# Evaluating Psychometric Properties of the Stress Measurement Instrument (the Operational and Organizational Police Stress Questionnaires) with the Application of Rasch Model in the Indonesian Nasional Police (INP)

**Abi Risa Bayu Argo[1], Heri Yulianto[2], Dwi Nuryanto[3]**

Psychology Bureau of the Indonesian National Police, Indonesia[1,2]
Faculty of Psychology, University of Indonesia, Indonesia[3]

heripsi@polri.go.id
(Corresponding author)

## Abstract

Police Stress Questionnaire (PSQ) is a questionnaire developed by McCreary and Thompson (2006) to measure stress exercised by the police. The development of this questionnaire is based on the reason that the previous stress measuring instruments only measure general stressors and can't describe specific stressors, especially in types of work that exert high level of stress. This study aims to evaluate the Police Stress Questionnaire (PSQ) instrument consists of two instruments, namely the Operational Police Stress Questionnaire (PSQ-Op) and the Organizational Police Stress Questionnaire (PSQ-Org), each containing 20 items with seven-point Likert scale for police officers in Indonesia. Respondents in this study were 313 police officers who served in the National Police Headquarters work unit (32.9%), Regional Police (38.3%), Resort Police (19.5%) and Sector Police (9.3%). The data collection technique used non-probability sampling with the help of the google form application. The method used is the Rating Scale Model (RSM). The results show that the Indonesian version of the Police Stressor Questionnaire (PSQ) is proven to meet unidimensional assumptions and the reliability analysis for person and items shows a strong level of reliability. However, the Rasch RSM analysis found violations of the assumption of local independence and problematic discrimination at specific thresholds (threshold disorder) in the seven response categories used. Furthermore, the results of the application of the Rasch Model Rating Scale model show that the psychometric facts of the two research instruments are very good and precise, as well as the suitability of the items to the model. Implications and suggestions for future research are also presented in the discussion.

**Keywords**: Rasch models, rating scale model, police stress questionnaire.

## Introduction

In general, stress is an internal state caused by physical demands of the body or environmental and social conditions that are considered potentially harmful, uncontrollable or exceed the ability of individuals to cope (Lazarus & Folkman, 1984). Stress can occur in workplace when a person gets excessive workload, feelings of difficulty and emotional tension that hinder performance (Robbins, 2004).

Some researchers (e.g., Holt, 1993; Spector, 1997) report that work stress is one of the causes of low levels of job satisfaction. These findings have important implications for organizations because low job satisfaction can predict low levels of commitment and an increase of likelihood to quit jobs (Hellman, 1997; Tett & Meyer, 1993).

High work stress can be physically, psychologically and socially damaging, and burnout can occur when all three aspects peak. One example of the impact of employees experiencing burnout is avoiding work, not wanting to deal with everyday tasks or being completely involved in work and ignoring other aspects of life (Yulianto, 2020).

Work stress can be caused by several factors such as intrinsic factors such as uncomfortable working environment conditions, non-ergonomic work environment, work using shift system, high risk and dangerous work, excessive workload, the use of new technology, and so on. In addition to factors in the work several other factors can also cause the onset of stress such as the role of individuals in the organization of work, work relationship factors, career development factors, organizational structure factors and work atmosphere, as well as other factors coming from outside the work.

Police are one of the professions with high levels of work stress. Police officers are required to always work professionally and meet the demands of the community to have excellent work performance. Several previous studies on work stress in Indonesian police samples have been conducted and on average produced data that police have medium to high stress levels (Aulya, 2013; Muhammad, 2004; Jayanegara, 2007).

The American Institute of Stress stated that the police profession has been ranked in the top ten most stressful jobs in the U.S. and is categorized as one of the most stressful jobs in the world (Purda et al., 2012). This is in line with Nikam and Shaikh (2014) who stated that police work is very stressful because they always have to risk their lives in their daily work, where the police include six professions with high levels of stress and have an impact on health and low job satisfaction. The severity of the challenges and burdens of police duties can have a negative impact both physically and psychologically for the police (Queirós, et all., 2013).

The use of psychological tests to diagnose the work stress problems of police officers is indispensable to maintaining their emotional life, work performance and mental health. One of the instruments of measuring police work stress is the Police Stress Questionnaire (PSQ) developed by McCreary and Thompson in 2006 that has been considered valid and reliable.

In diagnosing the stress of police work, McCreary and Thompson, in 2006, developed two instruments based on two main sources of police work pressures namely the nature of police job (which is related to maintaining a balance between work and personal life) and the nature of police organizations (relating to how police perceptions of organizational demands negatively impact the families of police officers).

The statistical test that McCreary and Thompson used in their report in 2006 shows the Cronbach's Alpha for PSQ-Op was 0.93, and Cronbach's Alpha for PSQ-Org was 0.92. The total

correlation of corrected items for PSQ-Op ranges from 0.50 to 0.70, while for PSQ-Org it ranges from 0.41 to 0.73.

The PSQ-Op and PSQ-Org instruments were adapted into Malay by Abidin, et all. (2014) on 300 samples of traffic police officers in Kuala Lumpur, with the results showing Cronbach's Alpha coefficient of 0.93 for Operational PSQ (PSQ-Op) and 0.94 for PSQ Organizational (PSQ-Org). Meanwhile, Aziz (2020) has conducted research on 100 of the Brimob Police officers in Jakarta with PSQ-Op and PSQ-Org instruments that have been adapted into Indonesian language. This study reported Cronbach's Alpha coefficients of 0.93 for PSQ-Op and 0.92 for PSQ-Org. Further research on work stress levels in 200 police officers in West Java has been conducted by Hayati (2019) with the report of reliability test results PSQ-Op = 0.933 and PSQ-Org = 0.953.

Some work stress research on samples of police officers in Indonesia with the use of PSQ-Op and PSQ-Org instruments showed quite good results, but some of the above research are still analysed using the classic test theory approach. According to Retnawati (2017) there are some drawbacks in classical test theory, the first of which the measurement error score does not interact with the actual score. Second, error scores don't correlate with actual scores and error scores on other tests for the same test taker. The third assumption is that the average of these error scores is equal to zero. Thus, the use of analysis with the approach of classical test theory is considered to have some weaknesses.

To overcome the weaknesses of the classical psychometric approach, this study uses the modern theory approach (Item Response Theory and Rasch Model) considering the measuring instruments used is the Likert scale. One of the modern theory models used to analyse grain scores in the form of a Likert scale is the Rating Scale Model (RSM; Andrich, 1978) based on Rasch measurement model. Rasch model, as a measurement model, has advantages among others on the linearity of scale and objectivity where the estimated parameters of items and person can be separated and don't affect each other.

## Methods

### Participants

Sampling technique used in this study is non-probability sampling. Researchers use the help of google form application in disseminating research instruments through social media (Whatsapp). Data retrieval was conducted in July and August 2020. Respondents in this study are 313 police officers in Indonesia consisted of males (87.5%) and females (12.5%), who served in the task force of The National Police Headquarters (32.9%), Regional Police (38.3%), Resort Police (19.5%) and Police Sector (9.3%). The rank of research respondents consisted of middle officers (21.4%), first officers (37.7%), Petty Officers (40.6%) and Enlisted (0.3%). The service period of respondents became police officers are 0-8 years (16.6%), 9-16 years (28.4%), 17-24 years (45%), and over 24 years (10%). Background of respondents in high school research (20.4%), Associate Degree (2.2%), Bachelor Degree (47%), Master degree (30%) and Postgraduate (0.3%).

### Instrument

This study administers two instruments of Police Stress Questionnaires that measure the level of police work stress developed by (McCreary &Thompson, 2006), namely the Operational Police Stress Questionnaire (PSQ-Op) and the Organizational Police Stress Questionnaire (PSQ-Org). Both instruments measure the stress of police work that focuses on the interaction between work and family. PSQ-Op is an instrument that measures the stress of police work related to maintaining

a balance between work and personal life, as well as how to make the most of time for family and friends. PSQ-Org is an instrument that measures the stress of police work related to how police perceptions of organizational demands negatively impact the families of police officers.

The instruments used in this study are the Operational Police Stress Questionnaire (PSQ-Op) and the Organizational Police Stress Questionnaire (PSQ-Org) (McCreary &Thompson, 2006) instruments that have been adapted into Indonesian as well as in the working culture of the Indonesian police by involving four translators with basic English literature education and eight police psychologists using reference to the journal Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures (Beaton et al. , 2000). The adaptation process is as follows: (1) Initial Translation. At this stage, the research instrument is translated into Indonesian language. The translation process is done by two people. The first and second translators (P1&P2) have basic education in English literature and work in the same field. (2) Synthesis of Translations. After obtaining the results from the first and second translators (P1&P2), the research instruments were synthesized. When discrepancies are found between the two translations, the items are selected based on the meaning that best fits the initial scale. In the process, cultural factors become a consideration in choosing the translation results in accordance with the conditions of the police in Indonesia. (3) Back Translation. At this stage the translation is done back to the initial language scale. The retranslation process is carried out by two different translators (P3&P4) with English literature education background. This process is done to see if there are discrepancies in meaning if the scale in Indonesian language is translated into the initial language. If there are differences in meaning, the items are reviewed. (4) Expert Committee. After improving the translation by considering the results of back translation, the research instrument was discussed with experts in the field of statistics and experts in the field of psychology who understood the concept of the Police Stress Questionnaires (PSQ) instrument. (5) Test of the Prefinal Version. At this stage, the agreed scale through the results of the discussion was administered to 15 police officers. This step is to find out if the adapted items can be understood by the respondent. PSQ instrument to measure police work stress consists of two instruments (PSQ-Op and PSQ-Org) each consisting of 20 items with a Likert scale of one to seven. Scale of one means no stress at all, scale of four means moderate stress and scale of seven means a lot of stress.

**Application of Rasch Rating Scale Model**

A Rating Scale Model is a widely applied item response model used to model ordinally observed variables that are assumed to collectively reflect common latent variables (Adams et al., 2012). The model response item shows a probability relationship between the test taker's ability on the test item and the latent nature of the test taker. In the options of response options with multilevel scales sequentially, such as in a Likert scale, it is more appropriate to use a graded response model (GRM), polytomous item response model (PIRM), partial credit model (PCM), or rating scale model (RSM) (Muraki, 1990).

In the Likert scale, all test takers use sequential response categories of all scale statements in the same way so that they can be empirically tested. In the application of the Likert scale, there is a requirement that the response category must be given the same distance, the response category used can be arranged in a continuum line representing latent variables that are unidimensional (Cheung &Mooi, 1994; Rost, 2001). The basic formula of RSM used to perform analysis on the ordinal rating scale is:

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

Where $P_{nik}$ is the probability that the person $n$ who travels to item $i$ will be "observed" in category $k$, $P_{ni(k-1)}$ is the probability that the $n^{th}$ person will select category $k – 1$, $B_n$ is the trait level of the construct measured by the person $n$, $D_i$ is the item difficulty level of item $i$ and $F_k$ is the probability that the $k$ category will be selected depending on the $k – 1$ category. The estimated difficulty of the item ($D_i$) and the level of attitude towards the environment of the respondent ($B_n$) are expressed on a logit scale (Linacre, 2002; Wright &Mok, 2004).

RSM requires each response category (k) to have a minimum frequency of 10 (DiStefano et al., 2014; DiStefano &Morgan, 2010; Eckes, 2011). In addition, to be able to apply data analysis with RSM approach, it is required to meet its assumptions, namely unidimensionality, local independence, and monotonicity (Embretson &Reise, 2000; Hambleton et al., 1991; Wright &Stone, 1979). RSM is used to estimate the probability that a person will select a specific response category in the resulting rating scale when the respondent's "ability level" on the construct and the parameters of the item are known to be of magnitude.

### Data Analysis

The research instruments used in this study were designed using a Likert scale with seven answer response options with the same rating scale as the polytomous data form. This study used Rating Scale Model (RSM) which is part of Rasch Model. The data in this research analysis have qualified RSM namely that each response category (k) must have a minimum frequency of 10 (DiStefano et al., 2014; DiStefano &Morgan, 2010; Eckes, 2011).

Further data processing in this study used Winstep software (v. 3.65) (Linacre&Wright, 2019) to test the validity of PSQ-Op and PSQ-Org instruments. Person and item parameters are estimated using Joint Maximum Likelihood Estimation (JMLE). To obtain information on the psychometric characteristics of each instrument, the following analysis was conducted: (1) testing the unidimensionality assumptions of both research instruments using principal component analysis of the residuals (PCAR); (2) testing local independence assumptions using statistic Q3; (3) testing fit items with Rasch models, taking into account mean square (MNSQ) when using Rasch RSM; (4) reliability testing for persons and items; (5) testing the Rating of Scale Diagnostics to establish the functionality of each category and the discrimination of their rating scales; (6) displaying Wright map to find out the measurement results of PSQ-Op and PSQ-Org instruments, and (7) testing information function to determine the functioning of the test when given to individuals with trait levels obtained.

## Results and Discussion

### Unidimensionality

This study uses principal component analysis of residual (PCAR) method to test the unidimensionality instrument assumptions. This method was chosen because it is considered the most effective in testing the assumption of unidimensionality of measurement instruments (Wright &Mok, 2004). The criteria for using Rasch Principal Component Analysis of Residuals (PCAR) is when the raw variance value explained by measures of > 40% is good, while when the ≥ value is 30% and the minimum raw variance value explained by measures is acceptable is ≥ 20% has met the assumption of unidimensionality (Linacre, 2004; Pichardo et al., 2018; Reckase, 1979).

**Table 1.** Test Result of Principal Component Analysis of Residual (PCAR)

|  | Eigenvalues | Observed (%) | Expected (%) |
|---|---|---|---|
| Instrumen PSQ-Op |  |  |  |
| Total variance | 43.1 | 100.0% | 100.0% |
| Raw variance explained by measures | 23.1 | 53.6% | 55.4% |
| Raw variance explained by persons | 16.6% | 38.5% | 39.8% |
| Raw Variance explained by items | 6.5% | 15.1% | 15.6% |
| Raw unexplained variance (total) | 20.0% | 46.4% | 100.6% |
| Unexplained variance in 1st contrast | 2.9% | 46.4% | 100.0% |
| PSQ-Org |  |  |  |
| Total variance | 45.8 | 100.0% | 100.0% |
| Raw variance explained by measures | 25.8 | 56.3% | 57.5% |
| Raw variance explained by persons | 20.8% | 45.5% | 46.5% |
| Raw Variance explained by items | 5.0 | 10.8% | 11.1% |
| Raw unexplained variance (total) | 20.0 | 43.7% | 100.0% |
| Unexplained variance in 1st contrast | 2.7 | 6.0% | 13.6% |

Table 1 above shows the test results of unidimensionality assumptions. The PSQ-Op instrument that measures the level of work stress of police officers related to work and personal life obtained raw variance explained by measures of 23.1 in units of measure eigenvalues or in a percentage measure of 53.6%. That is, with raw variance explained by measures of 53.6% (> 20%), then 20 items that measure the stress of police work related between work and personal life are unidimensional. In the PSQ-Org instrument that measures the work stress of police officers related to the perception of police against the demands of private organizations obtained raw variance calculated explained by measures of 25.8 in units of measure eigenvalues or in a percentage a measure of 56.3%. That is, with raw variance explained by measures of 56.3% (> 20%), then 20 items that measure the stress of police work related to the perception of police against the demands of the organization is unidimensional.

The results of the analysis of two research instruments (PSQ-Op and PSQ-Org) showed raw variance explained by measures that are in percentage size above 20%. With the criteria that the size of 20% has met the assumption of unidimensionality (Reckase, 1979), the assumption of unidimensionality of both instruments in this study has been fulfilled and further analysis can be done.

**Local Independence**

The application of the Rasch Rating Scale Model is based on local independence assumptions. The assumption of local independence means that the response given by the respondent must be independent between each individual and the response given between the given items is not interrelated. This implies that a person's response to an item is not affected by its response to the previous item, nor will it affect the response to the next item (Meijer et al., 1990).

After fulfilling the unidimensionality assumption test, then conducted local independence assumption testing using Q3 statistics (Yen, 1984). The Q3 index is a correlation between the residual (the difference between the predicted probability of the estimated item parameters and the actual response of the person) of a pair of items (DeMars, 2003). With the Q3 index criteria that residual (raw) correlation between item pairs is not > 0.30 (Geldenhuys &Bosch, 2019). Violations of local independence assumptions indicate a response between items (item estimation bias) and have an impact on unidimensionality (Kunz et al., 2019).

The results of the analysis of local independence assumptions on both research instruments found three items (item 1, item 11 and item 15) on the PSQ-Op scale and 2 items (item 3 and item 17) on PSQ-Org instruments to have residual correlations between items >0.30 (do not meet local

independence assumptions). Furthermore, the five items are eliminated / discarded and then re-analyzed and the results obtained no more violations of local independence assumptions where there is residual correlation between items <0.30 (see Table 2). Thus, the assumption of local independence of the two instruments in this study has been fulfilled and further analysis can be done.

**Table 2.** Test Results of Local Independence Assumption PSQ-Op and PSQ-Org Instruments

| Raw Residual Correlation | Item Pairs |
|---|---|
| **Instrument PSQ-Op** | |
| -0.43 | Item 3 and item 14 |
| -0.35 | Item 4 and item 19 |
| -0.29 | Item 2 and item 9 |
| -0.26 | Item 3 and item 18 |
| -0.26 | Item 5 and item 17 |
| -0.25 | Item 2 and item 19 |
| -0.25 | Item 6 and item 12 |
| -0.24 | Item 12 and item 16 |
| -0.24 | Item 3 and item 13 |
| -0.24 | Item 4 and item 20 |
| **Instrument PSQ-Org** | |
| 0.26 | Item 4 and item 19 |
| 0.26 | Item 13 and item 20 |
| -0.35 | Item 4 and item 10 |
| -0.30 | Item 7 and item 8 |
| -0.29 | Item 5 and item 14 |
| -0.28 | Item 7 and item 11 |
| -0.27 | Item 1 and item 7 |
| -0.27 | Item 4 and item 12 |
| -0.26 | Item 14 and item 18 |
| -0.25 | Item 12 and item 19 |

**Item Fit**

In the Rasch Rating Scale model measurement, the fit index used is infit and outfit statistic MNSQ. Infit Mean Square (MNSQ) and Outfit Mean Square (MNSQ) values are used to identify data inaccuracies against models at the item level. In the Rasch Rating Scale Model the expectation value against Infit or Outfit for each item is 1.0, with an acceptable value range between 0.5 to 1.5. Values ranging from 0.5 to 1.5 are effective for a measurement (Andrich &Marais, 2019; Bond &Fox, 2015) and values that are out of bounds indicate a lack of accuracy between items and models (DiStefano &Morgan, 2010).

In this test will be presented an overview of psychometric characteristics of PSQ-Op and PSQ-Org instruments including fit statistics test, item difficulty level, and PTMEA correlation (point-measure). In the test of fit statistic items against both research instruments, there were six items that is showed as not fit with Rasch rating scale model measurement, because it has infit value and MNSQ outfit is not fit (outside the range of 0.5 to 1.5). Of the 17 items on the PSQ-Op instrument, there are three items that are not fit to measure the stress of police work related between work and personal life. The three items are item number two "Dinas sendirian di malam hari" (infit value MNSQ = 1.61 and outfit value MNSQ = 1.59), item number four "Risiko cedera/terluka saat bertugas" (outfit value MNSQ = 1.57) and item number six "Peristiwa-peristiwa traumatis (contoh: kecelakaan kendaraan bermotor, masalah rumah tangga, kematian, cidera)" (MNSQ outfit value = 1.61). While in the PSQ-Org instrument of 18 items there are three items that are not fit to measure the stress of the work of police officers related to the perception of the police against the demands of the organization. The three items are item number

eighteen "Berurusan dengan sistem peradilan umum" (infit value MNSQ = 2.21 and outfit value MNSQ = 2.54), item number ten "Adanya perintah untuk mengerjakan tugas di luar ketentuan dinas" (outfit value MNSQ = 1.65) and item number two" Merasa aturan yang ada tidak adil untuk semua personel (pilih kasih)" (outfit value MNSQ = 1.53).

The calibration results of the second instrument item obtained the results of 14 items on the PSQ-Op instrument and 15 items on the PSQ-Org instrument are fit with the measurement of Rasch Rating Scale Model with a range of acceptable infit and outfit MNSQ values between 0.5 to 1.5, as can be seen in Table 3.

**Table 3.** Item Calibration of PSQ-Op and PSQ-Org Instrument

| Item | Measure | Infit MNSQ | Outfit MNSQ | PTMEA Correlation |
|---|---|---|---|---|
| **PSQ-Op** | | | | |
| Item 10 | 1.00 | 1.42 | 1.22 | 0.39 |
| Item 16 | 0.54 | 1.22 | 1.20 | 0.49 |
| Item 1 | 0.45 | 1.10 | 1.21 | 0.52 |
| Item 9 | 0.23 | 1.05 | 1.38 | 0.56 |
| Item 20 | 0.14 | 0.85 | 0.84 | 0.61 |
| Item 18 | 0.09 | 1.09 | 1.06 | 0.60 |
| Item 19 | -0.10 | 0.98 | 0.88 | 0.64 |
| Item 17 | -0.13 | 1.01 | 0.97 | 0.65 |
| Item 14 | -0.18 | 0.98 | 0.98 | 0.66 |
| Item 3 | -0.30 | 1.10 | 1.43 | 0.65 |
| Item 12 | -0.38 | 0.88 | 0.89 | 0.71 |
| Item 13 | -0.40 | 0.84 | 0.77 | 0.71 |
| Item 5 | -0.41 | 1.03 | 1.05 | 0.69 |
| Item 8 | -0.55 | 1.01 | 1.04 | 0.72 |
| **PSQ-Org** | | | | |
| Item 1 | 1.49 | 1.27 | 0.99 | 0.41 |
| Item 8 | 0.37 | 0.99 | 1.10 | 0.61 |
| Item 11 | 0.35 | 1.09 | 0.91 | 0.62 |
| Item 19 | 0.18 | 1.12 | 0.92 | 0.63 |
| Item 4 | 0.03 | 1.09 | 1.01 | 0.66 |
| Item 15 | -0.03 | 1.14 | 0.97 | 0.67 |
| Item 9 | -0.07 | 0.92 | 0.94 | 0.71 |
| Item 14 | -0.13 | 0.92 | 0.95 | 0.71 |
| Item 13 | -0.15 | 0.81 | 0.94 | 0.72 |
| Item 6 | -0.22 | 1.03 | 0.94 | 0.72 |
| Item 12 | -0.24 | 1.08 | 0.94 | 0.71 |
| Item 20 | -0.24 | 0.94 | 1.09 | 0.72 |
| Item 16 | -0.29 | 1.21 | 1.18 | 0.71 |
| Item 5 | -0.38 | 0.77 | 0.78 | 0.76 |
| Item 7 | -0.66 | 1.15 | 1.34 | 0.78 |

As can be seen in Table 3, items are sorted from the "most difficult to answer a lot of stress", up to the items that are "most easily answered a lot of stress ". On the PSQ-Op instrument the difficulty level was in the range of -0.55 to 1.00 and it was found that item number ten "tercukupinya makanan sehat saat bertugas" with the location at 1.00 logit was the most difficult item to get a "a lot of stress" response and item number eight "Tidak cukup waktu untuk dihabiskan bersama teman dan keluarga" with a

location of -0.55 was the easiest item to get a "a lot of stress" response. On the PSQ-Org instrument the difficulty level was in the range of -0.66 to 1.49 and it was found that item number one "Bekerja sama dengan rekan kerja" with a location at 1.49 logit was the most difficult item to get a "a lot of stress" response and item number seven "Aturan yang tidak jelas dan berbelit-belit" with a location of -0.66 were the easiest items to get a "a lot of stress" response.

To understand the matching aspect of the data model, it is also necessary to look at the PTMEA Correlation value, where if the value is negative it means that the item does not have the right score and also does not work as it should (Linacre, 2018). In this study all items on both research instruments showed a positive correlation. In PSQ-Op instrument, the correlation size ranges from 0.39 to 0.72 and in PSQ-Org instrument the correlation size ranges from 0.41 to 0.78. The correlation amount of all items on both instruments passes the criteria of 0.30. These findings suggest that all items in this measuring instrument function well in the same direction as what is theorized.

**Person and Item Separation Reliability**

In Rasch RSM's analysis, reliability is estimated for person and items. The reliability coefficient criteria used > 0.70, which indicates that the instrument has good internal consistency. Then the separation index criteria for person and items used (> 1.5) is considered sufficient to perform comparative analysis at the group level (Tennant & Conaghan, 2007).

**Table 4.** Person and Item Separation Reliability of PSQ-Op and PSQ-Org instrument

| Instrument | Person separation reliability | Person separation index | Item separation reliability | Item separation index |
|---|---|---|---|---|
| PSQ-Op | 0.73 | 1.66 | 0.98 | 7.24 |
| PSQ-Org | 0.80 | 2.01 | 0.99 | 8.70 |

In the analysis of person and item reliability (Table 4), the results showed that person separation reliability that estimates how well the instrument distinguishes person on measured variables (Wright &Masters, 1982) obtained a value of >0.70 for both instruments (PSQ-Op= 0.73, PSQ-Org= 0.80). The findings indicate that the two instruments in this study are good for distinguishing person (Duncan et al., 2003; Kunz et al., 2019). While the person separation index is an estimated spread from respondents whose value is >1.5 (PSQ-Op= 1.66, PSQ-Org= 2.01).

Item separation reliability for both instruments is >0.90 (PSQ-Op= 0.98, PSQ-Org= 0.99) and item separation indexes of 7.24 and 8.70, respectively. These findings suggest that the psychometric characteristics of both research instruments are excellent (Duncan et al., 2003; Kunz et al., 2019). Separation index criteria of 1.5 is sufficient to conduct individual level analysis (Tennant &Conaghan, 2007), and those criteria have been met in this study.

**Rating Scale Diagnostics**

The application of the Likert scale generally produced distances between categories whose magnitude is not the same compared to the assumption of the same distance between answer options, numbers in the Likert scale have meaning as psychological distances (Wakita et al., 2012). Using the Rasch model, researchers were able to understand the responses given by respondents using a rating scale and determine the actual distance applicable to respondents in choosing the existing options.

Diagnostic testing with RSM is used to evaluate how well the seven categories that make up a response assembly serve to create an interpretable size. The threshold of each category on the non-cognitive test measuring instrument can be seen in Table 5 below:

**Table 5.** Rating Scale Diagnostics of PSQ-Op and PSQ-Org Instrument in Seven Response Categories

| Categories | Threshold | Observed count (%) | Observed average | Infit | Outfit |
|---|---|---|---|---|---|
| **PSQ-Op** | | | | | |
| 1 (No Stress At All) | NONE | 1420 (40) | -1.86 | 0.96 | 0.98 |
| 2 | -0.70 | 690 (20) | -1.02 | 1.00 | 0.90 |
| 3 | -0.69 | 593 (17) | -0.66 | 0.89 | 0.93 |
| 4 (Moderate Stress) | 0.16 | 320 (9) | -0.29 | 1.01 | 1.29 |
| 5 | 0.29 | 204 (6) | 0.06 | 0.79 | 0.76 |
| 6 | 0.21 | 182 (5) | 0.25 | 0.89 | 0.89 |
| 7 (A Lot of Stress) | 0.73 | 119 (3) | 0.20* | 1.49 | 2.14 |
| **PSQ-Org** | | | | | |
| 1 (No Stress At All) | NONE | 1144 (30) | -1.76 | 1.38 | 1.22 |
| 2 | -0.80 | 668 (18) | -1.06 | 0.61 | 0.72 |
| 3 | -0.73 | 700 (18) | -0.46 | 0.81 | 0.83 |
| 4 (Moderate Stress) | 0.24 | 424 (11) | -0.02 | 0.79 | 0.79 |
| 5 | 0.46 | 286 (8) | 0.29 | 0.66 | 0.64 |
| 6 | 0.38 | 273 (7) | 0.61 | 0.74 | 0.79 |
| 7 (A Lot of Stress) | 0.44 | 315 (8) | 0.55* | 1.47 | 1.68 |

Rasch RSM measurement expected adequate performance of the seven categories of measuring instrument response, by showing that: 1) the estimated average size for each response category increases monotonously and in the expected direction when the response category moves from the lower category to the higher; 2) the threshold of adjacent response categories increases monotonously and in the expected direction, and 3) each of the seven response categories shows acceptable MNSQ infit and outfit MNSQ statistics (Tennant &Conaghan, 2007).

The analysis of Scale Diagnostic Rating shows that no category has a response frequency of 0 (zero). This test also found that the threshold disorder that is expected in Rasch RSM threshold analysis that separates category six and seven should be higher than the threshold separating category five and six, threshold separating category five and six should be higher than threshold separating category four and five, and so on. The findings of this threshold disorder are an indication that respondents did not use the response category well (Houghton et al., 2017).

Upon obtaining the findings of the response category of the two research instruments that are irregular (threshold disorder), the researchers test the Diagnostic Scale Rating again by collapsing the response category by combining the reverse category with the previous category (collapsing categories). To confirm the accuracy of the use of response categories, researchers conducted five tests with a number of different response categories.

Testing the first collapsing categories, response categories are collapsed into six categories, five categories and four categories. The test obtained the response format that best suits the data, which is the format of four response categories (category one fixed, category two and three combined into category two, category four, five and six combined into category three and category seven into category four). Diagnostic Scale Rating test results with four response categories can be seen in Table 6.

**Table 6.** Rating Scale Diagnostics of PSQ-Op and PSQ-Org Instrument in Four Response Categories

| Categories | Threshold | Observed count (%) | Observed average | Infit | Outfit |
|---|---|---|---|---|---|
| **PSQ-Op** | | | | | |
| 1 (No Stress At All) | NONE | 1420 (40) | -3.06 | 1.00 | 1.00 |
| 2 | -2.07 | 1283 (36) | -1.40 | 0.85 | 0.89 |
| 3 | -0.12 | 706 (20) | 0.05 | 0.87 | 0.98 |
| 4 (A Lot of Stress) | 2.18 | 119 (3) | 0.40 | 1.49 | 1.61 |
| **PSQ-Org** | | | | | |
| 1 (No Stress At All) | NONE | 1144 (30) | -2.77 | 1.19 | 1.18 |
| 2 | -2.02 | 1368 (36) | -1.14 | 0.74 | 0.77 |
| 3 | 0.02 | 983 (26) | 0.58 | 0.68 | 0.69 |
| 4 (A Lot of Stress) | 2.00 | 315 (8) | 1.04 | 1.41 | 1.45 |

The findings of the diagnostic rating scale of Rasch measurements on PSQ-Op and PSQ-Org instruments with four categories of responses obtained observed count with positive skewness, in which each value obtained no more than 9 % answering a "a lot of stress" response to a set of items measuring stress levels. Then, information about obtained thresholds is very precise and good, where thresholds obtained appropriate values from negative order to positive order in four categories of responses tested. With four response categories, the MNSQ infit index and MNSQ outfit are no larger than two (Linacre, 2010), meaning that the scale diagnostics rating shows the accuracy of the measurements in the responses tested. It can be concluded that the findings of scale diagnostic rating with four categories of response on PSQ-Op and PSQ-Org instruments function appropriately. Diagnostic Scale Rating analysis procedure by collapsing this response category was also conducted in previous studies (Geldenhuys &Bosch, 2019; Houghton et al., 2017; Bond &Fox, 2007).

Furthermore, the assumption that needs to be met is the monotonic increasing threshold that the distance from one to the other should be 1.4-1.5 logit (Linacre, 1999) in this study by shrinking the distance response category from that assumption is well met, so that the probability curve of the response category shows the optimal shape (see Figure 1).
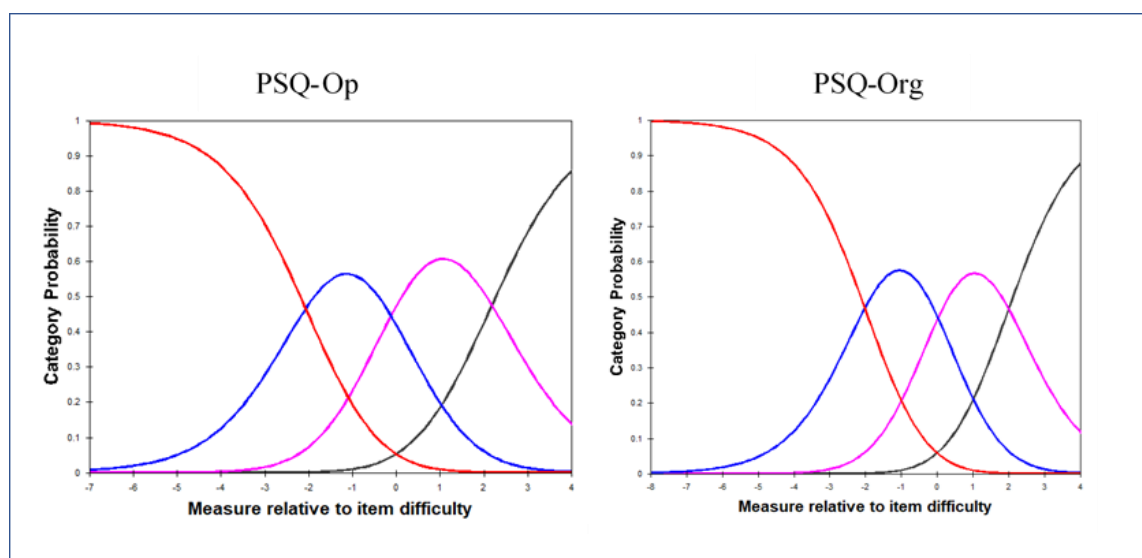


**Figure 1.** Category Response Curve of PSQ-Op and PSQ-Org Istrument

Collapsing seven response categories into four response categories indicates that monotonicity assumptions are met which means that all assumptions from RSM implementation are fulfilled. The fulfilment of this assumption indicates that the measurement process that has been done is no longer disturbed by errors in measurement within unnatural limits so that the data that fit against RSM even to the level of response category is very important to be fulfilled. In this study, further analysis was conducted in the format of four response categories.

### Wright Map

To answer the sixth study question, researchers evaluated items in the study instrument to determine which conditions were "most difficult" for stressed police officers. The Rasch Rating Scale Model (RSM) sets the validity of the construct according to the hierarchy of observable items in the Wright Map (Pichardo et al., 2018). This folder describes the difficulty items on the right and the ability person on the left. Items are presented on a continuum scale where the items with the lowest difficulty are at the bottom and the items with the highest difficulty are at the top. The Wright Map of the instrument analysis of attitude towards the environment can be seen in Figure 2 below:
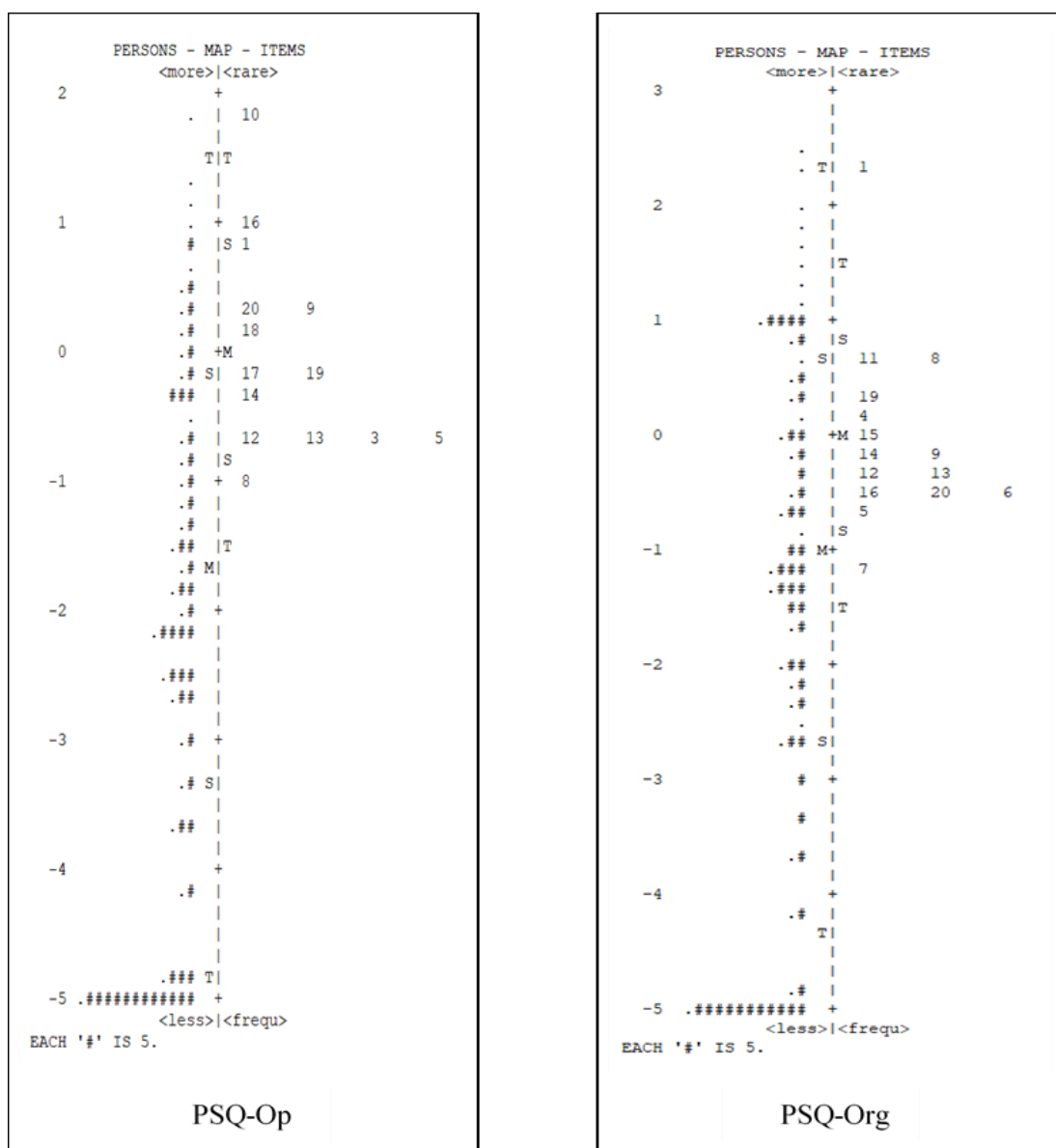


**Figure 2.** Wright Map of PSQ-Op and PSQ-Org Instrument

From Figure 2 on the PSQ-Op instrument, it can be seen that the hardest level of difficulty of the item to get a "a lot of stress" response is item number ten "tercukupinya makanan sehat saat bertugas" and the difficulty level of the item that is easy to get a response "a lot of stress" is item number eight "Tidak cukup waktu untuk dihabiskan bersama teman dan keluarga". The average ability person (person measure) of -2.58 logit (Standard Deviation = 2.24) is much lower than the average item measure of 0 (zero). This suggests that the respondents' tendency does not have stress levels related to work and personal life. This can be seen from the number of respondents who are in position two Standard Deviation below the average item.

On the PSQ-Org instrument, it can be seen that the difficulty level of the item that is most difficult to get a response "a lot of stress" is item number one "Bekerja sama dengan rekan kerja" and the difficulty level of the item that is easy to get a response "a lot of stress" is item number seven "Aturan yang tidak jelas dan berbelit-belit". The average ability person (person measure) of -1.97 logit (Standard deviation = 2.51) is much lower than the average item measure of 0 (zero). This indicates that respondents' tendency does not have a level of work stress related to perception of the demands of the organization. This can be seen from the number of respondents who are in position one Standard Deviation below the average item.

### Test Information Function

In addition to the above information, information is also produced in the form of test information function (TIF) that describes the value of information for each level of trait measured along with their respective error standards. Test Information Function (TIF) shows the functionality of the test when given to individuals with the level of trait obtained. Trait in question are items that measure non-cognitive test constructs obtained by individuals tending to be low, moderate, or, up to high. The better the item is targeted at the person, the more information the item provides about the person's parameters. Test information function that is expected to peak will be obtained reference test criteria tested and reference test sample model shows normal. With this, TIF is an effective test measurement range (Linacre, 2018). The TIF of PSQ-Op and PSQ-Org instruments can be seen in Figure 3.
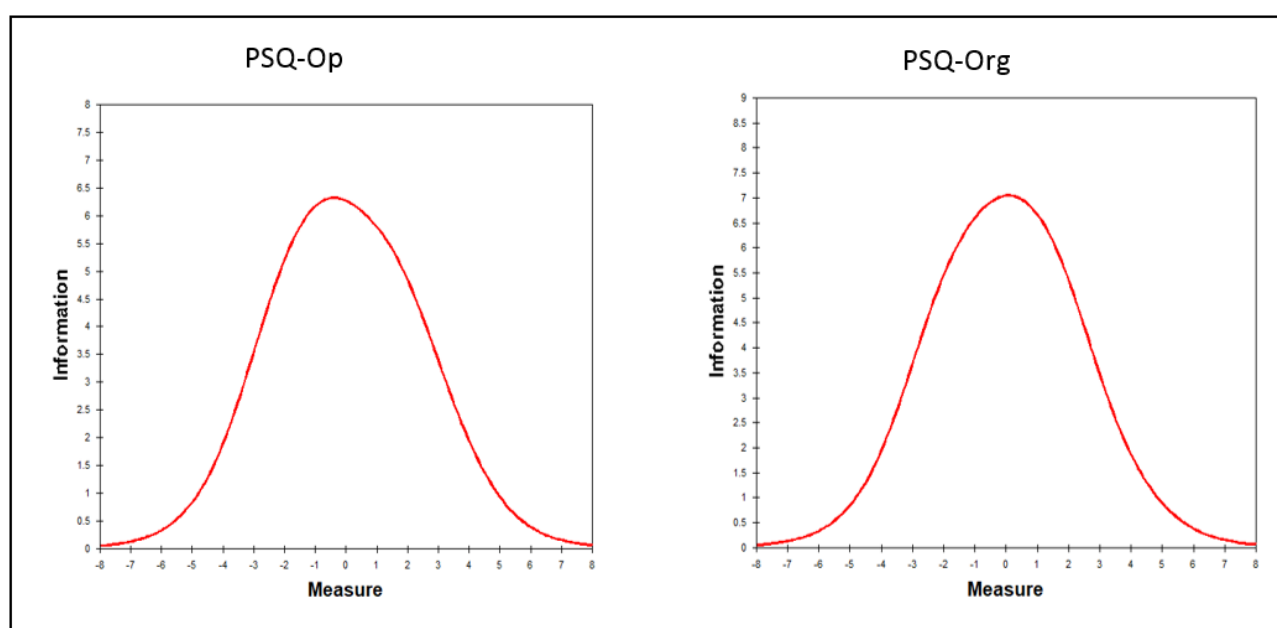


**Figure 3.** Test Information Functions of PSQ-Op and PSQ-Org Instrument

Figure 3 above shows the results of TIF analysis on PSQ-Op instruments with the curve peaks are at -0.40 logit points and on PSQ-Org instruments the curve peaks are at the 0.08 logit point, meaning that both instruments will function optimally to be administered to respondents who have a medium downward stress level. This means that information about tests on a set of items that measure stress constructs in police officers is very precise and optimal for people who have relatively low to moderate stress levels.

## Conclusion

Psychometric test evaluation results using Rasch RSM against both research instruments (PSQ-Op and PSQ-Org) can be concluded as follows. In the test of unidimensional assumptions, it was reported that the results of the analysis using the principal component analysis of residual (PCAR) methods of both instruments meet unidimensional assumptions, meaning that the items in both research instruments are unidimensional (single). However, although the items in both instruments were shown to measure the same dimension, local independence assumptions testing found some violations. Q3 statistics (Yen, 1984) with criteria 0.30 on PSQ-Op instrument found three items (item number one, item number eleven and item number fiveteen) and on the PSQ-Org scale found two items (item number three and item number seventeen) that have residual correlation between items >00.30, then those items are eliminated/discarded so that local independence assumptions are met.

Rasch RSM's analysis found that there were three items (item number two, item number four and item number six) and three items (item number eighteen, item number ten and item number two) on PSQ-Op and PSQ-Org instruments that had unacceptable MNSQ infit and outfit statistical values (out of the range of 0.5 to 1.5). Furthermore, the six items were eliminated/discarded because they were not fit to measure the stress of police work related to work and personal life as well as measuring the work stress of police officers related to the perception of the police to the demands of the organization.

The next finding is that both instruments, PSQ-Op and PSQ-Org have reliability for items and persons that are acceptable and fall into the category of good (>0.70). Even a perfect unidimensional scale would not be useful in practical terms if the resulting scale score had very low reliability. Thus, it can be interpreted that Rasch's reliability is acceptable and that PSQ-Op and PSQ-Org instruments can create useful scores in practice.

The next interesting finding in this study was the use of seven response categories in PSQ-Op instruments and PSQ-Org occurred threshold disorder. This is possible because respondents cannot distinguish response category two with response category three and response category five with response category six. But after the response category is collapsed into four response categories (response category one remains, response category two and three are combined into category two, response category four, five and six are combined into category three and category seven remains category four) response category performance results can work well, by showing that: (1) the average estimated size for each response category increases monotonously and in the expected direction when the response category moves from low to high category , (2) the threshold of adjacent response categories increase monotonously and in the expected direction, and 3) each of the five response categories is displayed acceptable infit statistics and MNSQ outfits. Thus, it is expected that the next study will conduct research using four categories of response (no stress at all- somewhat stress - enough stress - a lot of stress) to complete the findings in this study.

As for the limitations of this study, first the data collection of respondents using the help of google form application with non-probability sampling techniques may not be able to provide an accurate representation of the police population as a whole. However, the use of Rasch RSM analysis has provided a solution in testing the validity of stress measurement constructs in the Operational Police

Stress Questionnaire (PSQ-Op) and Organizational Police Stress Questionnaire (PSQ-Org) instruments as the methodology does not depend on the sampling involved, thus allowing the generalization of effective measurement properties of both stress measurement instruments. Second, further research is needed to fully understand the gender differences and differences in the function of police (staff and operational) assignments that play a role when assessing and considering the needs of individuals who play a role in stress measurement.

Overall, this model can be applied to future research and can provide a technical overview of the stages of data analysis for the application of the same analysis method. This study can be a reference for researchers in the field of psychology to conduct analysis with the Rating Scale Model (RSM) method.

## References

Abidin, Z., Shahrul, A., & Isha, N. (2014). *A psychometric properties of the Malay-version Police stress questionnaire*. 42–50.

Adams, R. J., Wu, M. L., & Wilson, M. (2012). The Rasch rating model and the disordered threshold controversy. *Educational and Psychological Measurement*, *72*(4), 547–573.

Andrich, D., & Marais, I. (2019). *A course in Rasch measurement theory: measuring in the educational, social and health sciences.* Springer Singapore

Aulya, D. (2013). Faktor-faktor yang berhubungan dengan stres kerja pada polisi lalu lintas di Polres Metro Jakarta Pusat Bulan April – Agustus 2013. Skripsi. Jakarta: FKM UIN.

Aziz, M. A. (2020). Uji klinis tingkat stres personil Polri: pengaruh hormon stres terhadap halitosis yang disebabkan bakteri patogen periodontal pada anggota Korps Brimob Polri. *Jurnal Litbang Polri*, *23*(1), 1–20.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). *Guidelines for the process of cross-cultural adaptation of self-report measures*. 3186–3191.

Bond, T., & Fox, C. (2015). Applying the Rasch model: fundamental measurement in the human sciences (3rd ed.). Mahwah, NJ: Lawrence Erblaum & Associates.

Bond, T. G., & Fox, C. M. (2007). Applying the Rasch model: fundamental measurement in the human sciences (2nd ed.). Mahwah, NJ: Lawrence Erlbaum

Cheung, K. C., & Mooi, L. C. (1994). A comparison between the rating scale model and dual scaling for likert scales. *Applied Psychological Measurement*, *18*(1), 1–13.

DeMars, C. E. (2003). Detecting multidimensionality due to curricular differences. *Journal of Educational Measurement*, *40*(1), 29–51.

DiStefano, C., Greer, F. W., Kamphaus, R. W., & Brown, W. H. (2014). Using Rasch rating scale methodology to examine a behavioral screener for preschoolers at risk. *Journal of Early Intervention*, *36*(3), 192–211.

DiStefano, C., & Morgan, G. B. (2010). Evaluation of the BESS TRS-CA using the Rasch rating scale model. *School Psychology Quarterly*, *25*(4), 202–212.

Deviana, T., Hayat, B., & Suryadi, B. (2020). Validasi instrumen social provision scale pada siswa Sekolah Menengah Atas: studi dengan penerapan Rasch model. *Indonesian Journal of Educational Assessment,*3(1), 1-15.

Duncan, P. W., Bode, R. K., Min Lai, S., Perera, S. (2003). Rasch analysis of a new stroke-specific outcome scale: the stroke impact scale. *Archives of Physical Medicine and Rehabilitation,* 84 (7), 950-963.

Eckes, T. (2011). Item banking for C-tests: A polytomous Rasch modeling approach. *Psychological Test and Assessment Modeling*, *53*, 414–439.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. L. Erlbaum Associates.

Geldenhuys, M., & Bosch, A. (2019). A Rasch adapted version of the 30-item bem sex role inventory (BSRI). *Journal of Personality Assessment*, 1–12. https://doi.org/10.1080/00223891.2018.1527343

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. SAGE.

Hayati, U. (2019). Pengaruh stres kerja terhadap kesejahteraan subjektif yang dimediasi oleh resiliensi pada polisi reserse kriminal di kepolisian daerah Jawa Barat (*Doctoral Dissertation*, Universitas Pendidikan Indonesia).

Hellman, C. M. (1997). Job satisfaction and intent to leave. *Journal of Social Psychology*, 137, 677– 689.

Hidayati, R., Purwanto, Y., & Yuwono, S. (2011). Kecerdasan emosi, stres kerja dan kinerja karyawan. *Jurnal Psikologi*, *2*(1).

Holt, R. R. (1993). Occupational stress. In L. Goldberger & S. Breznitz (Eds.), Handbook of stress: Theoretical and clinical aspects (2nd ed.) (pp. 342–367). New York: Free Press.

Houghton, S., Wood, L., Marais, I., Rosenberg, M., Ferguson, R., & Pettigrew, S. (2017). Positive mental well-being: a validation of a Rasch-derived version of the warwick-edinburgh mental well-being scale. *Assessment*, *24*(3), 371–386. https://doi.org/10.1177/1073191115609995

Jayanegara. (2007). Stres kerja dan coping pada Polisi Indonesia. Jakarta: Fakultas Psikologi Universitas Indonesia.

Kunz, S., Fellinghauer, C., & Peter, C. (2019). Measuring posttraumatic growth and depreciation after spinal cord injury: A Rasch analysis. *Rehabilitation Psychology*, *64*(4), 407–424. https://doi.org/10.1037/rep0000288

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping* (11. [print.]). Springer.

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement*, *3(2)*, 103–122.

Linacre, J. M. (2004). Rasch model estimation: further topics. *Journal of Applied Measurement*, *5 (1)*, 95–110.

Linacre, John M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3 (1)*, 85–106.

Linacre, John M. (2010). *Predicting responses from Rasch measures. 11 (1)*, 1–10.

Linacre, John M, & Wright, B. D. (2019). *A user's guide to BIGSTEPS: Rasch-model computer program*. Mesa Press.

McCreary, D. R., & Thompson, M. M. (2006). Development of two reliable and valid measures of stressors in policing: The operational and organizational police stress questionnaires. *International Journal of Stress Management*, *13*(4), 494–518. https://doi.org/10.1037/1072-5245.13.4.494

Meijer, R. R., Sijtsma, K., & Smid, N. G. (1990). Theoretical and empirical comparison of the mokken and the Rasch approach to IRT. *Applied Psychological Measurement*, *14*(3), 283–298. https://doi.org/10.1177/014662169001400306

Muhammad, A.N. (2004). Gambaran hubungan faktor-faktor dengan stres kerja pada polisi lalu lintas di kawasan terminal Kampung Melayu Jakarta 2004. Skripsi. Jakarta: FKM UI.

Muraki, E. (1990). Fitting a polytomous item response model to likert-type data. *Applied Psychological Measurement*, *14*(1), 59–71. https://doi.org/10.1177/014662169001400106

Nikam, P.K. & Shaikh, M. (2014). *Occupational stress, burnout and coping in police personel: finding from a systematic review. American InternationalJournal of Research in Humanities, Arts and Social Scienes.* ISSN (Print): 2328-3734

Pichardo, M. C., Cano, F., Garzón-Umerenkova, A., de la Fuente, J., Peralta-Sánchez, F. J., & Amate-Romera, J. (2018). Self-regulation questionnaire (SRQ) in Spanish adolescents: factor structure and Rasch analysis. *Frontiers in Psychology*, *9*, 1370. https://doi.org/10.3389/fpsyg.2018.01370

Purda, Valeria, N., Liliana, & Amelia. (2012). Burnout – an effect of professional stress in the police environment. *International Conference on Economics, Trade and Development IPEDR*, *vol.36.*

Queiros, C., Kaiseller, M & Da Silva, A.N. (2013). Burnout as predictor of aggressivity among police officer. *European Journal of Policing Studis, 1 (2), 110-135.* ISSN 2034-760

Rahayu, W., Putra, M. D. K., Iriyadi, D., Rahmawati, Y., & Koul, R.K. (2020). A Rasch and factor analysis of an Indonesia version of the student perception of opportunity competence development (SPOCD) questionnaire. Cogent Education, 7

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics*, *4*(3), 207–230.

Retnawati, H. (2017). *Validitas reliabilitas & karakteristik butir.* Yogyakarta: Parama Publishing.

Robbins, S. P. (2004). Teori organisasi, struktur, desain, dan aplikasi. (Alih bahasa: tim indeks). New Jersey: Prentice Hall.

Rost, J. (2001). *The growing family of Rasch models. In essays on item response theory*. Springer - Verlag New York.

Spector, P. E. (1997). Job satisfaction: Application, assessment, causes, and consequences. Thousand Oaks, CA: Sage.

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism*, *57*(8), 1358–1362. https://doi.org/10.1002/art.23108

Tett, R. P. & Meyer, J. P. (1993). Job satisfaction, organizational commitment, turnover intention, and turnover: Path analyses based on meta-analytic findings. Personnel Psychology, 46, 259 –293.

Wakita, T., Ueshima, N., & Noguchi, H. (2012). Psychological distance between categories in the likert scale: comparing different numbers of options. *Educational and Psychological Measurement*, *72*(4), 533–546. https://doi.org/10.1177/0013164411431162

Wright, B. D., & Mok, M. M. C. (2004). An overview of the family of Rasch measurement models. *Introduction to Rasch Measurement,* 1–24.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA Press.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, *8*(2), 125–145

Yulianto, H. (2020). Maslach burnout inventory-human services survey (MBI-HSS) versi bahasa Indonesia: studi validasi konstruk pada anggota Polisi. *JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)*, *9*(1), 19-29.

**Appendix 1.** *Organizational Police Stress Questionnaire (PSQ-Org)* (*English Version*)

## Organizational Police Stress Questionnaire

Below is a list of items that describe different aspects of being a police officer. After each item, please circle how much stress it has caused you over the past 6 months, using a 7-point scale (see below) that ranges from "No Stress At All" to "A Lot Of Stress" :

| No Stress At All | | Moderate Stress | | | A Lot of Stress | |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** |

| | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1. | Dealing with co-workers | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | The feeling that different rules apply to different people (e.g. favouritism) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 3 | Feeling like you always have to prove yourself to the organization | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 4 | Excessive administrative duties | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 5 | Constant changes in policy/legislation | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 6 | Staff shortages | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 7 | Bureaucratic red tape | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | Too much computer work | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 9 | Lack of training on new equipment | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 10 | Perceived pressure to volunteer free time | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 11 | Dealing with supervisors | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 12 | Inconsistent leadership style | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 13 | Lack of resources | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 14 | Unequal sharing of work responsibilities | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 15 | If you are sick or injured your co-workers seem to look down on you | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 16 | Leaders over-emphasise the negatives (e.g. supervisor evaluations, public complaints) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 17 | Internal investigations | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 18 | Dealing the court system | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 19 | The need to be accountable for doing your job | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 20 | Inadequate equipment | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

**Appendix 2.** *Operational Police Stress Questionnaire (PSQ-Op)* (*English Version*)

## Operational Police Stress Questionnaire

Below is a list of items that describe different aspects of being a police officer. After each item, please circle how much stress it has caused you over the past 6 months, using a 7-point scale (see below) that ranges from "No Stress At All" to "A Lot Of Stress":

| No Stress At All | | | Moderate Stress | | | A Lot of Stress |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

| | | |
|---|---|---|
| 1. | Shift work | 1 2 3 4 5 6 7 |
| 2 | Working alone at night | 1 2 3 4 5 6 7 |
| 3 | Over-time demands | 1 2 3 4 5 6 7 |
| 4 | Risk being injured on the job | 1 2 3 4 5 6 7 |
| 5 | Work related activities on days off (e.g. court, community events) | 1 2 3 4 5 6 7 |
| 6 | Traumatic events (e.g. MVA, domestics, death, injury) | 1 2 3 4 5 6 7 |
| 7 | Managing your social life outside of work | 1 2 3 4 5 6 7 |
| 8 | Not enough time available to spend with friends and family | 1 2 3 4 5 6 7 |
| 9 | Paperwork | 1 2 3 4 5 6 7 |
| 10 | Eating healthy at work | 1 2 3 4 5 6 7 |
| 11 | Finding time to stay in good physical condition | 1 2 3 4 5 6 7 |
| 12 | Fatigue (e.g. shift work, over-time) | 1 2 3 4 5 6 7 |
| 13 | Occupation-related health issues (e.g. back pain) | 1 2 3 4 5 6 7 |
| 14 | Lack of understanding from family and friends about your work | 1 2 3 4 5 6 7 |
| 15 | Making friends outside the job | 1 2 3 4 5 6 7 |
| 16 | Upholding a "higher image" in public | 1 2 3 4 5 6 7 |
| 17 | Negative comments from the public | 1 2 3 4 5 6 7 |
| 18 | Limitations to your social life (e.g. who your friends are, where you socialize) | 1 2 3 4 5 6 7 |
| 19 | Feeling like you are always on the job | 1 2 3 4 5 6 7 |
| 20 | Friends / family feel the effects of the stigma associated with your job | 1 2 3 4 5 6 7 |

**Appendix 3.** *Organizational Police Stress Questionnaire (PSQ-Org)* (*Indonesian version*)

### Organizational Police Stress Questionnaire (PSQ-Org)

Di bawah ini adalah daftar pernyataan yang menggambarkan berbagai aspek yang berbeda dari tugas polisi. Pada masing-masing pernyataan, pilihlah seberapa besar tekanan yang Anda rasakan dari situasi tersebut dalam 6 bulan terakhir dengan melingkari satu dari tujuh skala sesuai kondisi yang Anda rasakan. Rentang skala dari "**Tidak Stres Sama Sekali**" sampai "**Sangat Stres**"

| Tidak Stres Sama Sekali | | | Stres Sedang | | | Sangat Stres |
|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** |

| No | Pernyataan | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1. | Bekerja sama dengan rekan kerja | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | Merasa aturan yang ada tidak adil untuk semua personel (pilih kasih) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 3 | Merasakan Anda selalu harus membuktikan diri Anda kepada organisasi | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 4 | Tugas-tugas administratif (pembuatan laporan, pemberkasan kasus) setiap hari yang terlampau banyak (di luar batas normal) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 5 | Kebijakan yang selalu berubah-rubah dan/atau tumpang tindih (*overlapping*) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 6 | Keterbatasan jumlah Personel/anggota | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 7 | Aturan yang tidak jelas dan berbelit-belit | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 8 | Terlalu banyak pekerjaan yang harus diselesaikan dengan komputer | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 9 | Kurang pelatihan dan sosialisasi peralatan/teknologi baru | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 10 | Adanya perintah untuk mengerjakan tugas di luar ketentuan kerja | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 11 | Menghadapi pengawasan dan pemeriksaan (wasrik) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 12 | Tipe/gaya pemimpin, yang sering kali berganti (tidak konsisten) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 13 | Keterbatasan sarana/prasarana dalam melakukan pekerjaan | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 14 | Tanggung jawab pekerjaan yang tidak terbagi rata dengan rekan lain | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 15 | Sekiranya Anda sakit/cidera rekan kerja seolah meremehkan | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 16 | Atasan terlalu mempermasalahkan hal-hal kecil/sepele | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 17 | Berurusan dengan pemeriksa/sidang disiplin | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 18 | Berurusan dengan sistem peradilan umum | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 19 | Standar pertanggungjawaban kinerja menyelesaikan tugas (misal: reserse 1x24 jam harus bisa merespon laporan pengaduan) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 20 | Keterbatasan/kelayakan peralatan (tidak *up-to-date*) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

**Appendix 4.** *Operational Police Stress Questionnaire (PSQ-Op)* (*Indonesian version*)

### Operational - Police Stress Questionnaire (PSQ-Op)

Di bawah ini adalah daftar pernyataan yang menggambarkan berbagai aspek yang berbeda dari tugas polisi. Pada masing-masing pernyataan, pilihlah seberapa besar tekanan yang Anda rasakan dari situasi tersebut dalam 6 bulan terakhir dengan melingkari satu dari tujuh skala sesuai kondisi yang Anda rasakan. Rentang skala dari "**Tidak Stres Sama Sekali**" sampai "**Sangat Stres**".

| Tidak Stres Sama Sekali | | | Stres Sedang | | | Sangat Stres |
|---|---|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** | **6** | **7** |

| | | |
|---|---|---|
| 1. | Kerja menggunakan sistem shift | 1 2 3 4 5 6 7 |
| 2 | Bekerja sendirian di malam hari | 1 2 3 4 5 6 7 |
| 3 | Dituntut kerja lembur | 1 2 3 4 5 6 7 |
| 4 | Risiko cedera saat bekerja | 1 2 3 4 5 6 7 |
| 5 | Kegiatan terkait pekerjaan di hari libur (misalnya: pengadilan, acara komunitas) | 1 2 3 4 5 6 7 |
| 6 | Peristiwa-peristiwa traumatis (contoh: kecelakaan kendaraan bermotor, masalah rumah tangga, kematian, cidera) | 1 2 3 4 5 6 7 |
| 7 | Mengelola kehidupan sosial Anda di luar pekerjaan | 1 2 3 4 5 6 7 |
| 8 | Tidak cukup waktu untuk dihabiskan bersama teman dan keluarga | 1 2 3 4 5 6 7 |
| 9 | Melakukan pekerjaan yang berhubungan dengantulis menulis / administratif | 1 2 3 4 5 6 7 |
| 10 | Makan yang sehat di tempat kerja | 1 2 3 4 5 6 7 |
| 11 | Meluangkan waktu untuk tetap dalam kondisi fisik yang baik | 1 2 3 4 5 6 7 |
| 12 | Kelelahan (misal : kerja shift, lembur) | 1 2 3 4 5 6 7 |
| 13 | Masalah kesehatan terkait pekerjaan (misal : nyeri punggung) | 1 2 3 4 5 6 7 |
| 14 | Kurangnya pemahaman dari keluarga dan teman-teman terkait pekerjaan Anda | 1 2 3 4 5 6 7 |
| 15 | Berteman di luar pekerjaan | 1 2 3 4 5 6 7 |
| 16 | Menunjukkan sosok teladan di depan umum | 1 2 3 4 5 6 7 |
| 17 | Komentar-komentar negatif dari masyarakat | 1 2 3 4 5 6 7 |
| 18 | Keterbatasan dalam kehidupan sosial Anda (misal. Siapa teman Anda, dimana tempat Anda bersosialisasi) | 1 2 3 4 5 6 7 |
| 19 | Merasa Anda selalu bekerja | 1 2 3 4 5 6 7 |
| 20 | Rekan/keluarga merasakan efek stigma berkaitan dengan pekerjan Anda. | 1 2 3 4 5 6 7 |