JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)
DOI: http://dx.doi.org/10.15408/jp3i.v14i1.38023
http://journal.uinjkt.ac.id/index.php/jp3i

Measuring the Intention for Innovative Work Behavior: Scale Development for Indonesian Employees

Arum Etikariena[1], Aflaha Hadini Kusuma[1], Devika Nur Shabrina[1], Anak Agung Istri Ratna Kumala Dewi[1], Rana Hafizhah[1], Zidan Raleto Seno[1]

Faculty of Psychology, Universitas Indonesia, Indonesia[1]

arum.etikariena@ui.ac.id

Abstract

The current business era is dynamic and filled with uncertainty, and innovating is no longer an option but a solution to survive and compete. Innovation in the organisational context is displayed by employees in their work, known as Innovative Work Behaviour (IWB). However, the measurement tool for innovative work behavior focuses on past behavior, whereas innovative behavior is characterized by uncertainty. Previous behavior may not necessarily predict future behavior. Meanwhile, intention is the closest factor to predicting future behavior. Some researchers have adjusted the measurement of innovative work behavior into the intention for innovative work behavior. But its reliability and validity have not been tested, and it has not been adapted to the Indonesian language. Therefore, this quantitative study aims to develop a measurement tool of intention for innovative work behavior in Indonesia based on the Su et al. Scale (2019). Using the convenience sampling method, 1,071 participants who had worked for at least 1 (one) year in their organisation participated in this study. The results show that the I-IWB Scales had good reliability ($\alpha = 0.840$). Confirmatory factor analysis also revealed that the scale fit the one-factor model (RMSEA = 0.054 < 0.070; 90% CI 0.044, 0.066), (CFI = 0.970 > 0.950), (TLI = 0.955 > 0.900). Construct validity shows the model fit as a unidimensional scale. Thus, based on the analysis results, it can be said that the I-IWB Scale is still reliable and valid for measuring the construct of intention for innovative work behaviour for Indonesian employees.

Keywords: construct validity, confirmatory factor analysis, intention for innovative work behaviour, model fit, reliability

Abstrak

Era bisnis saat ini begitu dinamis dan dipenuhi oleh ketidakpastian, sehingga berinovasi bukan lagi pilihan, melainkan solusi untuk dapat bertahan dan bersaing. Inovasi pada konteks organisasi yang dilakukan oleh karyawan di dalam pekerjaannya dikenal sebagai perilaku kerja inovatif atau Innovative Work Behaviour (IWB). Hanya saja, alat ukur perilaku kerja inovatif saat ini lebih berfokus pada perilaku yang sudah dilakukan sebelumnya. Padahal, perilaku inovatif merupakan perilaku yang penuh dengan ketidakpastian, sehingga perilaku sebelumnya belum tentu dapat memprediksi perilaku individu ke depannya secara akurat. Intensi merupakan faktor terdekat untuk memprediksi perilaku individu di masa datang. Beberapa peneliti telah melakukan penyesuaian pengukuran perilaku kerja inovatif menjadi intensi perilaku kerja inovatif, namun belum teruji reliabilitas dan validitasnya, serta belum diadaptasi ke dalam Bahasa Indonesia. Penelitian kuantitatif ini bertujuan untuk mengembangkan alat ukur intensi perilaku kerja inovatif dalam Bahasa Indonesia berdasarkan skala yang disusun oleh Su, et al. (2019). Dengan teknik sampling convenience sampling, sebanyak 1.071 partisipan yang sudah bekerja selama minimal 1 (satu) tahun di organisasinya berpartisipasi dalam penelitian ini. Hasil menunjukkan reliabilitas yang baik dari Skala Intensi Perilaku Kerja Inovatif sebesar $\alpha = 0.840$. Confirmatory factor analysis menunjukkan bahwa model yang ditampilkan juga fit (RMSEA = 0.054 < 0.070; 90% CI 0.044, 0.066), CFI (CFI = 0.970 > 0.950), TLI = 0.955 > 0.900). Karenanya, validitas konstruk juga menunjukkan model fit sebagai skala untuk mengukur konstruk yang bersifat unidimensional. Dengan demikian, dapat dikatakan bahwa Skala Intensi Perilaku Kerja Inovatif masih reliabel dan valid untuk mengukur konstruk intensi perilaku kerja inovatif bagi partisipan karyawan di Indonesia.

Kata kunci: *confirmatory factor analysis, intensi perilaku kerja inovatif, model fit, reliabilitas, validitas konstruk*

Introduction

In the era of dynamic and uncertainty-filled business digitalisation, most organisations consider innovation not merely an option but a definite step to ensure survival (Millar, 2018). Innovation also helps organisations adapt and face challenges, overcome stagnation, support organisational growth, and differentiate the business from competitors (Boyles, 2022). Survey results conducted by Banholzer et al. (2023) indicate that taking significant risks in innovation appears to be a safer approach than gradual investment in innovation. Data illustrates that 46% of companies that underwent substantial innovation and organisational changes from 2021 to 2022 managed to enhance sales and partnerships. Innovation in workplace settings consists of three levels of innovation: at the organisational level, at the group or work unit level, and at the individual level or, known as innovative work behaviour (IWB) (Kanter, 1988; Scott & Bruce, 1994).

Employees have important role in initiating innovation in an organisation. Employees often encounter business processes and products directly to detect potential opportunities for organisational improvement and development (Bos-Nehles et al., 2017). Therefore, innovative work behaviour from employees allows organisations to maintain competitive advantages and organisational sustainability (AlEssa & Durugbo, 2021). Therefore, employees capable of fostering innovation become crucial for organisational sustainability (Salvador & Sting, 2022). Innovative work behaviour significantly impacts organisational effectiveness and sustainability. Employees with innovative work behaviour can align the company's desired business vision and strategy with evolving technological and operational system changes (Muchiri et al., 2020).

Innovative Work Behaviour (IWB) is an intentional behaviour within roles, workgroups, or organisations to gain performance benefits for roles, groups, or organisations. Innovative work behaviour consists of three stages: idea generation, the initial stage where individuals identify problems and generate new ideas or adaptive solutions. The next stage is idea promotion, which involves seeking support to realise these ideas. The final stage is idea realisation, which is implementing and applying these ideas within the team or organisation. These stages form a unified whole, making them inseparable from one another. Hence, the innovative work behaviour scale is unidimensional (Janssen, 2000; Scott & Bruce, 1994). Studies on innovative work behaviour have become a prominent topic among researchers. Not surprisingly, research on innovative work behaviour has increased over the last 5 years (Srirahayu et al., 2023).

Innovative work behaviour is a unidimensional construct measured by rating the innovative work behaviour previously performed by individuals. In its development, the concept and measurement tools of innovative work behaviour developed by Janssen (2000) are frequently used in research (AlEssa & Drugbo, 2021). However, the measurement tool developed by Scott & Bruce (1994) and continued by Janssen (2000) focuses only on previous behaviour. There are some scales to measure innovative work behaviour. Several researchers in their studies have modified Janssen's (2000) measurement tool for innovative work behaviour into intentions, including studies by Lu and Luh (2013) and Su et al. (2019).

Previous behaviour can indeed predict the tendency for subsequent behaviour. However, this is more evident when the behaviour has become a habit or routine activity, as employees frequently engage in such actions (Ouellette & Wood, 1998). In contrast, innovative work behaviour is a high-risk activity and is not commonly performed by most employees (Williams et al., 2020). For some reason, Ajzen (1991) has demonstrated that when the activity is uncertain in its implementation due to its high risk, intentions play a more significant role in predicting changes in a person's behaviour. Therefore, measuring the intention to perform the behaviour can be more effective in improving the predictors and the behaviour rather than the behaviour itself. Therefore, a measurement tool for innovative work behaviour is needed to gauge employees' intentions to understand their tendency to engage in innovative behaviour in the future.

Su, et al. (2019) made the adjustment by adding "I would" to each item. The scale is presented in Chinese to align with the study participants. However, that study still used a frequency scale or how often the previous behaviour occurred. In Indonesia, the measurement tool for innovative work behaviour by Janssen (2000) has been adapted to Indonesian by Etikariena and Muluk (2014), but similar to Su et al. (2019), the scale still uses a frequency for the occurrence of previous behaviour. However, when intending to measure intentions, it is advisable to use an attitude scale (Ajzen, 1991). Therefore, in this study, several sentence adjustments were made to each item to turn them into an intention statement and the attitude scale according to the Theory of Planned Behaviour by Ajzen (1991) to ensure that the responses align with the intended meaning of intentions. The scale interpretation is as follows: scale 1 (one) is unsuitable to scale 6 (six), which means it is very suitable. These adjustments are expected to produce a measurement tool for innovative work behaviour adapted into Indonesian to further test its validity and reliability. Accordingly, the language has also been translated into Indonesian. This procedure followed the Guidelines for the Cross-Cultural Adaptation Process proposed by Beaton (2007). Hopefully, the adaptation process will minimise bias and enhance the reliability and validity of the scale.

Methods

Participants

This study is a quantitative cross-sectional study. Participants in this research are Indonesian employees who have been working for at least one year in their respective organisations. Participants were selected using the convenience sampling technique, where sampling is based on participants' availability to engage in this study (Cohen, 2013). Researchers sought participants by disseminating broadcast messages on various social media platforms. The total number of participants successfully collected for this study was 1,257. However, after data cleaning because of outlier ot incocistent data, the final total of usable participants became 1,071. The participants criteria is being employees who had worked for at least one year in their organisations, predominantly from edtech, marketplace/e-commerce, and manufacturing companies. Data cleaning involved using confounding variables such as demographic data and data with extreme values. The minimum ideal number of participants for factor analysis varies depending on the study's purpose. There is 200 participants (Williams, et al., 2010) or 300 participants (DeVellis, 2017) for scale development purposes. Nevertheless, to use the CFA, 315 participants are the minimum criteria for sample size to avoid missing and non-normal data (Muthen & Muthen, 2002). Therefore, the participants in this research have already fulfilled the minimum number of participants.

Instruments

The measurement tool for Intention to Innovative Work Behaviour (I-IWB) was developed based on the concept of Scott & Bruce (1994) regarding the unidimensional construct consisting of three stages: idea generation, idea promotion, and idea realisation. The development process of the measurement tool is divided into several stages, namely the preparation and field data collection stages. In the preparation stage, for adjusting the scale of intention to IWB, the researchers obtained approval from the ethics review committee from the Faculty of Psychology, Universitas Indonesia, to conduct research related to the Intention to IWB measurement tool (No: 196/FPsi.Komite Etik/PDP.04.00/2023). After obtaining approval, the researchers modified the measurement tool by adding the phrase "I will" to the nine items of innovative work behaviour adapted into Indonesian by Etikariena and Muluk (2014). We also adjusted the filling instructions and Likert scale categories.

In the next stage, as part of the preparation phase, the adapted Intention for IWB measurement tool underwent expert judgment and parallel data collection testing. Two lecturers from the Faculty of Psychology, Universitas Indonesia, provided expert judgment. The first reviewer gave minor revisions to some items in the intention of the IWB measurement tool, such as item number 9, changing from "I will review the usefulness of new ideas I have proposed at work" to "I will evaluate the usefulness of ideas I have conveyed." Unfortunately, the second reviewer did not give specific feedback. Therefore, the

researchers decided to decline the opinion. Furthermore, based on the results of the parallel data collection test with 184 participants, Cronbach's α reliability value of 0.92 was obtained, indicating that the measurement tool is reliable. In the next stage, the researcher proceeded with field data collection.

Data Collection

The online questionnaire distribute to participants from February to March 2023. Participants who were willing to participate first filled in demographic data such as initials, age, education level, type of company, total length of employment, current employment in the organisation, and position. The age, gender, tenure and education of the participants were asked in the questionnaire because those data were needed in the innovative work behaviour study (Etikariena, 2018). Then, the participants marked (V) in the willingness field to indicate their willingness to participate in the study. Participants were then asked to complete the I-IWB questionnaire, consisting of 9 items, according to their current perceived conditions in the organisation. At the end of the questionnaire, the researcher provided information regarding rewards for selected participants who were willing to join the lucky draw process.

Data Analysis

The collected data then undergoes data cleansing based on the criteria. The data analysis method used to test validity and reliability in this research used the SEM (Structural Equation Models). The researchers analyzed Confirmatory Factor Analysis (CFA) and analyses the relationships between indicators and latent variables to test the model fit (Hair et al., 2011). Goodness-of-fit (GOF) indicators determine whether the measurement tool fits the data. There are four fit measurement indices: absolute fit index, incremental fit index, parsimonious fit index, and predictive fit indices (Kline, 2011). This study's fit measurement indices include chi-squared, RMSEA, GFI, CFI, NFI, and SRMR (Table 1). These indices are chosen because they provide the most fundamental indication of how well the proposed theory fits the data (Hooper et al., 2008).

Subsequently, reliability testing also conducted by examining Cronbach's Alpha values. This technique is chosen as one of the most used methods for testing the reliability of a measurement tool (Bryman & Bell, 2011). The reliability criterion used is based on Kaplan and Saccuzzo (2005), which states that a measurement tool is considered reliable when the Cronbach's Alpha value is ≥ 0.70.

We also did the invariance analysis measurement that compares the Multi Group Confirmatory Factor Analysis (MGCFA) and Differentiate Item Functioning (DIF). The basic idea in measuring invariance or non-invariance in a latent class model is simple: first, a model in which all response probabilities are allowed to vary across groups is examined; then, a model where these probabilities are constrained to be equal is examined, and the two models are compared using a difference likelihood ratio test. If the constrained model is significantly worse than the unconstrained model, invariance across classes is lacking. At the last stage, individual indicators are examined for uniform and non-uniform DIF. The difference between the two types of DIF exists in the shape of the item response functions. Uniform DIF exists when the item response pattern is related to the group at all levels of the latent class variable and does not depend on the latent ability level, whereas in non-uniform DIF, the IRFs across groups cross (Tsaousis, Sideridis, & AlGhamdi, 2020). The entire analysis process uses IBM SPSS Statistics Version 26 and Jamovi for CFA and other analyses.

Results and Discussion

Demographic Result

A total of 1.257 employees filled out the questionnaire distributed online via Google Form, and only 1,071 respondents could proceed with the analysis process because of the outliers or inconsistent data. In terms of gender, the most respondents were female (N=564; 53%). Furthermore, most respondents were between 26 - 30 years old (N=469; 43%). Moreover, many participants held bachelor's degrees (N=801; 74%) and originated from marketplace/e-commerce companies (N=518; 48%). Additionally, most

participants had total work experience ranging from 3-5 years (N=452; 42%), with their current tenure in the company being between 1 - 2 years (N=546; 51%). The workforce predominantly comprised employees in Staff positions (N=705; 65.8%). The complete demographic result is shown in Table 1.

Table 1. Participant Demographic Characteristics

Demographic Characteristic		Number	Percentage
Gender	Male	507	47
	Female	564	53
Age	20 – 25 years	310	29
	26 – 30 years	469	43
	31 – 35 years	195	18
	36 – 40 years	71	6
	41 – 45 years	17	2
	46 – 50 years	8	0.7
	51 – 55 years	1	0.09
Education	SMA/SMK/Sederajat	76	7.1
	Diploma	149	13.9
	Sarjana 1	801	74.8
	Sarjana 2	45	4.2
Company	Edtech	269	25
	Online Training Provider	18	1.6
	Marketplace / E-commerce	518	48
	Manufacture	266	25.4
Total Work Experience	< 1 year	49	4.6
	1 – 2 years	312	29.1
	3 – 5 years	452	42.2
	6 – 10 years	200	18.7
	> 10 years	58	5.4
Current Tenure	<1 year	96	9
	1 – 2 years	546	51
	3 – 5 years	328	30.6
	6 – 10 years	77	7.1
	> 10 years	24	2.3

http://journal.uinjkt.ac.id/index.php/jp3i

Demographic Characteristic		Number	Percentage
Current Position	Staff	705	65.8
	Supervisor	261	24.3
	Manager	92	8.5
	Others	13	1.3
Total		1.071	100

Source: Personal Data (2025).

Construct Validity Test

To analyse the construct validity, first, we reffered the model fit criteria mentioned in Table 2:

Table 2. Model Fit Criteria

Fit Indices	Estimation	Criteria
Chi-Square X^2 (p-value)	$p > 0.050$	Good fit
Root Mean Square of Approximation (RMSEA)	< 0.030 < 0.070	Excellence fit Good fit
Tucker - Lewis Index (TLI)	> 0.900	Good fit
Comparative Fit Index (CFI)	> 0.950 ≥ 0.900	Good fit Moderate fit

Sources: Hopper, Coughlan, & Mullen (2008).

For analysing the scale's model fit, we did the confirmatory factor analysis. The researchers conducted Chi Square, CFI, TLI, and RMSEA tests as listed in Table 3.

Table 3. Confirmatory Factor Analysis Result

Chi-Square	CFI	TLI	RMSEA	RMSEA 90% CI	
				Lower	Upper
101, 24; $p < 0.01$	0.970	0.955	0.054	0.044	0.066

Source: Author's work.

Based on the analysis results, The Intention for Innovative Work Behavior Scale had an acceptable fit with indices meeting the three out of four cut-off criteria (RMSEA = 0.054 < 0.070; 90% CI 0.044, 0.066), CFI (CFI = 0.970 > 0.950), TLI = 0.955 > 0.900), only the Chi-Square coefficient is not fit with the criteria (X^2(101,24; $p < 0.01$). Eventhough, in general, it can be concluded that combinational rules based on the two-index presentation strategy committed less sums of Type 1 and Type II error rates than the single-index presentation strategy (Hu & Bentler, 1999). So, by using four determining criteria that have been carried out by researchers, this has fulfilled the reason for not determining conclusions based on only one criterion. In addition, Hopper, Coughlan, & Mullen (2008) suggest that based on Bentler and Bonnet, (1980); Jöreskog and Sörbom (1993), the Chi-Square statistic is in essence a statistical significance test that is sensitive to sample size, which means that the Chi-Square statistic nearly always rejects the model when large samples are used. Large sample is 200 participants or more (Field, 2009). This study has 1,071 partipants, thus this study has large sample to be tested. Regarding

this situation, Hu & Bentler (1999) recommend CFI, TLI and RMSEA to test the model fit in rather than the Chi Square criteria.

For the RMSEA coefficient, this study can not meet the excellent fit criteria at 0.030. However, with (RMSEA = 0.054 < 0.070; 90% CI 0.044, 0.066) it still has good fit criteria, because it still below 0.070 regarding Hopper, Coughlan, & Mullen criterion (2008). The other study by Bagheri et, al (2014) used the minimum criteria of RMSEA at 0.060. This condition refers to Hu and Bentler (1999) that used the 0.060 coefficient for the minimum criteria for the RMSEA. Thompson (2004) that mention that the *root-mean-square error of approximation* (RMSEA) will estimate how well the model parameters will do at reproducing the population covariances. A model estimated to reproduce exactly the population covariances would have an RMSEA of zero. Values of roughly 0.060 or less are generally taken to indicate reasonable model fit. Therefore, the model fit meet the minimum criteria that common used in the prior study. Thus, I-IWB is a unidimensional construct consistent with the theory of innovative work behaviour by Scott & Bruce (1994).

Measurement of Invariance

Multi-item surveys are frequently used to study scores on latent factors, such as attitude, human values or behaviour. Such studies often compare specific groups of individuals or residents of different countries at one or multiple points (i.e., a cross-sectional or a longitudinal comparison or both). If latent factor means are to be meaningfully compared, the measurement structures of the latent factor and their survey items should be stable; that is, "invariant." Van De Schoot et al (2015). Therefore, this study that develops the scale to measure someone's intention to innovate also relates to the latent factors. Therefore, we continue measuring invariance, and in this study, we compare participants' gender, age, and educational background. The study by GonzaÂlez-Blanch, et al. (2018) divides the group into gender, age, marital status, level of education, and employment situation. Another study by Dewi and Widhiarso (2010) divided only by gender. This decision depends on the focus of the study and the distribution of the data obtained in the research. Therefore, for this situation, we grouped the participants into gender, age and educational background of the participants.

There are some methods of analysis to measure invariance between groups. The common method that is used to analyse is Multi-Group Confirmatory Factor Analysis (Alatli, 2020).

http://journal.uinjkt.ac.id/index.php/jp3i

Table 4. Invariance analysis with MGCFA for the Gender, Age and Education

Model	χ^2 (df)	$\Delta\chi^2$ (df)	p	CFI	TLI	RMSEA	ΔCFI	ΔRMSEA	AIC	BIC	Decision
Gender											
Configural Invariance	140.93 (54)			0.97	0.97	0.06			22022	22290	Baseline Model
Metric Invariance	145.42 (62)	4.49 (8)	0.81	0.97	0.96	0.05	0.00	0.00	22010	22239	Supported
Scalar Invariance	152.52 (70)	7.10 (8)	0.53	0.97	0.97	0.05	0.00	0.00	22001	22190	Supported
Age											
Configural Invariance	182.35 (81)			0.96	0.95	0.06			22026	22429	Baseline Model
Metric Invariance	209.29 (97)	26.94 (16)	0.14	0.96	0.95	0.06	0.00	0.00	22021	22345	Supported
Scalar Invariance	222.81 (113)	13.52 (16)	0.63	0.96	0.96	0.05	0.00	0.00	22003	22247	Supported
Education											
Configural Invariance	243.01 (108)			0.95	0.93	0.07			22001	22539	Baseline Model
Metric Invariance	275.41 (132)	32.40 (24)	0.12	0.94	0.94	0.06	0.00	0.00	21986	22404	Supported
Scalar Invariance	304.48 (156)	29.07 (24)	0.22	0.94	0.96	0.06	0.00	0.00	21967	22265	Supported

Source: Personal Data (2025).

From the multi-group CFA analysis, there are no significant issues regarding the response of each group in this study. There was no differentiated item functioning according to gender, age, or educational background. The Scalar Invariance for gender is χ^2 (df) = 152.52 with CFI = 0.970, TLI = 0.970, and RMSEA = 0.05 that supported the invariance latent variable. The Scalar invariance for age is χ^2 (df) = 222.81 with CFI = 0.960, TLI = 0.960, and RMSEA = 0.054, which is supported the invariance latent variable. The Scalar invariance for education is χ^2 (df) = 304.48 with CFI = 0.940; TLI = 0.960, RMSEA = 0.060 supported by the invariance latent variable. Therefore, this results support that from the MGCFA analysis, there is no differentiate item functioning (DIF) for I-IWB scale.

To confirm the results from the MGCFA analysis, we also analysed the differentiated item functioning (DIF) Mantel-Haenszel analysis. We conducted the DIF analysis to find possible biases in item functioning across different demographic groups to ensure that the measurement really measures the intended constructs fairly and accurately for all participants (Scott et al., 2010). This method allows us to detect discrepancies in responses that may arise. Below is the DIF test result.

Table 5. Invariance Analysis using DIF Mantel-Haenszel base in Gender, Age and Education

Number of Item	Gender					Age					Education				
	Stat	p	OR	ΔMH	DIF Class	Stat	p	OR	ΔMH	DIF Class	Stat	p	OR	ΔMH	DIF Class
Item 1	0.64	0.42	0.79	0.56	A	0.24	0.63	0.84	0.41	A	0.37	0.54	1.26	-0.55	A
Item 2	1.80	0.18	1.54	-1.01	B	1.42	0.23	0.65	1.01	B	0.00	0.97	1.05	-0.12	A
Item 3	0.56	0.45	0.74	0.69	A	2.57	0.11	0.51	1.60	C	0.00	0.98	1.06	-0.14	A
Item 4	0.88	0.35	1.38	-0.76	A	0.00	1.00	1.05	-0.12	A	1.11	0.29	1.51	-0.96	A
Item 5	2.24	0.13	0.64	1.05	B	1.79	0.18	1.53	-1.00	B	5.89	0.02*	0.36	2.38	C
Item 6	0.29	0.59	1.22	-0.46	A	0.00	0.99	0.95	0.12	A	0.03	0.87	0.88	0.30	A
Item 7	0.34	0.56	1.24	-0.51	A	0.07	0.79	0.86	0.35	A	0.42	0.52	0.74	0.72	A
Item 8	0.09	0.76	0.89	0.28	A	0.01	0.91	1.01	-0.03	A	0.76	0.38	1.42	-0.83	A
Item 9	0.00	1.00	1.05	-0.11	A	5.83	0.02*	2.30	-1.96	C	0.10	0.75	1.22	-0.47	A

Notes: Effect size code: 'A': negligible effect; 'B': moderate effect; 'C': large effect.

Source: Personal Data (2025).

The DIF Mantel-Haenszel analysis indicates varying degrees of differential item functioning (DIF) across gender, age, and education groups. Most items exhibit a negligible effect size (Class A), suggesting minimal bias across groups, especially regarding gender bias. However, moderate DIF (Class B) is observed in the gender group for the item 2 which relates to asking for approval regarding innovative ideas (OR = 1.54, ΔMH = -1.01), and item 5, addressing making direct supervisors enthusiastic about new ideas (OR = 0.64, ΔMH = 1.05). Similarly, in the education group, item 5 shows a statistically significant result ($p < 0.05$), suggesting that respondents from different educational backgrounds interpret or respond to this item differently (OR = 3.46, ΔMH = 2.38). In the age group, a statistically significant result is found in Item 9 ($p < 0.05$), indicating potential disparities in how age groups perceive the item related to making direct supervisors enthusiastic about new ideas (OR = 2.30, ΔMH = -1.96). Overall, while most items show negligible DIF, the results highlight items 2, 5, and 9 that may need further attention in terms of measurement equivalence across demographic groups.

Reliability Test

Following the process, the researchers analysed the reliability of the scale that calculations for the total score and each stage are listed in Table 6.

Table 6. Reliability Test Result

Scale	N	Cronbach's α
Intention to Innovative Work Behaviour (I-IWB)	1.071	0.84
Intention to Idea Generation Stage	1.071	0.62
Intention to Idea Promotion Stage	1.071	0.70
Intention to Idea Realization Stage	1.071	0.62

Source: Personal Data (2025).

Then, referring to the theory that innovative work behaviour is a unidimensional construct (Janssen, 2000; Scott & Bruce, 1994), the calculation of reliability tests was conducted overall and per stage to check whether each stage has good reliability or not, with the results listed in Table 7. Based on Table 7, with Kaplan and Sacuzzo's reliability minimum criteria, that is, alpha coefficient $\alpha > 0.700$ (2005), it can be observed that the I-IWB measurement tool, in terms of total scores, exhibits good reliability ($\alpha = 0.840$). This indicates that the items within the I-IWB measurement tool consistently measure the intended construct, namely the intention for innovative work behaviour. For further analysis, we also examined the reliability of each I-IWB item to ascertain whether this measurement tool could be multidimensional or should remain unidimensional. The results showed that only the intention for the idea promotion stage has a good reliability ($\alpha = 0.700$). The reliability coefficient for intention for the idea generation stage and intention for the idea realisation stage have low-reliability coefficients (each $\alpha = 0.620$). Therefore, it will be a potential issue if we analyse the data from each stage. Besides, this result also supports the theory of innovative work behaviour from Janssen (2000) and Scott and Bruce (1994), who propose that if the (intention for) innovative work behaviour is multi-stages, it will be seen as a unidimensional construct.

http://journal.uinjkt.ac.id/index.php/jp3i

Table 7. Reliability Test Result for each item

No	Item	Corrected inter-item total correlation		Cronbach's Alpha if the item is deleted	
		Construct	Stage	Construct	Stage
1	*Saya akan mencari metode kerja yang baru di pekerjaan saya*	0.51	0.42	0.83	0.54
2	*Saya akan mencari persetujuan dari pihak terkait atas ide-ide inovatif yang saya sampaikan.*	0.60	0.54	0.82	0.58
3	*Saya akan mewujudkan ide-ide inovatif menjadi aplikasi atau program yang dapat dilaksanakan.*	0.51	0.43	0.83	0.52
4	*Saya akan memperkenalkan ide inovatif dengan cara yang sistematis.*	0.54	0.41	0.82	0.55
5	*Saya akan membuat pimpinan saya antusias dengan ide-ide baru yang saya sampaikan.*	0.54	0.49	0.82	0.64
6	*Saya akan memberikan solusi yang belum pernah digunakan orang lain untuk mengatasi masalah yang ada.*	0.54	0.41	0.82	0.55
7	*Saya akan mencari ide baru untuk mengembangkan tugas atau organisasi saya.*	0.55	0.46	0.82	0.48
8	*Saya akan mencari dukungan dari pihak terkait untuk ide inovatif yang saya sampaikan.*	0.58	0.52	0.82	0.60
9	*Saya akan mengevaluasi kegunaan ide yang pernah saya sampaikan.*	0.58	0.45	0.82	0.49

Source: Personal Data (2025).

The above data suggested that the correlation values of each item have better reliability when included in the total score as a construct compared to when broken down into each stage. This also indicates that the I-IWB items are more reliable when the items are parts of a unidimensional intention construct for innovative work behaviour.

Table 8. Factor Loading Coefficient for each Item

I-IWB Item	Estimates	Std. Estimates	SE	z	p
Item 1	0.505	0.559	0.028	18.4	<.001
Item 2	0.585	0.669	0.026	22.9	<.001
Item 3	0.456	0.559	0.025	18.3	<.001
Item 4	0.486	0.590	0.025	19.6	<.001
Item 5	0.535	0.601	0.027	20.0	<.001
Item 6	0.506	0.592	0.026	19.6	<.001
Item 7	0.499	0.604	0.025	20.1	<.001
Item 8	0.573	0.645	0.026	21.9	<.001
Item 9	0.55	0.641	0.025	21.7	<.001

Source: Personal Data (2025).

After confirming that the unidimensional model is accepted (fits the data), we also calculated and confirmed the parameter values obtained and tested their significance. Usually reported in the form of a table of factor loading coefficients for each item, accompanied by standard error, z or t value, and probability (significance). As usual, a t or z (absolute) value greater than 1.96 is a significant minimum at the 5% level. It is best that what is reported here is a "standardised" coefficient so that its magnitude can be compared directly (Umar & Nisa, 2020). Table 6 shows that each item in the I-IWB Scale has met the criteria, in the standardised coefficient between 0.559 – 0.669 with $p < .001$.

Discussion

The study aimed to develop the intention for innovative work behaviour (I-IWB) measurement among Indonesian employees. The development study of the Intention for Innovative Work Behaviour (I-IWB) scale indicates its reliability and validity, making it suitable for studies that will utilise intention to display innovative work behaviour as its variable. As a scale, it is unidimensional because all three stages must emerge to explain the intention of innovative work behaviour. Statistical analysis shows that when analysed separately, the reliability for each stage becomes lower than the reliability of the whole scale. Kaplan and Sacuzzo (2005) proposed the reliability minimum criteria for alpha coefficient $\alpha > 0.70$. These criteria supported by DeVellis (2017) that the ranges for research scales are as follows: below .60, unacceptable; between .60 and .65, undesirable; between .65 and .70, minimally acceptable; between .70 and .80, respectable; between .80 and .90, very good; and much above .90. This shows that the intention to idea generation stage and intention to idea realisation stage that has $\alpha = 0.62$ will have potential issues for the reliability. Therefore, this study supports the construct of intention for innovative work behaviour to be measured on a unidimensional scale. Therefore, the three stages of intention behaviour simultaneously consist of intention to display idea generation, idea promotion, and idea realisation. This research has contributed to the development of new measurements of intention for innovative work behaviour that were previously developed in English or other languages, providing the possibility of potential bias if given to employee participants in Indonesia. As Ajzen (1991) mentioned, when the activity is uncertain in its implementation due to its high risk, intentions play a more significant role in predicting changes in a person's behaviour. Therefore, not only by measuring previous behaviour, we could also integrate the one's intention in predicting the emergence of innovative work behaviour.

http://journal.uinjkt.ac.id/index.php/jp3i

In this study, the invariance analysis used gender, age, and educational background to identify group participants. This approach is a consideration for researchers because, from several previous studies, it is known that gender bias can be one thing that must be considered when developing a measuring instrument or test tool. A study by Parker et al. (2016) on medical students found that the issue of gender bias needs to be taken into consideration so that it must be ensured in the measuring tools used in evaluating medical students' learning outcomes. In their review, Dewi and Widhiarso (2010) also stated that there was a need for a separate analysis of the Social Support Scale that they compiled to ensure that the items they developed could pay attention to which items might contain gender bias because the construct of social support is more easily felt by female participants than male participants. Therefore, in developing the I-IWB scale, invariance analysis was also carried out based on participants by gender groups. A study by Etikariena (2018) showed no significant difference in innovative work behaviour between male and female participants. Therefore, this study also supports the result. This is different from a study conducted by Luksyte, et al (2018), which found that the construct of innovation was more nuanced for men than women. In the DIF analysis, we found that two items have significant differences in demographic groups. Therefore, future research should exercise caution when utilising this measurement, carefully examining potential gender and age biases to ensure fairness and accuracy in assessment outcomes. This can also be checked using the item response theory (IRT)-based approaches to get more comprehensive results (Scott et al., 2010). Similar conditions can be found in health-related quality of life assessments (Scott et al., 2010), which also contain items with DIF across age, gender, and cultural groups. However, this measurement is still utilised due to the high overall reliability (Scott et al., 2010). However, from the MGCFA analysis, there are no special considerations for using this scale because there are no significant DIFs to be considered.

However, the study also needs some consideration. First, the participants in this study are limited to employees from the manufacturing, Edtech, and e-commerce sectors, which could not represent the overall business sector. This refers to the convenience sampling methods with limitations to getting a more representative sample. Hence, the researchers may list the sectors to ensure the representation of overall sectors in Indonesia. This method will require more effort to gain participants, but that method increases the possibility of generalising the results of the study (Gozar et al., 2022). Second, the content validity checking process in this study was reviewed by only one expert. Although we initially approached two experts, the second reviewer provided no relevant feedback. As a result, we decided to exclude the second reviewer's input from the process. This action aligns with DeVellis (2017), who emphasised the importance of exercising caution during the review process and carefully making decisions, particularly if the process does not improve the items or the scale. We recommend that future research allocate more time and gather comprehensive information about the experts' backgrounds to enhance the review process. The validity of the results may lack objectivity if only one reviewer is involved.

Additionally, further exploration is needed to assess the scale's internal reliability. Regarding the scale being valid and reliable, the two items with different responses among the different genders and ages were still considered to be used as its purpose (Dewi & Widhiar, 2010). Therefore; in this study, the scale purpose is to measure the intention to innovate work behaviour. However, the researchers must caution the potential bias caused by the gender or age response. This study suggests that the researchers consider this situation and continue the analyses with invariance analysis. Even though the conventions and reporting on measurement invariance are still in flux, researchers are often left with limited understanding and inconsistent advice (Putnik & Bornstein, 2016). For other limitations, we suggest that future research employ concurrent validity by correlating the results with other scales (Mohajan, 2017), such as the Innovative Work Behaviour Scale, to ensure validity. Another approach is to measure composite reliability, which provides the actual score of the coefficient reliability. This metric reflects the total actual score variance relative to the total scale score variance (Brunner & Süß, 2005).

Conclusion

The I-IWB Scale has been validated as a reliable and effective tool for measuring the intention to engage in innovative work behaviour. The results confirm that the I-IWB construct is unidimensional, as demonstrated by the model fit outcomes. Factor loading for each item also supports this result. This finding supports this study's original theory, reinforcing that the intention for innovative work behaviour is a unidimensional construct. Furthermore, the I-IWB Scale proves valuable for assessing employees' intentions to exhibit innovative work behaviour. It provides actionable data for organisations and HR departments to enhance and foster innovative work behaviour among employees, particularly in the Indonesian context.

Acknowledgment

We would like to express our sincere gratitude to all the participants in this study.

Conflict of Interest

All authors do not have any conflicts of interest to declare.

Authors Contribution

AE proposed the main ideas, led the research process, reviewed the results and interpretation, wrote the manuscript and checked the English grammar. AHK, AAI, DNS, and RH conducted literature studies, devised the research methodology, collected data, and wrote the manuscript. ZRS performed data analysis and translated the manuscript into English.

References

Ajzen, I. (1991). The Theory of Planned Behavior. *Organisational Behavior and Human Decision Processes*, 50(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-T

AlEssa, H. S., & Durugbo, C. M. (2021). Systematic review of Innovative Work Behaviour Concepts and contributions. *Management Review Quarterly*, 72(4), 1171–1208. https://doi.org/10.1007/s11301-021-00224-x

Alatli, B. (2020). A Systematic review of research articles on measurement invariance in education and psychology. *International Journal of Assessment Tools in Education*. 2020, 7(4), 607–630. https://doi.org/10.21449/ijate.738560

Andrade, C. (2020). Sample Size and its Importance in Research. *Indian Journal of Psychological Medicine*, 42(1), 102–103. https://doi.org/10.4103/IJPSYM.IJPSYM_504_19

Bagheri Z, Jafari P, Tashakor E, Kouhpayeh A, & Riazi H. (2014). Assessing whether measurement invariance of the KIDSCREEN-27 across child-parent dyad depends on the child gender: a multiple group confirmatory factor analysis. *Global Journal Health Science.* 6(5):142-53. https://10.5539/gjhs.v6n5p142.

Banholzer, M., Birshan, M., Doherty, R., & LaBerge, L. (2023). Innovation: Your solution for weathering uncertainty. *McKinsey & Company*. https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/innovation-your-solution-for-weathering-uncertainty

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

Bos-Nehles, A., Renkema, M., & Janssen, M. (2017). HRM and Innovative Work Behaviour: A Systematic Literature Review. *Personnel Review*, 46(7), 1228-1253. https://doi.org/10.1108/PR-09-2016-0257.

Boyles, M. (2022). Innovation in Business: What it is & Why it's so Important. Harvard Business School Online. https://online.hbs.edu/blog/post/importance-of-innovation- in-business

Brunner, M., & Süß, H.-M. (2005). Analysing the Reliability of Multidimensional Measures: An Example from Intelligence Research. *Educational and Psychological Measurement*, 65(2), 227–240. https://doi.org/10.1177/0013164404268669

Bryman, A. & Bell, E. (2011*). Business Research Methods*. 3rd Edition, Oxford University Press, Oxford. https://doi.org/10.1093/hebz/9780198869443.003.0012

Cohen, J. (2013). *Statistical Power Analysis for The Behavioural Sciences*. Routledge Academic. https://doi.org/10.4324/9780203771587

Dewi, A.N & Widhiar, W. (2010). Keberfungsian butir diferensial dalam pengukuran dukungan sosial. *Public Release*. Fakultas Psikologi UGM.

DeVellis, R. F. (2017). Scale development: Theory and applications (4th ed.). Thousand Oaks, CA: Sage. https://doi.org/10.2307/2075704

Etikariena, A., & Muluk, H. (2014). Correlation between organisational memory and innovative work behaviour. *Hubs-Asia*, 18(2): 77-88. https://doi.org/10.7454/mssh.v18i2.3463

Etikariena, A. (2018). Perbedaan Perilaku Kerja Inovatif Berdasarkan Karakteristik Individu Karyawan. *Jurnal Psikologi*, *17*(2), 107–118. https://doi.org/10.14710/jp.17.2.107-118

Field, A. (2009). *Discovering statistic and using SPSS (and sex and drugs and rock'n roll)*. 3rd Edition. Sage Publication Ltd. London.

Golzar, J; Noor, S & Tajik, O. (2022). Convenience Sampling. Sampling Method | Descriptive Research. *International Journal of Education and Language Study*. https://www.ijels.net/article_162981_06fc448b8914d42316b8d44829c88188.pdf

GonzaÂlez-Blanch C, Medrano LA, Muñoz-Navarro R, RuõÂz-RodrõÂguez P, Moriana JA, Limonero JT. (2018) Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary care patients in Spain. *PLoS ONE* 13(2):e0193356. https://doi.org/10.1371/journal.pone.0193356

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19(2), 139–152. https://doi.org/10.2753/mtp1069-6679190202

Hooper, D., & Coughlan, J. & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. In *The Electronic Journal of Business Research Methods*, 6, 53-60.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Janssen, O. (2000). Job demands, perceptions of effort-reward fairness and innovative work behaviour. *Journal of Occupational and Organizational Psychology*, *73*, 287-302. https://doi.org/10.1348/096317900167038

Kanter, R. M. (1988). When a thousand flowers bloom: Structural, collective, and social conditions for innovation in organisations. *Knowledge Management and Organisational Design*, 93–131. https://doi.org/10.1016/b978-0-7506-9749- 1.50010-7

Kaplan, R.M & Saccuzo, D.P. (2005). *Psychological Testing, Principles, Applications. And Issues*. (6th ed.). Thomson Wadsworth, Belmont USA. https://doi.org/10.1007/978-3-642-39229-0_57

Lu, C.C., & Luh, D.-B. (2013). Innovative Behavioural Intention and Creativity Achievement in Design: Test of an Integrated Model. *Lecture Notes in Computer Science*, 535–544. https://doi.org/10.1007/978-3-642-39229-0_57

Luksyte, A; Unsworth, K.L & Avery, D.R.(2018). Innovative work behaviour and sex-based stereotypes: Examining sex differences in perceptions and evaluations of innovative work behaviour. *Journal of Organizational Behaviour*. Vol.39 (3). https://doi.org/10.1002/job.2219

Mohajan, H. K. (2017). Two Criteria for Good Measurements in Research: Validity and Reliability. Annals of Spiru Haret University*, Economic Series*, 17, 59-82. https://doi.org/10.26458/1746

Muchiri, M. K., McMurray, A. J., Nkhoma, M., & Pham, H. C. (2020). Mapping antecedents of Innovative work behaviour: A conceptual review. *The Journal of Developing Areas*, 54(4). https://doi.org/10.1353/jda.2020.0047

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, 599-620. https://doi.org/10.1207/S15328007SEM0904_8

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behaviour predicts future behaviour. *Psychological Bulletin*, 124(1), 54–74. https://doi.org/10.1037/0033-2909.124.1.54

Parker, RB; Parker, PD; Larkin, T & Cockburn, J. (2016). A psychometric evaluation of the gender bias in medical education scale. *BMC Medical Education*. (2016) 16:251. https://10.1186/s12909-016-0774-2

Putnik, DL & Bornstein, M.H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Dev Rev*. 2016 September ; 41: 71–90. https://10.1016/j.dr.2016.06.004

Salvador, F., & Sting, F.J. (2022). *How your company can encourage innovation from all employees. Harvard Business Review.* https://hbr.org/2022/09/how-your-company-can-encourage-innovation-from-all-employees

Scott, S. G. & Bruce, R. A. (1994). Determinants of innovative behaviour: A path model of individual innovation in the workplace. *Academy of Management Journal*, *37*, 580-607. https://doi.org/10.2307/256701

Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., de Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M. A., & Sprangers, M. A. G. (2010). Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. *Health and Quality of Life Outcomes,* 8(81). https://doi.org/10.1186/1477-7525-8-81

Srirahayu, D. P., Ekowati, D., & Sridadi, A. R. (2023). Innovative work behaviour in public organisations: A systematic literature review. *Heliyon*, 9(2). https://doi.org/10.1016/j.heliyon.2023.e13557

Su W, Lin X & Ding H (2019) The Influence of Supervisor Developmental Feedback on Employee Innovative Behaviour: A Moderated Mediation Model. *Frontier Psychology*. 10:1581. https://doi.org/10.3389/fpsyg.2019.01581

Thompson, B. (2004). *Exploratory and confirmatory factor analysis : Understanding concept & application*. Published by American Psychological Association.

Tsaousis, I, Sideridis, G.D & AlGhamdi, H.M (2020) Measurement invariance and differential item functioning across gender within a latent class analysis framework: Evidence from a high-stakes test for university admission in saudi arabia. *Frontier Psychology*. 11:622. https://10.3389/fpsyg.2020.00622

Umar, J & Nisa, Y.F (2020). Uji validitas konstruk dan pelaporannya. *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)*, 9(2), 2020, 1-11. http://dx.doi.org/10.15408/jp3i.v9i2.16964

Van De Schoot R, Schmidt P, De Beuckelaer A, Lek K and Zondervan-Zwijnenburg M (2015) Editorial: Measurement Invariance. *Frontier Psychology*. 6:1064. https://10.3389/fpsyg.2015.01064

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care*, 8(3), 1-13. https://doi.org/10.33151/ajp.8.3.93