

Nikolina KENIG

UDK: 005.953.2:159.9.072.42

005.953.2:303.622

Original research paper

## INITIAL STUDY OF THE PSYCHOMETRIC CHARACTERISTICS OF INDIVIDUAL WORK PERFORMANCE QUESTIONNAIRE

### *Abstract*

*This study is an initial analysis of the functioning of the brief form of the Individual Work Performance Questionnaire (IWPQ) translated in Macedonian language. IWPQ measures three dimensions of job performance across wide variety of jobs and by assessing work performance, contextual performance and counterproductive behaviour. The sample was comprised of 404 employees (55.7% women) from 14 occupational backgrounds, with average age  $M = 38$  ( $SD = 12.1$ ) and average working experience  $M = 12.6$  ( $SD = 10.4$ ) years. All participants voluntarily answered the online based instrument which included demographic data, IWPQ, OLBI and the UWES, sent at their personal e-mail addresses.*

*Descriptive statistics, item-analysis, confirmatory factor analysis and correlations with two other included constructs were performed in order to explore the psychometric characteristics of the instrument in this particular cultural context. Results confirmed the tridimensional structure of IPWQ with marginal fit indicators, while the Alpha coefficients suggested adequate reliability of the scales ( $\alpha=0.73-0.86$ ). The construct dimensions exhibit significant associations with other measures expected to be related to work performance. It was concluded that although the IWPQ version translated in Macedonian is adequate to be used as a measure of job performance, further research, particularly adjustment of the content of items might contribute to improvement of its validity.*

**Key words:** *individual work performance, psychometric characteristics*

### **Introduction**

A significant body of research in the field of industrial-organizational psychology oriented towards the effects and prerequisites of individual work performance (IWP) has not been preceded by many efforts to generate psychometrically well-founded measures. The most commonly used definition of the individual work performance is that it represents different "behaviours or actions that are relevant to the goals of the organization" (Campbell, 1990, p.704). This definition suggests that the individual work performance focuses on behaviours of employees, rather than the results of these actions, which

are under their own control, free of any constraints that might come from the environment (Koopmans et al, 2014). The overview of the conceptual backgrounds of the individual work performance in the literature identified 17 generic frameworks that apply across all occupations and another 18 that are job-specific (Koopmans et al., 2011). The same study proposes a heuristic framework of IWP in which it is represented through four dimensions: (1) task performance, (2) contextual performance, (3) counterproductive work behaviour and (4) adaptive performance. All of them together capture the complete range of behaviours that constitute individual work performance in occupation, although the extent of relevance of these dimensions might vary from one to another occupational context.

The *task performance*, being the most frequently mentioned and examined dimension, has been described as the competence with which employees perform the tasks central to their job (Campbell, 1990). The activities and behaviours that add to the organization's effectiveness in many ways that shape the psychological and social context of the organization, which are considered catalysts for the operations and activities of the mission, comprise the second dimension, named as contextual performance (Borman & Motowidlo, 1997). It includes performing tasks beyond job duties, initiative and enthusiasm. Although there were ideas that this dimension might have two aspects, one regarding behaviours directed toward the work and another toward the people, meta-analytic studies have found that its best interpretation is through one unidimensional construct (Hoffman et al., 2007). The third dimension of IWP is counterproductive work behaviour and it refers to activities that harm the image and the well-being of the organization (Rotundo & Sackett, 2002). Adaptive performance is manifested through both behaviours and the capacity to cope with change. It consists of actions like taking responsibility with uncertain or unpredictable work situations, learning new skills and the potential for adjusting to different physical surroundings or cultures or to personalities of the co-workers (Daderman et al., 2020). This last dimension is considered as a separate one by some researchers, while others include it in contextual performance.

### **Description of Individual Work Performance Questionnaire (IWPQ)**

The main intention behind IWPQ development was to surpass the limitations in the previously launched scales that measure the construct. Koopmans et al. (2014) discuss several problems inherent to formerly used instruments intended to operationalize this variable. Most remarkably, none of them captures all dimensions of individual work performance, which results into decreased efficacy in its measurement. On the other hand, when several scales that measure different dimensions are administered simultaneously, leads to using items overlapping in content. In addition, due to differences in conceptual definitions or target populations, previously used scales use different

operationalization of the same dimensions which decreases the comparability across different studies.

The IWPQ incorporates all identified dimensions of IWP. One of the most important advantages is that the operationalization of the dimensions has been proposed considering workers in all types of occupations, without inclusion of items that overlap across different dimensions. All these together, make the scale both functional for workers in all types of occupations, also for all identified aspects of work performance and without redundant indicators of the construct.

At the beginning, it was launched as a 27-item scale (Koopmans, et al., 2013), has a shorter version with 18 items, which was used in this study. The short-form scale was created by implementing Rasch analysis (Koopmans, et al. 2014) and it is referred to as ISPW 3.0. Although the framework proposed 4 dimensions of work performance, the included items refer to only three dimensions of job performance: 1. Task performance (5 items), 2. Contextual performance (8 items), and 3. Counterproductive work behaviour (5 items). As it is a self-administered questionnaire, respondents are asked to recall period of the last three months and report their feelings and behaviour related to work at 5-point rating scale (0 = *seldom* to 4 = *always* for task and contextual performance; and 0=*never* to 4=*often* for counterproductive work behaviour).

Koopmans et al. (2014) report convincing convergent validity and excellent discriminative validity of the IWPQ 3.0. and acceptable internal consistency of the subscales: Task performance had  $PSI=0.82^1$ , for Contextual performance it was 0.90 for, and 0.79 for Counterproductive work behaviour, Overall IWPQ reliability, similarly to overall score, cannot be calculated as their valid calculation requires one-dimensionality of the variable being measured.

Several cross-cultural validation studies have shown that IWPQ could be successfully used across different cultures. It was successfully translated and adapted from Dutch (the original language) to English (Koopmans et al., 2016), Swedish (Daderman, Ingelgard, & Koopmans, 2020), Spanish (Ramos-Villagrasa et. al., 2019) and it is used as a valid instrument in Argentina (Gabini & Salessi, 2016), Indonesia (Widyastuti & Hidayat, 2018) and South Africa (Van der Vaart, 2021).

The aim of this initial study was to provide translation of IWPQ from English to Macedonian language, along with information on the questionnaire's internal consistency, validity and goodness of fit with the suggested three performance dimensions.

---

<sup>1</sup> The PSI assesses how adequately the set of items can distinguish subjects on different levels of the scale and its required is value  $\geq 0.70$  (Pallant & Tennant, 2007).

## Proposed analyses

The analysis of gathered data will start with reporting the averages for the included items, as well as for the three dimensions. An internal criterion (correlations with the subscale) will be used to check the discriminability of the items of the IWPQ scale. The reliability is to be determined by calculating the Cronbach's alpha coefficients. The comparison of the internal structure of IWPQ scale with the model offered by the authors of the instrument will be performed by confirmatory factor analysis. In line with previous research (e.g. Koopmans et al., 2014a), two hypotheses that aimed to support the IWPQ's validation process will be tested by calculating Pearson coefficients: *Hypothesis 1*: Task performance as well as Contextual performance are negatively correlated with exhaustion and disengagement and positively connected with, dedication, vigor and absorption and *Hypothesis 2*: Counterproductive work behaviour is positively correlated with exhaustion and disengagement and negatively connected with dedication, vigor and absorption.

## METHOD

### Participants

The convenient sample consisted of 404 employees (55.7% women) of 14 professions divided in three occupational sectors: blue collar (eg. manual and factory workers - 34.2%), pink collar (eg. service workers like nurses and teachers - 31.9%), and white collar (eg. office workers, like lawyers, architects and It engineers - 33.9%). Their working experience ranged from 1 to 40 years ( $M = 12.6$ ;  $SD = 10.4$ ), whereas the average age was  $M = 38$ ;  $SD = 12.1$ . All respondents were formally employed in 17 different towns in the country. Vast majority of participants (89.4%) self-declared themselves as ethnic Macedonians. Participation was voluntary, anonymous and not compensated. The recruitment of participants was facilitated by the students at the Institute of Psychology at the Faculty of Philosophy.

### Instruments used to test hypotheses of IWPQ's validity

**The Oldenburg Burnout Inventory (OLBI)** was developed by Demerouti and Nachreiner in 1998 (Demerouti & Bakker, 2008) as a measure of burnout. It contains 16 items (8 reversed) and covers two dimensions: exhaustion and disengagement. It has been extensively used in research to measure job and academic burnout as well as its impact on well-being and mental health. The internal consistency of this scale obtained with this particular sample was  $\alpha=0.82$ .

The **Utrecht Work Engagement Scale (UWES)** was designed to assess the extent of work engagement. The 17-item version measures three aspects of this variable – vigor, dedication and absorption (Schaufeli & Bakker, 2003). Work engagement is characterized by a high level of work related enthusiasm and identification with one's work, meaning that it could be regarded as being opposite of work burnout operationalized through a low level of energy and low identification with the one's work. For this particular sample, the internal consistency of the scale is similar ( $\alpha=0.94$ ) to the coefficients reported in previous studies (e.g. Goliath-Yarde, & Roodt, 2011).

### **Procedure of translating and administering the instruments**

In the forward phase, the instruments were translated from English by a professional interpreter. Then, in consultation with the author of the study, certain phrases were changed in order to sound more colloquially. This version was translated back to English by an independent translator. The two versions were then compared and found to be similar to each other.

All included instrument along with the demographic questions were administered online, by contacting the potential respondents via their personal e-mail addresses. A letter explaining the purpose of the research, the guarantee of anonymity and confidentiality and the right to withdraw from the study without foreseen negative consequences, accompanied the questionnaire. The responses were collected in April 2022.

### **Statistical analysis**

All calculations, with the exception of the CFA that was performed by JASP 16, were completed by using the SPSS-26 statistical package. We used the standard criteria for goodness of fit: RMSEA values of 0.01, 0.05 and 0.08 indicate excellent, good and mediocre fit respectively, while CFI/GFI values greater than 0.90 excellent and from 0.80 to 0.90, marginal fit.

### **Findings**

Table 1 presents descriptive statistics of the IWPQ subscales (dimensions). Compared to others obtained so far, these averages are higher for all included dimensions<sup>2</sup>. All Alpha coefficients are within the range of being sufficiently high.

---

<sup>2</sup> For example, in a sample from the Swedish translation, the average for WP is  $M = 2.62$  for CP is  $M=2.91$  and for CPB,  $M= 0.98$  (Daderman et al., 2020).

**Table 1.**  
*Descriptive data and reliability of the IWPQ subscales*

IWPQ dimensions	n	Min.	Max	M*	SD	Kurt**	Skw***	$\alpha$
WP – Work performance	5	7	20	3.48	0.50	1.65	-1.22	.73
CP – Contextual performance	8	5	48	3.13	0.71	.97	-1.07	.84
CPB – Counter-productive behaviour	5	0	20	1.24	1.04	.12	.78	.86

\*M is computed by adding the scores of the items of each subscale and then dividing the total on the number of items in the subscale,\*\*SE = .242, \*\*\*SE = .121

The findings of analysis of items is presented at Table 2. All item-to-scale correlations were statistically significant, ranging between 0.28 and 0.88. The range of these values is considerably wider than in the other studies<sup>3</sup>, similarly as the averages of the items are higher than those reported in other studies. This table also shows that only item 5 improves the value of Cronbach's alpha if deleted.

**Table 2.**  
*Descriptive statistics of the IWPQ items (total and across occupational sectors)*

IWPQ item	r	$\alpha$ if item deleted	M <sub>wh</sub> (SD)	M <sub>bl</sub> (SD)	M <sub>pnk</sub> (SD)	M <sub>all</sub> (SD)
<b>WP – In the past three months...</b>						
1. I managed to plan my work so that I finished it on time	.78**	.644	3.35 (.98)	3.43 (.81)	3.56 (.67)	3.45 (.80)
2. I kept in mind the work result I needed to achieve	.80**	.622	3.45 (.71)	3.36 (.84)	3.56 (.64)	3.45 (.74)
3. I was able to set priorities	.74**	.642	3.53 (.67)	3.41 (.84)	3.62 (.63)	3.52 (.72)
4. I was able to carry out my work efficiently	.70**	.674	3.57 (.86)	3.49 (.72)	3.64 (.57)	3.57 (.62)

<sup>3</sup> For instance, the Argentinian version reported values from  $r=0.59-0.84$  (Gabini & Salessi, 2016) and the Swedish adaptation (Daderman et al., 2020) identified values from  $r=0.40$  to  $0.69$ .

5. I managed my time well	.38**	.800	3.43 (.74)	3.43 (.73)	3.46 (.66)	3.44 (.71)
<hr/>						
CP – In the past three months...						
<hr/>						
6. On my own initiative, I started new task when my old tasks were completed	.24**	.829	2.96 (1.10)	2.88 (1.26)	3.09 (1.08)	2.97 (1.54)
7. I took on challenging tasks when they were available	.31**	.819	2.92 (1.12)	2.86 (1.23)	3.02 (1.12)	2.93 (1.14)
8. I worked on keeping my job-related knowledge up-to-date	.41**	.816	3.14 (1.01)	3.11 (.99)	3.31 (.80)	3.18 (.95)
9. I worked on keeping my work skills up-to-date	.42**	.826	3.21 (.89)	3.01 (1.12)	3.37 (.87)	3.19 (.98)
10. I came up with creative solutions for new problems	.28**	.821	3.26 (.96)	3.29 (1.03)	3.49 (.72)	3.34 (.91)
11. I took on extra responsibilities	.34**	.820	3.17 (.94)	3.12 (1.07)	3.35 (.82)	3.21 (.95)
12. I continually sought new challenges in my work	.35**	.816	3.01 (1.02)	2.91 (1.13)	3.11 (.97)	3.00 (1.04)
13. I actively participated in meetings and/or consultations	.32**	.837	3.16 (1.09)	2.91 (1.26)	3.48 (.92)	3.18 (1.12)
<hr/>						
CWB– In the past three months...						
<hr/>						
14. I complained about minor work-related issues at work	.71**	.863	1.33 (1.23)	1.17 (1.39)	1.31 (1.32)	1.27 (1.29)
15. I made problems at work bigger than they were	.80**	.829	1.04 (1.17)	1.02 (1.32)	.92 (1.22)	1.02 (1.21)

16. I focused on the negative aspects of situation at work instead of the positive aspects	.77**	.840	1.09 (1.19)	1.09 (1.33)	1.94 (1.12)	1.04 (1.27)
17. I talked to colleagues about the negative aspects of my work	.88**	.802	1.30 (1.13)	1.56 (1.51)	1.29 (1.37)	1.39 (1.40)
18. I talked to people outside the organization about the negative aspects of my work	.85**	.813	1.43 (1.22)	1.67 (1.42)	1.40 (1.36)	1.50 (1.34)

r= item-to-scale correlation,  $\alpha$ =Cronbach alpha if item deleted, Mwh=M of white collars, Mbl=M of blue collars, Mpnk=M of pink collars

CFA was run with an assumption that the factors are correlated. The tested model corresponds to the version of IWPQ with three dimensions. The model fit the data was marginal ( $\chi^2= 3249.7$ ,  $df =153$ ,  $p<.001$ , CFI = 0.86, RMSEA = 0.08, SRMR = 0.08), since CFI, values were less 0.90, SRMR and the RMSEA values were 0.08, being higher than the cut off values for even mediocre fit.

The item „*I could work efficiently*“ has the lowest factor loading and according to the misfit plot, it exhibits high misfit with the item “*I was able to set my priorities*”. Another high misfit was registered between the items “*On my own initiative, I started new task when my old tasks were completed*” and “*I took on challenging tasks when they were available*”. Further examination of the relations among items has shown that the ability to work efficiently correlates negatively with initiative and taking challenging tasks. Withdrawal of these items from the model did not contribute to improve the fit.

In this model, task performance and contextual performance correlated at 0.55; task performance and counterproductive work behaviour correlated at -0.13; and contextual performance and counterproductive work behaviour correlated at -0.07. These correlations are not very different from the ones assessed of the same model in previous research.

The results regarding the tested associations between the IWPQ dimensions and aspects of burnout and work engagement are as expected (with only one exception for the relationship between counterproductive behaviour and disengagement) and provide basic evidence for the instrument’s nomological validity. The correlation coefficients are presented in Table 4.

**Table 4.***Correlations of IWPQ with job burnout (OLBI) and work engagement (UWES)*

	WP	CP	CPB
OLBI – Exhaustion	-.23**	-.42**	.53**
OLBI- Disengagement	.51**	.74**	-.08
UWES - Vigor	.52**	.87**	-.21**
UWES - Absorption	.49**	.92**	-.12*
UWES - Dedication	.50**	.89**	-.15**

\*\* $p < .01$  \* $p < .05$

**Discussion**

The version of the individual work performance (IWP) scale that was translated in Macedonian has good psychometric performances from the perspective of high reliability coefficients and nomological validity. However, the fit model of the construct is only marginal, suggesting that it needs further refinement in translation and adjustment to the cultural context. The data from studies conducted with the same instrument and methodology show mediocre fit. For example, the Spanish (Ramos-Villagrasa et al., 2019) and the Indonesian (Ramdani, 2019) version reported similar values of indicators of fit as this one. This might indicate that some of the items are not equivalent in meaning in different cultural contexts and that further adjustment is needed based on how the particular culture construes certain meanings. In this specific case, it is likely that the translation itself is the major challenge. The items that exhibited misfit, particularly “I managed my time well” and “I took on challenging tasks when they were available” needed to be translated differently from the original English version (more descriptively), because of lack of terms that are exact equivalents that would sound enough conversantly.

On the other hand, poor fits might stem from the methodology used to assess the equivalence of models. As it has been previously argued, in CFA, the factor loadings are estimated with an assumption that each item will load on the expected factor only (Ramos-Villagrasa et al., 2019) and when this assumption is not met, the fit is substantially reduced.

In conclusion, this initial study provides evidence that the IWPQ could be used in Macedonian language, although not without caution. It shows similar factor structure as in the original language and acceptable internal reliability. Future research should focus on whether the ESEM fit would perform better than CFA fit, as well as whether further adjustments of the items that were shown to have poor loadings might improve the validity.

**Limitations**

In addition to the fact that the sample was convenient, this initial study has other limitations. It was based only on self-report measures that might include considerable bias. Thus, further research should analyse whether the findings could be replicated with measurements that employ different ratters, such as supervisors or peers. In addition to that, the instrument was administered online which adds risks to data integrity due to the impossibility to control the overall process of responding. Finally, another source of uncontrolled variance might be the period of instrument administration. All the items require recalling the last three months, which for many respondents, especially those from the white-collar occupations might have been the months when they still worked from home due to the anti-Covid measures.

## References

- Borman, W. C., & Motowidlo, S. J. (1997). Task Performance and Contextual Performance: The Meaning for Personnel Selection Research. *Human Performance*, 10(2), 99–109. [https://doi.org/10.1207/s15327043hup1002\\_3](https://doi.org/10.1207/s15327043hup1002_3)
- Campbell, J.P.(1990). Modeling the performance prediction problem in industrial and organizational psychology. In: Dunnette, M.D., Hough, L.M., (Eds.) *Handbook of industrial and organizational psychology*, (pp. 687-732). Consulting Psychologists Press.
- Daderman, A.M., Ingelgard, A. & Koopmans, L. (2020). Cross-cultural adaptation, from Dutch to Swedish language, of the Individual Work Performance Questionnaire. *Work*, 65 (1), 97-109.
- Demerouti, E., & Bakker, A. B. (2008). The Oldenburg Burnout Inventory: A good alternative to measure burn-out and engagement. In J. Halbesleben (Ed.), *Stress and burnout in health care* (pp. 65–78). Nova Sciences.
- Gabini, S. & Salessi, S. (2016). Validation of the job performance scale in Argentinean workers, *Revista Evaluar* 16(1), 10–26. <https://doi.org/10.35670/1667-4545.v16.n1.15714>
- Goliath-Yarde, L, & Roodt, G. (2011). Differential item functioning of the UWES-17 in South Africa. *SA Journal of Industrial Psychology*, 37(1), 01-11.
- Hoffman, B. J., Blair, C. A., Meriac, J. P., & Woehr, D. J. (2007). Expanding the criterion domain? A quantitative review of the OCB literature. *Journal of Applied Psychology*, 92, 555-566. <https://doi.org/10.1037/0021-9010.92.2.555>
- Lubbadeh, T. (2021). The Phenomenon of Job Burnout and its Relation to Employees' Performance: Empirical Evidence From The Jordanian Banking Sector With International Perspective, *Doctoral thesis* defended at the University of PÉCS, Lubbadeh Tareq Bassam Ramez\_Disszertáció.pdf (pte.hu)
- Koopmans, L., Bernaards, C., Hildebrandt, V., Schaufeli, W., De Vet, H. & van der Beek, A. (2011). Conceptual Frameworks of Individual Work Performance. *Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine*. 53. 856-66. 10.1097/JOM.0b013e318226a763.
- Koopmans, L., Bernaards, C., Hildebrandt, V., van Buuren, S., van der Beek, A.J. and de Vet, H.C.W. (2013). "Development of an individual work performance questionnaire", *International Journal of Productivity and Performance Management*, Vol. 62 No. 1, pp. 6-28. <https://doi.org/10.1108/17410401311285273>
- Koopmans, L., Bernaards, C., Hildebrandt, V., Buuren, S., van der Beek, A., & De Vet, H. (2014). Improving the Individual Work Performance Questionnaire using Rasch Analysis, *Journal of applied measurement*. 15. 160-75. 10.1136/oemed-2013-101717.51.
- Koopmans, L., Coffeng, J. K., Bernaards, C. M., Boot, C. R., Hildebrandt, V. H., de Vet, H. C., & van der Beek, A. J. (2014a). Responsiveness of the

- individual work performance questionnaire. *BMC public health*, 14, 513. <https://doi.org/10.1186/1471-2458-14-513>
- Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., Lerner, D., De Vet, H., C. W., & Van Der Beek, A. J. (2016). Cross-cultural adaptation of the Individual Work Performance Questionnaire. *Work*, 53, 609-619 <https://doi.org/10.3233/WOR-152237>
- Pallant J.F., Tennant A. (2007). An introduction to the Rasch measurement model: an example using the hospital anxiety and depression scale (HADS). *Journal of Clinical Psychol.*; 46/ 1,1–18.
- Ramdani, Z. (2019). The individual work performance scale: a psychometric study and its application for employee performance, *Humanities & Social Sciences Reviews*, 7/5, 405-415. <https://doi.org/10.18510/hssr.2019.7545>
- Ramos-Villagrasa, P. J., Barrada, J. R., Fernández-del-Río, E., & Koopmans, L. (2019). Assessing job performance using brief self-report scales: The case of the individual work performance questionnaire. *Journal of Work and Organizational Psychology*, 35, 195-205. <https://doi.org/10.5093/jwop2019a21>
- Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: a policy capturing approach. *Journal of Applied Psychology*, 87(1), 66–80. <https://doi.org/10.1037/0021-9010.87.1.66>
- Schaufeli, W.B., & Bakker, A. (2003). *Utrecht Work Engagement Scale*. Utrecht University, Occupational Health Psychology Unit.
- Van der Vaart, L. (2021). The performance measurement conundrum: Construct validity of the Individual Work Performance Questionnaire in South Africa, *South African Journal of Economic and Management Sciences* 24(1), a3581. <https://doi.org/10.4102/sajems.v24i1.3581>
- Widyastuti, T. & Hidayat, R., (2018). Adaptation of individual work performance questionnaire (IWPQ) into Bahasa Indonesia, *International Journal of Research Studies in Psychology* 7(2), 101–112. <https://doi.org/10.5861/ijrsp.2018.3020>