JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)
DOI: http://dx.doi.org/10.15408/jp3i.v14i1.39333
http://journal.uinjkt.ac.id/index.php/jp3i

The Development of The Situational Judgement Test (SJT) for Workforce Agility

Avin Fadilla Helmi[1], Ramadhan Dwi Marvianto[1], Arif Budiman Al Fariz[1], Indahinsani Purnasari Anggoro[1], Wahyu Jati Anggoro[1]

Faculty of Psychology, Gadjah Mada University, Indonesia[1]

E-mail: Ramadhan.dwi.m@ugm.ac.id

Abstract

This research aimed to develop workforce agility measuring tool with the situational judgment test (SJT) format and evaluate its psychometric properties. This research included 886 respondents with higher education. The data were analyzed using confirmatory factor analysis (CFA) and multigroup confirmatory factor analysis (MGCFA). The CFA results showed that 18 items have satisfactory psychometric properties, thus forming a fit factor structure and providing valid evidence based on internal structure. Besides, the MGCFA results showed that this measuring instrument has measurement invariance at the strict invariance level to be used as validity evidence based on test consequences. The Omega reliability of this instrument was estimated, and the required criteria were met. The implication is that this measuring instrument has at least three of the five sources of validity to allow solid interpretations of the measuring results.

Keywords: Workforce agility, situational judgment test, multigroup confirmatory factor analysis, validity, psychometrics.

Abstrak

Penelitian ini bertujuan untuk mengembangkan alat ukur workforce agility dalam format situational judgement test (SJT) serta melakukan evaluasi properti psikometris dari alat ukur tersebut. Penelitian ini mengikutsertakan 886 responden yang memiliki pendidikan tinggi. Data yang didapatkan dianalisis menggunakan confirmatory factor analysis (CFA) dan multigroup confirmatory factor analysis (MGCFA). Hasil CFA menunjukkan bahwa 18 butir memiliki propertis psikometris yang memuaskan sehingga membentuk struktur faktor yang fit dengan data dan menjadi bukti validitas berdasarkan struktur internal (validity based on internal structure). Selain itu, MGCFA menunjukkan bahwa alat ukur ini memiliki measurement invariance dalam tataran strict invariance sehingga hasil ini dapat menjadi validity evidence based on test consequence dari alat ukur ini. Reliabilitas Omega diestimasi untuk instrument ini dan hasilnya sesuai dengan kriteria yang disyaratkan. Implikasinya, alat ukur ini memiliki setidaknya dua dari lima sumber validitas sehingga interpretasi dari alat ukur ini dinilai kuat.

Keywords: *Workforce agility, situational judgment test, multigroup analisis faktor konfirmatori, validitas, psikometris.*

Introduction

Changes in organizations are a common thing. The changes may happen due to internal factors such as downsizing, team addition or reshuffle, and external factors such as new government regulations or a global pandemic (Aamodt, 2010). One example of a source of change is the unpredictable Coronavirus Disease 2019 (Covid-19) pandemic. The pandemic caused marked changes in every aspect of life, including learning and work, business, and consumer behaviors (Alshurideh et al., 2021). Based on reports from McKinsey and Company, among marked changes in the world of work that have occurred in the working system is the change from working from the office to working from home and a combination of working from home and the office (hybrid) (Lund et al., 2021). Other domains of change due to the pandemic in the world of work include virtual teamwork, social distancing and the resulting loneliness, virtual management, and the perceived threat of layoff and becoming unemployed (Kniffin et al., 2021). With unpredictable changes in many domains in organizations, workers must be able to adapt to these changes to remain productive at work (Karaca et al., 2022).

Employee responses to changes are very diverse. They are not always able to adapt quickly. Adaptation includes stages from denial, defense, discarding, and adaptation to internalization (Aamodt, 2010). Those who can quickly adapt to various changes and uncertainties are called agile individuals (Alavi et al., 2014; Petermann & Zacher, 2022). According to Junior and Saltorato (2021), agility is the ability to respond quickly and exploit opportunities by effectively and efficiently configuring resource strategies. Agility at work leads to increased productivity and an ability to handle higher targets effectively (Alavi, 2016). Organizations need agile workers due to the ever-changing work environment and increasing competition. Organizations can enhance employee agility through training programs, fostering an agile culture, or promoting learning (Junior & Saltorato, 2021).

Workforce agility enables us to achieve an agile workplace. According to Petermann and Zacher (2022), workforce agility is significantly and positively correlated with innovative performance, task performance, job satisfaction, and organizational citizenship behavior. In another study, workforce agility positively predicts flexibility, especially external manufacturing flexibility (new product flexibility, mixed flexibility, and volume flexibility) (Alavi, 2016). The antecedents of workforce agility include risk-taking, anticipation and planning, job-related curiosity, learning from past mistakes, job self-efficacy, trust, and ambiguity tolerance (Storme et al., 2020). These findings suggest that workforce agility creates an agile work environment and is important to develop.

Research on agility has a general focus such as workforce agility or agile workforce (Alavi, 2016; Alavi et al., 2014; Breu et al., 2001; Petermann & Zacher, 2022) or specific focus such as agile software and development (Gupta et al., 2019; Misra et al., 2012) and agile supply chain (Shashi et al., 2020; Tarafdar & Qrunfleh, 2016). This present study focused on workforce agility to measure employee agility in the workplace. Workforce agility is indicated by employee competency to be more responsive and more able to adapt to rapid and unpredictable environmental changes (Breu et al., 2001). However, researchers still have different views on defining workforce agility as a specific ability, attitude, or behavior required from a worker in an unpredictable business world (Muduli & Pandaya, 2018). Then, rather than a personality or attribute, workforce agility can be viewed as behavioral performance observable in the workplace through several aspects, namely proactive, adaptive, and resilient behaviors, especially in dealing with unpredictable changes (Sherehiy & Karwowski, 2014).

Research studies on workforce agility have been dominated by those investigating multidimensional workforce agility (Junior & Saltorato, 2021; Muduli & Pandaya, 2018). For example, Petermann and Zacher's (2022) research was on a multidimensional workforce agility measuring tool development. They found ten aspects: acceptance of changes, decision-making, transparency creation, collaboration, reflection, user-centricity, iteration, testing, self-organization, and learning. In addition, four dimensions of workforce agility are interrelated. They are proactivity, flexibility and adaptability, resilience, and competence (Junior & Saltorato, 2021), but they have not been field tested. Meanwhile, other researchers

that develop and measure workforce agility are dominated by models with three dimensions: proactivity, adaptability, and resilience (Alavi, 2016; Cai et al., 2018; Sherehiy & Karwowski, 2014).

Previous developments in workforce agility measuring tools were still limited to using the Likert scale. Some experts argue that the Likert scale is unsuitable for measuring work behaviors because of its susceptibility to social desirability (Hough & Oswald, 2008; Kreitchmann et al., 2019; Lee et al., 2019). On the other hand, the Situational Judgment Test (SJT) item format has a low faking rate (Kasten et al., 2020) and tends to be suitable to be used in employee selection contexts or organizational contexts (Corstjens et al., 2017; Tiffin et al., 2020). Besides, participants' responses to tests with SJT format can also reflect the results of the justification process for behavior that might occur if the context of the situation presented occurred (Olaru et al., 2019). The SJT item format concerning workforce agility construct measuring tools has not been widely developed. While several previous studies on developing measuring instruments focused on the Likert scale test format, the workforce agility measuring tool development in this present study used a multidimensional model with three dimensions using the SJT item format and tested its psychometric properties.

Workforce agility

Research on developing the workforce agility construct has been carried out in various contexts. The workforce agility scale was first empirically tested in an Information Technology (IT) context by developing five characteristics of agile employees, including intelligence (interpreting changes and being responsive to customer and market needs), competencies (quickly developing new abilities and searching for information), collaboration (effectively collaborating with various people), culture (empowering agility to make independent decisions), and information systems (helping IT infrastructure) (Breu et al., 2001).

The construct of workforce agility was empirically researched by Sherehiy (2008) in a general context using work adjustment theory, according to which individuals develop three aspects in dealing with uncertain situations, i.e., proactivity, adaptability, and resilience. Moreover, Sherehiy and Karwowski (2014) provided proactivity, adaptability, and resilience. First, they defined proactivity as the ability to initiate activity that positively affects environmental changes (Griffin & Hesketh, 2003). Then, adaptability refers to the ability to change one's behavior or other people's behavior to environmental changes, such as learning new skills and dealing with new people (Sherehiy, Karwowski & Layer, 2007). Last, resilience refers to the ability to behave effectively when under pressure, face environmental changes, and tolerate stressful or unpredictable situations by developing stress coping (Sherehiy & Karwowski, 2014; Sherehiy, Karwowski & Layer, 2007).

In terms of measurement instruments, the developed workforce agility scale by Sherehiy (2008) consists of three dimensions, each with a different number of items: 9 items for proactivity, 11 for adaptability, and 9 for resilience. The measurement used a Likert scale with seven response options. Furthermore, empirically testing workforce agility was carried out with five items of just one factor (Braun et al., 2017). Braun et al. (2017) state that workforce agility refers to individual skills, such as remaining proactive in overcoming challenges and seeing opportunities in obstacles. Another empirical research on workforce agility measuring instruments used a ten-dimensional multidimensional model. The ten dimensions include accepting changes, decision-making, creating transparency, collaboration, reflection, user-centricity, iteration, testing, self-organization, and learning (Petermann & Zacher, 2022). This measuring tool consists of 30 items, three for each dimension, and uses a 5-response option scale.

In this study, the authors measured workforce agility according to cultural context using a three-dimensional workforce agility model. The three dimensions were proactivity, adaptability, and resilience (Sherehiy, 2008; Sherehiy & Karwowski, 2014). This model is very appropriate to the context of research that considers the construct of workforce agility as an ability and aspect used following the work change theory. This model has been widely tested in various countries and found suitable (Alavi, 2016; Braun et al., 2017; Cai et al., 2018; Sherehiy & Karwowski, 2014).

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

Situational judgment test

The situational judgment test (SJT) has been widely used in various academic and practical contexts. Lievens et al. (2008) described SJT as a measurement method that presents test takers with situations related to their work and response options related to each situation. The test taker's task is to choose which response they will most likely make when faced with the situation. The SJT also presents dilemmas or problems, and the test takers should choose which problem they will prioritize facing. By doing so, test takers will determine which response choice represents their behavior in the given situation.

SJT can be used to evaluate individual performance at knowledge, skills, and abilities (Sorrel et al., 2016), applied skills such as leadership, and basic personality tendencies such as integrity (Christian et al., 2010). In addition, recent research shows that SJT can be used to measure competencies, including employees' basic work abilities (Sala-Roca et al., 2021), teachers' ability to regulate emotions (Koschmieder & Neubauer, 2021), service competence towards patients (Johannsen et al., 2020), general or global technical competence (Andrea et al., 2020), teacher socio-emotional competence (Aldrup et al., 2020), professional competence (Smith et al., 2020), tactical ability-decision making in military personnel (Männiste et al., 2019) and ability-based emotional intelligence (Fajrianthi & Zein, 2017).

Then, the use of SJT includes selection and mapping contexts, such as previous research using SJT to carry out national selection in the fields of pharmacy (Paudyal et al., 2021) and medical schools (Greatrix & Dowell, 2020) in England and prospective pilots at AFQT or Air Force Officers. Qualifying Test (Barron et al., 2021). The popularity of SJT can be used in various domains and for different purposes because SJT has a low risk of test takers faking both good and bad (Kasten et al., 2020). Apart from that, SJT can also enable testing to be carried out in contexts with high levels of diversity (Barron et al., 2021; Juster et al., 2019).

Considering the research above, this study employed the SJT response format to develop a new workforce agility measurement. Furthermore, given that measurements of workforce agility are limited to the scale developed by Sherehiy (2008), the development of workforce agility using the SJT response format could be a novelty presented in this study.

Methods

Test Development Procedures

This research referred to the procedure of Patterson et al. (2015) for developing a workforce agility (WFA) measurement tool. First, this research examined previously developed WFA concepts. The results of this study led to the use of the WFA concept coined by Sherehiy (2008) to develop measurement indicators. This research construct definition was derived from the workforce agility construct developed by Sherehiy (2008) as follows:

1. ***Workforce Agility***

   Workforce behavior emerges as a positive response to adapt quickly and flexibly, thus enabling individuals to face unpredictable changes with certain characteristics, such as having problem-solving and negotiation abilities, creating new ideas, having a positive attitude toward learning new things, being able to work under stress, and being ready to accept new roles and responsibilities.

2. ***Proactive***

   Taking the initiative to carry out activities that can have a positive effect on the changing environment.

3. ***Adaptive***

   Modifying oneself or changing one's behavior to adapt to changes or a new environment.

4. Resilience

Carrying out work efficiently even under pressure due to changing environments or when implemented strategies and solutions fail.

The next step was determining test specifications, from several items, type of SJT scenario, length of answer choices, scenarios, and scoring procedures to the blueprint as the basis for developing the test items.

The development of the measuring instrument continued with item writing by the authors and three experienced associates in SJT writing, producing an initial number of 60 items. The items were then reviewed by a WFA expert and a psychometric expert to ensure that each item contains an appropriate dilemma, has answer choices matching the context's effectiveness level and scoring weight, and substances representing behavioral indicators. This process provided validity evidence based on the latest concept of the American Educational Research Association (AERA) et al. (2014): validity evidence based on test content. The process also produced items that were ready to be tested. The instrument trial results were then analyzed, and the items were arranged to become a WFA SJT measuring tool.

Research Participants

The data used in this study was secondary data owned by Unit Pengembangan Kualitas Manusia (UPKM) of the faculty of Psychology UGM Yogyakarta in the form of assessment results that measured the construct of workforce agility. The use of secondary data and reporting in this research complies with the principles of existing standards (Trzesniewski & Donnellan, 2011). Then, the data used in this research included 886 respondents with an average age (M) = 34 years and SD = 11. The majority of respondents were aged between 20 to 30 years (51.76%), followed by participants aged 41 to 50 years (20.77%), 31 to 40 years (20.10%), and 51 to 60 years (7.15%). The proportion of men and women was not much different, namely 44.95% and 55.05% for men and women, respectively. All respondents have a higher education, with most graduates of master's programs (50.74%), followed by undergraduates (34.05%) and doctoral graduates (14.42%).

Analysis Procedures

The data collected was analyzed using confirmatory factor analysis (CFA) to get valid evidence based on the internal structure of the factor structure that underlies this test. In addition, the results of this analysis were also used to select items with satisfactory psychometric properties. Then, measurement invariance testing on this factor structure was carried out to obtain valid evidence based on test consequences by grouping gender variables. The analysis also produced omega reliability estimates to demonstrate the internal consistency of field trial results using this test. The entire analysis was carried out using Rstudio software (R Core Team, 2023) with the help of the lavaan package (Rosseel, 2012).

Results and Discussion

Results

Item Pool and Review

Initially, this research developed 60 items prepared based on the Workforce Agility Scale (Sherehiy, 2008) that has been adapted into Bahasa Indonesia by a team from the Mind, Brain, and Behavior Psychology Laboratory of the Faculty of Psychology, Universitas Gadjah Mada. All items were then reviewed by one WFA concept expert and one psychometric expert to examine the item's representation of its behavioral indicators and the score sequence for each answer choice. The review process resulted in 19 items accepted without revision and 41 other items accepted with direct revision by both reviewers. This process was proof of validity evidence based on item content. Then, the sixty questions were included in the testing process for the psychometric properties.

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

Descriptive Statistic of Items

First, this research conducted descriptive statistical testing using means, standard deviations (SD), skewness, and proportion of answer choices (option scores 1 to 5) to check the normality of the distribution of each item. The analysis results show that almost all the items are typically distributed (see Appendix 1, Appendix 2, and Appendix 3 with skewness values not exceeding the value of 2, either positive or negative (Kim, 2013). However, several items had high skewness (> 2), so the Confirmatory Factor Analysis in this research used the Robust Maximum Likelihood (MLR; Huber, 1981) estimator.

Confirmatory Factor Analysis

The CFA was carried out with a single-factor model for every 20 items in the adaptive (Model A), proactive (Model C), and resilience (Model F) dimensions using a robust maximum likelihood estimator (Huber, 1981). The analysis results in Table 1 show that models A, C, and E have parameters that do not meet the criteria of CFI and TLI > .95 (Hu & Bentler, 1999) and GFI > .90 (McDonald & Ho, 2002), RMSEA < .06 (Hu & Bentler, 1999), and SRMR < .08 (Hu & Bentler, 1999). Therefore, model modification was carried out by dropping items with factor loading < .40 (Hair et al., 1998) and inter-correlated item residuals.

Table 1. Summary of statistical and goodness of fit indices

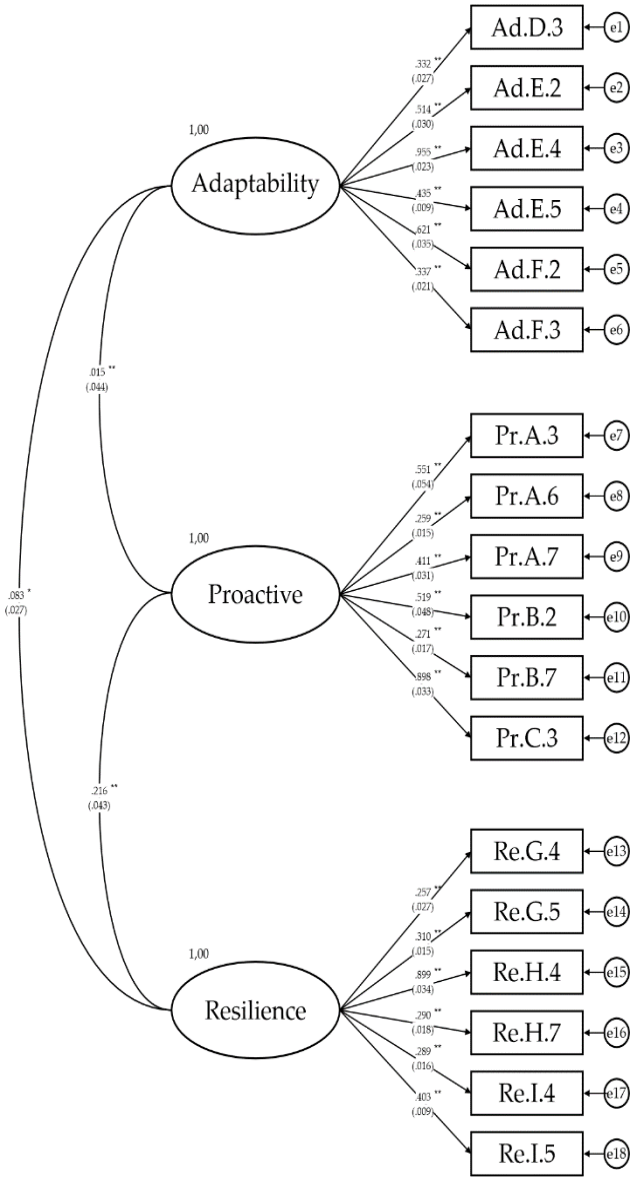
Model	χ^2	df	CFI	TLI	GFI	RMSEA	SRMR	BIC
A	835.372***	170	.787	.761	.918	.066	.054	35,576.984
B	40.537***	9	.981	.969	.986	.063	.033	11,845.565
C	880.984***	170	.772	.745	.930	.069	.043	35,732.444
D	28.156***	9	.976	.960	.990	.049	.029	12,632.269
E	884.382***	170	.765	.737	.923	.069	.052	25,767.350
F	128.539***	9	.942	.903	.958	.122	.045	8,367.464

*Notes. χ^2 = Chi-squared; df = degree of freedom; *** = p < .001; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; GFI = Goodness of Fit Index; RMSEA = Root Mean Square of Error Approximation; SRMR = Standardized Root Mean Square of Residual; BIC = Bayesian Information Criteria.*

Sources: Personal Data (2024).

The model modification produced model B for the adaptive dimension, model D for the proactive dimension, and model F for the resilience dimension. In Table 1, the modification model shows satisfactory CFI, TLI, and GFI values ranging from .903 to .990. Meanwhile, the RMSEA shows a minor change in the adaptive dimension (models A and B). The RMSEA value that becomes smaller in the proactive dimension (.49) suggests that the modification made the model fit better. However, in the resilience dimension, the model modification produces a higher RMSEA while producing a smaller SRMR value. Apart from that, the BIC value in the model after modification shows a significant value below the initial model. Therefore, the modified model fits better than the initial model.

Next, the three modified models were modeled into a broader multidimensional model (see Figure 1). The analysis results showed that this broader model had satisfactory fit indices with CFI =.962, TLI =.956, GFI =.964, RMSEA = .038, and SRMR = .042 and statistical fit using $\chi2 = 302.827$ (df = 132, p < .001). Apart from that, most of the factor loading values from this multidimensional model were satisfactory, i.e., above .40, although several items such as Ad.D.3, Ad.F.4, Re.G.4, and Re.H.7 had a factor loading value within the tolerance limit > .03 (Furr, 2022). On the other hand, several items have a value above .25, namely Pr.A.6, Pr.B.7, Re.G.4, and Re.I.4. Furthermore, the correlations between dimensions showed that the adaptive dimension had an insignificant correlation with the proactive dimension (r = .015; p > .05) but had a significant correlation even though relatively small, r = .083 (p< .05), with the resilience dimension, while the proactive dimension had a significant correlation with the resilience dimension (r= .216; p< .01).



Sources: Personal data (2024).

Figure 1. Factor structure of multidimensional SJT Workforce Agility

Reliability Estimation

CFA analysis produces reliability values using the McDonald's omega (ω; McDonald, 1999) formula estimation. The analysis yielded reliability coefficients of .760, .700, and .760 for the adaptive, proactive, and resilience dimensions.

Measurement Invariance between Male and Female

The measurement invariance testing process used model testing with certain restrictions on the male and female gender subgroups. This process is summarized in Table 2. First, the testing with a single-group solution for each gender showed that the factor structure in both models had a good fit. Then, configural invariance testing showed satisfactory results where this model met the cutoff for all indices (CFI). Furthermore, this study provided constraints on the factor loadings to make the two subgroups equal, resulting in a satisfactory fit index, and the change in chi-square was not significant (p = .972). This means that this measuring instrument has metric invariance. Subsequently, this study added constraints to equalize the intercept between subgroups, resulting in a satisfactory fit index, and the chi-squared change was not significant (p = .481). This result indicated that this measuring instrument has

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

scalar invariance. In addition, this study added constraints to equalize the residual variance between subgroups, resulting in satisfactory fit values, but the change in the chi-square was significant ($p < .001$). This result suggests that this measuring tool only achieved scalar invariance. However, with the use of AFIs (Alternative Fit Indices) criteria, it can be concluded that relatively small changes in fit indices (ΔRMSEA, ΔSRMR, ΔCFI) mean that the constraint-added model is no worse than the previous model, meaning that this measuring instrument has strict invariance.

Table 2. Summary of measurement invariance analysis between male and female groups

Model	χ^2	df	Ref.	p-value	RMSEA	SRMR	CFI	TLI	BIC
Single group solution									
M0a	220.113	132	-	-	.037	.047	.964	.959	18307.898
M0b	201.672	132	-	-	.036	.05	.966	.961	14413.611
Measurement invariance									
M1	421.785	-	-	-	.037	.046	.965	.960	33020.123
M2	428.433	15	M1	.972	.035	.047	.967	.964	32925.038
M3	442.840	15	M2	.481	.034	.048	.967	.966	32837.712
M4	519.328	18	M3	< .001	.039	.052	.954	.955	32792.121

Notes. χ^2 *= Chi-squared; df = degree of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square of Error Approximation; SRMR = Standardized Root Mean Square of Residual; BIC = Bayesian Information Criteria.*

Sources: Personal data (2024).

Discussion

First, this research aimed to develop a WFA measuring tool using the SJT format to minimize faking by test participants. The process of defining measuring constructs and reviewing items followed Patterson et al.'s (2015) procedures by considering the American Educational Research Association (AERA) et al.'s (2014) guidelines. This process found that the items written in the item pool and assessed by the experts represent the WFA construct in each dimension while conforming to the SJT question format. The implication is that this process provides valid evidence based on item content in the SJT WFA measurement interpretation.

Then, before carrying out a confirmatory factor analysis of the field test results, the data normality assumption was tested using several indicators, namely skewness and kurtosis. The standard value of normal skewness and kurtosis is close to O and is in the range of -1 to +1 (Tabachnick & Fidell, 2018). However, values between -2 to +2 for skewness and between -7 to +7 for kurtosis are still acceptable (Finney & DiStefano, 2013; Kim, 2013). Using these looser standards, most of the items in this study had a normal distribution. However, some items' skewness and/or curtosis do not meet these standard values, thus violating the normality assumption. Items that violate the normality assumption included seven items on the adaptive dimension, four on the proactive dimension, and six on the resilience dimension (see Appendix 1 - 3).

The implication of discovering several items that are not normally distributed is that the requirement for factor analysis using the MLR estimator is not met (Bollen, 1989; Satorra, 1990). However, we can still use the MLR estimator on data showing mild to moderate normality assumption violations (Li, 2016). This estimator is also unbiased when used in large samples (Bollen, 1989). Future research can use the Weighted Least Square Mean-Variance Adjusted (WLSMV) estimator in the analysis because it can better test factor structures of non-normally distributed and asymptotic data (Li, 2016; Wirth & Edwards, 2007).

Furthermore, the MLR estimator's CFA analysis revealed that the three-factor/multidimensional model fits the field data. This model resulted from the unidimensional model testing for each dimension

to identify items that have satisfactory psychometric properties or at least meet the criteria of 0.30 (Brown, 2015; Wang & Wang, 2020). Based on model testing for each dimension, the results show that each dimension is classified as fit with field data, which means that all items can represent the underlying dimensions or latent factors. These results thus could be used as an argument for validity evidence based on the internal structure of the SJT WFA measuring tool.

In addition, this study found a factor structure that fits a multidimensional model with three dimensions. This finding confirmed Al-Faouri et al. (2014), who revealed the multidimensionality of the WFA construct. This study tested the factor structure of the WFA construct using the three dimensions of adaptability, proactiveness, and resilience and found this model fits both the multidimensional and second-order models (Sherehiy & Karwowski, 2014). The multidimensional model of the WFA construct was also found in another study, although the WFA construct used was a 10-dimensional model (Petermann & Zacher, 2022). Several other studies consistently used multidimensional models with three or more dimensions aligned with this study (Alavi, 2016; Cai et al., 2018; Junior & Saltorato, 2021; Muduli & Pandaya, 2018). Several findings from previous studies confirm that the WFA construct is consistent with this research and similarly obtained model fit in multidimensional models despite different theories used. However, while previous WFA measurement studies used a self-report item model with a Likert scale, this research used a situational judgment test (SJT) model.

This study used McDonald's omega (ω) reliability estimates. Using ω reliability analysis can produce estimates almost identical to using CFA to estimate item factor loadings and error variance. The ω reliability analysis is also very suitable for CFA analysis models (Hayes & Coutts, 2020; McDonald, 1999). The reliability estimate ω is more recommended than α particularly because it does not require non-tau-equivalent items or congeneric models (Zhang & Yuan, 2016). This study's reliability estimates for adaptability, proactivity, and resilience dimensions were .760, .700, and .760, respectively. This coefficient is above or equivalent to the reliability requirement ≥ 0.70 as the rule of thumb for reliability (Bean & Bowen, 2021; McDonald, 1999).

Then, the measurement invariance test results showed slightly different results between the p-value and AFI criteria. The implication of using the p-value of the chi-square difference ($\Delta\chi^2$) is that a significant difference indicates that the constrained model is significantly worse than the previous model. However, using the p-value statistic of $\Delta\chi^2$ is sensitive to sample size, especially large ones (Chen, 2007). In addition, the sample size exceeding 700 strengthens not to use $\Delta\chi^2$ (Putnick & Bornstein, 2016).

On the other hand, an alternative model was tested with other criteria, referred to as the AFIs (alternative fit indices), which include ΔCFI, ΔTLI, Δ RMSEA, and ΔSRMR. A ΔCFI value not exceeding .01 is the most widely used criterion (Cheung & Rensvold, 2002). This criterion means that differences in the CFI values between the two models below these criteria suggest that the model with other constraints is no worse than the previous model. The ΔCFI value and the ΔRMSEA < .015 and ΔSRMR < .030 can be used as stronger criteria (Chen, 2007). The results of the measurement invariance testing showed that this measuring instrument meets the strict invariance level criteria, namely having equality in form, factor loading, intercept, and residual variance. These results support the validity of the evidence-based on the test consequences by considering gender.

Conclusion

Based on the findings above, this research concludes that the development and evaluation of the psychometric properties of the SJT WFA show satisfactory results. The developed SJT WFA measuring instrument has psychometric properties that meet the criteria regarding goodness-of-fit indices and statistics on the model, factor loading, reliability, and measurement invariance. This research has at least three of the five pieces of evidence recommended by the American Educational Research Association (AERA) et al. (2014): validity evidence based on item content, internal structure, and test consequences.

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

On the other hand, this research has several limitations that can be improved in further research. First, the field test only involved highly educated respondents, with most having completed a master's program. The implication is that further research is highly recommended to investigate psychometric properties by involving respondents with other characteristics, such as respondents with secondary education, or even by evenly covering all levels of education. In addition, this research uses a classical approach (CFA) that still relies heavily on the samples, and the focus of analysis is at the item and test level. Further research is also recommended to examine the option scores, considering that the SJT has a unique response format for each response option. Besides, the researchers can use contemporary approaches such as item response theory (IRT) with the nominal response model method. Lastly, this research has obtained three of the five sources of validity, but further research is highly recommended to have other sources of validity evidence, such as conducting cognitive interviews to obtain validity evidence based on the response process and conducting correlational research between the SJT WFA and other theoretically related measuring instruments or using the same construct but with different format of questions to obtain validity evidence based on association with other variables.

Acknowledgment

Our gratitude goes to the Faculty of Psychology, Universitas Gadjah Mada, Indonesia, for providing us with a research grant. We also would like to thank the Human Quality Development Unit, Faculty of Psychology, Universitas Gadjah Mada, for helping us collect the data for this research.

Conflict of Interest

The authors have no potential conflicts of interest to disclose related to this study.

Authors Contribution

AVH provided the conception, framework, and research design, supervision, and final approval. RDM conceptualized the methodology, conducted data analysis, interpreted the results, and wrote the article. ABALF contributed to the writing of the article and discussion, performed data analysis, and edited the final paper. IPA assembled the data and assisted in the preparation of the draft. WJA contributed to the preparation of instruments and participated in the drafting of the article. All authors gave their approval for the final manuscript.

References

Aamodt, M. G. (2010). *Industrial organizational psychology: an applied approach* (6th ed.). Wadsworth Cengage Leaning.

Alavi, S. (2016). The influence of workforce agility on external manufacturing flexibility of Iranian SMEs. *International Journal of Technological Learning, Innovation and Development*, *8*(1), 111. https://doi.org/https://doi.org/10.1504/IJTLID.2016.075185

Alavi, S., Wahab, D. A., Muhamad, N., & Shirani, B. A. (2014). Organic structure and organizational learning as the main antecedents of workforce agility. *International Journal of Production Research*, *52*(21), 6273–6295. https://doi.org/10.1080/00207543.2014.919420

Aldrup, K., Carstensen, B., Köller, M. M., & Klusmann, U. (2020). Measuring Teachers' Social-Emotional Competence: Development and Validation of a Situational Judgment Test. In *Frontiers in Psychology, 11*, 1-20. https://doi.org/10.3389/fpsyg.2020.00892

Al-Faouri, A. H., Al-Nsour, M. M., & Al-Kasasbeh, M. M. (2014). The impact of workforce agility on organizational memory. *Knowledge Management Research and Practice*, *12*(4), 432–442. https://doi.org/10.1057/kmrp.2013.19

Alshurideh, M. T., Hassanien, A. E., & Masa'deh, R. (Eds.). (2021). *The effect of coronavirus disease (COVID-19) on business intelligence: studies in system, decision and control*. Springer International Publishing. https://doi.org/http://dx.doi.org/10.1007/978-3-030-67151-8

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Andrea, M., K., J. B., & Allison, G. (2020). Development of Global Engineering Competency Scale: Exploratory and Confirmatory Factor Analysis. *Journal of Civil Engineering Education*, *146*(2), 4019003. https://doi.org/10.1061/(ASCE)EI.2643-9115.0000006

Barron, L. G., Aguilar, I. D., Rose, M. R., & Carretta, T. R. (2021). Development of a situational judgment test to supplement current US air force measures of officership. *Military Psychology*, *36*(1), 33-48. https://doi.org/10.1080/08995605.2021.1997500

Bean, G. J., & Bowen, N. K. (2021). Item Response Theory and Confirmatory Factor Analysis: Complementary Approaches for Scale Development. *Journal of Evidence-Based Social Work*, *18*(6), 597–618. https://doi.org/10.1080/26408066.2021.1906813

Braun, T. J., Hayes, B. C., DeMuth, R. L. F., & Taran, O. A. (2017). The Development, Validation, and Practical Application of an Employee Agility and Resilience Measure to Facilitate Organizational Change. *Industrial and Organizational Psychology*, *10*(4), 703–723. https://doi.org/10.1017/iop.2017.79

Breu, K., Hemingway, C. J., Strathern, M., & Bridger, D. (2001). Workforce agility: the new employee strategy for the knowledge economy Workforce agility: the new employee strategy for the knowledge economy. *Journal of Information Technology*, *17*(1), 21–31. https://doi.org/10.1080/02683960110132070

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Cai, Z., Huang, Q., Liu, H., & Wang, X. (2018). International Journal of Information Management Improving the agility of employees through enterprise social media : The mediating role of psychological conditions. *International Journal of Information Management*, *38*(1), 52–63. https://doi.org/10.1016/j.ijinfomgt.2017.09.001

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology*, *63*(1), 83–117. https://doi.org/10.1111/j.1744-6570.2009.01163.x

Corstjens, J., Lievens, F., & Krumm, S. (2017). Situational Judgement Tests for Selection. *The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention*, 226–246. https://doi.org/10.1002/9781118972472.ch11

Fajrianthi, & Zein, R. A. (2017). Development of a psychological test to measure ability-based emotional intelligence in the Indonesian workplace using an item response theory. *Psychology Research and Behavior Management*, *10*, 339–352. https://doi.org/10.2147/PRBM.S143113

Greatrix, R., & Dowell, J. (2020). UKCAT and medical student selection in the UK-what has changed since 2006? *BMC Medical Education*, *20*(1). https://doi.org/10.1186/s12909-020-02214-1

Griffin, B., & Hesketh, B. (2003). Adaptable behaviours for successful work and career adjustment. *Australian Journal of Psychology*, *55*(2), 65–73. https://doi.org/10.1080/00049530412331312914

Gupta, M., George, J. F., & Xia, W. (2019). Relationships between IT department culture and agile software development practices: an empirical investigation. *International Journal of Information Management*, *44*, 13–24. https://doi.org/10.1016/j.ijinfomgt.2018.09.006

Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (1998). *Multivariate Data Analysis* (Fifth Ed.). Prentice-Hall.

Hayes, A. F., & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But…. *Communication Methods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Hough, L. M., & Oswald, F. L. (2008). Personality Testing and Industrial–Organizational Psychology: Reflections, Progress, and Prospects. *Industrial and Organizational Psychology*, 1(3), 272–290. https://doi.org/10.1111/j.1754-9434.2008.00048.x

Hu, L. T., & 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

Huber, P. J. (1981). *Robust statistics*. Wiley.

Johannsen, L. M., Frerichs, W., Inhestern, L., & Bergelt, C. (2020). Assessing competencies of healthcare professionals caring for parents with cancer: The development of an innovative assessment tool. *Psycho-Oncology*, *29*(10), 1670–1677. https://doi.org/10.1002/pon.5507

Junior, G. T., & Saltorato, P. (2021). Workforce agility : a systematic literature review and a research agenda proposal. *Innovar*, *31*(81), 155–168. https://doi.org/https://doi.org/10.15446/innovar.v31n81.95582

Juster, F. R., Baum, R. C., Zou, C., Risucci, D., Ly, A., Reiter, H., Miller, D. D., & Dore, K. L. (2019). Addressing the Diversity-Validity Dilemma Using Situational Judgment Tests. *Academic Medicine*, *94*(8), 1197–1203. https://doi.org/10.1097/ACM.0000000000002769

Karaca, A., Aydogmus, M. E., & Gunbas, N. (2022). Enforced remote work during covid-19 and the importance of technological competency: a job demands-resources perspective. *European Review of Applied Psychology, 73*(6),1–52. https://doi.org/10.1016/j.erap.2022.100867

Kasten, N., Freund, P. A., & Staufenbiel, T. (2020). "Sweet little lies": An in-depth analysis of faking behavior on Situational Judgment Tests compared to personality questionnaires. *European Journal of Psychological Assessment*, *36*(1), 136–148. https://doi.org/10.1027/1015-5759/a000479

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52. https://doi.org/10.5395/rde.2013.38.1.52

Kniffin, K. M., Narayanan, J., Anseel, F., Antonakis, J., Ashford, S. P., Bakker, A. B., Bamberger, P., Bapuji, H., Bhave, D. P., Choi, V. K., Creary, S. J., Demerouti, E., Flynn, F. J., Gelfand, M. J., Greer, L. L., Johns, G., Kesebir, S., Klein, P. G., Lee, S. Y., … Vugt, M. van. (2021). COVID-19 and the workplace: implications, issues, and insights for future research and action. *American Psychologist*, *76*(1), 63–77. https://doi.org/10.1037/amp0000716

Koschmieder, C., & Neubauer, A. C. (2021). Measuring emotion regulation for preservice teacher selection: A theory-driven development of a situational judgment test. *Personality and Individual Differences*, *168*. https://doi.org/10.1016/j.paid.2020.110363

Kreitchmann, R. S., Abad, F. J., Ponsoda, V., Nieto, M. D., & Morillo, D. (2019). Controlling for response biases in self-report scales: forced-choice vs. Psychometric modeling of Likert items. *Frontiers in Psychology*, *10*, 1–12. https://doi.org/10.3389/fpsyg.2019.02309

Lee, P., Joo, S. H., & Lee, S. (2019). Examining stability of personality profile solutions between Likert-type and multidimensional forced choice measure. *Personality and Individual Differences*, *142*, 13–20. https://doi.org/10.1016/j.paid.2019.01.022

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*, *37*(4), 426–441. https://doi.org/10.1108/00483480810877598

Männiste, T., Pedaste, M., & Schimanski, R. (2019). Situational judgment test for measuring military tactical decision-making skills. *Military Psychology*, *31*(6), 462–473. https://doi.org/10.1080/08995605.2019.1664366

McDonald. (1999). *Test theory: A unified treatment*. Lawrence Earlbaum Associates.

McDonald, R. P., & Ho, M. H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, *7*(1), 64–82. https://doi.org/10.1037/1082-989X.7.1.64

Misra, S., Kumar, V., Kumar, U., Fantazy, K., & Akhter, M. (2012). Agile software development practoces: evolution, principles, and criticisms. *International Journal of Quality & Reliability Management*, *29*(9), 972–980. https://doi.org/10.1108/02656711211272863

Muduli, A., & Pandaya, G. (2018). Psychological empowerment and workforce agility. *Psychological Studies*, *63*(3), 276–285. https://doi.org/10.1007/s12646-018-0456-8

Olaru, G., Burrus, J., MacCann, C., Zaromb, F. M., Wilhelm, O., & Roberts, R. D. (2019). Situational judgment tests as a method for measuring personality: development and validity evidence for a test of dependability. *PLoS ONE*, *14*(2), 1–19. https://doi.org/10.1371/journal.pone.0211884

Patterson, F., Zibarras, L., & Ashworth, V. (2015). Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. *Medical Teacher, 38*(1), 1–15. https://doi.org/10.3109/0142159X.2015.1072619

Paudyal, V., Gunner, E., Bowen, M., Dehele, I., Shamim, A., Khan, K., & Fleming, G. (2021). National recruitment system for pre-registration pharmacist trainees in England and Wales: Context, process and outcomes. *International Journal of Pharmacy Practice*, *29*(3), 285–290. https://doi.org/10.1093/ijpp/riab004

Petermann, M. K. H., & Zacher, H. (2022). Workforce Agility: development and validation of a multidimensional measure. *Frontiers in Psychology*, *13*, 1–15. https://doi.org/10.3389/fpsyg.2022.841862

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. In *Developmental Review 41*, 71-90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. https://www.R-project.org/

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Sala-Roca, J., Jariot-Garcia, M., Arnau-Sabatés, L., & Doval, E. (2021). Situational judgment test of basic employability competences development: Validation of a psychoeducative instrument for socioeducative intervention. *Pedagogia Social*, *37*, 39–52. https://doi.org/10.7179/PSRI_2021.37.02

Shashi, Centobelli, P., Cerchione, R., & Ertz, M. (2020). Agile supply chain management: where did it come from and where will it go in the era of digital transformation? *Industrial Marketing Management*, *90*, 324–345. https://doi.org/10.1016/j.indmarman.2020.07.011

Sherehiy, B. (2008). *Relationships between agility strategy, work organization and workforce agility* [Dissertation]. University of Louisville.

Sherehiy, B., & Karwowski, W. (2014). The relationship between work organization and workforce agility in small manufacturing enterprises. *International Journal of Industrial Ergonomics*, *44*(3), 466–473. https://doi.org/10.1016/j.ergon.2014.01.002

Sherehiy, B., Karwowski, W., & Layer, J. K. (2007). A review of enterprise agility: Concepts, frameworks, and attributes. International Journal of industrial ergonomics, 37(5), 445-460. https://doi.org/10.1016/j.ergon.2007.01.007

Smith, K. J., Flaxman, C., Farland, M. Z., Thomas, A., Buring, S. M., Whalen, K., & Patterson, F. (2020). Development and validation of a situational judgement test to assess professionalism. *American Journal of Pharmaceutical Education*, *84*(7), 985–992. https://doi.org/10.5688/ajpe7771

Sorrel, M. A., Olea, J., Abad, F. J., de la Torre, J., Aguado, D., & Lievens, F. (2016). Validity and Reliability of Situational Judgement Test Scores: A New Approach Based on Cognitive Diagnosis Models. *Organizational Research Methods*, *19*(3), 506–532. https://doi.org/10.1177/1094428116630065

Storme, M., Suleyman, O., Gutlib, M., & Lubart, T. (2020). Who is agile? an investigation of the psychological antecedents of workforce agility agile. *Global Business and Organizational Excellence*, 1–11. https://doi.org/10.1002/joe.22055

Tarafdar, M., & Qrunfleh, S. (2016). Agile supply chain strategy and supply chain performance: complementary roles of supply chain practices and information systems capability for agility. *International Journal of Production Research*, *55*(4), 925–938. https://doi.org/10.1080/00207543.2016.1203079

Tiffin, P. A., Paton, L. W., O'Mara, D., MacCann, C., Lang, J. W. B., & Lievens, F. (2020). Situational judgement tests for selection: traditional vs construct-driven approaches. *Medical Education*, *54*(2), 105–115. https://doi.org/10.1111/medu.14011

Trzesniewski, K. H., Donnellan, M. B., & Lucas, R. E. (2011). Secondary data analysis: An introduction for psychologists. In Secondary data analysis: An introduction for psychologists. American Psychological Association. https://doi.org/10.1037/12350-000

using Mplus (D. J Wang, J., & Wang, X. (2020). *Structural equation modeling: applications*. Balding, N. A. C. Cressie, G. Fitzmaurice, & H. Goldstein (eds.); 2nd ed.). John Wiley & Sons. https://doi.org/10.1002/9781119422730

Zhang, Z., & Yuan, K. H. (2016). Robust coefficients alpha and omega and confidence intervals with outlying observations and missing data: methods and software. *Educational and Psychological Measurement*, *76*(3), 387–411. https://doi.org/10.1177/0013164415594658

Appendix

Table A. Sample Responses to the Rolnow Survey

Item	Mean	SD	Skewness	Kurtosis	Proportion of responses				
					Score 1	Score 2	Score 3	Score 4	Score 5
Ad.D.1	4.449	1,041	-1.867	5.177	.015	.108	.005	.158	.714
Ad.D.2	3.228	1,192	-.006	2.314	.089	.141	.436	.121	.213
Ad.D.3	4.547	0,913	-1.964	5.797	.008	.053	.086	.090	.763
Ad.D.4	3.547	0,866	-.418	3.385	.023	.064	.375	.420	.119
Ad.D.5	3.53	1,462	-.064	1.098	.009	.438	.045	.029	.479
Ad.D.6	**3.976**	**0,350**	**-5.248**	**45.587**	**.007**	**.008**	**.009**	**.955**	**.021**
Ad.D.7	3.331	0,922	-.570	1.641	.297	.093	.594	.017	NA
Ad.E.1	2.775	1,312	.575	2.124	.132	.411	.198	.069	.191
Ad.E.2	4.284	1,021	-1.394	4.192	.023	.052	.128	.214	.584
Ad.E.3	2.730	1,206	.842	2.465	.068	.535	.156	.082	.159
Ad.E.4	2.957	1,058	.630	2.083	.003	.46	.222	.203	.111
Ad.E.5	3.392	0,498	.499	1.539	.001	.609	.386	.003	NA
Ad.E.6	**3.003**	**0,235**	**1.259**	**35.735**	**.001**	**.019**	**.958**	**.018**	**.003**
Ad.F.1	**3.859**	**0,610**	**-2.515**	**12.041**	**.019**	**.021**	**.087**	**.826**	**.046**
Ad.F.2	4.183	1,218	-1.393	3.892	.072	.024	.161	.134	.608
Ad.F.3	3.394	0,657	.623	3.900	.006	.015	.624	.291	.064
Ad.F.4	**4.053**	**0,327**	**.059**	**15.377**	**.006**	**.011**	**.907**	**.076**	**NA**
Ad.F.5	**2.949**	**0,283**	**-1.707**	**24.268**	**.003**	**.053**	**.938**	**.002**	**.003**
Ad.F.6	**4.900**	**0,329**	**-3.767**	**21.049**	**.002**	**.002**	**.089**	**.906**	**NA**
Ad.F.7	**2.010**	**0,476**	**1.284**	**12.202**	**.091**	**.818**	**.085**	**.006**	**NA**

Notes. Items in bold do not meet the skewness -2 to 2 and/or kurtosis -7 to 7 normality assumptions.

Sources: Personal Data (2024).

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

Table B. Descriptive statistic of Proactive dimension's items

Item	Mean	SD	Skewness	Kurtosis	Proportion of responses				
					Score 1	Score 2	Score 3	Score 4	Score 5
Pr.A.1	3.599	1.148	-.336	2.612	.069	.017	.475	.124	.315
Pr.A.2	3.169	1.021	.282	2.033	.008	.308	.309	.256	.119
Pr.A.3	3.182	1.4	-.286	1.935	.21	.047	.322	.193	.228
Pr.A.4	3.538	1.107	-1.063	3.275	.091	.096	.113	.582	.117
Pr.A.5	**4.011**	**.196**	**.43**	**35.372**	**.001**	**.01**	**.965**	**.024**	**NA**
Pr.A.6	3.786	.455	-.904	3.454	.002	.227	.754	.017	NA
Pr.A.7	3.01	.844	.093	1.72	.001	.335	.328	.323	.012
Pr.B.1	2.903	.826	-.288	5.442	.103	.027	.793	.018	.059
Pr.B.2	4.333	1.161	-1.942	5.753	.078	.019	.037	.223	.642
Pr.B.3	3.586	1.03	.03	2.643	.038	.003	.582	.086	.29
Pr.B.4	3.826	1.091	-.503	2.334	.023	.094	.28	.243	.361
Pr.B.5	**1.078**	**.336**	**6.744**	**67.053**	**.932**	**.064**	**.003**	**NA**	**NA**
Pr.B.6	4.436	.814	-1.631	5.558	.005	.043	.052	.314	.587
Pr.B.7	3.605	.518	-.159	1.768	.001	.406	.579	.014	NA
Pr.C.1	3.334	.909	-.021	3.839	.05	.017	.619	.179	.135
Pr.C.2	3.293	1.048	.336	1.953	.001	.26	.366	.192	.182
Pr.C.3	2.086	1.155	.589	1.978	.442	.205	.185	.158	.009
Pr.C.4	2.655	.523	.044	4.009	.002	.356	.634	.001	.007
Pr.C.5	**3.02**	**.227**	**3.786**	**59.01**	**.002**	**.002**	**.975**	**.014**	**.007**
Pr.C.6	**2.045**	**.34**	**6.283**	**46.097**	**.005**	**.973**	**.018**	**.005**	**NA**

Notes. Items in bold do not meet the skewness -2 to 2 and/or kurtosis -7 to 7 normality assumptions.

Sources: Personal Data (2024).

Table C. Descriptive statistic of Resilience dimension's items

Item	Mean	SD	Skewness	Kurtosis	Proportion of responses				
					Score 1	Score 2	Score 3	Score 4	Score 5
Re.G.1	**2.113**	**.467**	**4.178**	**19.649**	**.940**	**.012**	**.042**	**.006**	**NA**
Re.G.2	3.679	1.14	-.357	2.324	.043	.077	.0374	.0172	.034
Re.G.3	3.632	1.267	-.513	2.392	.095	.036	.0376	.0129	.036
Re.G.4	4.061	1.113	-.951	2.919	.027	.082	.0181	.0222	.049
Re.G.5	2.818	.392	-1.65	4.107	.001	.181	.0817	.001	NA
Re.G.6	**4.017**	**.18**	**-2.021**	**108.999**	**.001**	**.001**	**.0976**	**.021**	**NA**
Re.G.7	**3.001**	**.111**	**-4.11**	**168.614**	**.001**	**.002**	**.0991**	**.006**	**NA**
Re.H.1	3.903	.899	-1.04	4.345	.024	.054	.0152	.0535	.0235
Re.H.2	3.902	1.023	-.449	2.001	.115	.231	.029	.0363	NA
Re.H.3	3.560	.843	-1.623	5.71	.060	.019	.0255	.0633	.033
Re.H.4	3.665	1.116	-.769	2.575	.036	.192	.053	.0509	.021
Re.H.5	**3.009**	**.184**	**6.761**	**111.426**	**.001**	**.002**	**.099**	**.007**	**NA**
Re.H.6	**3.027**	**.182**	**6.289**	**51.663**	**.001**	**.973**	**.024**	**.002**	**NA**
Re.H.7	2.233	.565	.403	3.432	.058	.665	.0265	.012	NA
Re.I.1	4.116	1.196	-.745	1.81	.148	.219	.002	.0631	NA
Re.I.2	4.398	1.369	-1.98	5.061	.129	.014	.003	.04	.081
Re.I.3	3.65	1.111	-.482	2.15	.019	.198	.0144	.0392	.025
Re.I.4	2.773	.445	-.913	4.257	.002	.229	.0764	.002	.002
Re.I.5	2.744	.495	-.312	3.76	.001	.277	.0703	.016	.003
Re.I.6	**3.002**	**.082**	**3.973**	**147.449**	**.002**	**.993**	**.005**	**NA**	**NA**

Notes. Items in bold do not meet the skewness -2 to 2 and/or kurtosis -7 to 7 normality assumptions.

Sources: Personal Data (2024).