondents are more likely to choose scale 3 (undecided/neutral). This indicates that customers already understand games but are still confused about using them when set in mobile banking, so this can provide benefits for banks and users. Meaningful customer engagement can be achieved through easy and fun and interactive game features. On the other hand, customer involvement in mobile banking applications can strengthen relationships with customers [39].

4.1.2 Data collection

In this research, a quantitative methodology is adopted, utilizing non-experimental designs to explore diverse phenomena. The collection of measurement data is of paramount importance as it enables the identification of relationships between empirical observations and quantitative mathematical expressions. To ensure the acquisition of primary data, an online questionnaire was meticulously designed and subsequently distributed via the user-friendly Google Forms platform. The questionnaire was thoughtfully conducted in the Indonesian language (Bahasa Indonesia) to specifically cater to the targeted participants from Indonesia. Table 3 provides an overview of the respondents' demographic profiles, encompassing factors such as gender, age, education, occupation, frequency of mobile banking usage (on a weekly basis), and weekly hours spent on the internet, social networks, video games, and similar activities. The sample exhibits a nearly equal distribution of male and female respondents, with the majority falling within the age range of 10 to 24 years old.

Table 3. The profile of research respondents

Category	Details	Person (*n*)	Proportion (%)
Gender	Male	238	52,7%
	Female	213	43,7%
Age (in years)	18-24	272	60,2%
	25-34	117	25,9%
	35-44	45	10%
	45-54	17	4%
Education	Senior High School	263	58,2%
	Diploma	28	6,2%
	Bachelor	147	32,5%
	Masters	13	2,9%
Occupation	Student	224	49,6%
	Government employees	22	4,9%
	BUMN	11	2,4%
	Private	87	19,2%
	Self-employed	74	16,4%
	Housewife	33	7,5%
Ability to use mobile banking services (within 1 week)	Very low	22	4,9%
	Low	72	15,9%
	Medium	262	58%
	Tall	62	13,7%
	Very high	33	7,5%
Hours in 1 week using the internet/social networks/video games and the like	1-3 hours	101	22,3%
	3-6 hours	80	17,7%
	6-9 hours	90	19,9%
	9-11 hours	70	15,5%
	Above 11 hours	110	24,6%

4.1.3 Data analysis

Linacre [78] propose an extensive set of indicators based on the Rasch model that can be applied to both individuals and items. These indicators encompass a range of psychometric properties, including mean square outfit (MNSQ), Z-standardized outfit (ZSTD), and point measurement correlation (PT-Measure Corr.). The evaluation process of the model commences by examining the MNSQ outfit value, which ideally falls within the 0.5 to 1.5 range, indicating a good fit for measurement. If the MNSQ value deviates from this range, attention is then shifted to the corresponding outfit ZSTD value, which should ideally be between -1.9 and 1.9, indicating reasonable data predictability. Internal reliability consistency, reflecting the average correlation among the items in the instrument, is assessed using Cronbach's α coefficient. A value approaching 1 signifies a favorable level of internal measurement consistency.

The study data was efficiently arranged using Microsoft Excel software and examined utilizing Winstep software version 3.73. After establishing suitable measurement intervals and satisfying validity and reliability criteria based on the Rasch model, the data underwent binary logistic regression analysis employing SPSS software. This particular methodology was selected to address the dichotomous nature of the questions answered by the participants.

4.1.4 Instrument reliability test

Table 4 displays the statistical overview of the instrument, encompassing person reliability and item reliability. The consistent findings confirm the accuracy and reliability of the measurements. The analysis yields two separate outputs, specifically person-output and item-output. The person table evaluates the statistical alignment of respondents with the data, whereas the item table assesses the compatibility of the instrument's items. As per Linacre [78], the mean value represents the signal-to-noise ratio within the data, with the mean coefficient indicating the square root of the ratio between the true person variance and the error variance present in the data. Boone et al. [79] stated that the index of the mean person and the mean item is a very important addition to the evaluation of the function of the measuring instrument.

Table 4. Statistical results of person and item reliability instruments

Person	**451 Input**		**451 Measured**		**INFIT**		**OUTFIT**	
	Total	Count	Measure	Realse	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	178.6	48.0	1.92	.60	1.04	-1.2	.99	-1.3
S.D.	55.4	.0	3.62	.63	.96	4.6	.95	4.5
Real RSME .87. TRUE SD			3.51 SEPARATION		4.03 PERSON RELIABILITY			.94
Item	48 Input		48 Measured		INFIT		OUTFIT	
	Total	Count	Measure	Realse	IMNSQ	ZSTD	OMNSQ	ZSTD
MEAN	1678.1	451.0	.00	.08	1.01	-.1	.99	-.3
S.D.	76.3	.0	.48	.00	.22	2.9	.30	2.9
Real RSME .08		True SD	.48 SEPARATION 5.67 ITEM RELIABILITY .97					

Based on the information presented in Table 4, the average person measure is recorded as 1.92. An average value above 0 indicates that the individuals or respondents possess higher abilities relative to the difficulty level of the items within the instrument. The individual reliability index is 0.94, while for items, it reaches 0.95. The responses from the subjects demonstrate strong consistency, and the overall reliability of the items is considered excellent, following the guidelines provided in standard Table 1 for person and item dependability.

The optimal values for INFIT and OUTFIT MNSQ approach 1, signifying a good fit, while the desirable values for INFIT and OUTFIT ZSTD are those close to 0, indicating reasonable predictability. After conducting a comprehensive analysis of the person and item tables, it is evident that the mean values of INFIT and OUTFIT MNSQ, as well as INFIT and OUTFIT ZSTD, align with the desired range. A higher separation value is preferred because it identifies a broader spectrum of subjects (those who are able versus unable) and item groups (difficult versus easy). Table 1 indicates a Separation Person value of 4.03, indicating that the instrument items are highly sensitive (very good) in covering the entire range of respondent abilities. The Separation Item value is 5.67, suggesting considerable variability among respondents and excellent detection of instrument item performance. The findings of the person and item validity tests, along with the MNSQ and PT-measure values, are depicted in Table 5. All

items have exceeded the critical values for MNSQ and PT-Measure, thus affirming their validity. Linacre [78] suggests that the statistical output's sensitivity to outliers facilitates identifying and resolving any issues related to the appropriateness or compatibility between the data and the model. According to Boone et al. [79] Outfit Means-square, Outfit z-standard and Point Measure correlation values are the criteria used to determine item suitability. It is advisable to get the item repaired or replaced if it does not satisfy the requirements. The Outfit Mean Square (MNSQ) value received is 0.5 MNSQ 1.5, which is the criterion used to determine whether an item is appropriate. Table 5 displays item statistics in summary.

Table 5. Summary of item statistics

Entry Number	Total Score	Total Count	Measure	Model S.E.	Infit		Outfit		Pt-measure			Exact match		Item
					MNSQ	ZSTD	MNSQ	ZSTD		CORR.	EXP.	OBS%	EXP%	
1	1506	451	1.05	.08	1.33	4.0	1.88	7.4	A	.84	.86	38.6	53.1	UE1
25	1591	451	.55	.08	1.39	4.6	1.68	6.0	B	.82	.85	65.0	55.1	LD2
5	1533	451	.89	.08	1.31	3.8	1.63	5.7	C	.84	.86	62.5	53.1	AC1
4	1596	451	.52	.08	.96	- .5	1.48	4.4	D	.86	.85	68.1	55.3	SE2
13	1808	451	- .86	.08	1.13	1.6	1.41	2.7	E	.81	.82	66.1	60.5	RW1
45	1666	451	.09	.08	1.34	4.0	1.40	3.5	F	.82	.84	62.2	57.7	IN1
44	1650	451	.19	.08	1.35	4.2	1.26	2.4	G	.82	.85	40.3	57.1	EM2
2	1493	451	1.13	.08	1.17	2.2	1.30	2.9	H	.86	.86	66.9	53.4	UE2
46	1700	451	- .13	.08	1.28	3.4	1.18	1.6	I	.82	.84	67.5	58.3	IN2
11	1762	451	- .54	.08	1.28	3.3	1.17	1.4	J	.81	.83	64.7	59.3	PO1
23	1761	451	- .53	.08	1.26	3.1	1.13	1.1	K	.82	.83	69.2	59.3	FE2
48	1680	451	.00	.08	1.26	3.1	1.16	1.5	L	.82	.84	67.8	58.0	RW2
43	1651	451	.18	.08	1.24	2.9	1.21	2.0	M	.83	.85	64.7	57.1	EM1
38	1614	451	.41	.08	1.23	2.9	1.15	1.5	N	.84	.85	69.2	53.1	AV2
47	1732	451	- .34	.08	1.18	2.2	1.08	.7	O	.82	.84	70.8	56.0	RW1
35	1750	451	- .46	.08	1.17	2.1	1.14	1.2	P	.82	.83	71.1	59.0	PE1
22	1715	451	- .22	.08	1.13	1.6	1.17	1.5	Q	.83	.84	66.4	58.5	FE1
29	1643	451	.23	.08	1.12	1.6	1.02	.2	R	.84	.85	68.9	57.0	TM2
27	1807	451	- .85	.08	1.12	1.4	.96	- .3	S	.81	.82	67.8	60.3	HO2
36	1779	451	- .65	.08	1.09	1.1	.97	- .2	T	.82	.83	70.8	59.6	PE2
3	1576	451	.64	.08	.93	- .9	1.09	.9	U	.87	.85	63.6	54.7	SE1
34	1623	451	.36	.08	1.07	.9	1.04	.4	V	.84	.85	68.6	56.3	SI2
26	1789	451	- .72	.08	1.03	.4	.90	- .8	W	.82	.83	71.7	59.6	HO1
33	1608	451	.45	.08	1.03	.4	.98	- .2	X	.85	.85	67.5	55.7	SI1
6	1646	451	.21	.08	1.02	.3	.97	- .3	x	.85	.85	64.4	57.0	AC2
10	1677	451	.02	.08	1.02	.2	.97	- .2	w	.84	.84	60.6	57.9	AN2
28	1649	451	.20	.08	1.01	.2	.92	- .8	v	.85	.85	69.7	57.1	TM1
14	1781	451	-.67	.08	1.00	.0	.85	- 1.2	u	.82	.83	73.1	59.6	RW2
31	1724	451	-.28	.08	.97	- .3	.89	- 1.0	t	.84	.84	71.1	58.8	LV2
30	1724	451	- .28	.08	.96	- .4	.88	- 1.1	s	.84	.84	73.3	58.8	LV1
12	1740	451	- .39	.08	.91	- 1.1	.83	- 1.5	r	.83	.83	68.6	59.1	PO2
42	1784	451	- .69	.08	.90	- 1.2	.76	- 2.0	q	.83	.83	75.8	59.7	VR2
40	1729	451	- .32	.08	.88	- 1.6	.77	- 2.1	p	.84	.84	74.2	58.8	LT2
8	1583	451	60	.08	.87	- 1.8	.85	- 1.6	o	.87	.85	65.6	54.8	BE2
7	1579	451	.62	.08	.86	- 1.9	.86	-1.5	n	.88	.85	62.8	54.7	BE1
37	1665	451	.10	.08	.86	- 1.9	.78	- 2.3	m	.86	.85	70.8	57.7	AV1
24	1629	451	.32	.08	.83	- 2.4	.84	- 1.7	l	.86	.85	66.7	56.5	LD1
21	1718	451	- .24	.08	.79	- 2.9	.68	- 3.2	k	.86	.84	72.5	58.8	TS1
16	1663	451	.11	.08	.78	- 3.0	.70	- 3.2	j	.86	.85	73.3	58.8	RK2
18	1679	451	.01	.08	.76	- 3.4	.69	- 3.3	i	.86	.84	75.8	58.8	BD2
9	1591	451	.55	.08	.74	- 3.8	.69	- 3.6	h	.88	.85	68.1	58.8	AN1
39	1731	451	- .33	.08	.73	- 3.9	.64	- 3.6	g	.85	.84	73.1	58.8	LT1
15	1724	451	- .28	.08	.71	- 4.1	.63	- 3.8	f	.86	.84	74.2	58.8	RK1
41	1743	451	- .41	.08	.70	- 4.3	.63	- 3.6	e	.85	.83	74.7	58.8	VR1
32	1675	451	.03	.08	.69	- 4.4	.63	- 4.1	d	.87	.84	74.4	58.8	SH1
20	1702	451	- .14	.08	.65	- 5.1	.57	- 4.7	c	.86	.84	71.4	58.8	SC2
17	1682	451	- .01	.08	.60	- 6.0	.58	- 4.7	b	.87	.84	73.9	58.8	BD1
19	1696	451	- .10	.08	.60	- 6.1	.56	- 4.9	a	.87	.84	76.1	58.8	SC1

Table 6. Rating scale category measurement results

Category		Observed		Obsvd	Sample	Infit	Outfit	Andrich	Category	
LABEL	SCORE	COUNT	%	AVRGE	EXPECT	MNSQ	MNSQ	THRESHOLD	MEASURE	
1	1	2037	9	-2.26	-2.59	1.66	1.81	NONE	(-3.52)	1
2	2	1825	8	-1.26	- .96	.76	.77	-2.20	-1.80	2
3	3	4768	22	.29	.29	.75	.89	-1.30	- .10	3
4	4	4534	21	1.74	1.70	.77	.67	1.02	1.79	4
5	5	8484	39	3.55	3.55	1.28	1.25	2.47	(3.72)	5

Analysis of the discussion of results and Table 5. Summary of Statistical Items can provide insight into the psychometric properties and performance of gamified customer engagement instruments in the banking sector. The table shows that the INFIT Mean square (MNSQ) value for each instrument item lies between 0.5-1.5 as shown in column 6. This shows that the items are all appropriate and do not need to be replaced, so it can be concluded that there is no there is an obvious disturbance or impact to the model while performing the statistical analysis.

4.1.5 Development of an instrument rating scale

The instrument's rating scale will be evaluated in the following stage. analysis on the results of this measurement is used to verify whether the choice rating used confuses the response or not. To see whether the rating scale function is running well, it can be seen through the measurement results in Table 6.

Based on the information presented in Table 6, the average person measure is recorded as 1.92. An average value above 0 indicates that the individuals or respondents possess higher abilities relative to the difficulty level of the items within the instrument. The individual reliability index is 0.94, while for items, it reaches 0.95. The responses from the subjects demonstrate strong consistency, and the overall reliability of the items is considered excellent, following the guidelines provided in standard Table 1 for person and item dependability.

The optimal values for INFIT and OUTFIT MNSQ are those that approach 1, signifying a good fit, while the desirable values for INFIT and OUTFIT ZSTD are those close to 0, indicating reasonable predictability. After conducting a comprehensive analysis of the person and item tables, it is evident that the mean values of INFIT and OUTFIT MNSQ, as well as INFIT and OUTFIT ZSTD, align with the desired range. A higher separation value is preferred because it allows for the identification of a broader spectrum of subjects (those who are able versus unable) and item groups (difficult versus easy). Table 1 indicates a Separation Person value of 4.03, indicating that the instrument items are highly sensitive (very good) in covering the entire range of respondent abilities. The Separation Item value is 5.67, suggesting considerable variability among respondents and excellent detection of instrument item performance. The findings of the person and item validity tests, along with the MNSQ and PT-Measure values, are depicted in Table 6. All items have exceeded the critical values for MNSQ and PT-Measure, thus affirming their validity. Linacre (2012) suggests that the statistical output's sensitivity to outliers facilitates the identification and resolution of any issues related to the appropriateness or compatibility between the data and the model [79]. According to Boone, et al. (2014), Outfit Means-square, Outfit z-standard and Point Measure correlation values are the criteria used to see the level of item suitability [78]. The Outfit Mean Square (MNSQ) value received is 0.5 MNSQ 1.5, and this number is the criterion used to determine whether an item is appropriate. Table 5 displays item statistics in summary.

Analysis of the discussion of results and Table 5. Summary of Statistical Items can provide insight into the psychometric properties and performance of gamified customer engagement instruments in the banking sector. The table shows that the INFIT Mean square (MNSQ) value for each instrument item lies between 0.5-1.5 as shown in column 6. This shows that the items are all appropriate and do not need to be replaced, so it can be concluded that there is no there is an obvious disturbance or impact to the model while performing the statistical analysis

4.1.6 Development of an instrument rating scale

The instrument's rating scale will be evaluated in the following stage. analysis on the results of this measurement is used to verify whether the choice rating used confuses the response or not. To see whether the rating scale function is running well, it can be seen through the measurement results in Table 6.

Table 6 presents the measurement outcomes for each category in the rating scale. Category 1 corresponds to a score of 1 and consists of 2037 observations, accounting for 9% of the sample. The average measure for this category is -2.26, with an expected measure of -2.59. The Infit MNSQ is 1.66, and the Outfit MNSQ is 1.81. No threshold measures are available for this category. Category 2 represents a score of 2 and encompasses 1825 observations, making up 8% of the sample. The average measure is -1.26, while the expected measure is -0.96. The Infit MNSQ is 0.76, and the Outfit MNSQ is 0.77. The threshold measure for this category is -2.20. Category 3, with a score of 3, comprises 4768 observations, accounting for 22% of the sample. Both the average measure and expected measure for this category are 0.29. The Infit MNSQ is 0.75, and the Outfit MNSQ is 0.89. The threshold measure is -1.30. Category 4 represents a score of 4 and includes 4534 observations, making up 21% of the sample. The average measure is 1.74, while the expected measure is 1.70. The Infit MNSQ is 0.77, and the Outfit MNSQ is 0.67. The threshold measure is 1.02. Lastly, Category 5 corresponds to the highest score of 5, with 8484 observations, accounting for 39% of the sample. Both the average measure and expected measure for this category are 3.55. The Infit MNSQ is 1.28, and the Outfit MNSQ is 1.25. The threshold measure is 2.47. These findings provide a summary of the distribution and measurement characteristics for each category of the rating scale employed in the study.

The scale categories in Table 5 column 5 in the Category measure section show that each category is going well because on a scale of 1-5 the Category measure value increases proportionally and sequentially from -3.52 to 3.72. The existence of this monotonic increase indicates that the measurement has been going well. It is apparent that responders can discriminate between the answer options from very improper to very appropriate because there are no similar values in these five selections. Not all studies directly get good rating measurement results like this. Several studies related to Rasch analysis show that there are disproportionate conditions or unbalanced conditions, sometimes even the measurement of the rating scale is not shown as in research [80]. This can also be seen based on the probability of a person's ability as shown in Figure 3. Figure 3 demonstrates that if a person's ability is low, there is a high likelihood that they will respond with a rating of 1, while their ability is high, there is a higher likelihood that they would respond with a rating of 5. Figure 3 below depicts an ideal Rasch analysis model that can be used to compare the instrument's applicability as a whole.

Within a section of Figure 4, the scores of the instrument items are compared to the difficulty level measurement of the items, denoted by the blue line. The figure depicts the pattern of tested instrument item scores, which generally adheres to the ideal model of Rasch analysis, represented by the red line. While not flawless, the chart demonstrates that the instrument closely approximates the ideal Rasch pattern. This alignment

between the instrument item scores and the Rasch ideal model is an encouraging discovery, indicating a reasonable level of agreement between the instrument and the targeted construct or variable. The closer the alignment between the instrument item scores and the Rasch ideal model, the stronger the measurement properties of the instrument.

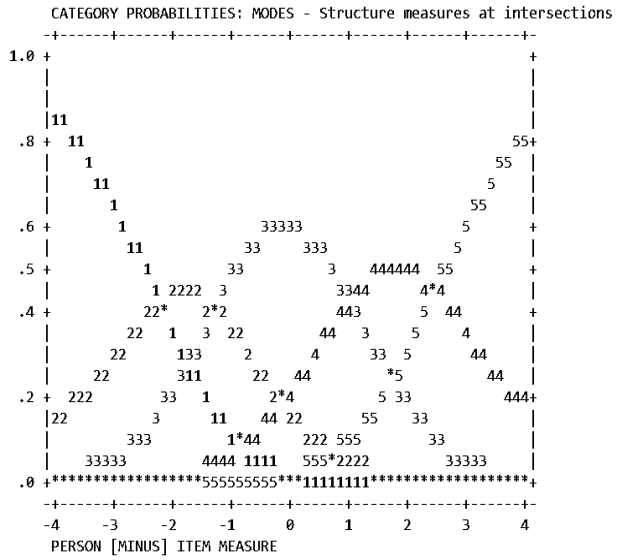

Figure 3. Probability of person's ability

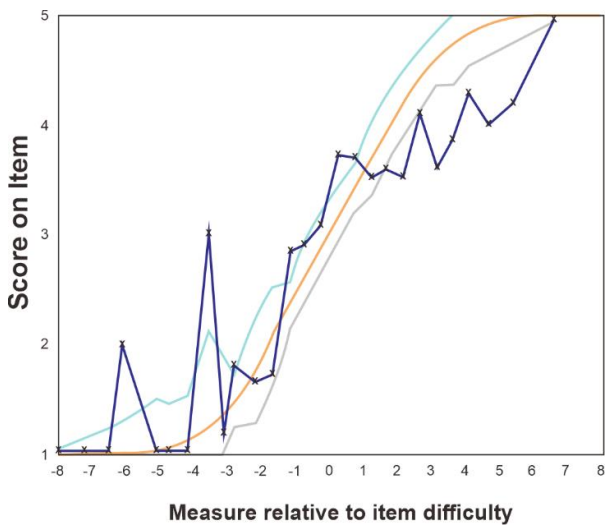

Figure 4. Overall instrument suitability results

4.2 Discussion

Comparing these results with previous research conducted by Lailiyah et al. [81], in analyzing self-assessment instrument items it shows that there are response patterns that are too far from the ideal model line curve. Therefore, additional analysis must be done, namely by gradually deleting things from the analysis that do not fit until the findings show that all of the items are consistent with the model. In Kadaryanto's study, which focuses on the design of teacher professional development programs utilizing the Rasch model, the analysis indicates a lack of alignment between the Item Characteristic Curve (ICC) and the desired pattern represented by the Red line (ideal model). In simpler terms, there are outliers observed in the distribution of respondents' answers to the survey questions. Ideally, the measured values should closely resemble the ideal model line. This suggests that some respondents may have provided careless or thoughtless responses to the survey questions [82]. Zansen's research utilized the Rasch measure to examine the fairness of the listening part across gender subgroups of a nationwide high-stakes computerized English examination. The NUDIF research' findings reveal that 12 items have gender-based DIF. The difference in ICC is not always discernible when pupils are separated into high-and low-achieving boys and girls [82].

The validity of the instruments used in further research can be assessed by examining the log values of each item or person. Deviations in logit values indicate a potential problem with a particular individual or item, indicating a need for revision or replacement. Ideally, the logit or size of each person and item should be closer to zero, indicating a better fit to the construct being measured. Another instrument validity indicator is the standard error of measurement (S.E. Measurement). A low standard error of measurement (less than 0.5) indicates high accuracy for an item or person, meaning that the measurements are reliable and precise. A standard measurement error between 0.5 and 1 is considered acceptable or moderately accurate, while a value exceeding 1.0 indicates poor accuracy for the person or item being measured. In addition to the standard error of measurement, the Outfit Mean Square (OUTFIT MNSQ) value provides insight into the instrument's validity. The ideal range for the OUTFIT MNSQ value is typically between 0.5 and 1.5. Values in this range indicate a good fit between the instrument items and the underlying construct. Departing from this range could suggest non-compliance or an issue with measurement that requires attention. The OUTFIT ZSTD metric is an additional means of evaluating the instrument's reliability. Values falling within the -2 to +2 range are deemed favorable, signifying a satisfactory alignment between the observed and expected responses derived from the Rasch model. Values outside this range may indicate potential problems with the measurement pattern or responses that deviate significantly from the expected model. Finally, Point Size Correlation is crucial in assessing the instrument's validity. Ideally, the Point Size Correlation should fall within the range of 0.4 to 0.85 to demonstrate sound validity. This correlation coefficient represents the degree of connection between the item's score on the instrument and the underlying construct being measured. Stronger correlations indicate more robust associations and offer supporting evidence of the instrument's proficiency in effectively measuring the intended construct [81].

By considering these validity indicators, researchers can assess the quality and feasibility of the instruments used in the study. If the instrument exhibits favorable log values, low standard error of measurement, suitable clothing mean squared values, clothing ZSTD values within the desired range, and satisfactory point size correlation, it indicates good validity. Conversely, if these indicators fall outside the recommended ranges, further investigation and potential instrument modification may be required to increase their validity.

Overall, the discussion emphasizes the significance of assessing the reliability of the research tools by considering numerous elements and criteria. Validity assessment ensures that the instrument accurately measures the intended construct and provides reliable and meaningful data for research purposes.

This study's limitations include its reliance on quantitative methods without thoroughly investigating respondents' diverse backgrounds. It leaves a gap in empirical insights for

banks targeting specific community segments for their marketing efforts.

5. CONCLUSION

In conclusion, this study highlights the overwhelmingly positive assessments provided by respondents regarding the gamification approach in the banking sector, as evidenced by the frequent selection of the highest rating on the Likert scale. This positive inclination indicates a favorable perception and active engagement with the gamified elements implemented in mobile banking applications. The Rasch analysis further confirms that respondents possess a strong understanding of the instrument items, rendering the instrument suitable without needing replacement. The research successfully captures and evaluates multiple aspects of customer engagement, thereby validating the instrument's reliability and validity. The results highlight the potential effectiveness and attractiveness of gamified elements in boosting customer involvement and improving user experiences, ultimately resulting in higher customer satisfaction.

However, it is crucial to acknowledge that these results are based on a specific sample of 451 participants. Therefore, further research is necessary to examine the generalizability of these findings to a broader population and diverse banking contexts. Qualitative data and follow-up studies can provide deeper insights into the specific aspects of gamification that resonate most strongly with customers. Bank managers can derive significant benefits from the implications of this research by increasing awareness, aligning with customer preferences, and enhancing human resources, services, and infrastructure within their organizations. Proactively anticipating customer behavior through excellent service management, product adjustments, and soliciting feedback are critical considerations for successful implementation.

This study emphasizes how crucial it is for bank executives and governmental organizations, such as the OJK, to increase public knowledge of and preference for banking by fostering extensive stakeholder collaboration. It also emphasizes how important it is for Islamic banks to improve infrastructure, streamline services, and increase the quantity and caliber of their human resources. It also emphasizes the need for banks to aggressively seek community feedback, adjust product offers, and improve service management in response to shifts in customer behavior.

Future research endeavors should explore bank customer profiles, identify target market segments, and develop tailored marketing strategies. Combining quantitative and qualitative analyses while minimizing biases in measurement tools can provide more comprehensive results. This study offers valuable insights into measuring customer engagement with gamification in mobile banking applications. The instrument exhibits good comprehension, reliability, and validity, establishing its suitability as an effective data-collection tool. The implications of this research can serve as a guiding resource for bank managers in improving their strategies to address customers' evolving needs within the digital banking landscape.

ACKNOWLEDGMENT

The author would like to thank the Doctoral Program in Information Systems, Universitas Diponegoro, Semarang and the Faculty of Information Technology, Universitas Mercu Buana Yogyakarta, Yogyakarta, for all their support for our research. Lastly, we would like to express our gratitude to the editors, reviewers, and publishing journal personnel for their thoughtful consideration and direction during the peer review and publication processes. Their professionalism and knowledge have been crucial to the completion of this project.

REFERENCES

[1] Lee, Y.J., Ghasemkhani, H., Xie, K., Tan, Y. (2021). Switching decision, timing, and app performance: An empirical analysis of mobile app developers' switching behavior between monetization strategies. Journal of Business Research, 127: 332-345. https://doi.org/10.1016/j.jbusres.2021.01.027

[2] Chibba, M. (2009). Financial inclusion, poverty reduction and the millennium development goals. The European Journal of Development Research, 21: 213-230. https://doi.org/10.1057/ejdr.2008.17

[3] Correa, T., Pavez, I. (2016). Digital inclusion in rural areas: A qualitative exploration of challenges faced by people from isolated communities. Journal of Computer-Mediated Communication, 21(3): 247-263. https://doi.org/10.1111/jcc4.12184

[4] Statista (2020). Number of mobile app downloads worldwide from 2016 to 2019. https://www.statista.com/statistics/271644/worldwide-free-and-paid-mobile-app-store-downloads/.

[5] Localytics (2019). 25% of users abandon apps after one use. https://uplandsoftware.com/localytics/resources/blog/25-of-users-abandon-apps-after-one-use/.

[6] Cechetti, N.P., Bellei, E.A., Biduski, D., Rodriguez, J.P.M., Roman, M.K., De Marchi, A.C.B. (2019). Developing and implementing a gamification method to improve user engagement: A case study with an m-Health application for hypertension monitoring. Telematics and Informatics, 41: 126-138. https://doi.org/10.1016/j.tele.2019.04.007

[7] Friedrich, J., Becker, M., Kramer, F., Wirth, M., Schneider, M. (2020). Incentive design and gamification for knowledge management. Journal of Business Research, 106: 341-352 https://doi.org/10.1016/j.jbusres.2019.02.009

[8] Hofacker, C.F., De Ruyter, K., Lurie, N.H., Manchanda, P., Donaldson, J. (2016). Gamification and mobile marketing effectiveness. Journal of Interactive Marketing, 34(1): 25-36. https://doi.org/10.1016/j.intmar.2016.03.001

[9] Koivisto, J., Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. International Journal of Information Management, 45: 191-210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013

[10] Ho, M.H.W., Chung, H.F. (2020). Customer engagement, customer equity and repurchase intention in mobile apps. Journal of Business Research, 121: 13-21. https://doi.org/10.1016/j.jbusres.2020.07.046

[11] Kamboj, S., Rana, S., Drave, V.A. (2020). Factors driving consumer engagement and intentions with gamification of mobile apps. Journal of Electronic

Commerce in Organizations (JECO), 18(2): 17-35. https://doi.org/10.4018/JECO.2020040102.
[12] Suh, A., Wagner, C., Liu, L. (2018). Enhancing user engagement through gamification. Journal of Computer Information Systems, 58(3): 204-213. https://doi.org/10.1080/08874417.2016.1229143
[13] Jia, Y., Xu, B., Karanam, Y., Voida, S. (2016). Personality-targeted gamification: A survey study on personality traits and motivational affordances. In Proceedings of The 2016 CHI Conference on Human Factors in Computing Systems, pp. 2001-2013. https://doi.org/10.1145/2858036.2858515
[14] Orji, R., Nacke, L.E., Di Marco, C. (2017). Towards personality-driven persuasive health games and gamified systems. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 1015-1027. https://doi.org/10.1145/3025453.3025577
[15] Tondello, G.F., Mora, A., Nacke, L.E. (2017). Elements of gameful design emerging from user preferences. In Proceedings of The Annual Symposium on Computer-Human Interaction in Play, pp. 129-142. https://doi.org/10.1145/3116595.3116627
[16] Baptista, G., Oliveira, T. (2017). Why so serious? Gamification impact in the acceptance of mobile banking services. Internet Research, 27(1): 118-139. https://doi.org/10.1108/IntR-10-2015-0295
[17] Rodrigues, L.F., Costa, C.J., Oliveira, A. (2017). How does the web game design influence the behavior of e-banking users? Computers in Human Behavior, 74: 163-174. https://doi.org/10.1016/j.chb.2017.04.034
[18] Nasirzadeh, E., Fathian, M. (2020). Investigating the effect of gamification elements on bank customers to personalize gamified systems. International Journal of Human-Computer Studies, 143: 102469. https://doi.org/10.1016/j.ijhcs.2020.102469
[19] Rodrigues, L.F., Oliveira, A., Costa, C.J. (2016). Does ease-of-use contributes to the perception of enjoyment? A case of gamification in e-banking. Computers in Human Behavior, 61: 114-126. https://doi.org/10.1016/j.chb.2016.03.015
[20] Rodrigues, L.F., Oliveira, A., Costa, C.J. (2016). Playing seriously–How gamification and social cues influence bank customers to use gamified e-business applications. Computers in human behavior, 63, 392-407. https://doi.org/10.1016/j.chb.2016.05.063
[21] Aydınlıyurt, E.T., Taşkın, N., Scahill, S., Toker, A. (2021). Continuance intention in gamified mobile applications: A study of behavioral inhibition and activation systems. International Journal of Information Management, 61: 102414. https://doi.org/10.1016/j.ijinfomgt.2021.102414
[22] Hamidi, H., Safareeyeh, M. (2019). A model to analyze the effect of mobile banking adoption on customer interaction and satisfaction: A case study of m-banking in Iran. Telematics and Informatics, 38: 166-181. https://doi.org/10.1016/j.tele.2018.09.008
[23] Dzandu, M.D., Hanu, C., Amegbe, H. (2022). Gamification of mobile money payment for generating customer value in emerging economies: The social impact theory perspective. Technological Forecasting and Social Change, 185: 122049. https://doi.org/10.1016/j.techfore.2022.122049
[24] Hergesell, A. (2022). Using Rasch analysis for scale development and refinement in tourism: Theory and illustration. Journal of Business Research, 142: 551-561. https://doi.org/10.1016/j.jbusres.2021.12.063
[25] Rasch, G. (1960). Probabilistic model for some intelligence and achievement tests. University of Chicago Press, Chicago, IL.
[26] Bond, T.G., Fox, C.M. (2007). Applying the Rasch Model: Fundamental Measurement in the Human Sciences, Second Edition (2nd ed.). Psychology Press. https://doi.org/10.4324/9781410614575
[27] The Standards for educational and psychological testing. American Educational Research Association, Washington D.C. https://www.apa.org/science/programs/testing/standards
[28] Andrich, D. (1988). Rasch models for measurement. Newbury Park London New Delhi: SAGE Publications.
[29] Rasch, G. (1993). Probabilistic models for some intelligence and attainment tests. MESA Press.
[30] Tennant, A., Conaghan, P.G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Care & Research, 57(8): 1358-1362. https://doi.org/10.1002/art.23108
[31] Nguyen, T.H., Han, H.R., Kim, M.T., Chan, K.S. (2014). An introduction to item response theory for patient-reported outcome measurement. The Patient-Patient-Centered Outcomes Research, 7: 23-35. https://doi.org/10.1007/s40271-013-0041-0
[32] Soh, S.E., Harris, I.A., Cashman, K., Heath, E., Lorimer, M., Graves, S.E., Ackerman, I.N. (2021). Implications for research and clinical use from a Rasch analysis of the HOOS-12 and KOOS-12 instruments. Osteoarthritis and Cartilage, 29(6): 824-833. https://doi.org/10.1016/j.joca.2021.02.568
[33] Boone, W.J., Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. Cogent Education, 4(1): 1416898. https://doi.org/10.1080/2331186X.2017.1416898.
[34] Brush, G.J., Soutar, G.N. (2022). A Rasch analysis of service performance in a tourism context. Journal of Business Research, 139: 338-353. https://doi.org/10.1016/j.jbusres.2021.09.038
[35] Stolt, M., Kottorp, A., Suhonen, R. (2022). The use and quality of reporting of Rasch analysis in nursing research: A methodological scoping review. International Journal of Nursing Studies, 132: 104244. https://doi.org/10.1016/j.ijnurstu.2022.104244
[36] Goswami, S., Barnard, M., Bynum, L.A., Thompson, S., Kang, M. (2022). Psychometric evaluation of the Health Care Provider HIV/AIDS Stigma Scale (HPASS) using Rasch analysis. Currents in Pharmacy Teaching and Learning, 14(9): 1143-1153. https://doi.org/10.1016/j.cptl.2022.07.035
[37] Soeharto, S., Csapó, B. (2022). Assessing Indonesian student inductive reasoning: Rasch analysis. Thinking Skills and Creativity, 46: 101132. https://doi.org/10.1016/j.tsc.2022.101132
[38] Gamble, J., Pallant, J., Creedy, D.K. (2022). Evaluation of the midwifery student evaluation of practice (MidSTEP) tool using Rasch analysis. Nurse Education Today, 108: 105174. https://doi.org/10.1016/j.nedt.2021.105174
[39] Prasetyaningrum, P.T., Purwanto, P., Rochim, A.F. (2022). Gamification on mobile banking application: A literature review. In 2022 1st International Conference on

Information System & Information Technology (ICISIT), Yogyakarta, Indonesia, pp. 222-227. https://doi.org/10.1109/ICISIT54091.2022.9873034
[40] Wright, B.D. (1992). Raw scores are not linear measures: Rasch vs. classical test theory CTT comparison. Rasch Measurement Transactions, 6(1): 208.
[41] Gothwal, V.K., Wright, T.A., Lamoureux, E.L., Pesudovs, K. (2009). Rasch analysis of the quality of life and vision function questionnaire. Optometry and Vision Science, 86(7): E836-E844. https://doi.org/10.1097/OPX.0b013e3181ae1ec7
[42] Hamari, J. (2013). Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. Electronic Commerce Research and Applications, 12(4): 236-245. https://doi.org/10.1016/j.elerap.2013.01.004
[43] Rahi, S., Othman Mansour, M.M., Alghizzawi, M., Alnaser, F.M. (2019). Integration of UTAUT model in internet banking adoption context: The mediating role of performance expectancy and effort expectancy. Journal of Research in Interactive Marketing, 13(3): 411-435. https://doi.org/10.1108/JRIM-02-2018-0032
[44] Rodrigues, L.F., Costa, C.J., Oliveira, A. (2016). Gamification: A framework for designing software in e-banking. Computers in Human Behavior, 62: 620-634. https://doi.org/10.1016/j.chb.2016.04.035
[45] Orji, R. (2014). Design for behaviour change: A model-driven approach for tailoring persuasive technologies. Doctoral Dissertation, University of Saskatchewan.
[46] Orji, R., Mandryk, R.L., Vassileva, J. (2014). Gender and persuasive technology: Examining the persuasiveness of persuasive strategies by gender groups. Persuasive Technology, 48-52.
[47] Oyibo, K., Orji, R., Vassileva, J. (2017). The influence of culture in the effect of age and gender on social influence in persuasive technology. In Adjunct Publication of The 25th Conference on User Modeling, Adaptation and Personalization, pp. 47-52. https://doi.org/10.1145/3099023.3099071
[48] Altmeyer, M., Lessel, P., Krüger, A. (2018). Investigating gamification for seniors aged 75+. In Proceedings of The 2018 Designing Interactive Systems Conference, pp. 453-458. https://doi.org/10.1145/3196709.3196799
[49] Denden, M., Tlili, A., Essalmi, F., Jemni, M. (2017). Educational gamification based on personality. In 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, Tunisia, IEEE, Hammamet, Tunisia, pp. 1399-1405. https://doi.org/10.1109/AICCSA.2017.87
[50] Garett, R., Young, S.D. (2019). Health care gamification: A study of game mechanics and elements. Technology, Knowledge and Learning, 24: 341-353. https://doi.org/10.1007/s10758-018-9353-4
[51] Hamari, J., Koivisto, J., Sarsa, H. (2014). Does gamification work?--A literature review of empirical studies on gamification. In 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, pp. 3025-3034. https://doi.org/10.1109/HICSS.2014.377
[52] Uskov, A., Sekar, B. (2015). Smart gamification and smart serious games. In Fusion of Smart, Multimedia and Computer Gaming Technologies, pp. 7-36.
[53] Berezan, O., Raab, C., Tanford, S., Kim, Y.S. (2015). Evaluating loyalty constructs among hotel reward program members using eWOM. Journal of Hospitality & Tourism Research, 39(2): 198-224. https://doi.org/10.1177/1096348012471384
[54] Zuo, L., Xiong, S., Iida, H. (2017). An analysis of hotel loyalty program with a focus on the tiers and points system. In 2017 4th International Conference on Systems and Informatics (ICSAI), Hangzhou, China, pp. 507-512. https://doi.org/10.1109/ICSAI.2017.8248345
[55] Zatwarnicka-Madura, B. (2015). Gamification as a Tool for Influencing Customers' Behaviour. International Journal of Economics and Management Engineering, 9(5): 1461-1464.
[56] Chen, C.M., Ming-Chaun, L., Kuo, C.P. (2023). A game-based learning system based on octalysis gamification framework to promote employees' Japanese learning. Computers & Education, 205: 104899. https://doi.org/10.1016/j.compedu.2023.104899
[57] Negruşa, A.L., Toader, V., Sofică, A., Tutunea, M.F., Rus, R.V. (2015). Exploring gamification techniques and applications for sustainable tourism. Sustainability, 7(8): 11160-11189. https://doi.org/10.3390/su70811160
[58] Dalpiaz, F., Snijders, R., Brinkkemper, S., Hosseini, M., Shahri, A., Ali, R. (2017). Engaging the crowd of stakeholders in requirements engineering via gamification. Gamification: Using Game Elements in Serious Contexts, 123-135. https://doi.org/10.1007/978-3-319-45557-0_9
[59] Perkasa, F.B.P., Emanuel, A.W.R. (2020). Gamifikasi sebagai metode untuk mendapatkan customer loyalty: Review Literatur. Aiti, 17(1): 11-21. https://doi.org/10.24246/aiti.v17i1.11-21
[60] Miller, Koivisto, J., Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. International journal of information management, 45: 191-210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013
[61] The Octalysis Framework for Gamification & Behavioral Design. https://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/.
[62] Werbach, K. (2014). (Re) defining gamification: A process approach. In Persuasive Technology: 9th International Conference, PERSUASIVE 2014, Springer International Publishing, Padua, Italy, pp. 266-272. https://doi.org/10.1007/978-3-319-07127-5_23
[63] Werbach, K., Hunter, D., Dixon, W. (2012). For the win: How game thinking can revolutionize your business. Philadelphia: Wharton digital press, 1.
[64] Chan, T.K., Zheng, X., Cheung, C.M., Lee, M.K., Lee, Z.W. (2014). Antecedents and consequences of customer engagement in online brand communities. Journal of Marketing Analytics, 2: 81-97. https://doi.org/10.1057/jma.2014.9
[65] Webster, J., Ahuja, J.S. (2006). Enhancing the design of web navigation systems: The influence of user disorientation on engagement and performance. Mis Quarterly, 30(3): 661-678. https://doi.org/10.2307/25148744
[66] Stokburger-Sauer, N. (2010). Brand community: Drivers and outcomes. Psychology & Marketing, 27(4): 347-368. https://doi.org/10.1002/mar.20335
[67] 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
[68] 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
[69] Ahmad, A., Saif, I., Safwan, N. (2010). An empirical investigation of Islamic banking in Pakistan based on perception of service quality. African Journal of Business Management, 4(6): 1185.
[70] Estiri, M., Hosseini, F., Yazdani, H., Javidan Nejad, H. (2011). Determinants of customer satisfaction in Islamic banking: Evidence from Iran. International Journal of Islamic and Middle Eastern Finance and Management, 4(4): 295-307. https://doi.org/10.1108/17538391111186546
[71] Echchabi, A., Nafiu Olaniyi, O. (2012). Malaysian consumers' preferences for Islamic banking attributes. International Journal of Social Economics, 39(11): 859-874. https://doi.org/10.1108/03068291211263907
[72] Amin, M., Isa, Z., Fontaine, R. (2013). Islamic banks: Contrasting the drivers of customer satisfaction on image, trust, and loyalty of Muslim and non-Muslim customers in Malaysia. International Journal of Bank Marketing, 31(2): 79-97. https://doi.org/10.1108/02652321311298627
[73] Zhavoronok, A. Popelo, O., Shchur, R., Ostrovska, N., Kordzaia, N. (2022). The role of digital technologies in the transformation of regional models of households' financial behavior in the conditions of the national innovative economy development. Ingénierie des Systèmes d'Information, 27(4): 613-620. https://doi.org/10.18280/isi.270411
[74] Engelhard Jr, G. (2013). Invariant measurement: Using Rasch models in the social, behavioral, and health sciences. Routledge. Invariant Measurement, 312. https://doi.org/10.4324/9780203073636
[75] Fisher Jr., W.P. (2007). Rating scale instrument quality criteria. Rasch Measurement Transaction, 21(1): 1095. https://www.rasch.org/rmt/rmt211m.htm.
[76] Cavanagh, R.F., Waugh, R.F. (2011). Applications of Rasch measurement in learning environments research. Springer Science & Business Media.
[77] Rasch, G. (1993). Probabilistic Model for some Intelligence and Achievement Tests. University of Chicago Press, Chicago. https://eric.ed.gov/?id=ED419814.
[78] Linacre, J.M. (2016). A user's guide to WINSTEPS® MINISTEP: Rasch-model computer programs (No Title), 719. https://cir.nii.ac.jp/crid/1370567187572937223.
[79] Boone, W.J., Staver, J.R., Yale, M.S. (2013). Rasch analysis in the human sciences. Springer Science & Business Media.. https://doi.org/10.1007/978-94-007-6857-4
[80] Setiawan, B., Panduwangi, M., Sumintono, B. (2018). A Rasch analysis of the community's preference for different attributes of Islamic banks in Indonesia. International Journal of Social Economics, 45(12): 1647-1662. https://doi.org/10.1108/IJSE-07-2017-0294
[81] Lailiyah, L., Supriyati, Y., Komarudin, K. (2018). Analysis of measures items in development of instruments self-assessment (Rasch Modeling Application). JISAE: Journal of Indonesian Student Assessment and Evaluation, 4(1): 1-9. https://doi.org/10.21009/jisae.v4i1.7819
[82] Alt, D., Raichel, N. (2020). Enhancing perceived digital literacy skills and creative self-concept through gamified learning environments: Insights from a longitudinal study. International Journal of Educational Research, 101: 101561. https://doi.org/10.1016/j.ijer.2020.101561