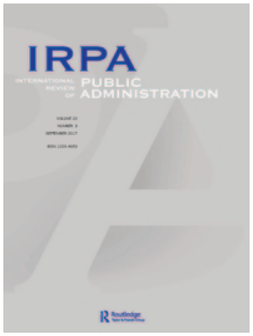

International Review of Public Administration

ISSN: 1229-4659 (Print) 2331-7795 (Online) Journal homepage: https://www.tandfonline.com/loi/rrpa20

Developing an item pool and testing measurement invariance for measuring public service motivation in Korea

Sangmook Kim

To cite this article: Sangmook Kim (2017) Developing an item pool and testing measurement invariance for measuring public service motivation in Korea, International Review of Public Administration, 22:3, 231-244, DOI: 10.1080/12294659.2017.1327113

To link to this article: https://doi.org/10.1080/12294659.2017.1327113

INTERNATIONAL REVIEW OF PUBLIC ADMINISTRATION, 2017
VOL. 22, NO. 3, 231–244
https://doi.org/10.1080/12294659.2017.1327113

Routledge
Taylor & Francis Group

Developing an item pool and testing measurement invariance for measuring public service motivation in Korea

Sangmook Kim

Department of Public Administration, Seoul National University of Science & Technology, Seoul, Korea

ABSTRACT

In a recent study, an international measure of public service motivation (PSM) failed to achieve measurement invariance across cultures and languages. However, since that research was able to confirm the four-dimensional structure of PSM, it can provide a starting point for studying PSM in a single country. This study aimed to develop an item pool for measuring PSM in Korea. Online survey data ($n = 1800$), collected from both the public and private sectors, were used to test measurement invariance to validate the use of the measure across genders and sectors. The results provide support for both the initial four-dimensional 29-item PSM model and the more concise 16-item PSM model, confirming that the dimensions have the same meaning and scaling across genders and sectors in Korea. This study may be the first to test the measurement equivalence of a PSM measure across different groups in a single country.

ARTICLE HISTORY
Received 1 November 2016
Accepted 25 April 2017

KEYWORDS
Public service motivation; measurement invariance; confirmatory factor analysis

Introduction

Public service motivation (PSM) represents a desire to serve the public. Perry and Wise (1990) proposed that PSM is characterized by rational, norm-based, and affective motives. Perry (1996) developed a 24-item measure to identify four dimensions in the PSM construct. Since then, many researchers around the world have used Perry's dimensions and items to measure PSM (Coursey & Pandey, 2007; Wright, 2008). In some cases, however, certain dimensions were excluded because they did not fit well (Lee & Kim, 2014), while in other cases dimensions were added to capture the full meaning of PSM (Kim, 2011; Kim & Vandenabeele, 2010). Such diversity limits the ability to replicate and build on previous findings (Wright, 2008).

In their effort to build international PSM research, Kim and Vandenabeele (2010) suggested that PSM is fundamentally rooted in self-sacrifice, which underpins three distinct motive categories: instrumental, value-based, and identification motives. Their proposal for developing an international PSM scale resulted in a new 16-item measure (Kim et al., 2013). However, the international measure has failed to achieve measurement invariance and therefore does not have the same psychometric properties across different cultures.

CONTACT Sangmook Kim smook@seoultech.ac.kr

Nevertheless, the international PSM measure can provide a starting point for single-country PSM research, as the four-dimensional structure of PSM was confirmed (Kim et al., 2013). The purpose of this study was to develop a PSM measure that is valid theoretically and empirically for a single country. This study had two specific goals. The first was to develop an item pool for measuring PSM in a single country (Korea). A previous study suggested that the PSM measure needs additional work on a country-by-country basis, and Korea is one such country that requires measurement modification (Kim et al., 2013). Perry and Vandenabeele (2015) also suggested that there is still a need for measurement improvements in the attraction-to-public-service dimension of PSM. Thus, we need to develop survey items that are more appropriate to reflect the theoretical components of PSM. The second goal was to test measurement invariance to validate using the measure across different groups. Kim et al. (2013) tested their international measure using only local government samples; thus, there was a need to consider the potential effect of sample bias in that study. If PSM measures do not have the same psychometric properties across different groups in a single country, then we cannot be confident about empirical findings comparing PSM and its relationships with other variables among different groups within that country. It still remains unknown in the PSM literature whether PSM has the same theoretical structure and psychological meaning across different groups in a single country. This study will contribute to the PSM research through developing a PSM measure that has the same psychometric properties across different groups in a single country.

The present study was structured as follows: first, we developed new items reflecting the institutional context of Korea and created an item pool to measure PSM. Second, we collected data (n = 1800) from both the public and private sectors in Korea and tested the four-dimensional PSM model. Finally, we validated the use of the PSM measure across different genders and sectors.

PSM and item generation

PSM is a public-focused concept, but it is not unique to government organizations (Perry, Hondeghem, & Wise, 2010). The stronger a person's PSM, the more likely he or she is to engage in behaviors that benefit the public, even when there is a lack of tangible individual rewards (Kim & Vandenabeele, 2010; Wise, 2000). PSM has been defined as 'an individual's predisposition to respond to motives grounded primarily or uniquely in public institutions and organizations' (Perry & Wise, 1990, p. 368). Brewer and Selden (1998, p. 417) described it as 'the motivational force that induces individuals to perform meaningful public service'. Rainey and Steinbauer (1999, p. 23) defined PSM as 'a general altruistic motivation to serve the interests of a community of people, a state, a nation or mankind'. PSM, moreover, has a 'common focus on motives and action in the public domain that are intended to do good for others and shape the well-being of society' (Perry & Hondeghem, 2008, p. 3). Such variety in the definitions of PSM means that greater care is needed when communicating conceptual and operational meanings (Bozeman & Su, 2015). However, even though the definitions vary, a commitment to the public interest, service to others, and self-sacrifice underlies the general understanding of PSM (Houston, 2006; Scott & Pandey, 2005). Therefore, this study defines PSM as an individual's predisposition to perform meaningful service for the country and the public, in pursuit of public values and the public interest rather than self-interest.

As mentioned earlier, Perry and Wise (1990) suggested that PSM has rational, norm-based, and affective motives. *Rational motives* are rooted in the maximization of individual utility. *Norm-based motives* relate to a desire to pursue the common good and further the public interest. *Affective motives* are rooted in human emotion. These various motives are primarily, or exclusively in some cases, associated with public service. Rational motives include participation in the policy-making process, commitment to a public program because of personal identification, and special- or private-interest advocacy. Norm-based motives include a desire to serve the public interest, a sense of duty and loyalty to the government, and an interest in social equity. Affective motives include 'the patriotism of benevolence' (Frederickson & Hart, 1985) and commitment to a program based on a genuine conviction about its social importance.

Kim and Vandenabeele (2010) assessed commonalities in the content of PSM internationally and slightly revised its concept. As noted above, they suggested that PSM is rooted in *self-sacrifice*, which underpins three motive categories (i.e. value-based, identification, and instrumental). This reconceptualization highlights the centrality of self-sacrifice in the construct and increases the distinctness of the theoretical components (Perry, 2014). *Value-based motives* concern the end public values that individuals want to achieve through their behaviors and actions. *Identification motives* relate to the people, groups, or objects that individuals want to serve. *Instrumental motives* consist of the means to perform meaningful public service. These three refined motivational components focus on value (for what), attitude (for whom), and behavior (how). Value-based motives relate to values and ethics, identification motives to attitude, and instrumental motives to behavior.

Using a revised PSM construct and data from an international survey, Kim et al. (2013) developed a 16-item measure of PSM that identified four dimensions in the PSM construct: attraction to public service (APS), commitment to public values (CPV), compassion (COM), and self-sacrifice (SS). This international measure of PSM produced the same four-factor structure (configural invariance), but failed to achieve the same factor loadings (metric invariance) and intercepts (scalar invariance) across the 12 studied countries. While Perry and Vandenabeele (2015) noted that the international PSM scale (Kim et al., 2013) improves upon the original instrument (Perry, 1996), its APS dimension is less theoretically compelling than a subscale rooted in the logics associated with loyalty to governance regimes. Governance regimes involve a set of actors and institutions as well as a collection of norms and procedures for pursuing common objectives at a country level (Bulmer, 1993; Weale, 1996). They suggested that the APS dimension needs to capture 'local' institutions and an individual's disposition toward the governance regime (Perry & Vandenabeele, 2015, p. 696).

Within the operational dimensions of PSM (Kim et al., 2013), the APS dimension concerns the disposition to serve the public, work for the common good, and participate in public policy processes. This dimension represents instrumental motives that correspond to Perry and Wise's (1990) rational motives from which self-serving motives are excluded. Instrumental motives concern the methods of performing meaningful public service (Kim & Vandenabeele, 2010). However, the current APS items are not likely to focus on behavior orientations, but rather they reflect general concerns regarding the public sector. This dimension can also be strengthened by developing more items associated with loyalty to democratic governance. The Republic of Korea's constitution calls for a liberal democratic political system (Kihl, 2015). Its principles are based on the people's sovereignty, the separation of powers among the three branches of government, the rule of law, the responsibility

to promote citizens' welfare, and the attainment of a peaceful unification of Korea. The APS dimension needs to focus more on democratic governance and legitimacy, public-service delivery, and working for the common good within Korea's institutional context (Jang & Shin, 2008; Kim, 2008). Accordingly, new items were developed, including those identifying loyalty to the democratic governance regime and those concerning one's disposition to contribute to the governance system. Thus, eight possible items were included in the APS dimension (Table 1). The four items (APS1, APS2, APS3, and APS4) are related to participating in the public policy process and working in the public sector (Kim & Vandenabeele, 2010; Perry & Wise, 1990), while the two items (APS5, APS6) are related to loyalty to the democratic governance regime (Perry & Vandenabeele, 2015) and the others (APS7, APS8) are related to serving the public and country (Kim & Vandenabeele, 2010; Kim et al., 2013).

Commitment to local institutions (Perry & Vandenabeele, 2015) is related to the CPV dimension. This dimension concerns the personal disposition to pursue the public interest and public values, representing value-based (Kim & Vandenabeele, 2010) or norm-based motives (Perry & Wise, 1990). Value-based motives concern the public values to which individuals want to contribute through their actions. The order of priority among common public values may differ across countries. A country's specific political institutions might affect people in terms of the public values they want to promote. The current items in the CPV dimension focus on more general values, but do not include important public values in the institutional context of Korea (Kim et al., 2013). Thus, several items focused on democracy and Korea's constitutional values were added to the CPV dimension along with the universal items in Kim et al. (2013). The three items (CPV2, CPV6, CPV8) are related to democracy and legitimacy (Perry & Vandenabeele, 2015), and the two (CPV3, CPV7) are related to human rights (Kim & Vandenabeele, 2010).

The COM dimension represents identification motives (Kim & Vandenabeele, 2010) or affective motives (Perry & Wise, 1990). This dimension focuses on affective bonding with the identified objects such as vulnerable people, the public, and society. Perry and Vandenabeele (2015) found that the COM subscale is more reliable than the original instrument (Perry, 1996). The two items pertaining to caring and empathy (COM1, COM6) were added to provide a wider range of choices in the COM dimension (Kim & Vandenabeele, 2010).

Public service requires an individual's self-sacrifice, which is 'the willingness to substitute service to others for tangible personal rewards' (Perry, 1996; p. 7). The items for the SS dimensions remained mostly unchanged from Perry's original items (1996) because they were not only the primary source of the items in the international measure (Kim et al., 2013), but were also generally acceptable in the Korean context. In total, 29 items were proposed to measure the 4 dimensions of PSM (Table 1). Among them, 11 items were also included in the 16-item international measure of PSM (Kim et al., 2013).

Data collection

This study aimed to develop and validate an item pool for measuring PSM in Korea. Two versions of the items – one in English and the other in Korean – were developed simultaneously. In the Korean-language questionnaire, a seven-point Likert-type scale (1 = strongly disagree, 7 = strongly agree) was used. The PSM items were placed randomly in the survey. The survey was exclusively designated for PSM research, including a total of 56 items: 35 PSM and related items, 17 questions on organizational context, and 4 demographic questions.

Table 1. Model estimations for the 29-Item PSM measure.

Dimensions and Items	Mean	SD	SFL
APS			
APS1: I like to discuss topics regarding government policies with others	4.45	1.332	.687
APS2: I like to participate in activities solving social problems	4.64	1.215	.799
APS3: I try to reflect my personal views onto policy issues	4.39	1.275	.762
APS4: I want to contribute to the societal development	4.92	1.165	.876
APS5: I want to contribute to realizing the constitutional principles in society	5.05	1.196	.869
APS6: For me it is of major concern to protect the democratic governance system	5.05	1.208	.805
APS7: Serving my country helps me realize myself	4.71	1.307	.778
APS8: I want to work to make my country better	4.98	1.208	.852
CPV			
CPV1: Equal opportunities for all citizens should be guaranteed*	5.80	1.138	.783
CPV2: Decisions regarding public policies should be democratic even if they take time and efforts	5.71	1.111	.810
CPV3: The dignity and well-being of all people should be the most important concerns in our society	5.71	1.133	.814
CPV4: The interests of future generations should be taken into account when making public policies*	5.65	1.067	.806
CPV5: To act ethically is essential for public servants*	5.99	1.083	.811
CPV6: Public servants must always be aware of the legitimacy of their activities	5.90	1.059	.806
CPV7: I personally support the protection of individual liberties and rights	5.74	1.035	.819
CPV8: We have to make every effort to pursue democracy	5.78	1.073	.859
COM			
COM1: It is difficult for me to contain my feelings when I see people in distress	5.06	1.141	.828
COM2: I feel sympathetic to the plight of the underprivileged*	5.40	1.091	.918
COM3: I empathize with other people who face difficulties*	5.47	1.086	.935
COM4: I get very upset when I see other people being treated unfairly*	5.70	1.086	.768
COM5: Considering the welfare of others is very important*	5.42	1.057	.744
COM6: I care much about other people	5.25	1.047	.790
SS			
SS1: I am prepared to make sacrifices for the public good of society*	4.67	1.197	.814
SS2: I believe in putting civic duty before self*	5.00	1.137	.776
SS3: I am willing to risk personal loss to help society*	4.50	1.264	.835
SS4: I think that people should give back to society more than get from it	4.66	1.275	.776
SS5: I would agree to a good plan to make a better life for the poor, even if it costs me money*	4.99	1.190	.812
SS6: Serving others would give me a good feeling, even if it makes me a loss	4.84	1.220	.842
SS7: Making a better society means more to me than personal achievements	4.52	1.263	.838
Measure of Fit (RMLE)	$SB\chi^2$ (df = 371) = 2749.9, $p < .05$; CFI = .9877; RMSEA = .0597, [.0576, .0618]; SRMR = .0647		

				Inter-dimensional Correlations			
	Alpha	CR	AVE	APS	CPV	COM	SS
APS	.925	.928	.649	(.806)			
CPV	.925	.940	.662	.634	(.814)		
COM	.914	.931	.695	.677	.777	(.834)	
SS	.920	.932	.669	.757	.501	.700	(.818)

Notes: Alpha = Cronbach's alpha; CR = composite reliability; AVE = average variance extracted; APS = attraction to public service; CPV = commitment to public values; COM = compassion; SS = self-sacrifice. Values in parentheses on the diagonal in the correlation matrix show the square root of AVE. All correlation coefficients are significant at $p < .01$.*indicates the item used in the 16-items international measure of PSM (Kim et al., 2013).

The data on which this study is based were collected from employees in both the public and private sector using an online survey. The sample was selected using purposive and quota sampling (Babbie, 2013). It consisted of 16 blocks based on two sector categories, four age categories, and two gender categories. Randomly selected from the panel were an equal number of males and females and an equal number of employees in the public and private sectors, all of whom were aged 20–59. When the total number of complete responses reached the target sample size that was set to 1800, the survey was closed. The

questionnaire was distributed to 6749 individuals who were randomly drawn from the panel and invited to complete this online survey, and 2214 participated (participation rate = 32.6%).[1] A total of 1842 questionnaires were completed (completion rate = 83.2%). Among these, 42 were removed because of unsatisfactory responses; thus, 1800 samples were ultimately obtained.

Online surveys usually produce less social desirability bias (Kreuter, Presser, & Tourangeau, 2008) and, if conducted properly, have significant advantages over other formats (Evans & Mathur, 2005), especially for methodological studies. However, the potential impact of sample selection bias needs to be considered for this survey (Winship & Mare, 1992). Sample selection bias is a systematic error causing some members of the population to be less likely to be included than others, and is therefore the likely reason for underrepresentation. Nevertheless, studies have suggested that the effects of such bias are overstated (e.g. Abraham, Maitland, & Bianchi, 2006). Given the nature of non-probability sampling, the estimates of the total samples should not be regarded as those of the population.

Of the total respondents, 49.6% were men and 50.4% were women. In terms of age, 19.4% were in their 20s, 30.3% in their 30s, and 30.0% in their 40s. As for the employment sector, similar numbers of respondents worked in the public and private sectors: 14.8% worked for government organizations; 6.6% for public corporations; 20.7% worked in public schools, universities, and R&D institutes; 2.2% in public welfare/health; 5.7% in other public-sector areas; 16.6% for private companies; 8.1% in private schools, universities, and R&D institutes; 3.7% in private welfare/health; and 21.6% in other private-sector areas.

Analyses

Descriptive statistics were computed for individual items using SPSS 22 (Table 1). Among the PSM dimensions, the respondents tended to report relatively high levels of CPV and COM but low levels of SS and APS. In the normality assessment, no item showed a skew or kurtosis value greater than the cutoffs of |3| or |8| recommended by Kline (2005). The inspection of item-total correlations showed that all correlation values were greater than .5. All Cronbach's alphas were greater than .9.

To test the hypothesized four-dimensional PSM model, confirmatory factor analyses (CFA) were conducted in LISREL 8.72 (Jöreskog & Sörbom, 2005) using the robust maximum likelihood estimation (RMLE) method (Yang-Wallentin, Joreskog, & Luo, 2010). The resulting CFA with RMLE showed that the initial 29-item four-dimensional model fit the data well ($SB\chi^2$ (df = 371) = 2749.9, $p < .05$; CFI = .9877; RMSEA = .0597; SRMR = .0647).[2] All items loaded significantly on their a priori dimension ($p < .001$), and the standardized factor loadings (SFL) ranged from .687 to .935. The composite reliability (CR) of the set of reflective indicators for each PSM dimension was greater than .9. The correlation estimates showed that all PSM dimensions were significantly correlated with each other ($p < .01$). The lowest correlation was .501 between CPV and SS, while the highest one was .777 between CPV and COM. It may reduce the possibility raised in the international measure that the CPV and COM dimensions were perfectly correlated in Korea (Kim et al., 2013). All average variance extracted (AVE) scores were above .6. The square root of the AVE of each factor was larger than the factor's correlation with any other factor in the model (Fornell & Larcker, 1981).[3] The results support the reliability, convergent validity, and discriminant validity of the model. Thus, the four-dimensional model with 29 items was confirmed.

Since shorter scales are generally preferred to reduce the respondents' workload, the more concise the measure, the more desirable it will be considered. The items for each dimension were 'reflective' because they represented reflections, or manifestations, of the dimension (Kim, 2011). The test results showed that all factor loadings were significant and rather high (all SFLs > .6). All correlation coefficients between items in each designated dimension were positive and statistically significant ($p < .01$) (see Appendix 1). This means the reflective items in this measure were relatively homogeneous and essentially interchangeable. Therefore, while adding or removing items could affect reliability, it would not change the essential nature of the underlying dimension (Diamantopoulos & Winklhofer, 2001). Any combination of items in a dimension can be regarded as equivalent. This means that any item in the APS and CPV dimensions could be used interchangeably with other items in the Korean context. When using structural equation modeling (SEM), each dimension needs to have at least three reflective indicators to achieve identification. Thus, any combination of at least three items in each dimension (at least 12 items in total) could be used as a measure of PSM in the Korean context.

To make the measure more concise, only 16 items were selected (Table 2). The items for the COM and SS dimensions were the same as in the international PSM measure (Kim et al., 2013) while newly developed items were mostly used for the APS and CPV dimensions. CFA showed that the overall model fit of the 16-item measure was quite strong ($SB\chi^2$ (df = 98) = 596.1, $p < .05$; CFI = .9921; RMSEA = .0532; SRMR = .0535). Cronbach's alpha, CR, and AVE exceed the criteria. The test results support the reliability, convergent validity, and discriminant validity of the 16-item measure.

Measurement invariance is the equivalence of a measured construct in two or more groups (Chen, 2008). It ensures that the same constructs are being assessed in each group. When the measures are comparable across different groups, the comparisons and analyses of the scores are valid, and subsequent interpretations are meaningful. The PSM research assumes that public employees are more likely to have higher levels of PSM than those employed in the private sector (Houston, 2006; Perry & Wise, 1990). Gender was the most frequently studied antecedent of PSM, and aggregate findings suggest that women tend to exhibit higher levels of PSM (Ritz, Brewer, & Neumann, 2016). However, we cannot have confidence that the differences are not merely artifacts of measurement if PSM measures do not have the same psychometric properties (Chen, 2008). Measurement invariance was tested across sector and gender in this study.

Multiple-group CFA is the most commonly used method for testing measurement invariance (Vandenberg & Lance, 2000). We tested the cross-gender and cross-sector equivalences of the 16-item measure, as done by Kim et al. (2013), focusing on the three types of measurement invariance most commonly used to validate using a measure across different groups (Chen, 2008; Meade, Johnson, & Braddy, 2008; Schmitt & Kuljanin, 2008; Vandenberg & Lance, 2000). First, configural invariance tests whether the measure has the same hypothesized factorial structure (the same number of factors and the same items loading on each factor) in each group. Configural invariance must be established for subsequent tests to be meaningful. Second, metric invariance tests whether the strength of the relationship between each factor and its associated items (reflected in the factor loadings) is the same across groups. When metric invariance is (at least partially) established, predictive relationships can be compared across groups. Third, scalar invariance tests whether the intercepts of the regression equations of the items on their hypothesized factors are equivalent across groups. Scalar invariance allows the factor means to be compared across groups.

Table 2. Model estimations for the 16-Item PSM measure.

Dimensions and Items	SFL
APS	
APS4: I want to contribute to the societal development	.871
APS5: I want to contribute to realizing the constitutional principles in society	.901
APS6: For me it is of major concern to protect the democratic governance system	.810
APS8: I want to work to make my country better	.849
CPV	
CPV1: Equal opportunities for all citizens should be guaranteed*	.728
CPV6: Public servants must always be aware of the legitimacy of their activities	.799
CPV7: I personally support the protection of individual liberties and rights	.844
CPV8: We have to make every effort to pursue democracy	.885
COM	
COM2: I feel sympathetic to the plight of the underprivileged*	.914
COM3: I empathize with other people who face difficulties*	.948
COM4: I get very upset when I see other people being treated unfairly*	.776
COM5: Considering the welfare of others is very important*	.764
SS	
SS1: I am prepared to make sacrifices for the public good of society*	.821
SS2: I believe in putting civic duty before self*	.792
SS3: I am willing to risk personal loss to help society*	.823
SS5: I would agree to a good plan to make a better life for the poor, even if it costs me money*	.789
Measure of Fit (RMLE)	$SB\chi^2$ (df = 98) = 596.1, $p < .05$; CFI = .9921; RMSEA = .0532, [.0491, .0573]; SRMR = .0535

				Inter-dimensional Correlations			
	Alpha	CR	AVE	APS	CPV	COM	SS
APS	.906	.918	.737	(.858)			
CPV	.857	.888	.666	.668	(.816)		
COM	.887	.915	.730	.672	.772	(.854)	
SS	.860	.881	.650	.760	.524	.710	(.806)

Notes: Alpha = Cronbach's alpha; CR = composite reliability; AVE = average variance extracted; APS = attraction to public service; CPV = commitment to public values; COM = compassion; SS = self-sacrifice. Values in parentheses on the diagonal in the correlation matrix show the square root of AVE. All correlation coefficients are significant at $p < .01$.
*indicates the item used in the 16-items international measure of PSM (Kim et al., 2013).

Testing the differences in nested models to determine measurement invariance requires comparing the differences in the chi-square and CFI fit indices. Measurement invariance is acceptable if the difference in the chi-square ($\Delta SB\chi^2$) between two models is not statistically significant ($p > .05$), and the difference in the CFI (ΔCFI) across the two models is less than .002 (Meade et al., 2008). Like the chi-square tests for overall model fit, however, the chi-square difference test is highly sensitive to sample size. Thus, we relied more on ΔCFI than $\Delta SB\chi^2$ when the results of two tests were different.[4]

First, the cross-gender equivalence of the 16-item PSM measure between female and male employees was tested. Table 3 shows the model fit statistics for each model and the model differences test for each model comparison. In testing for configural invariance, the separate models for female and male employees both exhibited acceptable levels of

Table 3. Measurement invariance tests for the 16-item PSM measure.

	Model fit measures				Model differences			
	$SB\chi^2$	df	*p*	CFI	$\Delta SB\chi^{2*}$	Δdf	*p*	ΔCFI
Female vs. male employees								
Female (*n* = 893)	507.7	98	.0	.9859				
Male (*n* = 907)	356.6	98	.0	.9926				
Configural invariance	857.4	196	.0	.9897				
Metric invariance	892.4	208	.0	.9893	13.6	12	.329	.0004
Scalar invariance	943.8	220	.0	.9887	59.2	12	.0	.0006
Public vs. Private sector								
Public (*n* = 900)	400.7	98	.0	.9915				
Private (*n* = 900)	423.8	98	.0	.9881				
Configural invariance	825.8	196	.0	.9900				
Metric invariance	859.1	208	.0	.9896	8.4	12	.751	.0004
Scalar invariance	924.9	220	.0	.9888	110.0	12	.0	.0008

*Computed by EXCEL macro file (Bryant & Satorra, 2013).

model fit, as did the combined multi-group CFA model ($SB\chi^2$ (df = 196) = 857.4, $p < .05$; CFI = .9897). This means the measure had the same hypothesized factorial structure in each group. Next, the test for metric invariance constrained each matching factor loading, resulting in the ΔCFI being .0004 (less than .002) and the $\Delta SB\chi^2$ being only 12.9 with 12 degrees of freedom, indicating a nonsignificant difference. Thus, the measure achieved full metric invariance across the female and male employees. The last test was for scalar invariance, which tests for the equality of the measured variable intercepts on the construct. The result showed that the ΔCFI was .0006 (less than .002), while the $\Delta SB\chi^2$ was 59.2 with 12 degrees of freedom, indicating a significant difference. Since ΔCFI statistics are considerably less sensitive to sample size than chi-square (Meade et al., 2008), it is reasonable to rely on ΔCFI rather than $\Delta SB\chi^2$. Thus, we can assume the conditions for scalar invariance were met. The 16-item measure for PSM was scalar invariant, and the factor means could be compared across genders in Korea.

Next, the cross-sector equivalence of the 16-item PSM measure between the public and private sectors was tested. The separate models for each sector exhibited acceptable levels of model fit. The results of the invariance test were similar to those of the cross-gender equivalence test. The configural invariance test indicated that the measure had the same hypothesized factorial structure in each sector. The metric invariance test result indicated that the strength of the relationship between ewach factor and its associated items was the same across sectors. The scalar invariance test showed that the ΔCFI was .0008 (less than .002) while the $\Delta SB\chi^2$ was 110.0 (df = 12, $p < .05$). We can say that scalar invariance was supported, relying on ΔCFI rather than $\Delta SB\chi^2$. Since invariance was achieved at both the factor loading and intercept levels, scores from different sectors and genders had the same unit of measurement (i.e. factor loading) as well as the same origin (i.e. intercept). Thus, predictive relationships (PSM's antecedents and consequences) and factor means can be meaningfully compared across sectors and genders in Korea.

This study produced several notable findings: first, the new items for measuring PSM were developed, based on the theoretical arguments and institutional context of Korea. Second, the four-dimensional 29-item PSM measure exhibited an acceptable level of model fit, and the test results provided support for the reliability, convergent validity, and discriminant validity of this measure. Third, all items in each dimension were significantly correlated

with each other, and all factor loadings were significant and quite high. Fourth, the items in each dimension can be used interchangeably with other items in the Korean context. Fifth, any combination of at least three items in each dimension (at least 12 items in total) can be used as a measure of PSM in Korea. Sixth, the 16-item PSM measure is recommended for its increased conciseness. Seventh, the 16-item measure exhibited an acceptable level of model fit, and the test results provided support for the reliability, convergent validity, and discriminant validity of the measure. Finally, the tests for measurement invariance across genders and sectors showed that the four-dimensional 16-item PSM measure met the criteria for configural invariance, metric invariance, and scalar invariance. As a result, any type of group comparison between female and male employees or between the public and private sectors in Korea can be made without concern that the differences are caused by differing measurement properties according to gender or sector.

Conclusion

Kim and Vandenabeele's (2010) proposal for an international scale resulted in a new instrument (Kim et al., 2013) that improved upon Perry's (1996) instrument in several ways. However, it is not a truly universal measure of PSM. Even though the 12 studied countries did not differ in terms of the four-dimensional composition, the measure failed to achieve metric invariance. Thus, the strength of the relationship between each dimension and its items was not the same across the 12 countries,[5] which means the exact meaning and scaling of the PSM dimensions were likely to differ across the different cultures and countries. This could have been attributable to the fact that the countries had various institutional backgrounds, including a newly independent country and a socialist state. Some countries had a presidential system of government while others had a parliamentary system, and the countries had varying degrees of democracy. In addition, Perry and Vandenabeele (2015) suggested that we need to improve current measures to better capture one's commitment to a governance regime. Thus, it is necessary to develop more items for measuring PSM and test the measurement instrument's equivalence for a single country.

This study developed an item pool for measuring PSM, and provided support for the initial 29-item and the more concise 16-item measure of PSM in Korea. The new items reflect the institutional context of Korea, and capture an individual's loyalty to the governance regime as well as the institutions that make public decisions (Perry & Vandenabeele, 2015). Since the items significantly correlate with others, all items in each dimension can be used interchangeably in the Korean context. The tests for measurement invariance showed that the factor loadings and intercepts were identical across genders and sectors. Thus, predictive relationships and factor means can be meaningfully compared across different groups in Korea.

As is true of any study, the research reported here has a number of limitations. The potential impact of sample bias needs to be considered in understanding this study (Winship & Mare, 1992). The respondents from the online panel are self-selected. Another weakness of online sampling is the underrepresentation of those individuals who are either very old or less educated than the general population. Thus, the estimates of the online samples should not be regarded as those of the population. This also means that the survey items for measuring PSM should be examined with different samples (Choi & Perry, 2010).

This study may be the first to show that the exact meaning and scaling of PSM dimensions are likely to be the same across genders and sectors in a single country. It is meaningful in that the items refined by the current study seem to be robust in the Korean context. It can encourage researchers to have confidence in that the predictive relationships between PSM and other variables can be meaningfully compared and generalized in Korea. It opens the path to use a shortened version of PSM by confirming that the items in each dimension can be used interchangeably. It can also encourage scholars in other countries to test the cross-group equivalence of the four-dimensional PSM measure in each country. It would be worthwhile to check whether this measure could be also used in the context of other countries. We can then consider how to develop more robust international measures for PSM.

Notes

1. The company *Invight Inc*. has an invitation-based panel of over 820,000 Internet users throughout Korea who have agreed to receive occasional e-mails asking them to participate in online surveys. Those who fill out survey questionnaires have a chance to win a gift of approximately $1. This online survey took 10 days to run in March 2016.
2. The Satorra-Bentler scaled chi-square statistic ($SB\chi^2$) is likely to reject model fit for large sample sizes. Following recent recommendations (Williams, Vandenberg, & Edwards, 2009), we assumed the model achieves an acceptable fit to the data when the comparative fit index (CFI) exceeds .95, the root mean square error of approximation (RMSEA) is below .08, and the standardized root mean residual (SRMR) is below .10.
3. AVEs above .50 are treated as indications of convergent validity, and Cronbach's alphas and CR above .70 are treated as indications of reliability. Each dimension meets the criterion in support of discriminant validity when the square root of the AVE of each dimension is larger than the dimension's correlation with any other dimension in the model.
4. Since the difference between the Satorra-Bentler scaled chi-square statistics ($\Delta SB\chi^2$) is not chi-square distributed, the nested chi-square test was calculated using the Satorra-Bentler scaling corrections outlined by Bryant and Satorra (2012).
5. The countries were Australia, Belgium, China, Denmark, France, Italy, Korea, Lithuania, the Netherlands, Switzerland, the United Kingdom, and the United States.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported by the Research Program funded by the SeoulTech (Seoul National University of Science and Technology).

Notes on contributor

Sangmook Kim is a professor in the Department of Public Administration at Seoul National University of Science and Technology, Seoul, Korea. His research focuses on public service motivation, organizational behavior, human resources management, public management reform, and gender-related issues. He is the author of *Public Service Motivation in Korea* (Seoul: Jipmoondang, 2013), and co-editor of *Public Administration and Policy in Korea: Its Evolution and Challenges* (Routledge, 2018).

References

Abraham, K. G., Maitland, A., & Bianchi, S. M. (2006). Nonresponse in the American time use survey: Who is missing from the data and how much does it matter? *Public Opinion Quarterly, 70*, 676–703.

Babbie, E. (2013). *The practice of social research* (13th ed.). Singapore: Wadsworth Cengage Learning.

Bozeman, B., & Su, X. (2015). Public service motivation concepts and theory: A critique. *Public Administration Review, 75*, 700–710.

Brewer, G. A., & Selden, S. C. (1998). Whistle blowers in the federal civil service: New evidence of the public service ethic. *Journal of Public Administration Research and Theory, 8*, 413–440.

Bryant, F., & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. *Structural Equation Modeling: A Multidisciplinary Journal, 19*, 372–398.

Bryant, F., & Satorra, A. (2013). EXCEL macro file for conducting scaled difference chi-square tests via LISREL 8, LISREL 9, EQS, and Mplus. Available from the authors.

Bulmer, S. J. (1993). The governance of the European Union: A new institutionalist approach. *Journal of Public Policy, 13*, 351–380.

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*, 1005–1018.

Choi, D. L., & Perry, J. L. (2010). Developing a tool to measure ethical sensitivity in public administration and its application. *International Review of Public Administration, 14*, 1–12.

Coursey, D. H., & Pandey, S. K. (2007). Public service motivation measurement: Testing an abridged version of Perry's proposed scale. *Administration & Society, 39*, 547–568.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*, 269–277.

Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet Research, 15*, 195–219.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

Frederickson, H. G., & Hart, D. K. (1985). The public service and the patriotism of benevolence. *Public Administration Review, 45*, 547–553.

Houston, D. J. (2006). "Walking the walk" of public service motivation: Public employees and charitable gifts of time, blood, and money. *Journal of Public Administration Research and Theory, 16*, 67–86.

Jang, J., & Shin, D. C. (2008). Democratizing economic policymaking in South Korea: Painfully slow progress toward an executive-legislative balance of power. *International Review of Public Administration, 13*, 59–70.

Jöreskog, K. G., & Sörbom, D. (2005). LISREL 8.72 [Computer Software]. Lincolnwood: Scientific Software International.

Kihl, Y. W. (2015). *Transforming Korean politics: Democracy, reform, and culture*. London: Routledge.

Kim, S. Y. (2008). The East Asian developmental state and its economic and social policies: The case of Korea. *International Review of Public Administration, 12*, 69–87.

Kim, S. (2011). Testing a revised measurement of public service motivation: Reflective versus formative specification. *Journal of Public Administration Research and Theory, 21*, 521–546.

Kim, S., & Vandenabeele, W. (2010). A strategy for building public service motivation research internationally. *Public Administration Review, 70*, 701–709.

Kim, S., Vandenabeele, W., Wright, B. E., Andersen, L. B., Cerase, F. P., Christensen, R. K., …De Vivo, P. (2013). Investigating the structure and meaning of public service motivation across populations: Developing an international instrument and addressing issues of measurement invariance. *Journal of Public Administration Research and Theory, 23*, 79–102.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys. *Public Opinion Quarterly, 72*, 847–865.

Lee, C., & Kim, S. (2014). Making the quality of government better: The possibility of public service motivation (PSM) change. *International Review of Public Administration, 19*, 107–125.

Meade, A. W., Johnson, E., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*, 568–592.
Perry, J. L. (1996). Measuring public service motivation: An assessment of construct reliability and validity. *Journal of Public Administration Research and Theory, 6*, 5–22.
Perry, J. L. (2014). The motivational bases of public service: Foundations for a third wave of research. *Asia Pacific Journal of Public Administration, 36*, 34–47.
Perry, J. L., & Hondeghem, A. (2008). *Motivation in public management: The call of public service*. Oxford: Oxford University Press.
Perry, J. L., & Vandenabeele, W. (2015). Public service motivation research: Achievements, challenges, and future directions. *Public Administration Review, 75*, 692–699.
Perry, J. L., & Wise, L. R. (1990). The motivational bases of public service. *Public Administration Review, 50*, 367–373.
Perry, J. L., Hondeghem, A., & Wise, L. R. (2010). Revisiting the motivational bases of public service: Twenty years of research and an agenda for the future. *Public Administration Review, 70*, 681–690.
Rainey, H. G., & Steinbauer, P. (1999). Galloping elephants: Developing elements of a theory of effective government organizations. *Journal of Public Administration Research and Theory, 9*, 1–32.
Ritz, A., Brewer, G. A., & Neumann, O. (2016). Public service motivation: A systematic literature review and outlook. *Public Administration Review, 76*, 414–426.
Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*, 210–222.
Scott, P. G., & Pandey, S. K. (2005). Red tape and public service motivation: Findings from a national survey of managers in state health and human services agencies. *Review of Public Personnel Administration, 25*, 155–180.
Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–70.
Weale, A. (1996). Environmental rules and rule-making in the European Union. *Journal of European Public Policy, 3*, 594–611.
Williams, L. J., Vandenberg, R. J., & Edwards, J. R. (2009). Structural equation modeling in management research: A guide for improved analysis. *The Academy of Management Annals, 3*, 543–604.
Winship, C., & Mare, R. D. (1992). Models for sample selection bias. *Annual Review of Sociology, 18*, 327–350.
Wise, L. R. (2000). The public service culture. In R. J. Stillman II (Ed.), *Public administration concepts and cases* (7th ed., pp. 342–353). Boston: Houghton Mifflin.
Wright, B. E. (2008). Methodological challenges associated with public service motivation research. In J. L. Perry & A. Hondeghem (Eds.), *Motivation in public management* (pp. 80–98). Oxford: Oxford University Press.
Yang-Wallentin, F., Joreskog, K. G., & Luo, H. (2010). Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling: A Multidisciplinary Journal, 17*, 392–423.

Appendix 1.

Table A1. Correlations between the Items in APS.

	APS1	APS2	APS3	APS4	APS5	APS6	APS7
APS2	.667**						
APS3	.647**	.734**					
APS4	.515**	.655**	.646**				
APS5	.512**	.607**	.592**	.765**			
APS6	.547**	.596**	.543**	.646**	.728**		
APS7	.438**	.530**	.517**	.627**	.600**	.561**	
APS8	.480**	.593**	.547**	.740**	.717**	.650**	.718**

** $p < .01$.

Table A2. Correlations between the Items in CPV.

	CPV1	CPV2	CPV3	CPV4	CPV5	CPV6	CPV7
CPV2	.684**						
CPV3	.656**	.721**					
CPV4	.560**	.544**	.590**				
CPV5	.558**	.555**	.534**	.634**			
CPV6	.527**	.546**	.542**	.600**	.774**		
CPV7	.536**	.590**	.589**	.593**	.579**	.610**	
CPV8	.570**	.613**	.616**	.631**	.638**	.645**	.733**

** $p < .01$.

Table A3. Correlations between the Items in COM.

	COM1	COM2	COM3	COM4	COM5
COM2	.754**				
COM3	.715**	.838**			
COM4	.530**	.598**	.670**		
COM5	.541**	.608**	.648**	.612**	
COM6	.620**	.637**	.655**	.546**	.635**

** $p < .01$.

Table A4. Correlations between the Items in SS.

	SS1	SS2	SS3	SS4	SS5	SS6
SS2	.563**					
SS3	.656**	.682**				
SS4	.554**	.568**	.629**			
SS5	.606**	.548**	.580**	.632**		
SS6	.645**	.572**	.644**	.594**	.715**	
SS7	.636**	.595**	.679**	.617**	.615**	.696**

** $p < .01$.