

The Indonesian Version of Digital Well-Being Scale: Validation Study using Rasch Model

Alfina Saskia Ananda^{1*}, Kezia Ayu Teena², Rizki Dharma Syahputra³, Wildah Azmah Azizah⁴,
Anna Armeini Rangkuti⁵

^{1,2,3,4,5}Universitas Negeri Jakarta

E-mail: ¹alfinasaskia20@gmail.com*, ²keziaayutee@gmail.com, ³dharmarizki10@gmail.com,
⁴wildahazmah@icloud.com, ⁵annarangkuti@unj.ac.id

Abstract. The limited research on digital well-being and the absence of a dedicated digital well-being instrument in Indonesia led to this study. The research adapted the Digital Well-Being Scale (DWBS) into an Indonesian version, focusing on achieving good validity and reliability. Digital well-being emerges as a solution to the problems and negative impacts of the digital era. It prioritizes well-being as a core focus for achieving happiness, and the search for meaning in life to ultimately reach self-actualization. This quantitative study utilized an instrument adaptation process with 78 early adult respondents aged 18-40 years. The Indonesian version of the DWBS demonstrated good validity and reliability. After a series of analyses using Rasch modeling, the final instrument consists of 10 items and 3 dimensions.

Keywords: *digital well-being, instrument adaptation, rasch.*

Introduction

Rapid technological advancements continue to develop and exert a significant impact on various aspects of human life, particularly in the information field, such as the internet (Kurniasanti et al., 2019). Digital transformation and rapid innovations in electronic devices have altered how humans interact, learn, seek entertainment, and even shape behaviors in social life (Kurniasanti et al., 2019). Individuals ranging from children to young adults have been accustomed to digital technology from an early age and are referred to as the digital generation, millennials, or Generation Y, living and growing up amid a swift flow of information and virtually unlimited access to technology (Kurniasanti et al., 2019).

This statement aligns with the survey results conducted by the Indonesian Internet Service Providers Association (APJII) in 2018, which revealed that the age group with the highest rate of internet usage was adolescents aged 15 to 19, reaching 91% (Tempo.co, 2019). Furthermore, in 2018–2019, Statistics Indonesia (BPS) recorded that the age range most frequently using the

internet was individuals aged 15–24, accounting for 83.58%, followed by the 25–64 age group at 46.83% (Badan Pusat Statistik, 2019). Meanwhile, in 2023, BPS reported that the majority of the Indonesian population using the internet within the last three months of that period were individuals aged 19–24 and 25 and above, with percentages of 13.74% and 60.36%, respectively (Badan Pusat Statistik, 2024). A similar phenomenon is observed in Indonesia, despite the growth rate of internet usage in the country being relatively slower compared to other developing nations (Atmadja & Kiswantomo, 2020). The Ministry of Communication and Information Technology (KOMINFO), in collaboration with UNICEF, conducted an in-depth study on internet usage behavior among Indonesian adolescents across both urban and rural areas (Atmadja & Kiswantomo, 2020).

In the current digital era, smartphones have become an indispensable part of daily life, utilized for various purposes ranging from communication and information searching to completing tasks (Wahyuni & Permana, 2022). In education, particularly higher education, smartphones play a positive role in supporting learning, enabling university students to access online materials, communicate with lecturers and peers, and utilize available educational applications (Wahyuni & Permana, 2022). On the other hand, technology offers various conveniences in education, such as broad access to learning materials, global collaboration opportunities, and flexibility in learning time (Hartati & Arisandi, 2025). However, excessive use of smartphones and technology can yield negative impacts. Excessive smartphone usage brings adverse effects, particularly on mental health and productivity (Wahyuni & Permana, 2022). Several studies indicate that high-intensity smartphone use can trigger psychological disorders and reduce focus (Wahyuni & Permana, 2022). Furthermore, overusing technology can result in negative outcomes such as digital fatigue, social anxiety, and sleep disturbances caused by excessive screen exposure (Hartati & Arisandi, 2025).

In recent years, studies on digital well-being have remained very limited (Gomes et al., 2023). Although the digital environment clearly influences daily life, researchers often rely on conventional well-being measurement tools due to the lack of dedicated scales for the digital context (Gomes et al., 2023). The digital environment itself can be defined as a virtual space where

technological systems are interconnected, enabling individuals to communicate and share information (Gomes et al., 2023). Therefore, it is critical to develop valid measurement tools specifically tailored to today's digital world (Gomes et al., 2023). With such tools, research on how the digital realm affects individual well-being can be conducted more accurately (Gomes et al., 2023). Based on the explanation above, it can be concluded that in this digital era, individuals across all age groups—both young and adults use smartphones freely. While smartphones facilitate various tasks and activities, they also give rise to several negative impacts, such as social anxiety, fatigue, sleep difficulties, and other related issues (Hartati & Arisandi, 2025).

Digital well-being then emerges as a solution to the issues and negative impacts of the digital era, where well-being serves as the main focal point for achieving happiness and the search for life meaning to ultimately reach self-actualization (Arslankara et al., 2022). According to prior literature, research on digital well-being remains very limited, which prompts researchers to adapt a digital well-being instrument to advance research in this area within Indonesia. Furthermore, based on the presented data, the age groups with high technology usage are 15–24 years and 25–64 years (Badan Pusat Statistik, 2019). A substantial portion of this combined age range falls into the early adulthood category, specifically ages 18–40 (Putri, 2018; Hurlock, 1980). Consequently, this leads the researchers to adapt the Digital Well-Being Scale by Arslankara et al. (2022) into Indonesian, involving early adults aged 18–40 as research subjects, given that this age group exhibits the highest rate of technology use.

Method

This study employed a quantitative survey to validate the measurement instrument through Rasch model. The research design utilizes an instrument adaptation procedure. Instrument adaptation involves a series of systematic steps taken to translate and adjust a measuring instrument from one language or cultural context into another that aligns with the target population's language and culture (Beaton et al., 2000). This process aims to yield a valid and reliable instrument. In this study, the cross-cultural adaptation guidelines proposed by

Beaton were adopted, which involve two languages (translators) to bridge cultural differences through systematic adjustments (Beaton et al., 2000).

The components addressed during adaptation included: 1) Semantic equivalence (ensuring similarity in sentence meaning); 2) Idiomatic equivalence (ensuring idioms retain equivalent meaning); 3) Experiential equivalence (ensuring relevance to the target population); and 4) Conceptual equivalence (ensuring consistency in conceptual meaning). The overall process of adapting the Digital Well-Being Scale by Arslankara et al. (2022) conducted in this study is illustrated in Figure 1.

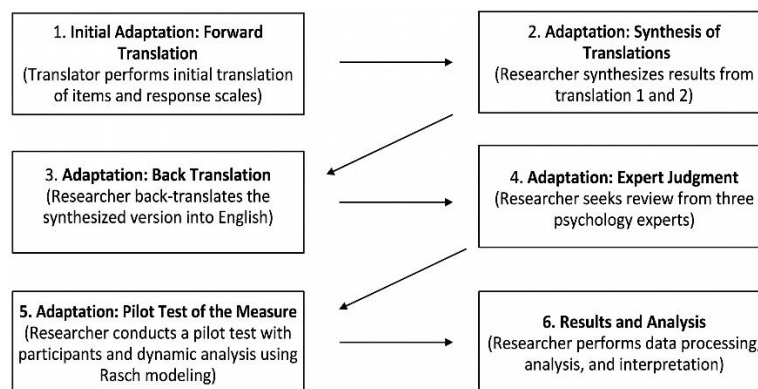


Figure 1. Process of Adapting the Digital Well-Being Scale

The study participants were characterized as individuals aged 18-40 years who possess social media accounts and actively engage on social media platforms. During the adaptation process, the sample size for the readability pilot test comprised 6 early adults ages 18-40, while 78 early adults aged 18-40 participates in the final item testing. The gender distribution results revealed that 21 respondents were male (26.9%) and 57 respondents were female (73.1%). The average age distribution of respondents was centered in their early 20s. The highest number of respondents was aged 24 ($n = 21$, 26.9%), followed by those aged 23 ($n = 9$, 11.5%), and those aged 18 ($n = 7$, 9.0%). All participants consented to the involvement after completing and agreeing to the informed consent form provided by the researcher.

The instrument adapted in this study was the Digital Well-Being Scale (DWBS) developed by Arslankara et al. (2022), define digital well-being as subjective well-being maintained in a virtual and technological environment characterized by an excessive increase in media communication. The DWBS consists of 12 items grouped into 3 sub-dimensions or 4 items each. This section is presented in full detail in the following table 1.

Table 1.
Scale Structure of DWBS

Sub-dimension	Items	Focus
F1. Digital Satisfaction	4	Harmony, pleasure, and positive feelings derived from digital tools and environments
F2. Safe and Responsible Behavior	4	Meaningful, cautious, respectful use of digital platforms
F3. Digital Wellness	4	Emotional relaxation/happiness associated with digital use (includes negatively worded items)

The scale uses a 5-point likert format: Fully reflects, reflects a lot, reflects moderately, reflects a little, does not reflect at all. Of the final 12 items, ten are positive and two negative (items 11 and 12). Negative items are reverse-coded (Arslankara et al., 2022).

To evaluate clarity, comprehensibility, and item interpretation, a preliminary pilot test was conducted with a small sample of 6 early adults. Serving strictly as a cognitive pre-test rather than a quantitative psychometric evaluation, the process confirmed that instructions and response categories were well-understood without major textual adjustments. Although limited in size, this pilot sample adequately served its purpose of confirming qualitative clarity before full-scale administration.

Table 2.
Result of Adaptation Process

Dimension	Original Item	Adaptation (Indonesian Version)
<i>Digital Satisfaction</i>	I can easily adapt to new technologies	Saya mudah beradaptasi dengan teknologi terbaru
	I enjoy spending time with digital technologies	Saya senang menghabiskan waktu dalam menggunakan teknologi digital
	I care about new digital experiences that can bring different experiences	Saya tertarik dengan pengalaman baru terkait dunia digital yang bisa memberi pengalaman berbeda.

Dimension	Original Item	Adaptation (Indonesian Version)
<i>Safe and Responsible</i>	In digital skills, I feel in harmony with the people around me	Terkait keterampilan digital yang saya miliki, saya merasa selaras dengan orang-orang di sekitar saya.
	I care about my digital reputation when using online platforms	Saya peduli dengan reputasi digital saya saat menggunakan platform daring.
	I take care not to exhibit behavior that disturbs other users on social media	Saya berhati-hati untuk tidak menunjukkan perilaku yang mengganggu pengguna lain di media sosial.
<i>Digital Wellness</i>	I use digital technology in purposeful meaningful ways	Saya menggunakan teknologi digital secara bermakna dan memang bertujuan.
	I always act cautiously against any harm that may come to me in the digital world	Saya berhati-hati terhadap segala bahaya yang mungkin saja saya alami di dunia digital.
	I feel comfortable knowing that someone will see my social media posts	Saya merasa nyaman saat mengetahui seseorang melihat postingan saya di media sosial.
	It makes me happy if the posts/stories/statuses I share are liked	Saya merasa senang jika postingan/cerita/ status yang saya bagikan disukai orang lain.
	A technological problem that I cannot solve makes me angry	Masalah teknologi yang tidak dapat saya selesaikan membuat saya marah.
	If I express myself freely on social media, I think that I will be ostracized by some people in my social networks.	Jika saya mengekspresikan diri secara bebas di media sosial, saya merasa saya akan dikucilkan oleh beberapa orang di jejaring sosial saya.

Parameter estimation was based on the Rasch Rating Scale Model (RSM), which maps both person ability and item difficulty onto a common interval logit continuum (Andrich, 2011). Unlike Classical Test Theory, Rasch analysis allows for invariant measurement and robust item-level diagnostics (Bond & Fox, 2015). Because all instrument item shared an identical Likert response structure, the RSM assumption of uniform threshold distances was met. The data analysis evaluated model fit, item/person fit statistics, reliability coefficients, dimensionality and response category performance. Parameter estimates were generated using Joint Maximum Likelihood Estimation (JMLE) implemented in Winsteps software version 3.73 (Linacre, 2011).

Results

Item fit was evaluated using a three-criteria approach: Outfit MNSQ (.50–1.50), Outfit ZSTD (-2.00 to +2.00), and Point-Measure Correlation (.40 – .85) (Linacre, 2011). In accordance with Rasch decision guidelines, an item is considered for deletion only if it fails at least two of these three criteria (Faradillah & Febriani, 2021). Items DW3 (Outfit ZSTD = 7.2, Outfit MNSQ > 1.50, PT-Measure = - .27) and DW4 (Outfit ZSTD = 8.1, Outfit MNSQ > 1.50, PT-Measure = - .12) failed all three criteria, confirming severe misfit and justification for removal. In contrast, items DS1 (Outfit ZSTD = - 3.4) and DS4 (Outfit ZSTD = - 2.7) deviated only in ZSTD, while maintaining acceptable Outfit MNSQ values (.81 and .85, respectively) and strong Point-Measure Correlations (> .50). Because ZSTD is highly sensitive to sample sizes, items meeting MNSQ and PT-Measure thresholds were retained to preserve content validity (Linacre, 2026). Consequently, the researcher removed these two items, resulting in improved outcomes.

The item measure test results indicate that all remaining 10 items met at least two of the outfit criteria (outfit ZSTD, MNSQ, and point-measure correlation). Therefore, the 10 remaining DWBS items proceeded to the next stage of analysis. The subsequent step involved examining the distribution of items and respondents using a Wright map. The respondent distribution maps for this analysis are shown in Figure 2 and Figure 3 below.

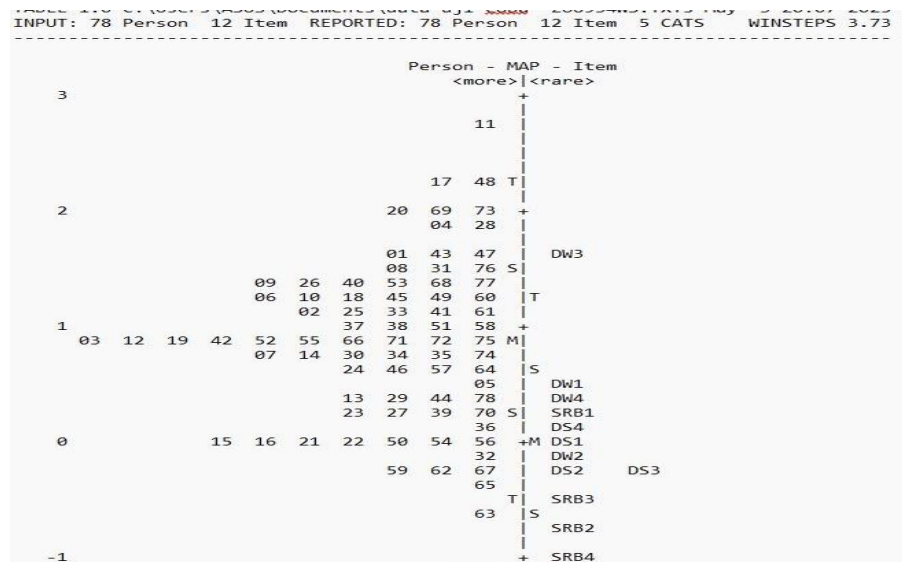


Figure 2. Wright Map Before Item Deleted

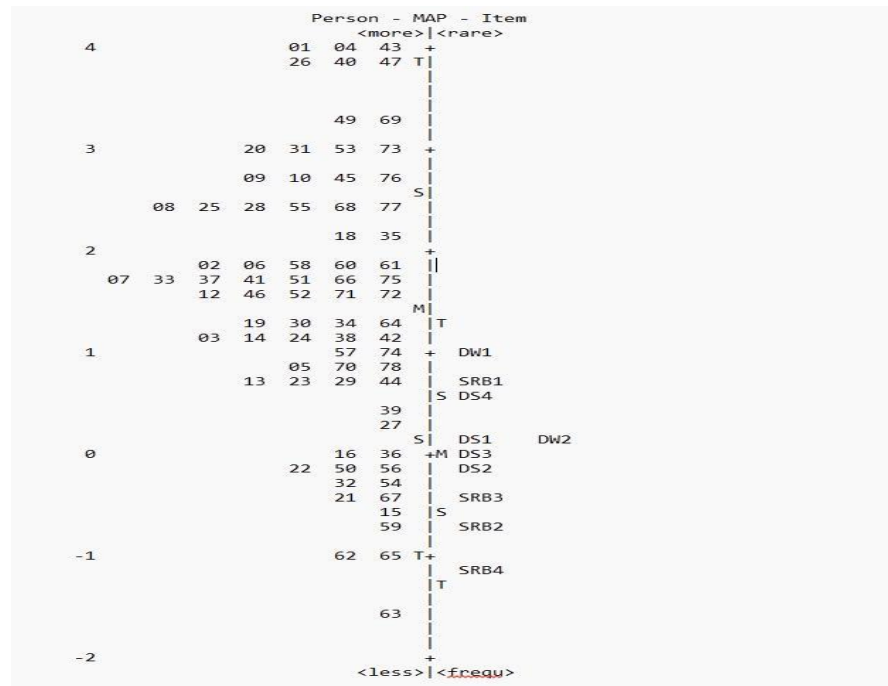


Figure 3. Wright Map After Item Deleted

In Rasch measurement, the .0 logit origin on the Wright Map represents the mean item difficulty calibrated to the center of the latent continuum, rather than a metric of response quality or seriousness (Bond & Fox, 2015). Person locations above .0 logit indicate higher latent levels of digital well-being, whereas locations below .0 indicate lower levels (Andrich, 2011).

The distribution map (Figure 2) shows that the most difficult item is DW3 (Digital Wellness), while the easiest is SRB4 (Safe and Responsible Behavior). Regarding respondent ability, the map indicates that almost all respondents possess comparable ability in answering the given items, as the majority are clustered near the .0 mark; however, a few respondents—specifically numbers 11, 17, and 48—showed signs of guessing. We considered excluding these respondents to improve the quality of the subsequent analysis. Additionally, as previously explained, items DW3 and DW4 were removed because they compromised the instrument's quality. The following shows the Wright map after removing the items and responses that negatively affected instrument quality, leaving a total of 75 respondents and 10 items.

Picture 3 displays the distribution map after the removal of items and respondents with poor quality metrics; it shows that item difficulty levels are now largely comparable and close to 0.0. While some respondents still appear far from the 0.0 point—indicating guessing or low ability—they were retained rather than subjected to a second round of exclusion, as removing them would have resulted in an excessively small sample size.

A crucial analysis result in Rasch modeling—specifically the unidimensionality test, assessed via the "raw variance explained by measure" indicates that the DWBS instrument meets the unidimensionality requirement; it yielded a raw variance explained by measure of 51.6%, a figure closely aligned with the expected (modeled) value of 51.2%. The instrument demonstrated a favorable outcome on the unidimensionality test, as the raw variance explained by measure exceeded the 40% threshold established for Rasch modeling standards (Hakiki et al., 2018). Consequently, the construct validity of the DWBS instrument is also confirmed, as the instrument effectively measures the intended construct. Table 3 presents the results of the unidimensionality test.

Table 3.
Unidimensionality Result

		<i>Empirical</i>		<i>Modeled</i>
<i>Total raw variance in observations</i>	=	20.6	100.0%	100.0%
<i>Raw variance explained by measures</i>	=	10.6	51.6%	51.2%
<i>Raw variance explained by persons</i>	=	5.8	28.0%	27.8%
<i>Raw variance explained by items</i>	=	4.9	23.5%	23.4%
<i>Raw unexplained variance (total)</i>	=	10.0	48.4%	100% 48.8%
<i>Unexplained variance in 1st contrast</i>	=	2.3	11.3%	23.4%
<i>Unexplained variance in 2nd contrast</i>	=	1.6	7.8%	16.1%
<i>Unexplained variance in 3rd contrast</i>	=	1.5	7.0%	14.5%
<i>Unexplained variance in 4th contrast</i>	=	1.1	5.4%	11.1%
<i>Unexplained variance in 5th contrast</i>	=	.8	4.1%	8.5%

independence criterion. Results from the local independence test on the DWBS instrument indicate that nearly all items are independent of one another, as they exhibit standardized residual correlation values below .2. Although two items—DW1 and DW2—showed interdependence with a correlation value of .52 (exceeding the .2 threshold), the items in the instrument are generally independent; the dependency between DW1 and DW2 is considered negligible because both belong to the same dimension—Digital Wellness—making it highly probable that they would influence each other. In conclusion, the local independence requirement for the DWBS instrument is met, as the items within each dimension do not significantly influence one another (Tennant & Küçükdeveci, 2023; Andrich & Marias, 2019). The results of the local independence test are presented in Table 4 below.

Table 4.
Local independence Result

<i>CORRELATION</i>	<i>ENTRY NUMBER</i>	<i>ITEM</i>	<i>ENTRY NUMBER</i>	<i>ITEM</i>
.52	9	DW1	10	DW2
-.45	1	DS1	10	DW2
-.41	3	DS3	10	DW2
-.33	4	DS4	10	DW2
-.33	5	SRB1	8	SRB4
-.33	3	DS3	9	DW1
-.31	1	DS1	8	SRB4
-.29	1	DS1	9	DW1
-.28	3	DS3	5	SRB1
-.27	4	DS4	9	DW1

Furthermore, the rating scale analysis for the DWBS instrument demonstrates a consistent rise in observed average values—ranging from -.43 to 2.57—across the response options, progressing from option 1 ("tidak menggambarkan diri saya sama sekali") to option 5 ("sepenuhnya menggambarkan diri saya"). The observed average values increase significantly as

the scale level rises, indicating that the rating scale for the DWBS instrument functions effectively. This suggests that the 1–5 scale used to measure the DWBS does not confuse respondents and is indeed appropriate for accurately measuring the construct. Additionally, the progression of Andrich Threshold values—moving from "NONE" (or from negative figures) to a positive value of 2.23—confirms that the scale enables respondents to effectively distinguish between the five provided response options (Andrich, 2011). The results of the rating scale analysis are presented in Table 5 below.

Table 5.
Rating Scale

<i>Category</i>		<i>Observed Count</i>	<i>%</i>	<i>Observed Average</i>	<i>Sample Expect</i>	<i>Infit MNS Q</i>	<i>Outfit MNS Q</i>	<i>Andrich Threshold</i>	<i>Category Measure</i>
<i>Lable</i>	<i>Score</i>								
1	1	18	2	-0.43	-1.03	1.70	2.28	NONE	(-3.17)
2	2	58	8	-0.43	-0.32	0.92	0.93	-1.85	-1.49
3	3	157	21	0.49	0.56	0.97	1.00	-0.88	-0.12
4	4	269	36	1.51	1.52	0.84	0.87	0.50	1.46
5	5	248	33	2.57	2.53	0.91	0.94	2.23	(3.44)

programs to demonstrate that the DWBS instrument's reliability is indeed robust and consistent. Initial results showed respondent reliability at .65 (weak category) and item reliability at .95 (excellent category), while Cronbach's Alpha for the entire instrument was .74 (adequate category). These categorizations are based on Rasch reliability standards as cited by Sari & Mahmudi (2024). As the initial reliability results were suboptimal, the researcher proceeded with a second reliability test after removing two problematic items and three respondents of poor quality. The final reliability test involved the remaining items and 75 respondents. Additionally, a reliability test using JASP was conducted to compare results with those obtained via Classical Test Theory (CTT). The JASP analysis yielded a Cronbach's Alpha of .710. Given this unsatisfactory result, the reliability test was repeated after removing items DW3 and DW4. The

final reliability test using Winstep software demonstrated improved consistency: respondent reliability was .81 (good category), and item reliability was .93 (indicating excellent reliability for the instrument's items). The Cronbach's Alpha for the entire instrument also fell into the good category at .88. Meanwhile, the JASP reliability test conducted after removing items DW3 and DW4 yielded a Cronbach's Alpha of .879. These results indicate good reliability, demonstrating that the DWBS instrument is consistent and suitable for use.

In addition to reliability testing, the separation values for this instrument can also be calculated using the following formulas (Wibisono, 2015; Misbah & Sumintono, 2014).

$$H = \frac{[(4 \times \text{SEPARATION}) + 1]}{3}$$

Note:

H = Item/Person Stratum Value

SEPARATION = Resulting separation value for persons/items

The calculated person and item strata values in this study indicate a person stratum value of 3.08, signifying that the respondents are divided into three categories—high, medium, and low (Sari & Mahmudi, 2024). Meanwhile, the item stratum value is 5.12, indicating that the items are distributed across five levels based on the difficulty perceived by the respondents (Sari & Mahmudi, 2024). Sari and Mahmudi (2024) also note that optimal person and item separation values fall within the 2–5 range. Thus, it can be concluded that both the person and item separation indices are satisfactory, as they are capable of accurately distinguishing between different groups of respondents and items.

Discussion

Data analysis in this study was executed using Rasch modeling, following prior literature emphasizing its capacity to focus on participant response patterns, maintain high accuracy, and construct standard error measurement values for the instrument

(Ekstrand et al., 2022). The Rasch paradigm asserts that comparisons between individuals remain independent of the specific test items used, just as item comparisons are independent of the particular sample evaluated; thus, person and item parameters operate independently (Asrijanty, 2014). This objective measurement approach aligns with Thurstone's principle of ideal measurement, which postulates that a robust instrument must yield measurements unaffected by the characteristics of the object being measured (Asrijanty, 2014).

The Rasch model analysis comprising item fit evaluation, person-item map (Wright map) examination, unidimensionality, local independence, rating scale diagnostics, reliability, and separation indices—demonstrated optimal results following the removal of two items from the digital wellness dimension, specifically DW3 and DW4. Item fit, Wright map, and reliability analyses were conducted in two stages: prior to and following the exclusion of these items. The contrast was substantial, with post-exclusion metrics showing marked improvement over initial scores. These initial discrepancies were partly attributable to respondents displaying guessing behavior; removing these misfitting responses yielded significantly enhanced psychometric properties across all three evaluations. This outcome aligns with Rasch measurement theory, which posits that key characteristics of Rasch modeling include the Item Characteristic Curve (ICC), item discrimination, objective comparison, and the capacity to detect response guessing (Asrijanty, 2014).

Based on item fit evaluation across three criteria outfit standard Z, outfit mean square and point-measure correlation two items failed to meet all threshold conditions. According to psychometric guidelines, an item should satisfy at least two of these criteria; failure to do so warrants consideration for item deletion (Faradillah & Febriani, 2021). Consequently, deleting these two items yielded improved model fit. Furthermore, the

item map distribution revealed that DW3 (Digital Wellness) was the most difficult item, whereas SRB4 (Safe and Responsible Behavior) was the easiest. In Rasch measurement, the 0.0 logit origin on the Wright Map represents the mean item difficulty calibrated to the center of the latent continuum, rather than a metric of response quality or seriousness (Bond & Fox, 2015). Person locations above 0.0 logit indicate higher latent levels of digital well-being, whereas locations below 0.0 indicate lower levels (Andrich, 2011). To identify aberrant responses or potential guessing, individual Outfit Mean Square (MNSQ) and Outfit ZSTD statistics were examined (Linacre, 2011). Three participants (Respondents 11, 17, and 48) demonstrated severe outfit MNSQ values (> 1.50) combined with ZSTD $> +2.00$, indicating unpredictable response patterns uncharacteristic of their overall trait levels, indicative of haphazard responding or guessing (Smith, 2000). Consequently, these three misfitting cases were removed to improve model fit.

Regarding unidimensionality, the minimum score for the unidimensionality test must be 20%, as observed from the raw variance explained by measure by examining the empirical and modeled values. If it falls below this threshold, the unidimensionality assumption is not met; conversely, if the obtained result exceeds 20% for instance, 40% it is categorized as good, and a result of 60% indicates exceptional unidimensionality of the instrument (Hakiki et al., 2018). In this study, the unidimensionality requirement was met, with the raw variance explained by measure reaching 51.6%, which does not differ significantly from its expected (modeled) value of 51.2%. Although the original DWBS was conceptualized into three sub-dimensions (Digital Satisfaction, Safe and Responsible Behavior, and Digital Wellness) (Arslankara et al., 2022), the present Rasch analysis supported unidimensionality, with the raw variance explained by measures reaching 51.6% (Hakiki et al., 2018). In Rasch measurement, unidimensionality indicates that a single primary latent construct—overall digital well-being—drives the responses across all items (Bond & Fox, 2015). This empirical unidimensionality does not invalidate the

theoretical sub-dimensions; rather, it suggests that while the sub-dimensions reflect distinct thematic facets, they function unitarily at the construct level (Linacre, 2026). Therefore, the adapted Indonesian DWBS is best interpreted as a composite measure yielding a single overall Digital Well-Being score, while sub-dimensions can still serve descriptive, qualitative purposes in sub-scale profiles.

Another requirement that must be fulfilled in Rasch modeling is local independence. Local independence in an item refers to a condition where a respondent's correct or incorrect answer does not affect the response of another respondent, or it can also be defined as the items being independent of one another (Tennant & Küçükdeveci, 2023; Andrich & Marias, 2019). Local independence can be tested mathematically using probability formulas and statistically through the Chi-Square test of independence (Tennant & Küçükdeveci, 2023; Andrich & Marias, 2019). The standard threshold for local independence is determined by the standardized residual correlation value, which should not exceed an absolute value of .2 (Andrich, 2011). Local independence assumes that responses to individual items are statistically independent after conditioning on the latent trait (Tennant & Küçükdeveci, 2023). A residual correlation between DW1 (Saya merasa nyaman saat mengetahui seseorang melihat postingan saya di media sosial.) and DW2 (Saya merasa senang jika postingan/cerita/ status yang saya bagikan disukai orang lain.) reached .52, exceeding the recommended (.20) threshold (Andrich, 2011). This residual dependency indicates item redundancy and shared trait variance specific to social media validation (Tennant & Küçükdeveci, 2023). While both items were retained in this preliminary adaptation to maintain content coverage of the Digital Wellness domain, the high correlation implies slight item over-fit and artificially inflated reliability (Baghaei, 2008). Future scale revisions should consider rephrasing or combining DW1 and DW2 to eliminate local dependency.

Furthermore, the rating scale criteria measured in this study can be evaluated through the observed values to determine whether they exhibit a consistent order from the first to the last option (Andrich, 2011). The Rating Scale Model (RSM) was selected over the Partial Credit Model (PCM) because the DWBS employs a uniform 5-point Likert response format across all items designed to measure a shared underlying latent continuum (Andrich, 2011). RSM assumes that the relative step difficulties (thresholds) between adjacent response categories are fixed across all items (Bond & Fox, 2015). Empirical diagnostic of category functioning supported this assumption: category measures progressed monotonically (-.43 to 2.57) and Andrich thresholds increased appropriately (-1.85, -.88, .50, 2.23) without category collapse (Linacre, 2011). Because threshold distances behaved uniformly across items without evidence of item-specific category distortion, RSM provided a more parsimonious model with higher degrees of freedom compared to PCM (De Ayala, 2009).

Construct validity was evaluated through multiple sources of evidence aligned with modern measurement standards (AERA, APA, & NCME., 2014). First, content validity was established prior to field testing using expert judgment evaluated via Aiken's V coefficient (Aiken, 1985). Second, internal structure validity was confirmed through Rasch analysis, demonstrating acceptable item fit (MNSQ), unidimensionality (51.6% variance explained), and proper ordering of response categories (Andrich Thresholds) (Andrich, 2011). However, while these analyses support internal structural integrity, further research is required to establish convergent and discriminant validity by correlating the Indonesian DWBS with external constructs such as psychological well-being, life satisfaction, or digital addiction (Gomes et al., 2023). A primary limitation of this study is the sample size (N=75 retained participants). While sample sizes between 50 and 100 are adequate for preliminary Rasch calibrations and item screening (Linacre,

2026), this relatively small sample limits the generalizability of the findings across broader demographic groups in Indonesia.

Conclusion

In conclusion, the 10-item Indonesian version of the DWBS demonstrates promising preliminary validity and reliability when evaluated via Rasch modeling. However, given the initial sample constraints, these findings should be interpreted as preliminary psychometric evidence. Future validation studies employing larger, randomly sampled, and demographically diverse populations are strongly recommended to establish national norms and cross-subgroup invariance.

Suggestion

From a practical perspective, the findings suggest that the Indonesian version of the DWBS can be used as a useful tool for assessing digital well-being behavior tendencies in organizational contexts. However, several limitations of this study include a relatively small sample size of respondents; Thus, future researchers are encouraged to expand the respondents. Additionally, subsequent researchers could focus on specific locations or companies to examine digital well being in a more tailored context.

References

- AERA, APA, & NCME. (2014). Standards for Education and Psychological Testing (Vol. 4, Issue 1). American Educational Research Association.
- Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45(1), 131–142.
- Andrich, D. (2011). Rating scales and Rasch measurement. *Expert Review of Pharmacoeconomics and Outcomes Research*, 11(5), 571–585. <https://doi.org/10.1586/erp.11.59>
- Andrich, D., & Marias, I. (2019). *A Course in Rasch Measurement Theory: Measuring in Educational, Social and Health Sciences*. Jerman: Springer.
- Arslankara, V. B., Demir, A., Öztaş, Ö., & Usta, E. (2022). Digital well-being scale validity and reliability study. *Journal of Teacher Education and Lifelong Learning*, 4(2), 263-274.
- Asrijanty. (2014). Model Rasch sebagai Kerangka Acuan Penyusunan Alat Ukur. *Jurnal Pendidikan Dan Kebudayaan*, 20(1), 109–123. <https://doi.org/10.24832/jpnk.v20i1.130>
- Atmadja, K., & Kiswantomo, H. (2020). Hubungan antara komponen-komponen subjective-well being dan internet addiction. *Humanitas (Jurnal Psikologi)*, 4(1), 27-42.

- Badan Pusat Statistik. (2019). Proporsi Individu yang Menggunakan Internet Menurut Kelompok Umur (Persen), 2018-2019. Bps.Go.Id. <https://www.bps.go.id/id/statistics-table/2/MTIyOCMy/proporsi-individu-yang-menggunakan-internet-menurut-kelompok-umur.html>
- Badan Pusat Statistik. (2024). Persentase Penduduk Usia 5 Tahun ke Atas yang Pernah Mengakses Internet dalam 3 Bulan Terakhir Menurut Kelompok Umur (Persen), 2022-2023. Bps.Go.Id. <https://www.bps.go.id/id/statistics-table/2/ODQwIzI=/persentase-penduduk-usia-5-tahun-ke-atas-yang-pernah-mengakses-internet-dalam-3-bulan-terakhir-menurut-kelompok-umur.html>
- Baghaei, P. (2008). Transactions of the Rasch Measurement SIG. 21(3).
- 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.
- Bond, T. G., & Fox, C. (2015). Applying The Rasch Model (Fundamental Measurement in The Human Science). Psychology Press
- De Ayala. (2009). The Theory and Practice of Item Response Theory. The Guilford Press.
- Ekstrand, J., Westergren, A., Årestedt, K., Hellström, A., & Hagell, P. (2022). Transformation of Rasch model logits for enhanced interpretability. *BMC Medical Research Methodology*, 1–10. <https://doi.org/10.1186/s12874-022-01816-1>
- Faradillah, A., & Febriani, L. (2021). Mathematical Trauma Students' Junior High School Based on Grade and Gender. *Infinity Journal*, 10(1), 53. <https://doi.org/10.22460/infinity.v10i1.p53-68>
- Mathew, Jain, et al. "Design and validation of the digital well-being scale." *Ricerche di Pedagogia e Didattica. Journal of Theories and Research in Education* 18.1 (2023): 239-251.
- Hakiki, A. W., Fitri, A. R., & Agung, I. M. (2018). Analisis Properti Psikometri Subtes Merkaufgaben (ME) dengan Rasch Model. *Jurnal Psikologi*, 14(1), 40. <https://doi.org/10.24014/jp.v14i1.4900>
- Hartati, S., & Arisandi, D. (2025). Studi pendahuluan: Psychological well-being dan positive education pada mahasiswa generasi Z di era teknologi. *Journal of Sustainable Education*, 2(1), 9-18.
- Hurlock, E. B. (1980). *Developmental Psychology: A life span approach*. (5th Edition). McGraw-Hill
- Kurniasanti, K. S., Assandi, P., Ismail, R. I., Nasrun, M. W. S., & Wiguna, T. (2019). Internet addiction: a new addiction?. *Medical Journal of Indonesia*, 28(1), 82-91.
- Linacre, J. (2011). A User's Guide to Winsteps Rasch Model Computer Programs.
- Linacre, J. M. (2026). Winsteps Help for Rasch Analysis 5.12.2.
- Putri, A. F. (2018). Pentingnya Orang Dewasa Awal Menyelesaikan Tugas Perkembangannya. *SCHOULID: Indonesian Journal of School Counseling*, 3(2), 35. <https://doi.org/10.23916/08430011>
- Sari, E. D. K., & Mahmudi, I. (2024). Analisis pemodelan rasch pada assessment pendidikan (analisis dengan menggunakan aplikasi winstep). PT. Pena Persada Kerta Utama.
- Smith, R. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement*, 1(2), 199–218.

- Tempo.co. (2019). Usia 15-19 Tahun Pengakses Internet Terbesar di Indonesia. <https://www.tempo.co/digital/usia-15-19-tahun-pengakses-internet-terbesar-di-indonesia-744330>.
- Tennant, A., & Küçükdeveci, A. A. (2023). Application of the Rasch measurement model in rehabilitation research and practice: early developments, current practice, and future challenges. *Frontiers in Rehabilitation Sciences*, 4, 1208670.
- Wahyuni, D., & Permana, I. S. (2022). Memantau Kesehatan Digital Melalui Aplikasi Digital Wellbeing : Implikasi Terhadap Disiplin dan Prestasi Akademik. *Prosiding Seminar Nasional Pascasarjana*, 6(1), 1125-1130
- Wibisono, S. (2015). Aplikasi Model Rasch Untuk Validasi Instrumen Pengukuran Fundamentalisme Agama Bagi Responden Muslim. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia (JP3I)*, 5(1). <https://doi.org/10.15408/jp3i.v5i1.9239>