

## Validation of the Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) with an Indonesian Student Sample: A Rasch Model Approach

Tina Deviana<sup>1</sup>, Nia Tresniasari<sup>2</sup>, Bahrul Hayat<sup>3</sup>, Irianto<sup>4</sup>

Psychology Study Program, Universitas Paramadina, Indonesia<sup>1</sup>

Faculty of Psychology, UIN Syarif Hidayatullah Jakarta, Indonesia<sup>2,3</sup>

Department General Education, Faculty of Resilience, Rabdan Academy, Abu Dhabi, United Arab Emirates<sup>4</sup>

[tina.deviana@paramadina.ac.id](mailto:tina.deviana@paramadina.ac.id)

### Abstract

Instagram is one of the most popular social media platforms in the world, especially among teenagers and young adults, allowing users to share visual content and interact socially online (Ting, 2014). As the psychological and behavioral impacts of Instagram use increase, it is important to assess not only its frequency, but also the behavior and emotional responses that accompany it through valid and reliable measurement tools such as the Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) (Paramboukis et al., 2016). This study aims to evaluate the psychometric characteristics of the Indonesian version of the Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) instrument using the Rasch Rating Scale model approach. The population in this study were Muslim women who wear the hijab. The characteristics of the population used were female students of UIN Syarif Hidayatullah Jakarta aged 17-25 years. Respondents actively use Instagram social media accounts. The selection of this age is in accordance with the developmental stages from late adolescence to early adulthood to Hurlock (1999). The sample in this study amounted to 335 people. The results of the analysis using the Rasch Rating Scale Model show that the psychometric characteristics of the IUBERQ in the Indonesian version have a good level of precision and good item suitability to the model. These findings support the use of IUBERQ as a reliable instrument in measuring aspects of Instagram use. Theoretical and practical implications and recommendations for further research are also discussed.

**Keywords:** instagram, psychometric properties, rasch model, social media platforms

### Abstrak

Instagram merupakan salah satu platform media sosial terpopuler di dunia, khususnya di kalangan remaja dan dewasa muda, yang memungkinkan penggunaannya untuk berbagi konten visual dan berinteraksi sosial secara daring (Ting, 2014). Seiring dengan meningkatnya dampak psikologis dan perilaku penggunaan Instagram, penting untuk menilai tidak hanya frekuensinya, tetapi juga perilaku dan respons emosional yang menyertainya melalui alat ukur yang valid dan reliabel seperti Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) (Paramboukis et al., 2016). Penelitian ini bertujuan untuk mengevaluasi karakteristik psikometrik instrumen Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) versi Indonesia dengan menggunakan pendekatan model Rasch Rating Scale. Populasi dalam penelitian ini adalah perempuan muslim yang berhijab. Karakteristik populasi yang digunakan adalah mahasiswa UIN Syarif Hidayatullah Jakarta yang berusia 17-25 tahun. Responden aktif menggunakan akun media sosial Instagram. Pemilihan usia ini sesuai dengan tahap perkembangan remaja akhir hingga dewasa awal dari Hurlock (1999). Sampel dalam penelitian ini berjumlah 335 orang. Hasil analisis menggunakan Rasch Rating Scale Model menunjukkan bahwa karakteristik psikometrik IUBERQ versi Indonesia memiliki tingkat presisi yang baik dan kesesuaian item terhadap model juga baik. Temuan ini mendukung penggunaan IUBERQ sebagai instrumen yang reliabel dalam mengukur aspek penggunaan Instagram. Implikasi teoritis dan praktis serta rekomendasi untuk penelitian selanjutnya juga dibahas.

**Kata kunci:** instagram, model rasch, platform media sosial, properti psikometrik

## Introduction

Instagram is a social media used worldwide. This application allows someone to share photos and videos with other users. Instagram was launched on October 6, 2010 and was acquired by Facebook in 2012 ([Ting, 2014](#)). In the sub-chapter entitled "Instagram" in the book "The SAGE Handbook of Research Methods" ([Sloan & Quan-Haase, 2017](#)) it is explained that the brief history of Instagram began when Kevin Systrom and Mike Krieger launched this application as a free application on the iPhone in 2010. Instagram provides three benefits to its users, namely (1) improving the appearance of ordinary photos by adding applications to "beautify" them (2) providing a place to share photos from various parts of the world directly (3) making it easier for users to upload directly ([Sloan & Quan-Haase, 2017](#)).

Social media's rapid development is considered an escape from boredom, filling free time and relaxing ([Casalo et al., 2017](#); [Agustiani & Gazi, 2021](#)). [Alhabash and Ma \(2017\)](#) stated that the eight most important aspects in using Instagram are all self-centered such as spending time, getting and sharing information, convenience, entertainment, self-expression, social interaction, and self-documentation. [Al-Kandari et al. \(2016\)](#) explained that someone can follow family, friends or strangers who show their daily life activities on Instagram social media. [Hu et al. \(2014\)](#) analyzed the selection of content and types of users on Instagram. Individuals upload a lot of selfie content to show themselves on Instagram ([Hu et al., 2014](#)). Research by [Jang et al. \(2015\)](#) showed that the use of likes, comments, and hashtags in adolescents shows self-expression in interacting on Instagram social media.

The importance of instruments that measure the construct of social media usage, especially Instagram, has grown widely along with the increasing role of social media in everyday life, especially among adolescents and young adults. One of the relevant instruments in measuring the use, behavior, and emotional reactions to Instagram usage is the Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) developed by ([Parambousis et al., 2016](#)). This instrument was developed based on the phenomenon that shows that the use of social media, including Instagram, not only reflects online activities but also affects the social behavior and psychological conditions of its users ([Hu et al., 2014](#); [Keles et al., 2020](#)).

Instagram as one of the most popular social media platforms, especially among students, has a significant influence on how individuals form self-image, establish social relationships, and process emotional experiences. Unlike other social media, Instagram emphasizes more expressive visual, aesthetic, and social connectedness aspects ([Sheldon & Bryant, 2016](#)). This makes users more susceptible to social comparison processes, seeking external validation, and mood swings due to online interactions ([Lup et al., 2015](#); [Huang, 2022](#)). Therefore, measuring Instagram usage is not enough to only look at the duration or frequency of use, but also includes the behavior and affective responses that accompany it ([Parambousis et al., 2016](#)).

Previous studies have used the IUBERQ to examine the relationship between Instagram use and mental health, psychological well-being, and self-image in adolescents and young adults ([Hu et al., 2014](#); [Mahmood & Malik, 2022](#)). However, no study has adapted and validated the IUBERQ specifically in Indonesia, especially in samples of students or college students who have different cultural characteristics, media habits, and social structures from the context in which the instrument was developed. For this group, Instagram use is not only a means of entertainment and communication but also a platform for identity expression, especially in terms of religiosity, modest fashion, and self-presentation in accordance with cultural and social expectations ([Casalo et al., 2017](#)). Muslim female students often navigate dual pressures; upholding religious values such as modesty while simultaneously engaging with contemporary digital culture, which makes the way they experience Instagram distinct from other populations ([Gulamhussein & Eaton, 2015](#)). Considering these cultural and gender-specific dynamics, it is crucial to ensure that the items of the IUBERQ function appropriately for this group. Furthermore, this study extends from a prior research framework that examined religiosity, peer attachment, Instagram use, and hedonistic behavior among female students wearing hijab ([Deviana, 2018](#)), where the IUBERQ was initially employed without modification by content (question statement). Thus, validating the instrument

with Muslim female college students ensures both cultural relevance and research continuity. Given the differences in cultural and social characteristics, local validation is needed so that this measuring instrument can be used appropriately and meaningfully ([Van de Vijver & Leung, 1997](#)).

On the other hand, the use of the Rasch model ([Rasch, 1960](#); [Wright, 1968](#)) in testing the validity of psychological instruments is now increasingly used because of its advantages in producing objective and stable measurements, through characteristics such as specific objectivity, parameter separation, and measurement invariance ([Fisher, 1987](#); [Bond & Fox, 2015](#); [Boone et al., 2014](#)). Although the IUBERQ was initially developed with a classical approach, no study has evaluated its psychometric characteristics using the Rasch Model, especially with student respondents in Indonesia. This is an important gap that needs to be filled so that this instrument is truly feasible and accurate for use in the local context ([Mair, 2018](#); [Linacre, 2002](#)). With this in mind, this study aims to evaluate the psychometric characteristics of the Indonesian version of the Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) using the Rasch Model. Evaluation is carried out on several important aspects, namely (1) unidimensionality, (2) local independence assumption, (3) item fit to the Rasch Model, (4) item and respondent reliability, (5) Wright map visualization, (6) test information function, and (7) diagnostic rating scale to assess the functioning and discrimination of each Likert scale category ([Bond & Fox, 2015](#); [Boone et al., 2014](#); [Linacre, 2012](#)). In addition, the analysis will include validation of the entire item as developed by [Pichardo et al. \(2018\)](#) to evaluate the psychometric quality of each dimension in the perspective of a modern measurement model.

## Methods

### Participants

The population in this study were Muslim women who wear the hijab, because this group represents a unique context where Instagram use is closely tied to identity expression, religiosity, and social expectations ([Ting, 2014](#)). Their experiences of self-presentation and social comparison may differ from other populations, making it important to validate the IUBERQ within this specific group, especially as the instrument was also employed in a previous undergraduate thesis examining religiosity, peer attachment, Instagram use, and hedonistic behavior among female students wearing the hijab ([Deviana et al., 2020](#)). The characteristics of the population used were female students of UIN Syarif Hidayatullah Jakarta aged 17-25 years. Respondents actively use Instagram social media accounts. The selection of this age is in accordance with the developmental stages of late adolescence to early adulthood from [Hurlock \(1999\)](#). The sample in this study amounted to 335 people. The sample size determination was based on the guidelines proposed by [Muthén and Muthén \(2002\)](#) in their article "How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power." According to the simulation results presented in **Table 1** of the article, for Confirmatory Factor Analysis (CFA) models with non-normal data and missing values, a minimum sample size of 315 participants is recommended to ensure adequate statistical power and stable parameter estimation. The sample size used in this study meets and exceeds this recommendation. Therefore, the determination of the sample size was based on Monte Carlo simulation guidelines rather than a general rule of thumb. In addition, the sampling method used in this study was non-probability sampling. This approach was chosen because the exact population size that met the study criteria was unknown. Consequently, 335 participants were recruited as the research sample. Data collection was conducted directly by the author across various faculties at UIN Syarif Hidayatullah Jakarta, ensuring that all participants met the predetermined population and sample criteria.

### Instruments

In this study, the measuring instrument used to measure Instagram is a modification of the Instagram usage, behavior, and emotional reaction (IUBRQ) measuring instrument from [Paramboukis et al \(2016\)](#). The adaptation process using [Beaton \(2000\)](#) procedure is carried out in four stages as follows: (1) Initial Translation. At this stage, the research instrument is translated into Indonesian. This translation process

is carried out by two people. The first translator has a background in English literature and works in the same field. The second translator is an expert in statistics and psychology. This aims to obtain comprehensive translation results because the results are a combination of linguists and experts who understand the concept of this scale. (2) Synthesis of Translations. After obtaining the results from translator one (P1) and translator two (P2), the research instrument is synthesized from the results obtained by P1 and P2. If differences are found between the two translation results, the items are selected which have the most appropriate meaning according to the initial scale. In this process, cultural factors are considered in choosing the translation results. (3) Back Translation. At this stage, a back translation is carried out into the original language of the scale. The back translation process was carried out by two translators (different from stage 1). This process was carried out to see if there were any differences in meaning when the scale in Indonesian was translated into the original language. If there were differences in meaning, the items were reviewed again. (4) Expert Committee. After correcting the translation by considering the results of the back translation, the research instrument was discussed with experts in statistics and experts in social psychology who understood the concept of the Instagram scale. The instrument measures Instagram usage in four dimensions, namely (1) Instagram Usage, which describes the use of Instagram which includes the intensity and quantity of the use of Instagram social media. (2) Instagram Behaviors, shows the activities carried out by someone in the form of likes, comments, uploaded content, and role models on Instagram accounts. (3) Instagram Attitudes, attitudes in assessing the importance of using Instagram in the form of likes, comments, uploaded content, and role models on Instagram accounts. (4) Instagram Emotional Reactions, describes a person's emotional reactions when using Instagram, such as giving emotional reactions when seeing other people's Instagram posts. This measuring tool for items 1 – 3 uses multiple choices, because it represents the use of Instagram which measures the intensity and amount of time spent uploading photos and videos with the same number of answer choices, namely 4 answer choices, and item 4 – item 16 uses a Likert scale with four answer choices, namely 1 = strongly disagree 4 = strongly agree (Paramboudis, et.al, 2016). Thus, because the answer choices are the same, the raw data results for questions 1-3 and 4-16 can be analyzed simultaneously using the Rasch Rating Scale Model (RSM) method which is in line with the RSM concept which requires the assumption that all questions have the same set of response categories. The blueprint for viewing the four dimensions of Instagram can be seen in **Table 1**.

**Table 1.** Blueprint Characteristics of using Instagram Social Media

No.	Aspects	Indicator	No. Item	Total
1.	<i>Instagram Usage</i>	<ul style="list-style-type: none"> <li>Length of time to access Instagram</li> <li>Number of uploads on Instagram</li> </ul>	1,2,3	3
2.	<i>Instagram Behavior</i>	<ul style="list-style-type: none"> <li>Show activity in the form of likes, comments, uploaded content, and role models on Instagram accounts.</li> </ul>	4,5,6,7,8	5
3.	<i>Instagram Attitudes</i>	<ul style="list-style-type: none"> <li>Assessing the importance of comments</li> <li>Assessing the importance of likes</li> <li>Assessing the importance of uploaded content, role models on Instagram</li> </ul>	9,10,11,12	4
4.	<i>Instagram Emotional Reactions</i>	<ul style="list-style-type: none"> <li>Give emotional reactions when seeing other people's posts</li> </ul>	13,14,15,16	4

Notes: Fav=*Favorable*; Unfav=*Unfavorable*.

## Application of the Rasch Rating Scale Model

The Rasch RSM is appropriate for assessing polytomous data using rating scales, such as those used in this study ([Andrich, 1978](#)). When applicable, response rating scales produce ordinal data that need to be converted to interval scales to be useful. A Rasch RSM generally follows the equation:

$$\log(P_{nik}/P_{ni(k-1)}) = B_n - D_i - F_k$$

Where  $P_{nik}$  is the probability that the  $n$ th person who takes item  $i$  will be “observed” in category  $k$ ,  $P_{ni(k-1)}$  is the probability that the  $n$ th person will choose category  $k - 1$ ,  $B_n$  is the trait level of the construct being measured (instagram usage level) of the  $n$ th person,  $D_i$  is the difficulty level of item  $i$  and  $F_k$  is the probability that category  $k$  will be chosen depending on category  $k - 1$ . In this study, estimates of item difficulty ( $D_i$ ) and respondent instagram usage level ( $B_n$ ) are expressed on a logit scale ([Linacre, 2018](#)). The mean of the logit is randomly set at 0, with a positive log indicating higher than the estimated mean and a negative log indicating lower than the estimated mean. The threshold indicates the location on the trait level scale where the respondent has a 50/50 chance of choosing a higher category than the current one ([Luo et al., 2009](#)). The Rasch Rating Scale Model requires tenable assumptions for accurate estimation, including (1) Construct Unidimensionality, (2) Local Independence, and (3) Monotonic scale (i.e., higher scores refer to higher levels of the latent construct) ([de Ayala, 2009](#)). In addition, the RSM the Rasch Rating Scale Model requires that rating scale categories increase with endorsement difficulty, and that thresholds for each item are ordered ([DiStefano & Morgan, 2010](#)).

## Data Analysis

In this study, data analysis was conducted using the RSM. The software used was the Winsteps (3.65) statistical package ([Linacre, 2008](#)) to test the validity of the Instagram usage, behavior, and emotional reactions questionnaire (IUBERQ) instrument. Person and item parameters were estimated using Joint maximum likelihood estimation (JMLE). To obtain information about the psychometric characteristics of each subscale, the reported analysis: (1) To check the expected unidimensionality in the RSM with the concurrent calibration method, calibrating all items on the instrument used in this study is the IUBERQ instrument without looking at the aspects measured as seen from the [Pichardo et al \(2018\)](#) procedure. (2) Testing the local independence assumption using the Q3 statistic, (3) Testing fit items with the Rasch model, considering the mean square statistic (MNSQ) when using the Rasch RSM; (4) Reliability testing for people and items; (5) Displaying Wright map to find out the measurement results of the instrument used in this study is the IUBERQ instrument by comparing people and items on the same scale; (6) Test Information Function to find out the function of the test when given to individuals with the level of trait obtained. (7) Testing Rating Scale diagnostics to determine the function of each category and the discrimination of their rating scales.

## Results and Discussion

### Results

#### *Unidimensionality*

The unidimensionality assumption test of the instrument was conducted using PCAR (Principal Component Analysis Residual) ([Chou & Wang, 2010](#); [Smith, 2002](#)). The criteria used to test unidimensionality in a construct being tested, including each aspect that measures social support, is seen when the raw variance explained by measures is  $> 30\%$  ([Linacre, 1998](#)), then it can be said that each aspect of the instagram usage construct is unidimensional. The results of the analysis of (1) First, concurrent calibration on the instrument used in this study is the IUBERQ instrument, which means that the Instagram construct is measured with 16 items regardless of the aspects contained there in, are listed in **Table 2**.

**Table 2.** Variances of standardized residuals (IUBERQ) instrument – N 335

	Eigenvalues	Observed (%)
<b>Instagram – initial analysis</b>		
Raw variance explained by measure =	12.7	44.2%

The results show that the instrument used in this study is the IUBERQ that measures the construct of characteristics of Instagram social media usage obtained a raw variance explained by measures of 12.7 in eigenvalues units which in percentage size is 44.2%. This means that with a raw variance explained by measures of 44.2% ( $> 30\%$ ), the 16 items that measure the characteristics of Instagram social media usage are unidimensional. This means that the 16 items from the Instagram social media usage characteristics construct meet the unidimensionality assumption.

#### *Local Independence*

In this study, the application of the Rasch rating scale model is based on the second assumption, namely local independence. In the assumption of local independence, a particular student or students from a given trait level, their work on one item does not depend on their work on other items (Mair, 2018). After the unidimensionality assumption has been proven to be met, the local independence assumption is tested using the Q3 statistic (Yen, 1984). When using the Q3 statistical index criteria where it is determined that the residual correlation between pairs of items is never  $> .30$  (Christensen et al., 2017; Das Nair, Moreton, & Lincoln, 2011) it means that no items are found to have local independence.

The results of the analysis of the Instagram usage, behavior, and emotional reactions questionnaire (IUBERQ) instrument are shown in **Table 4**.

**Table 4.** Identify Locally Dependent Items (IUBERQ) Instrument – N 335

	Highest Raw Residual Correlation	Pasangan Item
<b>Instagram</b>	.67	Item 2 dan Item 3

The results of the analysis show that the items with the highest residual correlation (raw residual correlation) are item pair 2 and item 3 of .67 with a value of  $> .30$ . This means that items with a residual correlation of .67 ( $> .30$ ) do not meet the assumption of local independence. The findings obtained are that item pair 2 which reads "I upload as many photos as..." is related to item 3 which reads "I upload as many videos as...". This means that when someone gets a high trait level to answer item 2, then the estimate of their ability also depends on when the individual answers item 3. Both measure very similar or even the same aspects in the construct to be measured, so that there is information redundancy between the two items.

Following up on the finding of a violation of the local independence assumption between Items 2 and 3, the researchers conducted a series of further analyses to determine the most appropriate handling strategy. This is important because if the violation is ignored, the validity of the measurement can be compromised (Marais & Andrich, 2008; Linacre, 1998). Several alternatives considered include: (1) deleting one of the problematic items, (2) combining both items into a single testlet (Wainer & Kiely, 1987; Yen, 1993), and (3) comparing the results of the analysis after these treatments against the fulfillment of the unidimensionality assumption, residual correlation, and the appropriateness of the item fit (Linacre, 2002). The following **Table 5** summarizes the results of the analysis experiments conducted under these four conditions.

**Table 5.** Results of Follow up Analyses to Address Local Independence Violation

Analysis conditions	Unidimensionality (% variance explained by measure)	Residual Correlation (Items 2–3)	Item Fit (INFIT/OUTFIT MNSQ)	Fit	Conceptual notes
(1) Item 2 removed	48.3%	< .30	Item 3 not meet the acceptable infit–outfit MNSQ criteria		Removing item 2 reduces local dependence, but item 3 becomes misfitting, suggesting potential construct underrepresentation.
(2) Item 3 removed	48.3%	< .30	Item 2 did not meet the acceptable infit–outfit MNSQ criteria		Removing item 3 eliminates residual correlation, but item 2 misfits, indicating instability and measurement distortion.
(3) Testlet (sum of Item 2 & 3)	57%	< .30	The testlet (item 2 and item 3) showed misfit (infit–outfit MNSQ > 1.5)		Combining items into a testlet improves unidimensionality and resolves local dependence, but some misfit remains at the testlet level.
(4) Both Item 2 & 3 removed	43.2%	< .30	Item 1 displayed misfit according to the infit–outfit MNSQ criteria		Removing both items weakens construct coverage and decreases explained variance, indicating a loss of essential content representation.

Based on the four analytical conditions, the most appropriate decision is to combine Item 2 and Item 3 into a single testlet ([Wang & Wilson., 2005](#); [Wainer & Kiely, 1987](#); [Yen, 1993](#)). This option yielded the highest unidimensionality (57%) while resolving the local independence issue, indicating that treating the two items as part of the same sub-construct provides the most stable measurement structure. Although the testlet displayed some misfit, this trade-off was considered more acceptable than the substantial decrease in unidimensionality observed when removing one or both items. Thus, the results of the analysis in other explanations will follow the results of the analysis of the testlet conditions applied.

#### *Item Fit*

Item fit testing such as infit and outfit statistics MNSQ can also be used to determine how well each item measures a construct as evidence that the construct is unidimensional. The MNSQ infit and outfit values are considered fit according to the Rasch Rating Scale Model with a range of .5 - 1.5 effective for a measurement ([Andrich & Marais, 2019](#); [Bond & Fox, 2015](#)). **Table 6a** presents the item fit statistics of

the Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) for the initial analysis, while **Table 6b** displays the results after applying the testlet approach (Condition 3). The item measure values (logits) reflect the relative difficulty of endorsing each item in relation to the latent trait of Instagram usage. Items with higher logit measures are more difficult to endorse, meaning that only individuals with higher levels of the underlying Instagram usage trait are likely to agree with them. Conversely, items with lower logit measures are easier, indicating that even individuals with lower trait levels tend to endorse them.

**Table 6a.** Rasch Item Statistics of the Instagram Usage, Behavior, and Emotional Reactions (IUBERQ – Item Measure Order – N335 – Initial Analysis

Item	Measure	Infit	Outfit	PTMEA	Interpretation
5	1.60	1.02	1.02	.36	V
2	1.42	1.16	2.32	.31	X
3	1.39	1.14	1.34	.31	V
9	.84	.89	.86	.53	V
10	.76	.79	.78	.62	V
4	.51	1.02	1.03	.42	V
12	.18	.84	.83	.59	V
1	.16	1.36	1.41	.20	V
7	.05	.94	.95	.52	V
8	-.03	.90	.89	.56	V
6	-.60	.93	.90	.50	V
16	-.63	1.01	.99	.46	V
15	-.65	1.18	1.22	.29	V
14	-1.44	1.05	1.05	.35	V
11	-1.47	.97	.97	.47	V
13	-2.09	.97	.97	.40	V

Notes: V, Fit; X, Not Fit.

In the initial analysis (**Table 6a**), Item 5 (logit = 1.60) was the most difficult item, suggesting that agreement with this statement requires a higher level of Instagram usage behavior. Meanwhile, Item 13 (logit = -2.09) was the easiest item, indicating that respondents across different trait levels tended to agree with it. After applying the testlet method (**Table 6b**), the overall item hierarchy remained stable, with Item 5 remaining the most difficult and Item 13 the easiest.

**Table 6b.** Rasch Item Statistics of the Instagram Usage, Behavior, and Emotional Reactions (IUBERQ – Item Measure Order – N335 – Analysis Condition (3))

Item	Measure	Infit	Outfit	PTMEA	Ket.
5	1.69	1.01	1.01	.36	V
2*3	1.26	1.56	3.86	.33	X
9	.94	.89	.85	.53	V
10	.86	.79	.78	.62	V
4	.61	1.02	1.03	.42	V
12	.29	.83	.83	.59	V
1	.28	1.34	1.39	.21	V
7	.17	.93	.94	.52	V
8	.08	.89	.88	.56	V
6	-.49	.92	.90	.50	V
16	-.51	1.01	.99	.46	V
15	-.54	1.17	1.20	.29	V
14	-1.33	1.05	1.04	.36	V
11	-1.35	.96	.96	.47	V
13	-1.97	.97	.97	.39	V

Notes: V, Fit; X, Not Fit.

All items, except Item 2 in the initial analysis and item testlet (2\*3), demonstrated acceptable infit and outfit MNSQ values within the recommended range (.5–1.5), indicating good model–data fit. This suggests that the items functioned well to measure a single latent construct. Interpreting these item difficulties within the construct of Instagram usage, individuals with higher trait levels are more likely to engage in more frequent or intense Instagram behaviors (e.g., uploading large amounts of content), whereas individuals with lower trait levels tend to agree more with general or less demanding behaviors.

#### *Reliability of person and items*

In the RSM analysis, reliability is estimated for both persons and items. The reliability criteria used in this study follow the commonly accepted cutoff of  $> .70$ , which indicates good internal consistency of the instrument. In addition, person and item separation indices are evaluated, where values greater than 2.5 are considered sufficient to support comparative analyses at the group level (Tennant & Conaghan, 2007).

The results of the initial analysis of the Instagram Usage, Behavior, and Emotional Reactions Questionnaire (IUBERQ) showed that the item reliability was .99 with an item separation index of 12.19, while the person reliability was .67 with a person separation index of 1.42. Under the analysis condition (3), the item reliability remained at .99 with an item separation index of 12.18, and the person reliability increased slightly to .73 with a person separation index of 1.63.

These findings indicate that the item separation reliability (ISR) is very high, meaning that the instrument contains items with a wide range of difficulties, allowing the Rasch model to distinguish multiple levels of item difficulty reliably. In other words, the items are well-targeted to the construct and can be consistently located on the measurement continuum. Meanwhile, a moderate person separation reliability (PSR) indicates that the instrument can distinguish between groups of respondents with different levels of the latent trait, although it may not be optimal for making fine-grained distinctions at

the individual level. Overall, these results suggest that the IUBERQ instrument demonstrates good measurement precision at the item level and acceptable precision at the person level.

### *Wright Map*

The RSM analysis establishes construct validity according to the hierarchy of items that can be observed in the Wright Map (Pichardo, 2018). This map depicts item difficulty on the right and person ability on the left. Items must form a continuum scale, where items with a low level of difficulty are located at the bottom, items with medium difficulty are in the middle, and items with a high level of difficulty are at the top. Likewise, persons are distributed according to their level of the measured attribute, in this study, IUBERQ.

The results of concurrent calibration on the IUBERQ instrument are presented in Figures 1a and 1b. From both figures, it is clear that the most difficult item to agree with is Item 5 (*"Saya mengunggah barang-barang yang baru saja saya beli."*), whereas the easiest item is Item 13 (*"Saya senang mendapatkan komentar positif dari orang lain."*). In Figure 1a, the average person ability is  $-.14$  logit ( $SD = .73$ ), which is lower than the average item difficulty set at zero. Similarly, in Figure 1b, the average person ability is  $-.03$  logit ( $SD = .72$ ), also below the average item difficulty level. This indicates that, on average, respondents have a lower level of the measured trait compared to the difficulty level of the items. The range of person distribution in Figure 1a ( $2.37$  to  $-2.81$  logits) is wider than the range of item distribution ( $1.60$  to  $-2.09$  logits). Likewise, in Figure 1b, the person distribution ( $2.38$  to  $-2.68$  logits) is also broader than the item distribution ( $1.69$  to  $-1.97$  logits).

Based on these statistics, the instrument can be classified as an off target test. An on target test is characterized by the alignment between the mean person ability and the mean item difficulty (ideally both centered around zero), indicating that the test items are well-targeted to the respondents' ability levels. Conversely, in an off target test, the mean person ability and mean item difficulty are not aligned either the test is too easy or too difficult for the respondents. In this study, the negative mean person measure values indicate that the items are slightly more difficult than the respondents' average ability levels. This misalignment has practical implications which will be explained in the discussion section.

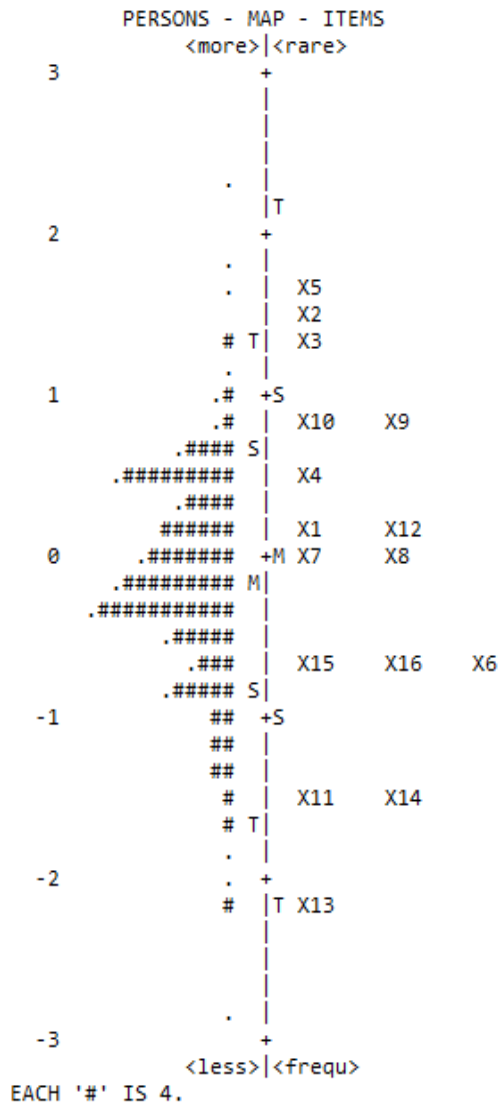


Figure 1a. initial analysis

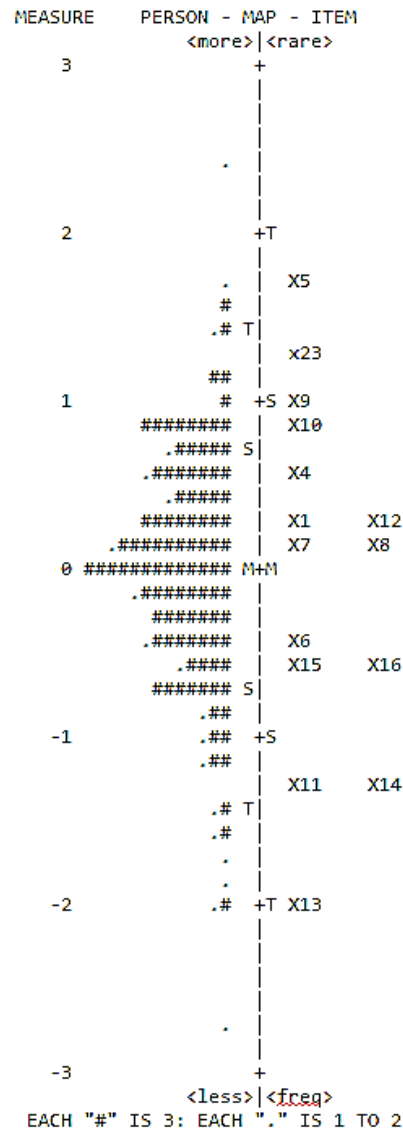
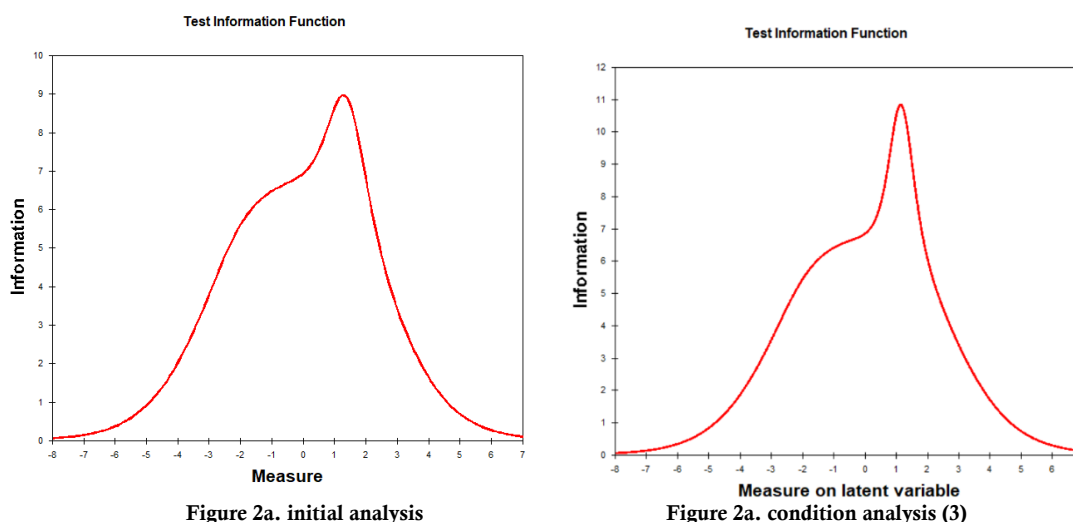


Figure 1b. condition analysis (3)

### Test Information Function

Test Information Function (TIF) shows the function of the test when given to individuals with the level of trait obtained. The trait in question is the items that measure the characteristic construct of Instagram social media use obtained by individuals tend to be low, moderate, or, up to high. The better the item is targeted at the person, the more Fisher information the item provides about the person's parameters. The expected test information function peaks, where the reference test criterion being tested will be obtained, and where the reference test sample model shows normality. With this, TIF is the effective test measurement range (Linacre, 2018).



[TIF. Instagram Instrument (IUBERQ) – N 335]

The results of the TIF analysis show that the Instagram usage, behavior, and emotional reactions scale that measures the Instagram construct has information about the functioning of the test at a relatively high level, and the trait size is also high. This means that information about the test on a set of items that measure the Instagram construct is very appropriate and optimal for people who have relatively high abilities. As illustrated in Figure 2a, the curve reflects the amount of information, or measurement precision, that the instrument provides at each ability level (logit). The peak of the curve indicates the ability level at which the instrument functions most effectively.

In the initial analysis (Figure 2a, left), the TIF reaches its maximum at approximately +1.5 logits, indicating that the instrument provides the greatest measurement precision for individuals with moderately high levels of the Instagram usage trait. Similarly, in the testlet condition analysis (Figure 2a, right), the maximum test information is observed at around +1.8 logits, demonstrating that the instrument continues to perform optimally at higher levels of the latent construct.

Overall, these findings indicate that the IUBERQ instrument is most precise for respondents with moderate to high levels of the Instagram usage trait, while the amount of information decreases for individuals with lower trait levels. This pattern suggests that the instrument is particularly well-targeted for populations with higher engagement levels on Instagram, whereas its precision is comparatively lower for those with less frequent usage.

#### *Rating Scale Diagnostics*

The rating scale in this study uses a Likert scale. The importance of information about the rating scale of an instrument is needed to understand the responses given by respondents who use the rating scale and determine the actual distance that applies to individuals in taking the existing response categories. Diagnostics with the rating scale model in Rasch measurement is used to evaluate how well the response set of the Instagram usage, behavior, and emotional reactions questionnaire functions to create interpretable measures (Kim & Kyllonen, 2006). The 4 categories of rating scales in this study are strongly disagree to strongly agree. The results of the analysis are shown in Table 9.

**Table 9.** Rating Scale Diagnostics Concurrent Calibration Instagram Instrument (IUBERQ) – N 335

Category	Threshold	Observed Count (%)	Observed Average	Infit	Outfit
<b>Initial analysis</b>					
Strongly Disagree	None	62 (19)	-.41	1.39	1.46
Disagree	-1.08	113 (34)	-.25	1.16	1.36
Agree	-.23	131 (39)	.10	1.01	1.00
Strongly Agree	1.80	29 (9)	-.22	1.74	1.76
<b>Condition analysis (3)</b>					
Strongly Disagree	None	62 (19)	-.30	1.38	1.44
Disagree	-.96	113 (34)	-.14	1.13	1.30
Agree	-.11	131 (39)	.21	1.01	.99
Strongly Agree	1.89	29 (9)	-.11	1.70	1.72

Note: Threshold: Rasch-Andrich thresholds: The relationship between adjacent categories, and correspond to the point where the probability curves of adjacent categories intersect.; Observed count and %: The number of occurrences of a category.; Observed average: The average ability of a person to answer each category; Infit & Outfit: a mean square statistic. The expected value is 1. A value of .5-1.5 is a good measure.

The rating scale diagnostics examine the functioning of each response category to ensure that they contribute meaningfully to measurement precision. According to [Linacre \(2010\)](#), the acceptable range for Infit and Outfit mean-square (MNSQ) statistics is typically .5 to 1.5, with values exceeding 1.5 indicating category misfit. The table clearly shows that the “strongly agree” category has an infit of 1.74 and an outfit of 1.76 (initial analysis), infit of 1.70 and an outfit of 1.72 (condition analysis 3), both of which exceed the upper threshold of 1.5. This indicates a poor fit for this specific category, meaning that responses in this category do not align well with the expected measurement model. Given these results, the most appropriate step is to collapse the “agree” and “strongly agree” categories. Collapsing categories is a recommended approach when extreme categories do not function well ([Linacre, 2002](#); [Andrich, 1996](#)). This adjustment can improve category functionality, reduce response noise, and improve model fit. Alternatively, future research could consider revising the wording of the “strongly agree” option to make it more discriminatory, or further examine whether extreme agreement reflects variance unrelated to the intended construct.

Substantially, the discrepancy in the “strongly agree” category suggests that respondents may not understand this category consistently. In the context of Instagram use, people who select “strongly agree” do not always come from groups with truly high levels of Instagram use. Some respondents may indeed have high levels of use, but others may choose this answer out of a desire to appear socially desirable or because they tend to over-endorse. As a result, the “strongly agree” category cannot effectively differentiate between respondents with truly high levels of Instagram usage and those who do not, thus reducing the accuracy of the scale's measurement at the upper end of the continuum construct.

In contrast, the categories “strongly disagree,” “disagree,” and “agree” have Infit and Outfit MNSQ values within the acceptable range ( $\leq 1.5$ ), indicating that these categories function well and provide accurate measurement information. In other words, these categories can differentiate the level of Instagram usage precisely among respondents and contribute to increasing the overall measurement accuracy of the instrument.

## Conclusion

In this study, the Rasch Rating Scale model analysis aims to validate the Instagram usage, behavior, and emotional reactions questionnaire instrument developed from the theoretical model proposed by [Paramboulis et al., \(2016\)](#). With the advantages of the Rasch rating scale model analysis in this study, an analysis was carried out with concurrent calibration. The results of the validity test using Rasch RSM showed that the IUBERQ instrument was proven to have a unidimensional factor structure. An interesting finding is when the unidimensional assumption was reported to be met based on the results of the analysis of all items (16 items), meaning that the Instagram usage construct had good results regardless of the fact that these aspects also measured the construct called Instagram (namely, usage, behavior, and emotional reactions). Then, the Instagram instrument also had reliability for items and persons that fell into the good category in both calibrations carried out. Even a perfect unidimensional scale would not be useful in practice if the resulting scale scores had very low reliability ([Gerbing & Anderson, 1988](#)); acceptable Rasch reliability and the IUBERQ instrument could produce scores that were useful in practice. From the results, it was found that the items that were not fit to measure IUBERQ were item 2, with statement 2. I upload as many photos as: a. < 2 photos/week; b. 2-4 photos/week; c. 5-7 photos/week; d. > 7 photos/week, and item 3, with statement 3, I upload as many videos as: a. < 2 videos / month. b. 2-4 videos / month, c. 5-7 videos / month, d. > 7 videos / month it is said to be underfit (inconsistent). This means that values that are outside these limits indicate a lack of conformity between the item and the person to the model ([Linacre, 2018](#)). Outfit values are sensitive to unexpected response patterns that are far from the person's "ability or location" ([Linacre, 2018](#); [Wang & Wilson, 2005](#)). If a item and person has a large Outfit, the person is indicated to have aberrant response behavior ([Karabatsos, 2003](#)). Thus, 14 items were obtained that can be used to measure individual activities when using Instagram according to the findings in the results (see **Table. 6**).

In addition, this study also conducted a third conditional analysis using a testlet approach to examine the assumption of local independence by combining items 2 and 3 into a single unit of analysis ([Wang & Wilson, 2005](#)). This approach was chosen because both items were considered to remain substantively important and content-related, so combining them into a testlet could address local dependencies without eliminating either item. The analysis results showed that the testlet condition provided relatively better results than the other analysis conditions, particularly in terms of increasing unidimensionality and reliability values. However, this finding did not significantly impact the item fit results, as in all analysis conditions, items 2 and 3 continued to show a misfit with the model, indicating that these two items were less representative in describing the construct measured by the IUBERQ.

In the Wright map results, a negative average value for each individual indicates that the items are slightly more difficult than the respondent's average ability level. This misalignment has practical implications; (1) respondents with lower ability levels may find several items too challenging, which could reduce measurement precision at lower levels of the trait. (2) the targeting mismatch may also affect the reliability and validity of score interpretations, particularly for individuals at the extremes of the distribution. For future use, adding items easier may help improve the match between item difficulty and person ability, thereby increasing test targeting and measurement precision.

The limitation of this study is that the population involved is specifically for Muslim women who wear the hijab who are active on social media on Instagram. Because initially, this instrument was used to answer research questions in the research of [Deviana et al., \(2020\)](#), which is associated with hedonic behavior. However, the use of the Rasch Rating Scale Model analysis has provided a solution in testing the validity related to the construct of social media use such as Instagram which is the novelty of the application of the Rasch Model, because the methodology does not depend on the sampling involved. Second, further research is needed to develop the use of Instagram with new features at this time. In short, the validation of the IUBERQ instrument by applying Rasch RSM helps to confirm the accuracy and usefulness of the scores on the instrument, because all individuals have their own standard errors, and provides this scale with additional psychometric support in the application of Rasch. The findings of

this study provide information for students, teachers, lecturers, and researchers, especially in the field of social psychology and their interest in technology and social media. With the IUBERQ instrument which was tested specifically in understanding how Instagram usage behavior can affect their emotional reactions and social dynamics, especially among students. This instrument allows the identification of social media usage patterns that have the potential to have positive or negative impacts on psychosocial well-being, such as feelings of being valued, socially accepted, to feelings of anxiety or social pressure due to social comparison. Furthermore, IUBERQ can be used by researchers, educators, and psychology practitioners to develop data-based interventions to reduce the negative impacts of social media use, encourage more conscious and healthy social media use, integrate assessment results into digital literacy programs or character education, especially for adolescents and young adults. In other words, IUBERQ is an important tool that can bridge the understanding between technology, user behavior, and its impact on the social and emotional aspects of individuals in the Indonesian context. Thus, this measuring instrument can be accessed through Appendix A and is expected to be a reference for further research and application in various related fields.

## References

- Agustiani, S., & Gazi, G. (2021). Pengaruh dukungan sosial dan perbandingan sosial terhadap kesejahteraan subjektif ibu muda pengguna media sosial. *Tazkiya: Journal of Psychology*, 9(2), 122–132. <https://doi.org/10.15408/tazkiya.v9i2.1754>
- Alhabash, S., & MA. (2017). A tale of four platforms: Motivations and uses of facebook, twitter, instagram, and snapchat among college students? *Journals.sagepub.com/home/sns*.
- Al-Kandari, A. J., Al-Hunaiyyan, A. A., & Al-Hajri, R. (2016). The influence of culture on Instagram use. *Journal of Advances in Information Technology*, 7(1), 54-57. <https://doi.org/10.12720/jait.7.1.54-57>
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561–573. <https://doi.org/10.1007/BF02293814>
- Andrich, D. (1996). Category ordering and their utility. *Rasch Measurement Transactions*, 9(4), Article 464.
- Andrich, D., & Marais, I. (2019). *A course in rasch measurement theory: Measuring in the educational, social and health sciences*. Springer Nature Singapore Pte. Ltd.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Bond, T., & Fox, C. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Lawrence Erlbaum & Associates.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. Springer. <http://dx.doi.org/10.1007/978-94-007-6857-4>
- Casaló, L. V., Flavián, C., & Ibáñez-Sánchez, S. (2017). Understanding consumer interaction on Instagram: The role of satisfaction, hedonism, and content characteristics. *Cyberpsychology, Behavior, and Social Networking*, 20(6), 369-375. <https://doi.org/10.1089/cyber.2016.0360>
- Chou, Y. T., & Wang, W. C. (2010). Checking dimensionality in item response models with principal component analysis on standardized residuals. *Educational and Psychological Measurement*, 70, 717–731. <https://doi.org/10.1177/0013164410379322>
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's  $Q_3$ : Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement*, 41(3), 178–194. <https://doi.org/10.1177/0146621616677520>

- Deviana, T. (2018). *Pengaruh religiusitas, peer attachment, dan karakteristik penggunaan media sosial Instagram terhadap perilaku hedonis mahasiswi* (Skripsi Sarjana, Universitas Islam Negeri Syarif Hidayatullah Jakarta). Retrieved from <https://repository.uinjkt.ac.id/dspace/bitstream/123456789/44510/1/TINA%20DEVIANA-FPSI.pdf>
- Deviana, T., Tresniasari, N., & Latifa, R. (2020). Effects of religiosity, peer attachment, characteristics of Instagram social media usage on hedonic behavior. *Proceedings of the 1st International Conference on Religion and Mental Health (ICRMH 2019)*, 351–364 <https://doi.org/10.4108/eai.18-9-2019.2293450>
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Press.
- DiStefano, C., & Morgan, G. B. (2010). Evaluation of the BESS TRS-CA using the Rasch rating scale model. *School Psychology Quarterly*, 25, 202–212. <https://doi.org/10.1037/a0021509>
- Fisher, R. P., Geiselman, R. E., Raymond, D. S., Jurkevich, L. M., et al. (1987). Enhancing enhanced eyewitness memory: Refining the cognitive interview. *Journal of Police Science & Administration*, 15(4), 291–297.
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25, 186–192. <https://doi.org/10.1177/002224378802500207>
- Gulamhussein, Q., & Eaton, N. R. (2015). Hijab, religiosity, and psychological wellbeing of Muslim women in the United States. *Journal of Muslim Mental Health*, 9(2), 25–40. <https://psycnet.apa.org/doi/10.3998/jmmh.10381607.0009.202>
- Hurlock, E. B. (1999). *Psikologi perkembangan: Suatu pendekatan sepanjang rentang kehidupan* (Edisi ke-5; Istiwidayanti, Penerj.). Erlangga.
- Hu, Y. L. & Manikonda L., Khambapati S. (2014). What we Instagram: A first analysis Instagram photo content and user types. *Proceeding of the Eighth International AAAI Conference on Weblogs and Social Media*, 1–4. [<https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewFile/8118/8087>].
- Huang, C. (2022). A meta-analysis of the problematic social media use and mental health. *International Journal of Social Psychiatry*, 68(1), 12–33. <https://doi.org/10.1177/0020764020978434>
- Jang, W, E., Ko, Y, J., Morris, J & Chang, Y. (2015). Scarcity message effects on consumption behavior: Limited edition product considerations. *Psychology & Marketing*. 32.(10), 989-1001. <https://doi.org/10.1002/mar.20836>.
- Karabatsos, G. (2003). Comparing the Aberrant response detection performance of thirty-six person-fit statistics. *Applied Measurement in Education*, 16(4), 277–298. [https://doi.org/10.1207/S15324818AME1604\\_2](https://doi.org/10.1207/S15324818AME1604_2)
- Keles, B., McCrae, N., & Grealish, A. (2020). A systematic review: The influence of social media on depression, anxiety and psychological distress in adolescents. *International Journal of Adolescence and Youth*, 25(1), 79–93. <https://doi.org/10.1080/02673843.2019.1590851>
- Kim, S., & Kyllonen, P. C. (2006). *Rasch rating scale modeling of data from the standardized letter of recommendation* (ETS Research Report RR-06 33). Princeton, NJ: Educational Testing Service
- Linacre, J. M. (1998). Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement*, 2, 266–283.

- Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85–106.
- Linacre, J. (2002) What Do Infit and Outfit, Mean-Square and Standardized Mean? *Rasch Measurement Transactions*, 16, 878.
- Linacre, J. M. (2008). *WINSTEPS Rasch measurement computer program* [Version 3.65.0]. Winsteps.com.
- Linacre, J. M. (2010). Predicting responses from Rasch measures. *Journal of Applied Measurement*, 11(1), 1–10.
- Linacre, J. M. (2018). *Winsteps® Rasch measurement computer program: User's guide*. Winsteps.com
- Lup, K., Trub, L., & Rosenthal, L. (2015). Instagram #instasad?: Exploring associations among Instagram use, depressive symptoms, negative social comparison, and strangers followed. *Cyberpsychology, Behavior, and Social Networking*, 18(5), 247-252. <https://doi.org/10.1089/cyber.2014.0560>.
- Luo, X., Cappelleri, J. C., Cella, D., Li, J. Z., Charbonneau, C., Kim, S. T., ... Motzer, R. J. (2009). Using the rasch model to validate and enhance the interpretation of the functional assessment of cancer therapy–kidney symptom index—disease-related symptoms scale. *Value In Health*, 12, 580–586. <https://doi.org/10.1111/j.1524-4733.2008.00473.x>
- Mahmood, S., & Malik, F. (2022). Impact of Instagram usage on young women's body esteem: Mediating role of beauty perception. *Pakistan Journal of Psychological Research*, 37(3), 351–364. <https://doi.org/10.33824/PJPR.2022.37.3.21>
- Mair, P. (2018). *Modern psychometrics with R*. Cham. Springer International Publishing AG.
- Marais, I., & Andrich, D. (2008). Effects of varying magnitudes of local dependence in the unidimensional Rasch model. *Journal of Applied Measurement*, 9(2), 105–124.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4), 599–620. [https://doi.org/10.1207/S15328007SEM0904\\_8](https://doi.org/10.1207/S15328007SEM0904_8)
- Paramboulakis, O., Skues, J., & Wise, L. (2016). An exploratory study of the relationships between narcissism, self-esteem and Instagram use. *Social Networking*, 5, 82-92. <http://dx.doi.org/10.4236/sn.2016.52009>
- Pichardo, M. C., Cano, F., Garzón-Umerenkova, A., de la Fuente, J., Peralta-Sánchez, F. J., & Amate-Romera, J. (2018). Self-Regulation Questionnaire (SRQ) in Spanish adolescents: Factor structure and Rasch analysis. *Frontiers in Psychology*, 9, 1370. <https://doi.org/10.3389/fpsyg.2018.01370>
- Rasch, G. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests*. Nielsen & Lydiche.
- Smith, E. V., Jr. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, 3(2), 205–231.
- Sheldon, P., & Bryant, K. (2016). Instagram: Motives for its use and relationship to narcissism and contextual age. *Computers in Human Behavior*, 58, 89–97. <https://psycnet.apa.org/doi/10.1016/j.chb.2015.12.059>
- Sloan, L., & Quan-Haase, A. (Eds.). (2017). *The SAGE Handbook of Social Media Research Methods*. SAGE Publications.

- 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 Rheum.*57(8):1358-62. [https://doi.org/ 10.1002/art.23108](https://doi.org/10.1002/art.23108)
- Ting, C. C. (2014). A study of motives, usage, self-presentation and number of followers on Instagram. *Discovery – SS Student E-Journal*, 3, 1-35.
- Van de Vijver, F., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Sage Publications.
- Wang, W. C., & Wilson, M. (2005). The Rasch testlet model. *Applied Psychological Measurement*, 29(2), 126–149. <https://doi.org/10.1177/0146621604271053>
- Wainer, H., & Kiely, G. L. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Educational Measurement*, 24(3), 185–201. <https://doi.org/10.1111/j.1745-3984.1987.tb00274.x>
- Wright, B. D. (1968). Sample-free test calibration and person measurement. In *Proceedings of the 1967 Invitational Conference on Testing Problems* (pp. 85–101). Educational Testing Service.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8, 125–145. <https://doi.org/10.1177/014662168400800201>
- Yen, W. M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement*, 30(3), 187–213. <https://doi.org/10.1111/j.1745-3984.1993.tb00423.x>

## Appendix A

### IUBERQ Research Instrument

#### Tina Deviana, et. al (2025)

Berilah tanda **SILANG** [X] pada pilihan jawaban yang sesuai dengan diri Anda dalam menggunakan Media Sosial Instagram.

1. Waktu yang Anda habiskan dalam menggunakan Instagram sebanyak :
  - a. <1 Jam/hari
  - b. 1-2 Jam/hari
  - c. 3-6 Jam/hari
  - d. Lainnya .....
2. Saya mengunggah foto sebanyak :
  - a. < 2 foto/minggu
  - b. 2-4 foto/minggu
  - c. 5-7 foto/minggu
  - d. > 7 foto/minggu
3. Saya mengunggah video sebanyak :
  - a. < 2 video/bulan
  - b. 2-4 video/bulan
  - c. 5-7 video/bulan
  - d. > 7 video/bulan

Pernyataan di bawah ini di isi sesuai **Petunjuk Pengisian** dengan empat pilihan jawaban berikut :

- **STS** : Sangat Tidak Setuju
- **TS** : Tidak Setuju
- **S** : Setuju
- **SS** : Sangat Setuju

NO	PERNYATAAN	STS	TS	S	SS
4.	Saya bercerita di unggahan foto Instagram saya.				
5.	Saya mengunggah barang-barang yang baru saja saya beli.				
6.	Saya menikmati unggahan foto/video dalam feed Instagram saya.				
7.	Saya memperhatikan komentar negatif orang lain pada post Instagram saya.				
8.	Saya memiliki <i>role model</i> pada akun Instagram saya.				
9.	Like dan Hastag sangat menunjang popularitas saya di Instagram.				
10.	Komentar orang lain menunjang popularitas saya di Instagram.				
11.	Penting dalam memilih foto/video yang diunggah pada akun Instagram saya.				
12.	Memiliki <i>role model</i> di Instagram membuat hidup saya lebih baik.				
13.	Saya senang mendapatkan komentar positif dari orang lain.				
14.	Saya suka memberikan <i>like</i> dan komentar pada akun Instagram orang lain.				
15.	Saya tidak suka melihat post Instagram orang lain yang isinya tidak menarik.				
16.	Saya senang berkomentar pada post Instagram orang lain yang isinya menarik.				

\*\*\*\*