AOSIS

The performance measurement conundrum: Construct validity of the Individual Work Performance Questionnaire in South Africa

Author:
Leoni van der Vaart[1] ◉

Affiliation:
[1]School of Industrial Psychology, Faculty Economic and Management Sciences, North-West University, Optentia Research Focus Area, Vanderbijlpark, South Africa

Corresponding author:
Leoni van der Vaart,
leoni.vandervaart@nwu.ac.za

Read online:
Scan this QR code with your smart phone or mobile device to read online.

Background: Performance remains high on the agenda of both researchers and practitioners, and rigorous measurement is central to evidence-based investigations. Several instruments have been developed to measure performance, of which one – the Individual Work Performance Questionnaire (IWPQ) – seems to be a scientifically rigorous instrument. Despite the growth in IWPQ validation studies, little is known about its psychometric properties in South Africa.

Aim: The current study seeks to address the construct validity of the IWPQ. The objective was to evaluate the construct (i.e. convergent, discriminant and nomological) validity of the instrument.

Setting: Information technology professionals were targeted in the current study, and the final sample consisted of 296 employees.

Method: Latent variable modelling was performed, using both independent cluster and exploratory structural equation model frameworks, with the mean- and variance-adjusted weighted least squares estimator. The analysis was complemented with the heterotrait-monotrait method.

Results: The results supported the construct validity of the instrument. Specifically, individual work performance was found to be a three-dimensional construct with acceptable convergent and discriminant validity. Job resources were also related to the three performance dimensions.

Conclusion: The factor structure of the IWPQ mirrors the factor structure found in the literature. Researchers and practitioners can use this instrument in South Africa to identify determinants and consequence of individual work performance, as well as to evaluate the effectiveness of performance improvement interventions.

Keywords: Individual Work Performance Questionnaire; autonomy; social support; coaching; opportunities for development; validation.

Introduction

Employee performance – 'behaviours and actions that support organisational goals' (Campbell 1990:67) – matters for both individuals and organisations. High-performing individuals are rewarded with bonuses and advancement opportunities, whereas organisations gain a competitive advantage and higher financial returns (Yeshitila & Beyene 2019). Information technology (IT) professionals are particularly under pressure to perform because of the expectations resulting from the digitisation and automation of work (i.e. fourth industrial revolution) (Van Zyl et al. 2019). Consequently, organisations are interested in ways to improve their (IT) employees' performance and often rely on the results of empirical investigations to make decisions relating to performance improvement (Koopmans et al. 2011, 2015). Reliable and valid research results are needed for organisational decision-making and can mainly be ensured through the use of psychometrically sound measuring instruments in investigations (see Ramos-Villagrasa et al. 2019).

Several instruments exist for measuring performance. One of these instruments, the Individual Work Performance Questionnaire (IWPQ) (Koopmans et al. 2013), was constructed using a conceptual framework developed from the results of a comprehensive review of the performance literature (Koopmans et al. 2011). The instrument not only has a solid theoretical underpinning, but its developers took an iterative approach in developing and improving the questionnaire before making it available for use. Several other advantages also exist: (1) as a multidimensional

instrument, it captures the full range and complexity of performance behaviours necessary in the contemporary world of work, (2) it is a generic instrument that can be used in any organisation, enhancing its practical usefulness and enabling the generalisation of findings across studies, and (3) it is also short (18 items) and easy to administer (Dåderman, Ingelgard & Koopmans 2020; Koopmans et al. 2013).

The initial study (Koopmans et al. 2013), as well as several follow-up studies (Koopmans et al. 2014a, 2014b), demonstrated the psychometric soundness of the questionnaire in the Netherlands: validity (i.e. face, structural, construct, convergent and discriminant) and reliability. The instrument has since been translated and validated in some countries outside of the Netherlands: Argentina (Gabini & Salessi 2016), Indonesia (Widyastuti & Hidayat 2018), North America (Koopmans et al. 2015), Spain (Ramos-Villagrasa et al. 2019) and Sweden (Dåderman et al. 2020). All these studies are in support of the psychometric soundness of the instrument.

Despite the growth in IWPQ validation studies, three noteworthy limitations still exist: firstly, published studies are scarce and mostly limited to the Netherlands. Secondly, to date, none of them was conducted in South Africa. Results of validation studies cannot be generalised indiscriminately because the importance of the different performance dimensions and the exact indicators associated with each dimension may be context-dependent (Koopmans et al. 2011). Self-report measures also require individuals to reflect on their lived experiences, and these experiences are often influenced by cultural elements (Lenz et al. 2018). Lastly, almost all of the previous studies were done using a confirmatory factor analysis (CFA) framework, despite more recent studies showing that the underlying assumptions of this framework may be too restrictive for the social sciences domain (Morin, Arens & Marsh 2016a; Howard et al. 2018).

The current study aims to address these limitations by investigating the psychometric properties (i.e. validity and reliability) of the instrument in a sample of IT professionals in South Africa. The current study also uses an exploratory structural equation (ESEM) approach, which relaxes some of the assumptions of the CFA framework (Howard et al. 2018). In doing so, the article contributes to the limited literature on the validity of the instrument outside of the Netherlands, using appropriate and sophisticated statistical techniques. The potential benefit from an organisational perspective is to provide organisations with a psychometrically sound performance measuring tool to identify the determinants and outcomes of performance behaviours and to evaluate the effectiveness of performance improvement interventions.

Literature review

Individual work performance questionnaire

A lack of consensus regarding the definition and operationalisation of performance resulted in several measuring instruments being developed to measure employee (or individual) performance. Using different definitions and instruments is problematic as it makes it challenging to identify and reach consensus about the determinants and consequences of individual performance and prevents us from evaluating the effectiveness of interventions (Koopmans et al. 2013, 2014a). A comprehensive definition, that most researchers will agree with, is the first step in facilitating the development of an instrument that is able to measure the construct optimally (Koopmans et al. 2011, 2013).

Guided by Campbell's widely accepted definition of work performance and a thorough systematic review of the occupational health, work and organisational psychology, as well as management and economics literature, Koopmans et al. (2011) performance operates by means of four dimensions: task performance, contextual performance, adaptive performance and counterproductive work behaviour. *Task performance* (TP) is defined as 'proficiency with which central job tasks are performed' (Koopmans et al. 2011:862). These types of behaviour may vary across jobs and are usually prescribed by the job description (Aguinis 2013). Task performance behaviours include, for example, completing job tasks, updating knowledge, and planning and organising (Koopmans et al. 2011). *Contextual performance* (CP) is defined as 'behaviors that support the organizational, social, and psychological environment in which the technical core must function' (Koopmans et al. 2011:862). These behaviours go beyond the job description and refer to, for example, taking initiative and being proactive (Koopmans et al. 2011). *Adaptive performance* (AP) is defined as 'behaviors in reaction to the work environment' (Koopmans et al. 2011:862). These behaviours include, for example, generating new and innovative ideas or being flexible and open-minded with others (Koopmans et al. 2011). *Counterproductive work behaviour* (CWB) is defined as 'behaviors that harm the well-being of the organization' (Koopmans et al. 2011:862) and included, for example, presenteeism, theft or substance abuse (Koopmans et al. 2011).

Following conceptualisation and operationalisation of work performance, Koopmans and colleagues followed an iterative process to develop, improve and validate the different versions of the instrument. Initially, Koopmans et al. (2013) developed a 47-item version of the questionnaire (IWPQ 0.1) based on selected indicators from the literature review. After piloting the instrument to evaluate face validity, clarity and readability, they administered the instrument among a representative sample of Dutch employees. Results of their study indicated, firstly, that three dimensions fit the Rasch model significantly better than four dimensions: AP should merge with CP (Koopmans et al. 2013). The findings, secondly, indicated which generic items fitted the Rasch model for all occupational sectors and resulted in a 14-item version of the scale (IWPQ 0.2) (Koopmans et al. 2013). Lastly, person-item threshold maps indicated that the discriminative ability of the instrument could be improved by including items that measure work performance at the higher (for TP) and lower (for CWB) ends of the scale (Koopmans et al. 2013).

Due to poor targeting of some of the IWPQ 0.2 items, more difficult items were added for TP and CP and easier items for CWB that resulted in a 27-item version (IWPQ 0.3) of the instrument. The items were again tested in a representative sample of Dutch employees, with a series of analyses indicating items that should be deleted in the final version to improve the targeting of the instrument. The final version (IWPQ 1.0) resulted in five items measuring TP, eight items measuring CP and five items measuring CWB, and results indicated a good fit to the Rasch model and acceptable reliability coefficients ranging from 0.74 to 0.85 (Koopmans et al. 2014b). The construct and discriminative validity of the IWPQ 1.0 was consequently supported in the Netherlands (Koopmans et al. 2014a).

Since these initial studies, the self-report scale has been validated in various countries outside of the Netherlands: North America (Koopmans et al. 2015), Indonesia (translated) (Widyastuti & Hidayat 2018), Spain (translated) (Ramos-Villagrasa et al. 2019), Argentina (Gabini & Salessi 2016) and Sweden (translated) (Dåderman et al. 2020). Results from these studies support the face, content, construct, convergent and discriminant validity of the instrument as well as its internal consistency.

Consistent with theory, for the present study, it is hypothesised that:

H1: Individual work performance is a three-dimensional construct.

H2: The IWPQ displays acceptable convergent validity.

H3: The IWPQ displays acceptable internal consistency.

All, but one (see Ramos-Villagrasa et al. 2019) IWPQ validation studies evaluate the construct (i.e. factorial) validity of the instrument using CFA. Confirmatory factor analysis is by far the most widely used method to evaluate construct validity in psychological research, when compared to ESEM. Few instruments meet the cut-off criteria for fit statistics and may then be deemed worthless (Howard et al. 2018). Consequently, researchers question the inherent independent cluster model (ICM) constraints of CFA. These constraints imply that cross-loadings between items and non-target factors should be constrained to zero (Howard et al. 2018). Although one cannot ignore the benefits of these constraints, that is, more parsimonious models and clearly defined constructs, the assumptions underlying these constraints may not be ideal for psychological measures as they often measure closely related constructs. Hence, the conceptual overlap between constructs results in items that are hardly ever uniquely related to a single construct (Howard et al. 2018; Morin et al. 2016a).

An alternative to CFA is ESEM, a novel approach that allows items to cross-load on non-target factors when assessing conceptually related constructs (Morin et al. 2016a, 2016b). Therefore, ESEM enables researchers to circumvent the restrictive assumptions inherent in ICM-CFA that often leads to over-inflated correlation coefficients. Exploratory structural equation modelling integrates exploratory factor analysis, CFA and structural equation modelling into a single model. This integration allows researchers to optimise the benefits of each in one model (for a detailed overview of CFA vs ESEM, refer to Howard et al. 2018). In addition to a CFA model, the current study will also specify an ESEM model as the three performance dimensions are conceptually related, resulting in items that may not be uniquely related to a particular performance dimension. This is particularly so for TP and CP. The lines, between what an employee is supposed to do as part of their job description (i.e. in-role) and what is considered additional tasks (i.e. extra-role), may become potentially more blurred in the current work environment (Koopmans et al. 2011).

Job resources' association with individual work performance

To further test for construct (i.e. discriminant validity and nomological) validity (Hair et al. 2014), the context of the job-demands resources (JD-R) theory was used (Demerouti et al. 2001) in which job resources affects work performance via a motivational process (Bakker & Demerouti 2017). Job resources can be defined as:

> … physical, psychological, social, or organizational aspects of the job that may […] be functional in achieving work goals, reduce job demands and its related costs, or stimulate personal growth and development. (Demerouti et al. 2001:501)

Job resources is frequently cited in work design models when explaining the impact of work on performance (Van Veldhoven et al. 2020) and meta-analytic results support the link between job resources and performance at work (Christian, Garza & Slaughter 2011). Resources, such as autonomy, social support, coaching and opportunities for development, potentially satisfy basic psychological needs and, therefore, exert a motivational effect that enhances performance. The self-determination theory (SDT; Ryan & Deci 2017), and more specifically the basic psychological need mini-theory (BPNT; Ryan & Deci 2017) of the SDT, holds that individuals have three basic psychological needs. *Autonomy* concerns the need to experience volition and willingness and results in one experiencing one's actions, emotions and thoughts as self-endorsed. *Competence* concerns the need to experience a sense of effectiveness while also mastering the environment. *Relatedness* involves the need to develop meaningful and satisfying relationships but also a sense that one is adding value to the lives of others (Vansteenkiste et al. 2020).

Within BPNT, it is argued that need satisfaction can be fostered through need supportive behaviour by key organisational figures (Ryan & Deci 2017). Leaders and co-workers that provide: (1) others with autonomy to satisfy employees' need for autonomy, (2) coaching and opportunities for development that are likely to enhance feelings of competence and relatedness and (3) social support to satisfy others' need for relatedness. Van den

Broeck et al. (2016) demonstrated the beneficial effects of experiencing autonomy, competence and relatedness on good quality motivation and performance. Taken together, the JD-R model and SDT provide evidence that job resources are associated with optimal performance (Van Wingerden et al. 2018).

Consistent with theory, in the present study, it is hypothesised that:

> **H4:** The IWPQ will display acceptable discriminant validity.

> **H5:** Autonomy, social support, coaching and opportunities for development are positively associated with TP (5a) and CP (5b) but negatively associated with CWB (5c).

Research design
Research approach

For the purposes of this study, a quantitative research approach was followed with a cross-sectional survey design. This approach and design enabled the exploration of the factor structure (or the dimensionality) of the instrument and the associations between the performance constructs and job resources at a specific moment in time.

Research method
Participants

A convenience sample of 296 IT professionals across various organisations in South Africa was included. Information technology professionals were defined as those employees who test, build, install, repair or maintain computer software systems. The population is deemed appropriate for validating the questionnaire for two reasons: IT professionals play an increasingly important role in the fourth industrial revolution (Van Zyl et al. 2019) and their work is considered highly complex knowledge work. Consequently, they must perform their core tasks (TP) well. At the same time, there are both the expectation and opportunity to generate innovative solutions (CP) as well as additional pressure that may make them more susceptible to negativity (CWB). The final sample comprised mainly men (74.3%; *n* = 220). The average age of the respondents was 37 years (SD = 9.81), and the average tenure in their current position was 6 years (SD = 5.85).

Measuring instruments

A *biographical questionnaire* was used to obtain information regarding the participants' age, gender and years of experience within their position to describe the sample.

Individual work performance was measured using the IWPQ 1.0 developed by (Koopmans et al. 2014b). Task performance consisted of five items (e.g. 'I kept in mind the results that I had to achieve in my work'), CP of eight items (e.g. 'I continually sought new challenges in my work') and CWB of five items (e.g. 'I talked to colleagues about the negative aspects of my work'). Participants were asked to reflect on how frequently they displayed a particular behaviour during the past three months ranging from 1 (Seldom) to 5 (Always) and 1 (Never) to 5 (Often) in the case of CWB (Koopmans et al. 2014b).

Autonomy, social support, coaching and opportunity for development was measured using the *Job Demands-Resource Questionnaire* (Bakker 2014). Participants were expected to rate their work situation on a five-point frequency scale ranging from 1 (Never) to 5 (Very often). Autonomy consisted of three items (e.g. 'Do you have control over how your work is carried out?'), social support consisted of three items (e.g. 'If necessary, can you ask your colleagues for help?'), coaching consisted of five items (e.g. 'I feel valued by my supervisor' and opportunities for development consisted of three items ('In my work, I can develop myself sufficiently') (Bakker 2014).

Research and ethics procedure

Ethics approval was granted by the research ethics committee of the university (NWU-HS-2017-0046). The following ethical considerations guided data collection: (1) participants were not put at risk unnecessarily; they were also respected at all times, (2) participants were provided with an information letter covering aspects such as inclusion criteria, purpose of the research and the possible publication of anonymous results, benefits for the participants, the expectations or requirements from participants, possible risks and mitigation thereof where possible, guarantee of anonymity and confidentiality, the right to withdraw from the study without foreseen negative consequences, and contact details of individuals in the event that the participant needs more details regarding the research, and (3) participants were asked to consent to participation before continuing with the questionnaire. Convenience sampling, in which potential participants are sampled because they are more available when a probability sample is not possible for sources of data (Creswell & Creswell 2018), was utilised. Data was collected via three avenues to ensure a more heterogeneous sample: (1) approaching IT-related organisations in the Johannesburg area, (2) a professional social media platform (i.e. LinkedIn) and (3) a data collection and research solutions company. A short introductory summary of what the study entailed, a Google Forms link to the consent form, and the questionnaire were sent. Regarding the data collection company, inclusion criteria for the sample were provided to source participants after which the same mail was sent to them.

Statistical analysis

For the descriptive statistics, RStudio version 1.2.5033 (RStudioTeam 2019) was used with R base-version 3.6.2 (RCoreTeam 2019). In RStudio, the 'psych' package's describe function was used to calculate the means and standard deviations for each of the factors in the model (Revelle 2018). Mplus 8.4 (Muthén & Muthén 1998–2019) was used to

perform the CFA and ESEM analyses and to calculate the reliability coefficients.

To evaluate construct validity, Hair et al. (2014) recommend one investigates convergent, discriminant and nomological validity. The first step was to evaluate the factor structure (or the dimensionality) of the IWPQ that will be used in subsequent analyses. This step entailed comparing different measurement models for IWPQ, guided by the literature and using both CFA and ESEM frameworks. The mean- and variance-adjusted weighted least squares (WLSMV) estimator was used due to the categorical nature of the data (Kline 2016). Acceptable model fit statistics, indicated by several goodness-of-fit indices, also indicate acceptable construct validity (Hair et al. 2014). The following cut-off criteria were used to evaluate the model fit (Kline 2016): root mean square error of approximation (RMSEA) $\leq$ 0.08, standardised root means square residual (SRMR) $\leq$ 0.10, Tucker-Lewis index (TLI) and the comparative fit index (CFI) $\geq$ 0.95.

Next, tests for convergent and discriminant validity were implemented. Convergent validity was evaluated based on (1) the standardised factor loadings of items $\geq$ 0.50 or ideally $\geq$ 0.70, (2) an average variance extracted (AVE) of $\geq$ 0.50 and (3) construct reliability (CR) of $\geq$ 0.70 (Hair et al. 2014). In order to evaluate the reliability of the measuring instrument, composite reliability coefficients (ρ) were calculated as they are deemed more appropriate for latent variables (Raykov 2009). Cronbach's coefficient alpha is also reported for potential future comparison by other researchers. For Cronbach's coefficients, the 'scaleReliability' function was used from the 'userfriendlyscience' R package (Peters 2018).

Discriminant validity was assessed by the: (1) AVE of each construct being greater than their shared variance with other variables (Hair et al. 2014) and (2) heterotrait-monotrait (HTMT) ratio of correlation with values close to 1.00 indicating a lack of discriminant validity. Using the HTMT as a criterion involves comparing it to a predefined threshold. If the value of the HTMT is higher than this threshold, one can conclude that there is a lack of discriminant validity (Henseler, Ringle & Sarstedt 2015). Some authors suggest a threshold of 0.85 (Kline 2016) whereas others suggest a threshold of 0.90 (Gold, Malhotra & Segars 2001; Teo et al. 2008). The 'semTools' package's 'htmt' function was used to calculate the HTMT values in this study (Jorgensen et al. 2019).

Nomological validity was evaluated by estimating the correlations between the performance constructs and autonomy, social support, coaching and opportunities for development. Previous studies suggest that more favourable evaluations of the latter should relate to more favourable evaluations of TP and CP, whereas the opposite would be expected for CWB. The standard cut-off criteria were used for effect sizes: r = 0.10–0.29 (small effect), r = 0.30–0.49 (medium effect) and $r \geq$ 0.50 (large effect) (Cohen 1992).

Results
Factorial structure of the Individual Work Performance Questionnaire

Six competing measurement models were estimated. In the first model, the items were allowed to load onto their respective a priori factors – TP (five items), CP (eight items) and CWB (five items) – in line with previous research (Koopmans et al. 2014a; Ramos-Villagrasa et al. 2019). The three factors were allowed to correlate. In the second model, the items loaded onto their respective a priori factors but the three performance factors were then allowed to load onto a second-order (performance) factor as recommended by (Koopmans et al. 2011). The third model was similar to Model 1, except the items were also allowed to load onto a general (performance) factor. All these models correspond to the ICM-CFA framework: cross-loadings were constrained to zero. Models 4 to 6 were the ESEM versions of the three ICM-CFA models. Exploratory structural equation modelling was specified using target rotation: all items were freely estimated on their a priori factors, but they were also allowed to cross-load. The cross-loadings were targeted to be close to zero, as is usual in ESEM modelling (Gomes, Almeida & Núñez 2017). Table 1 presents the goodness-of-fit indices for each of the estimated models.

All models demonstrated acceptable fit when compared to most of the goodness-of-fit cut-off criteria. When comparing the different ICM-CFA models with each other, the 90% confidence intervals for the RMSEA showed several overlaps between the solutions, indicating a low degree of differentiation between competing models. In each case, the CFI and TLI were above and the SRMR value below the required cut-off values. Consequently, none of the models seems to be superior compared to the others. When comparing the ICM-CFA models with their ESEM counterparts, the CFI and TLI values of Models 4 and 5 are

TABLE 1: Fit statistics for the measurement models.

Models	df	χ^2	Comparative fit index	Tucker-Lewis index	Root mean square error of approximation	90% confidence interval	Standardised root mean square residual
Model 1: First-order ICM-CFA	132	563.61*	0.95	0.95	0.11*	[0.09, 0.11]	0.06
Model 2: Second-order ICM-CFA	132	563.61*	0.95	0.95	0.11*	[0.09, 0.11]	0.06
Model 3: Bi-factor ICM-CFA	114	377.70*	0.97	0.96	0.09*	[0.08, 0.10]	0.04
Model 4: First-order ESEM	102	671.54*	0.94	0.91	0.14*	[0.13, 0.15]	0.04
Model 5: Second-order ESEM	102	562.59*	0.95	0.93	0.12*	[0.11, 0.13]	0.04
Model 6: Bi-factor ESEM	138	296.06*	0.98	0.96	0.09*	[0.08, 0.13]	0.03

ICM-CFA, independent cluster model confirmatory factor analysis; ESEM, exploratory structural equation modelling.
*, p < 0.001.

lower than Models 1 and 2, and their RMSEA values are higher. The ESEM counterparts have a poorer fit to the data than their ICM-CFA counterparts. Additionally, few of the cross-loadings in the ESEM model were significant (see Table 2) and the correlations between the performance variables were within acceptable bounds (see Table 3). Model 6 has the best CFI, TLI and SRMR values; however, its 90% confidence intervals for the RMSEA again overlaps with that of the ICM-CFA models indicating little differentiation. In this case, the most parsimonious model (with fewer parameters estimated) is preferred (Howard et al. 2018), that is, Model 1. This lends support for Hypothesis 1.

None of the models' root mean square error of approximation (RMSEA) values met the minimum criteria of ≤ 0.08 and, consequently, model improvement was performed guided by high modification indices (MI) values. The residual item-variances of CP3 ('I worked on keeping my job-related knowledge up-to-date') and CP4 ('I worked on keeping my work skills up-to-date') and CWB4 ('I talked to colleagues about the negative aspects of my work') and CWB5 ('I talked to people outside the organisation about the negative aspects of my work') were allowed to correlate. Correlated residuals are permitted when the items measure the same construct (Hair et al. 2014; Kline 2016) and can be a consequence of

similar prompts (i.e. updating of…) or overlapping item content (i.e. negativity). These modifications resulted in a model with a very good fit to the data: $\chi^2 = 393.25$ ($p < 0.001$), $df = 130$, CFI = 0.97, TLI = 0.97, RMSEA = 0.08 [0.08, 0.09], and SRMR = 0.05.

Convergent and discriminant validity

Based on the most parsimonious Individual Work Performance Questionnaire (IPWQ) measurement model, a full measurement model was constructed in which autonomy, social support, coaching and opportunities for development were included. The initial model presented with an ultra Heywood case – CWB3 had a standardised factor loading exceeding 1.00 and CWB5 had a standardised factor loading of less than 0.50. These two items were consequently removed from the model and the final model had a very good fit to the data: $\chi^2 = 1047.77$ ($p < 0.001$), $df = 383$, CFI = 0.95, TLI = 0.95, RMSEA = 0.07 [0.07, 0.08] and SRMR = 0.06. The model is depicted in Figure 1. The criteria for convergent validity was met. All factor loadings were significant and above 0.50 (most are above 0.70) (see Table 2). The AVE of each construct was above 0.50: for TP AVE was 0.70, for CP 0.63, for CWB 0.52 and all CR values exceeded 0.70 (see Table 3). This provides support for Hypotheses 2 and 3.

For discriminant validity, the AVE of each construct should be greater than their shared variance (Hair et al. 2014). The AVE for all the variables ranged from 0.52 to 0.82. The shared variance between the performance dimensions ranged from 0.04 to 0.47 and between the performance dimensions and autonomy, social support, coaching and opportunities for development, the shared variance ranged between 0.00 and 0.34. Furthermore, results of the HTMT analyses showed values ranging from 0.16 to 0.59 which is below the lowest recommended thresholds of 0.90 and 0.85. Therefore, discriminant validity was supported between all constructs suggesting that this performance measure is independent of other related constructs and that the performance dimensions are also independent of each other. This result supported Hypothesis 4.

Nomological validity

Table 3 contains the correlation coefficients between the performance constructs and autonomy, social support, coaching and opportunities for development. Most of the
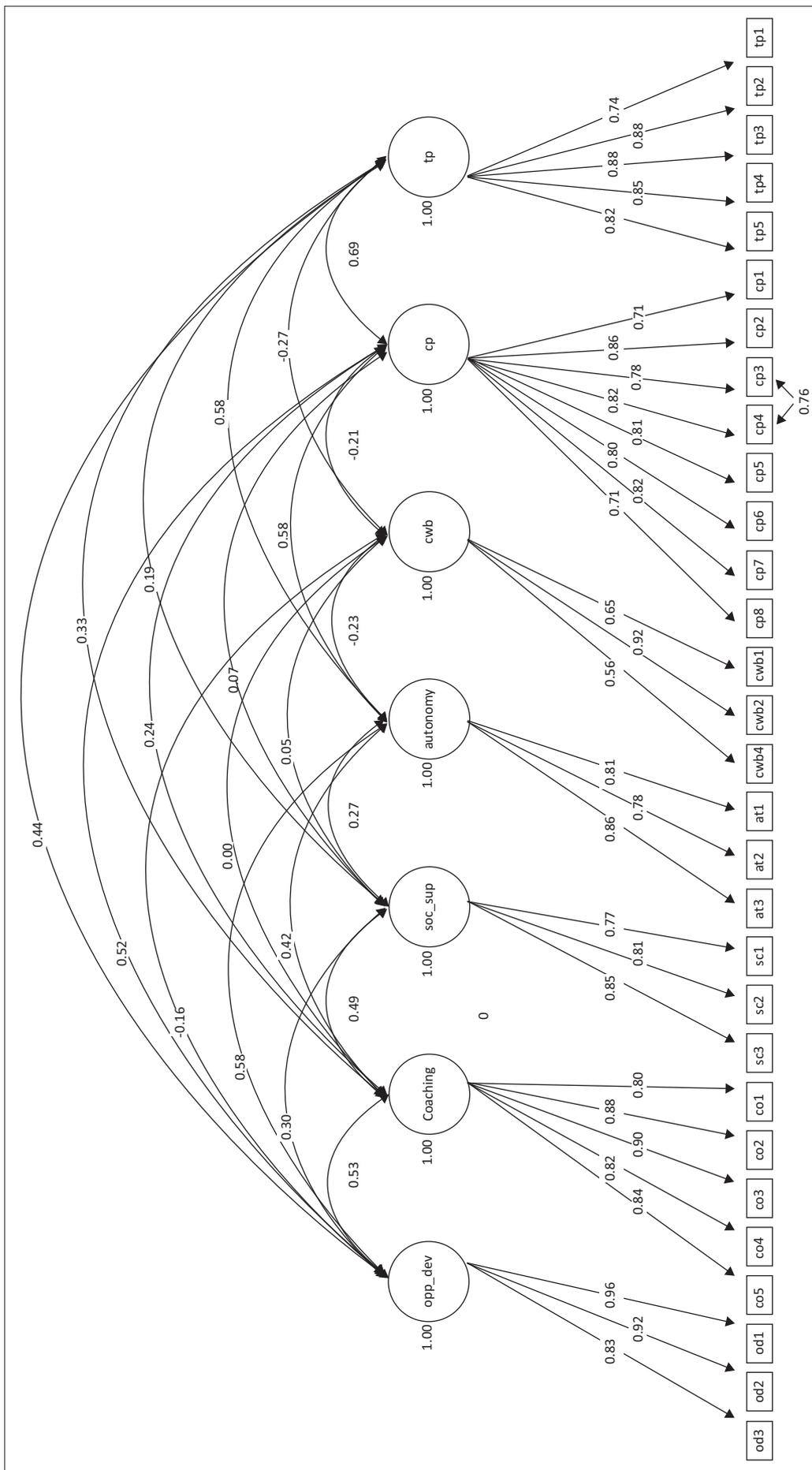
TABLE 2: Standardised factor loadings for independent and exploratory structural equation modelling.

Item	Loadings			
	Confirmatory factor analysis	Factor 1	Factor 2	Factor 3
Task performance				
TP1	0.74***	0.88***	-0.15***	0.03
TP2	0.88***	0.70***	0.20***	-0.02
TP3	0.88***	0.90***	-0.02	-0.06*
TP4	0.85***	0.83***	0.02	-0.01
TP5	0.82***	0.78***	0.07	0.00
Contextual performance				
CP1	0.71***	0.11	0.63***	0.03
CP2	0.86***	0.05	0.77***	-0.08*
CP3	0.78***	0.07	0.88***	0.06
CP4	0.83***	0.07	0.90***	0.05
CP5	0.81***	-0.12**	0.85***	-0.07*
CP6	0.80***	-0.12**	0.92***	0.04
CP7	0.82***	0.03	0.81***	-0.05
CP8	0.71***	-0.01	0.62***	-0.13*
Counter productive work behaviour				
CWB1	0.65***	0.05	0.06	0.69***
CWB2	0.92***	-0.14**	0.02	0.70***
CWB4	0.56***	0.13**	-0.01	0.83***

*, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$.

TABLE 3: Means, standard deviations, reliability and correlations.

Variable	Mean	SD	α	ρ	1	2	3	4	5	6
Task performance	3.64	0.85	0.89	0.92	-	-	-	-	-	-
Contextual performance	3.72	0.85	0.91	0.93	0.69**	-	-	-	-	-
Counterproductive work behaviour	2.02	0.75	0.68	0.76	-0.27**	-0.21**	-	-	-	-
Autonomy	3.68	1.07	0.81	0.86	0.58**	0.58**	-0.24**	-	-	-
Social support	3.40	0.98	0.74	0.85	0.19**	0.08	0.05	0.27**	-	-
Coaching	3.27	1.05	0.90	0.93	0.33**	0.24**	-0.00	0.42**	0.49**	-
Opportunities for development	3.66	1.02	0.89	0.93	0.44**	0.52**	-0.16*	0.59**	0.30**	0.53**

*, $p \leq 0.01$; **, $p \leq 0.001$.

FIGURE 1: Final confirmatory factor analysis model.

correlations are significant, and they are all in the expected direction providing support for the nomological validity of the instrument.

In terms of the correlations, all three performance factors correlated with each other: TP had a positive relationship with CP ($r = 0.69$; large effect) and both TP and CP had negative relationships with CWB ($r = -0.27$ and $r = -0.21$; small effect). Task performance had significant positive relationships with autonomy ($r = 0.58$; large effect), social support ($r = 0.19$; small effect), coaching ($r = 0.33$; medium effect) and opportunities for development ($r = 0.44$; medium effect). This supports Hypothesis 5a. Contextual performance had significant positive relationships with autonomy ($r = 0.58$; large effect), coaching ($r = 0.24$; small effect) and opportunities for development ($r = 0.52$; large effect) but was unrelated to social support ($r = 0.08$, *ns*). Counterproductive work behaviour was only significantly negatively related to autonomy ($r = -0.24$; small effect) and opportunities for development ($r = -0.16$; small effect). These results provide partial support for Hypotheses 5b and 5c.

Additionally, using the decision-making tree provided by Podsakoff et al. (2003), a latent common methods variance factor was added to the measurement model to test for common method bias (CMB). In this model, items load onto their a priori theoretical construct as well as onto the latent common methods variance factor. The factor loadings of the two models (i.e. with and without the latent common methods variance factor) are compared and a change in the factor loading values of more than 0.20 is deemed problematic: such items may be affected by CMB. Results indicated that one item of each of the autonomy, social support and opportunities for development scales and three items of the coaching scale may be affected by CMB. These items were removed from the measurement model (without the latent common methods variance factor) and the correlation coefficients of this model were compared to the model with the latent common methods variance factor. Results indicated that the relationships were almost identical, with no substantial deviations. Therefore, CMB was not considered a concern in this study.

Discussion

Given the academic and organisational importance of performance, scientifically rigorous measurement is imperative. The current study aimed to contribute to the limited body of knowledge on the psychometric properties of the IWPQ by validating (i.e. convergent, discriminant and nomological) the instrument in a South African context.

Results demonstrated that the instrument is valid for a sample of IT professionals in South Africa. This means that the constructs are accurately operationalised in this specific sample – the indicators measure the latent constructs (i.e. TP, CP, and CWB) as they are intended to. More specifically, results indicate that individual work performance comprises three separate but related constructs. This is in line with studies from other countries (Koopmans et al. 2014a; Ramos-Villagrasa et al. 2019). The findings indicate that researchers should refrain from calculating one performance score as important information relating to each performance dimension will be lost if we were to assume that performance can be summarised with a single score.

Even though the operationalisation of TP and CP, and for the biggest part of CWB, accurately reflects their respective constructs, two problematic items were highlighted in the CWB sub-scale: 'I focused on the negative aspects of situation at work instead of the positive aspects' and 'I talked to people outside the organization about the negative aspects of my work'. Pending replication of the current results, the finding suggests that the psychometric properties of these two items should be carefully monitored in future studies. From a methodological point of view, the results of the current study question the usefulness of more sophisticated analytical (i.e. ESEM) frameworks when modelling performance. Although some support (see Ramos-Villagrasa et al. 2019) is starting to emerge for ESEM models, more research is needed to provide conclusive evidence regarding the factor structure of the IWPQ.

The results of the current study not only show that the indicators are good reflections of their respective latent constructs, but they also converge or share a high proportion of variance (i.e. convergent validity). This is in line with previous research (Koopmans et al. 2014a; Ramos-Villagrasa et al. 2019) and indicates that the items measure the same underlying latent construct. At the same time, the results also indicate that the performance constructs are genuinely distinct from each other as well as from other constructs (i.e. discriminant validity), in line with previous research (Koopmans et al. 2014a). It is not surprising that the two positive forms of performance are more strongly correlated with one another than with the negative form. The lines between TP and CP are often blurred in modern organisations and meta-analyses have found modest correlations between the CWB and the other two performance dimensions (Koopmans et al. 2011). The instrument also measured performance consistently in the current study, which is in line with previous research (Koopmans et al. 2015; Ramos-Villagrasa et al. 2019) and lends further support for its convergent validity.

The results regarding the association between the IWPQ dimensions and job resources are mostly as expected and provide evidence for the instrument's nomological validity. The IWPQ TP positively associated with all of the job resources in line with JD-R theory (Bakker & Demerouti 2017) and the BPNT (Ryan & Deci 2017). Thus, having flexibility and control, receiving support from colleagues and supervisor, and having the opportunity to learn and develop oneself enable employees to perform their core job tasks well. The IWPQ CP was related to all, but one, job resource – social support. Although the other resources enable employees to

create a conducive environment in which they can perform their core functions, this is not dependent on support and validation from colleagues. The IWPQ CWB was only associated with autonomy and opportunities for development, but as expected in a negative direction. The findings suggest that support from colleagues and one's supervisor does not prevent behaviour that harms the well-being of the organisation – as these associations were non-significant. Although all the non-significant results were unexpected from a theoretical point of view, and although it could be the result of methodological artefacts, recent empirical studies challenge the universality of job resources (see Van Veldhoven et al. 2020 for an overview). In a series of papers, Van Veldhoven et al. (2020) argue that researchers should seek a more nuanced understanding of why, when and for whom job resources are beneficial.

Practical implications

The current study presents IT organisations with a psychometrically sound performance measuring tool that they can use to identify the determinants and outcomes of performance behaviours and to evaluate the effectiveness of performance improvement interventions. A scientifically rigorous instrument that goes beyond measuring prescribed job tasks is essential in modern-day organisations where employees are often expected to go beyond the scope of the tasks allocated to them (Carpini, Parker & Griffin 2017; Griffin, Neal & Parker 2007).

In improving TP and CP performance and lowering CWB, organisations can consider providing employees with more flexibility and control, encouraging team cohesion through team building activities, developing empowering leaders (supervisors) and providing employees with enough opportunities to learn and develop new skills.

Limitations and recommendations

The study is not without limitations. Most noteworthy is the use of a cross-sectional survey design. Although cross-sectional designs are still useful for exploratory studies where limited information is available (Spector 2019), they pose a limitation for the evaluation of predictive validity. The study also made use of self-report surveys. Although self-report surveys have several advantages in performance research (i.e. easily obtainable, employees have more opportunities to observe own behaviour, the halo effect is avoided and it ensures confidentiality and fewer missing values) (Koopmans et al. 2014a; Widyastuti & Hidayat 2018), one cannot ignore the fact that common method variance is a likely outcome of self-report surveys. Common method variance means bias is introduced in the ways in which constructs are measured (Podsakoff, Mackenzie & Podsakoff 2012). Due to this, relationships between constructs can be over-inflated (Spector et al. 2017). To counter the shortcomings (i.e. causal inferences and CMB) of cross-sectional research, future studies should employ longitudinal research designs. Researchers should, however, not employ

a longitudinal design blindly. Along this line, (Spector 2019) suggests that researchers should carefully consider which variables precede the others and how long one variable should be allowed to 'develop' or 'change' before measuring its outcomes and establishing causal inferences (Spector 2019). Furthermore, the current study employed statistical methods to test for CMB but several design strategies are also recommended: (1) using different sources (i.e. employees and leaders) to obtain predictor and outcome responses, (2) temporally separating the predictor and outcome variables (i.e. measure at different time points), (3) using different response scales for the different variables, (4) clearly formulating items and avoiding wording that enhances the likelihood of socially desirable answers, and (5) striking a balance between positive and negative items (Podsakoff et al. 2012).

Although we proactively implemented measures to reach a broader audience of IT professionals and the sampling strategy respects the voluntary participation of participants, we cannot ignore the possible biases introduced by self-selection. It is recommended that researchers use random sampling strategies in future studies in an attempt to obtain a representative sample to replicate the findings of the current study.

Apart from the methodological limitations and recommendations, an important limitation of the current study is the omission of dependent variables to evaluate the concurrent (and predictive) validity of the IWPQ. Consequently, future research is encouraged to include 'outcome' variables (e.g. objective organisational performance metrics) to evaluate the concurrent and predictive validity of the IWPQ.

Conclusion

The results of the current study provide sufficient evidence for the construct validity of the IWPQ in a South African context. Individual work performance consists of three related constructs: task performance, contextual performance, and counterproductive work behaviour, and the different types of performance sufficiently explain variance in their respective indicators (except for two of the CWB items). The different types of performance are also sufficiently different from each other as well as from other related constructs (i.e. job resources). In future, researchers (also in South Africa) can use the IWPQ questionnaire to develop a more mature and unified knowledge based on individual work performance.

Acknowledgements

Competing interests

The author has declared that no competing interest exists.

Author's contributions

The author conceptualised and wrote the manuscript and conducted the data analysis.

Ethical consideration

Ethics approval was granted by the research ethics committee of North-West University (NWU-HS-2017-0046).

Funding information

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability statement

Data sharing will be done in line with ethics requirements. The data and syntaxes are the intellectual property of North-West University and will be provided upon reasonable request. Contact author for more information.

Disclaimer

The views and opinions expressed in this article are those of the author and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References

Aguinis, H., 2013, *Performance management*, Pearson Prentice Hall, Upple Saddle River, NJ.

Bakker, A.B., 2014, *The Job Demands–Resources Questionnaire*, Rotterdam University, Rotterdam.

Bakker, A.B. & Demerouti, E., 2017, 'Job demands-resources theory: Taking stock and looking forward', *Journal of Occupational Health Psychology* 22(3), 273–285. https://doi.org/10.1037/ocp0000056

Campbell, J.P., 1990, 'Modeling the performance prediction problem in industrial and organizational psychology', in M.D. Dunnette & L.M. Hough (eds.), *Handbook of industrial and organizational psychology* (pp. 687–732), Consulting Psychologists Press, Palo Alto, CA.

Carpini, J.A., Parker, S.K. & Griffin, M.A., 2017, 'A look back and a leap forward: A review and synthesis of the individual work performance literature', *Academy of Management Annals* 11(2), 825–885. https://doi.org/10.5465/annals.2015.0151

Christian, M.S., Garza, A.S. & Slaughter, J.E., 2011, 'Work engagement: A quantitative review and test of its relations with task and contextual performance', *Personnel Psychology* 64(1), 89–136. https://doi.org/10.1111/j.1744-6570.2010.01203.x

Cohen, J., 1992, 'Statistical power analysis', *Current Directions in Psychological Science* 1(3), 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Creswell, J.W. & Creswell, J.D., 2018, *Research design: Qualitative, quantitative, and mixed methods approaches*, Sage, Thousand Oaks, CA.

Dåderman, A.M., Ingelgard, A. & Koopmans, L., 2020, 'Cross-cultural adaptation, from Dutch to Swedish language, of the individual work performance questionnaire', *Work* 65, 97–109. https://doi.org/10.3233/WOR-193062

Demerouti, E., Bakker, A.B., Nachreiner, F. & Schaufeli, W.B., 2001, 'The job demands-resources model of burnout', *The Journal of Applied Psychology* 86(3), 499–512. https://doi.org/10.1037/0021-9010.86.3.499

Gabini, S. & Salessi, S., 2016, 'Validación de la Escala de Rendimiento Laboral Individual en Tra- bajadores Argentinos [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

Gold, A.H., Malhotra, A. & Segars, A.H., 2001, 'Knowledge management: An organizational capabilities perspective', *Journal of Management Information Systems* 18(1), 185–214. https://doi.org/10.1080/07421222.2001.11045669

Gomes, C.M.A., Almeida, L.S. & Núñez, J.C., 2017, 'Rationale and applicability of exploratory structural equation modeling (ESEM) in psychoeducational contexts', *Psicothema* 29(3), 396–401.

Griffin, M.A., Neal, A. & Parker, S.K., 2007, 'A new model of work role performance: Positive behavior in uncertain and interdependent contexts', *The Academy of Management Journal* 50(2), 327–347. https://doi.org/10.5465/amj.2007.24634438

Hair, J.F., Black, W.C., Babin, B.J. & Andersen, R.E., 2014, *Multivariate data analysis*, Pearson, Essex.

Henseler, J., Ringle, C.M. & Sarstedt, M., 2015, 'A new criterion for assessing discriminant validity in variance-based structural equation modeling', *Journal of the Academy of Marketing Science* 43, 115–135. https://doi.org/10.1007/s11747-014-0403-8

Howard, J.L., Gagne, M., Morin, A.J.S. & Forest, J., 2018, 'Using bifactor exploratory structural equation modeling to test for a continuum structure of motivation', *Journal of Management* 44(7), 2638–2664. https://doi.org/10.1177/0149206316645653

Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A.M. & Rosseel, Y., 2019, *semTools: Useful tools for structural equation modeling*, R package version 0.5-2. Available from https://cran.r-project.org/web/packages/semTools/index.html

Kline, R.B., 2016, *Principles and practice of structural equation modeling*, The Guilford Press, New York, NY.

Koopmans, L., Bernaards, C.M., Hildebrandt, V.H., De Vet, H.C.W. & Van der Beek, A.J., 2014a, 'Construct validity of the individual work performance questionnaire', *Journal of Occupational and Environmental Medicine* 56(3), 331–337. https://doi.org/10.1097/JOM.0000000000000113

Koopmans, L., Bernaards, C.M., Hildebrandt, V.H., Lerner, D., De Vet, H.C.W. & Van der Beek, A.J., 2015, 'Cross-cultural adaptation of the individual work performance questionnaire', *Work* 53(3), 609–619. https://doi.org/10.3233/WOR-152237

Koopmans, L., Bernaards, C.M., Hildebrandt, V.H., Schaufeli, W.B., De Vet, H.C.W. & Van der Beek, A.J., 2011, 'Conceptual frameworks of individual work performance: A systematic review', *Journal of Occupational and Environmental Medicine* 53(8), 856–866. https://doi.org/10.1097/JOM.0b013e318226a763

Koopmans, L., Bernaards, C.M., Hildebrandt, V.H., Van Buuren, S., Van der Beek, A.J. & De Vet, H.C.W., 2013, 'Development of an individual work performance questionnaire', *International Journal of Productivity and Performance Management* 62(1), 6–28. https://doi.org/10.1108/17410401311285273

Koopmans, L., Bernaards, C.M., Hildebrandt, V.H., Van Buuren, S., Van der Beek, A.J. & De Vet, H.C.W., 2014b, 'Improving the individual work performance questionnaire using Rasch analysis', *Journal of Applied Measurement* 15(2), 160–175.

Lenz, A.S., Watson, J.C., Luo, Y., Norris, C. & Nkyi, A., 2018, 'Cross-cultural validation of four positive psychology assessments for use with a Ghanaian population', *International Journal for the Advancement of Counselling* 40, 148–161.

Morin, A.J.S., Arens, A.K. & Marsh, H.W., 2016a, 'A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality', *Structural Equation Modeling: A Multidisciplinary Journal* 23(1), 116–139. https://doi.org/10.1080/10705511.2014.961800

Morin, A.J.S., Arens, A.K., Tran, A. & Caci, H., 2016b, 'Exploring sources of construct-relevant multidimensionality in psychiatric measurement: A tutorial and illustration using the Composite Scale of Morningness', *International Journal of Methods in Psychiatric Research* 25(4), 277–288. https://doi.org/10.1002/mpr.1485

Muthén, L.K. & Muthén, B.O., 1998–2019, *Mplus user's guide*, Muthén & Muthén, Los Angeles, CA.

Peters, G., 2018, *Userfriendlyscience: Quantitative analysis made accessible*. Available from https://cran.r-project.org/web/packages/userfriendlyscience/index.html

Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. & Podsakoff, N.P., 2003, 'Common method biases in behavioral research: A critical review of the literature and recommended remedies', *Journal of Applied Psychology* 88(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Podsakoff, P.M., Mackenzie, S.B. & Podsakoff, N.P., 2012, 'Sources of method bias in social science research and recommendations on how to control it', *Annual Review of Psychology* 63, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

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', *Revista de Psicología del Trabajo y de las Organizaciones* 35(3), 195–205. https://doi.org/10.5093/jwop2019a21

Raykov, T., 2009, 'Evaluation of scale reliability for unidimensional measures using latent variable modeling', *Measurement and Evaluation in Counseling and Development* 42(3), 223–232. https://doi.org/10.1177/0748175609344096

RCoreTeam, 2019, *R: A language and environment for statistical computing [Software]*. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org/

Revelle, W., 2018, *Psych: Procedures for personality and psychological research (Version 1.8.12)*. Available from https://cran.r-project.org/web/packages/psych/index.html

RStudioTeam, 2019, *RStudio: Integrated development for R*. RStudio, PBC, Boston, MA. Available from http://www.rstudio.com/

Ryan, R.M. & Deci, E.L., 2017, *Self-determination theory: Basic psychological needs in motivation, development, and wellness*, Guilford Press, New York, NY.

Spector, P.E., 2019, 'Do not cross me: Optimizing the use of cross-sectional designs', *Journal of Business and Psychology* 34, 125–137. https://doi.org/10.1007/s10869-018-09613-8

Spector, P.E., Rosen, C.C., Richardson, H.A., Williams, L.J. & Johnson, R.E., 2017, 'A new perspective on method variance: A measure-centric approach', *Journal of Management* 45(3), 855–880. https://doi.org/10.1177/0149206316687295

Teo, T.S.H., Srivastava, S.C. & Jiang, L., 2008, 'Trust and electronic government success: An empirical study', *Journal of Information Management Systems* 25(3), 99–132. https://doi.org/10.2753/MIS0742-1222250303

Van Den Broeck, A., Ferris, D.L., Chang, C.-H. & Rosen, C.C., 2016, 'A review of self-determination theory's basic psychological needs at work', *Journal of Management* 42(5), 1195–1229. https://doi.org/10.1177/0149206316632058

Van Veldhoven, M., Van Den Broeck, A., Daniels, K., Bakker, A.B., Tavares, S.M. & Ogbonnaya, C., 2020, Challenging the universality of job resources: Why, when, and for whom are they beneficial?', *Applied Psychology* 69(1), 5–29. https://doi.org/10.1111/apps.12211

Van Wingerden, J., Derks, D. & Bakker, A.B., 2018, 'Facilitating interns' performance', *Career Development International* 23(4), 382–396. https://doi.org/10.1108/CDI-12-2017-0237

Van Zyl, L.E., Van Oort, A., Rispens, S. & Olckers, C., 2019, 'Work engagement and task performance within a global Dutch ICT-consulting firm: The mediating role of innovative work behaviors', *Current Psychology*, in press. https://doi.org/10.1007/s12144-019-00339-1

Vansteenkiste, M., Ryan, R.M. & Soenens, B., 2020, 'Basic psychological need theory: Advancements, critical themes, and future directions', *Motivation and Emotion* 44, 1–31. https://doi.org/10.1007/s11031-019-09818-1

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

Yeshitila, B. & Beyene, A., 2019, 'The mediating effect of employee engagement between transformational leadership and job performance in the commercial bank of Ethiopia', *European Journal of Business and Management* 11, 33–42.