

Running head: Unifactorial Citizenship Performance Measure

Development and Validation of a Unifactorial Measure of Citizenship Performance

Arthur E. Poropat¹ and Liz Jones²

¹Griffith University

²Griffith University

We are greatly indebted to Professor Walter Borman (University of South Florida), who gave permission to use the items that formed the basis for this research, and Professor Stephen Billett (Griffith University), for his invaluable support and guidance.

Word count (exc. figures/tables/references): 6,960

*Requests for reprints should be addressed to Arthur Poropat, Department of Management, Griffith University, Nathan, Queensland, Australia (e-mail: arthur.poropat@griffith.edu.au).

Abstract

Citizenship Performance has long been accepted as an important aspect of performance, but there has been little consensus on how best to measure this construct. Previous measures have assessed as many as five sub-factors, yet recent meta-analyses have indicated that Citizenship is largely unifactorial. At the same time, none of the previously developed unifactorial Citizenship Performance scales have been adequately validated. Consequently, this article reports the development of a unifactorial measure of Citizenship Performance, which was found to have good criterion validity. The unifactorial scale also had superior convergent and divergent validity, and better internal factor structure, than the commonly-used Job Dedication and Interpersonal Facilitation scales, developed by Van Scotter and Motowidlo (1996).

The last few decades have seen a significant research effort aimed at clarifying and elucidating the nature of workplace performance (Motowidlo, 2003). One aspect of workplace performance that has gained widespread support has had a variety of names, including Organizational Citizenship Behavior (OCB: Smith, Organ, & Near, 1983), Contextual Performance (Borman & Motowidlo, 1993) and Prosocial Organizational Behavior (Brief & Motowidlo, 1986), but has more recently been labeled Citizenship Performance (Borman, Penner, Allen, & Motowidlo, 2001; Coleman & Borman, 2000). Citizenship Performance has come to be seen as one of the major forms of work performance (Motowidlo, 2003), in part due to its observed and hypothesized relationships with important organizational outcomes such as productivity and profit (Schnake & Dumler, 2003).

Citizenship Performance is distinct from Task Performance, the specific tasks that people are employed for and that are formally monitored and recognized (Coleman & Borman, 2000; Organ & Paine, 1999). Instead, Citizenship Performance is seen as facilitating Task Performance by creating a supportive work context. As such, Citizenship Performance includes making an effort beyond what is formally required, volunteering, cooperating and supporting others, and following rules and procedures even when inconvenient (Borman, Penner et al., 2001). This range of activities is broad and as many as seven distinct aspects or themes of Citizenship have been proposed (Podsakoff, MacKenzie, Paine, & Bachrach, 2000), along with more than 40 different measures (LePine, Erez, & Johnson, 2002). A particularly thorough attempt at clarifying the underlying structure of Citizenship Performance was provided by Coleman and Borman (2000). On the basis of a literature review, these authors argued for a three-factor structure of Citizenship Performance and provided supporting evidence derived from inductive content sorting, exploratory factor analysis and multidimensional scaling. However, despite the quality of the developmental work the proposed three-factor structure was not supported by later research (Sun 2001, cited in Miles, Borman, Spector, & Fox, 2002).

Clearly, there have been problems with accurately assessing the underlying structure of Citizenship Performance.

Despite this diversity, two general models of Citizenship Performance have dominated within the literature: a two-factor and a five-factor model. The simpler, two-factor model of Citizenship was originally proposed by Smith et al. (1983), and is comprised of an interpersonal component, reflecting the extent to which staff assist each other with tasks, and a motivational component, expressed in the degree to which staff exceed standards. Later authors used a similar model (Van Scotter & Motowidlo, 1996; Williams & Anderson, 1991) even though the labels they applied to their scales differed from the original (Podsakoff et al., 2000). On theoretical grounds, this two-factor model was extended into five factors by Organ (1988), who proposed the dimensions of Altruism (supporting fellow-workers), Conscientiousness (diligent application to tasks), Sportsmanship (tolerating difficulties without complaint), Courtesy (aiming to prevent problems) and Civic Virtue (active participation in improvement activities). Several assessments have been developed to measure these dimensions, for example the scales developed by Podsakoff, MacKenzie, Moorman and Fetter (1990) and Lievens and Anseel (2004).

Unfortunately, both of these models have been subject to a range of concerns regarding their validity. For example, Van Scotter and Motowidlo (1996) presented two scales, Job Dedication and Interpersonal Facilitation, which roughly correspond to Organ's dimensions of Conscientiousness and Altruism respectively (Podsakoff et al., 2000). However several researchers, including the original authors, have concluded that Van Scotter and Motowidlo's Job Dedication scale overlaps with Task Performance and that it added no variance to the prediction of Overall Performance and individual difference variables such as personality (Barksdale & Werner, 2001). Conway (1999) argued for the empirical distinctiveness of Job Dedication based on ratings of managerial performance but his ratings were an aggregate of various ratings rather than a previously validated scale. Several of these ratings are clearly linked to Citizenship, but several, such as self-awareness, ethics and willingness to confront problem subordinates, are difficult to integrate with the conceptualizations described above. So, it is difficult to interpret the implications of Conway's results with respect to the distinctiveness of

Job Dedication. Therefore, there remains no clear evidence supporting the validity of the Job Dedication scale as a separate measure of Citizenship Performance.

The five-factor model of Citizenship has also been subjected to challenges. LePine et al. (2002) meta-analyzed the measures of Citizenship that had been developed using that model and found that they correlated at levels that were consistent with them acting as alternate forms of the same underlying construct. They also had similar correlations with criterion variables, so LePine et al. (2002) concluded that the various Citizenship scales may in fact be “not much more than equivalent indicators” (pp. 60-61) of a latent Citizenship construct. Consistent with this, Hoffman, Blair, Meriac & Woehr (2007) reported a confirmatory factor analysis (CFA) of Citizenship, based on a meta-analysis of Citizenship and Task Performance. This confirmed that a single latent construct was the best fitting model for the underlying structure of Citizenship Performance.

Even prior to these findings, several researchers had indicated that multi-dimensional measures of Citizenship may be, at the very least, unnecessary. For example, Koys (2001) developed a single-factor measure by selecting five items that corresponded with each of the most commonly discussed dimensions of Citizenship Performance. Even Podsakoff and MacKenzie, who have been responsible for developing several multidimensional measures of Citizenship (e.g., MacKenzie, Podsakoff, & Fetter, 1991, 1993; Podsakoff & MacKenzie, 1994), have argued that the correlations between their various Citizenship scales were high enough to justify combining them into one scale (MacKenzie, Podsakoff, & Ahearne, 1998). It appears that for most purposes there is little justification for using anything more than a single scale to measure Citizenship.

Despite the growing evidence that Citizenship Performance is unifactorial, little effort has been taken to validate unifactorial measures of Citizenship. Although they used unifactorial measures, both Koys (2001) and MacKenzie et al. (1998) only reported internal consistency reliabilities for their scales. It is sometimes assumed that Cronbach's alpha is a sufficient measure of a scale's internal structure, but scales can have an apparently satisfactory alpha and still have a complex underlying structure, effectively measuring more than one factor (Shevlin, Miles, Davies, & Walker, 2000). In addition, the MacKenzie et al. (1998) scale had a correlation with in-role performance ($r = .03$) that was substantially less than correlations reported elsewhere (e.g., LePine et al. (2001) cited correlations

of .36 when ratings came from different raters, and .62 when both ratings came from supervisors). In summary, neither of these scales was thoroughly validated, nor does there appear to be well-validated alternatives.

The preceding review of measures of Citizenship Performance suggests two important points. First, for most purposes Citizenship Performance should be treated as a unifactorial measure, a point that has been tacitly accepted by previous researchers. This is not to say that the evidence cited above nullifies the theoretical arguments and empirical evidence for multiple dimensions summarized by Podsakoff et al.(2000). Admittedly, Podsakoff et al.(2000) did not test the statistical significance of the differences they reported and the covariates with the greatest range of correlations with Citizenship dimensions (i.e., Job Satisfaction, Leader Support, & Organizational Commitment) were considered in the meta-analyses of LePine et al. (2001) and Hoffman et al. (2007). It nonetheless remains possible that future researchers examining the dimensionality of Citizenship Performance will find significant differences in correlations of Citizenship dimensions with other constructs, perhaps on the basis of further refinement of multifactorial measures. At this stage, however, the value of using multiple Citizenship Performance scales for other research purposes, appears to be limited. Yet, unlike personality variables such as Conscientiousness, for which the underlying dimensionality is still a matter for research despite general acceptance of a variety of valid measures of the overarching construct (Roberts, Bogg, Walton, Chernyshenko, & Stark, 2004), the foregoing review found that no existing unifactorial measures of Citizenship Performance had been adequately validated. So the second point is that there is a need for a new, well-validated, unifactorial measure of Citizenship Performance.

The research reported in this article was conducted in order to develop just such a measure of Citizenship Performance. We obtained from Walter Borman (Personal communication, September, 1999) a set of items designed to measure the three components of Citizenship Performance identified in Coleman and Borman's (2000) review. The measure reported here was refined from the Borman items using exploratory factor analysis but, as has been argued, in order to validate a scale it is necessary to verify its underlying dimensionality, using tools such as CFA (McDonald, 1999). Consequently, a CFA was used to assess scale structure in this research. In addition to having a

reliable, simple structure, a valid scale should be correlated with other variables in a manner that is consistent with theoretical expectations. When it was first proposed as a dimension of performance, Citizenship was theorized to be that component of performance most responsive to attitudinal, motivational and dispositional variables (Borman & Motowidlo, 1993; Smith et al., 1983). This was because Citizenship was seen as largely volitional behavior, in contrast with the role-prescribed behaviors included within Task Performance (Borman & Motowidlo, 1993). Hence, whether an individual engaged in Citizenship was due to their willingness to perform, while an individual's Task behaviors was more closely linked with their ability to perform. Personality has long been accepted as a major component of a person's willingness to perform (Blumberg & Pringle, 1982). Previous research has found that personality variables based on the Five Factor Model of personality (FFM) are reliably correlated with Citizenship (Hurtz & Donovan, 2000). Specifically, the sample-weighted average of the corrected correlations reported between Citizenship Performance measures and the FFM factors, reported by Hurtz and Donovan (2000) were – Extraversion: .08; Agreeableness: .16; Conscientiousness: .19; Emotional Stability: .16; Openness: .03. A valid measure of Citizenship Performance should reflect this pattern of correlations, so the first hypothesis of this research was that:

Hypothesis 1: A reliable Citizenship Performance scale, reflecting a single latent factor, will have a similar pattern of correlations with FFM personality measures to those observed in previous research.

A valid measure should also have stronger correlations with measures to which it is conceptually linked and weaker correlations with measures that are conceptually linked to independent constructs. As described above, Citizenship Performance has been conceived of as reflecting motivation and effort. while ability factors, including cognitive ability and job knowledge, should not have a close relationship with Citizenship Performance, instead having a direct effect on Task Performance (Motowidlo, Borman, & Schmit, 1997). So, a valid measure of Citizenship Performance should have a comparatively lower correlation with a measure of job knowledge than with a measure of effort

Hypothesis 2: A reliable Citizenship Performance scale, reflecting a single latent factor, will be more strongly correlated with a measure of effort than with a measure of job knowledge.

A further distinction between the constructs of Task and Citizenship Performance is that they have been defined to include different behaviors and outcomes. One of the features of Citizenship that was highlighted by Borman, Penner et al. (2001) was an employee's willingness to follow rules and procedures, even when inconvenient, while Task Performance is associated with the technical core of a job, in particular the quality of work undertaken. Therefore, it should be expected that a measure of Citizenship Performance should be more closely associated with independent measures of compliance than with measures of work quality.

Hypothesis 3: A reliable Citizenship Performance scale, reflecting a single latent factor, will be more strongly correlated with a measure of compliance with rules and procedures than with a measure of quality of work.

Finally, it would be surprising if a valid scale was not correlated with other measures of the same construct. Despite the problems with existing measures of Citizenship Performance that were outlined earlier in this article, they nonetheless appear to reflect the underlying latent Citizenship factor. Although previously developed measures of Citizenship are based on models that assume an underlying structure that is not empirically supported, they should nonetheless reflect the latent Citizenship factor (Hoffman et al., 2007), which means that any new measure of Citizenship Performance should be strongly associated with previously developed measures.

Hypothesis 4: A reliable Citizenship Performance scale, reflecting a single latent factor, will be highly correlated with pre-existing measures of Citizenship Performance.

In conclusion, although multifactor models of Citizenship Performance have long been advocated, the dominant models for measuring Citizenship Performance have faced problems. Some measures appear to reflect constructs other than Citizenship while the most reliable evidence, summarized by Hoffman et al. (2007) in their meta-analysis, is consistent with Citizenship being a single latent factor. Single-factor measures of Citizenship Performance used in previous research have not been thoroughly evaluated, so the research reported in this article was designed to develop and validate a single-factor measure.

Two studies are reported in this article. The first study involved the development of a single-factor scale, based on a principal axis factor analysis of the Borman items, as well as an initial test of the reliability of the scale's structure and a comparison of the scale with measures of personality. The second study was undertaken to test the reliability of this unifactorial Citizenship Performance scale's structure, using CFA, and whether the new scale demonstrated predicted relationships with measures of effort, job knowledge. The second study also permitted a comparison of the unifactorial Citizenship Performance scale with previously developed measures of Citizenship Performance.

STUDY 1: DEVELOPMENT AND INITIAL VALIDATION OF SCALE

The most important decision in scale development is the selection of an appropriate item pool as the basis for analysis (Reise, Waller, & Comrey, 2000). Reise et al. (2000) recommend that the item pool should be developed based on a thorough theoretical review of the construct to be measured and the selection of items should be over-inclusive. On the basis of a thorough theoretical and empirical review, Coleman and Borman (2000) identified 27 behaviors that encompassed the broad range of previously developed Citizenship Performance measures. A content sorting of these behaviors, undertaken by a sample of professional psychologists, was then used as the basis for exploratory factor analysis, multidimensional scaling analysis and cluster analysis, which ultimately produced a three-component structure. The Borman items used in this study consisted of three scales to measure these components, with four items per scale. Reise et al. (2000) argued that one of the major reasons for revising a measure is because its 'factor structure is "not as advertised" by the original authors' (p. 288). So, the fact that later research did not support the Coleman and Borman three factor model (Sun 2001, cited in Miles et al., 2002) indicated that the structure of the Borman items needed to be reconsidered. An analysis of these items has been reported here, prior to describing the development and initial validation of the unifactorial Citizenship Performance scale.

Method

Participants

Students in two semesters of a first-year undergraduate management course at an Australian university were the participants in Study 1. There are strong parallels between study and work, such as the emphasis on extrinsic rewards, external assessment of performance (marking and performance appraisals), and the degree of alienation from the outcomes of performance (Bowles, Gintis, & Meyer, 1999). Success in both fields is reliably associated with intellectual ability (Strenze, 2007) and the personality dimension of Conscientiousness (Lounsbury, Gibson, Sundstrom, Wilburn, & Loveland, 2004). Such are the similarities that some have argued that study *is* work (Marshall, 1988; Munson & Rubenstein, 1992). Consequently, although most research on Citizenship Performance has been conducted with employees, several researchers have examined Citizenship Performance with student samples (e.g., Allen & Rush, 1998; LePine & Van Dyne, 2001). Completion of the various rating scales included in this study was a formal part of the teaching process but inclusion within Study 1 was voluntary, with students able to decide whether or not they wished their ratings to be included in this research. Of the 454 students who were invited to participate, 389 agreed to do so, but ratings were only obtained for 298 of these students, an overall participation rate of 66%. Participants ranged in age from 17 to 53 years with an average age of 21.7, while 43% were male.

Measures

A hand scored version of the Borman items was used in this study as the basis for developing the unifactorial Citizenship Performance scale. This version had four items for each of three dimensions, namely: Personal Support, or the amount of help and cooperation provided to colleagues; Organizational Support, or the degree to which people comply with rules and show loyalty to the organization; and Conscientious Initiative, or the level of persistence and initiative demonstrated. For each item raters were presented with a short statement and asked to decide how accurately they believed the item described the individual being rated. Ratings for each item were made on a five-point, verbally-anchored scale, ranging from 'not at all accurately' (1) to 'very accurately' (5).

Saucier's (1994) Mini-Markers were used to assess personality in this study. The Mini-Markers are an abbreviated version of Goldberg's (1992) FFM markers, but have similar reliability

and validity, fewer difficult items and lower inter-scale correlations (Dwight, Cummings, & Glenar, 1998). Being quicker to administer, the Mini-Markers were a better choice for use in a time-constrained setting (Gosling, Rentfrow, & Swann Jr, 2003). Each item in the Mini-Markers consists of a single adjective, which participants were asked to rate on a nine-point, verbally anchored scale, from 'extremely inaccurate' (1) to 'extremely accurate' (9).

Procedure

Participants completed the Mini-Markers during class in the first two weeks of semester and were provided with feedback on their scores as part of a teaching exercise. After completing an assessed team project that lasted four weeks, participants were rated on the Borman items by other members of their teams. Most teams had three or more members, which meant that more than one rating was obtained for two-thirds of the participants. Ratings were completed during class-time and students were provided with guidance on how to score their ratings as well as on providing feedback to each other, as part of a teaching exercise linked to the team project. Few, if any, of the students had known each other prior to the commencement of this team project, so the ratings were largely based on students' behavior during the four week project. Although ratings were collected during class, all ratings were stored confidentially and the researchers were not able to access ratings until after the completion of the course, to reassure students that ratings would not affect end of semester marks.

Results

The descriptive statistics, reliabilities and inter-correlations for the Borman Citizenship Performance scales are presented in Table 1. These results show that the individual scales were relatively low on internal consistency reliability (none of the three scales had an alpha value over .7), and the scales were quite variable in terms of their inter-rater reliability. The three scales had relatively high inter-correlations, which when corrected for internal consistency reliability approximated 1.0. Thus, these results are consistent with Hoffman et al.'s (2007) conclusion that different scales for measuring Citizenship Performance are alternate forms for assessing the same, latent factor.

Insert Table 1 around here

Consequently, the Borman items were factor analyzed to identify their underlying factor structure. Worthington and Whittaker (2006) recommend that data should be assessed for factorability prior to any factor analysis. In this case, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .83, which exceeds Tabachnik and Fidell's (2001) recommended cutoff of .60. A principal axis factoring of the items revealed three factors with eigen-values greater than one (4.01, 1.63 and 1.10). Although it is common to retain all factors with eigen-values greater than one this can lead to spurious results, so decisions about factor retention should be made using more accurate methods, such as the scree test (Cortina, 2002) or parallel analysis (Lance, Butts, & Michels, 2006). The scree test showed that all factors had eigen-values that were less than 50% larger than the next largest factor, except for the first factor, which had an eigen-value that was 2.5 times the size of the eigen-value of the next extracted factor. A parallel analysis was conducted using the procedure developed by O'Connor (2000). This showed that only the first factor had an eigen-value greater than that obtained for corresponding factors derived from similar random datasets (i.e., 3.47, 2.13 and 1.93). Hence, only one factor was retained from the analysis.

Items for which at least a third of their variance (factor loadings of greater than .57) was accounted for by the first factor were chosen as the basis of the new scale, with three items coming from the original Conscientious Initiative sub-scale, two from Personal Support and one from the Organizational Support. The resulting scale is presented in an appendix to this article. Not surprisingly, the unifactorial Citizenship Performance scale correlated highly with the first factor ($r = .94$), but it also had internal reliability (Cronbach's alpha = .81) at a level that Ponterotto and Ruckdeschel (2007) argued was excellent for a six-item scale and a sample of less than 300, while maintaining good inter-rater reliability (.54).

While it is inappropriate to use one sample to develop a model using exploratory techniques and the same sample to conduct a CFA of that model (Gerbing & Hamilton, 1996), the current study had more than one rating of most participants. Alternative ratings were available for 197 of the

participants and these were used as the basis for a CFA. Along with the other CFAs reported in this article, this was performed using AMOS 6.0 (Arbuckle, 2005). In line with the recommendations of Hu and Bentler (1999), Chi-Square, SRMR and CFI were used to test models that are presented in this article. RMSEA is also reported because it has become a standard in social science research (Curran, Bollen, Chen, Paxton, & Kirby, 2003), but unfortunately, RMSEA tends to over-reject models in studies with samples of 250 or less (Hu & Bentler, 1999). The cutoff criteria recommended by Hu and Bentler (1999) were used to assess the fit of the factor model ($SRMR \leq .08$; $CFI \geq .95$; $RMSEA \leq .06$). The results of the CFA analysis of the unifactorial Citizenship Performance scale showed good fit of the factor structure ($\chi^2 = 15.6$; $df = 9$; $p = .076$; $SRMR = .039$; $CFI = .979$; $RMSEA = .061$). Although the value for RMSEA slightly exceeds the cutoff criterion, this is inconsistent with the non-significant Chi-Square, suggesting that the unsatisfactory RMSEA is due to the sample size. Hence, it was concluded that the factor model provided a satisfactory representation of the underlying structure of the unifactorial Citizenship Performance scale.

To test Hypothesis 1, the scores on the unifactorial Citizenship Performance scale were compared with the Five-Factor Model of personality. Table 2 presents the descriptive statistics for the five scales of the Mini-Markers and the unifactorial Citizenship Performance scale (with both sets of raters), as well as the inter-correlations between these measures. Corrected correlations are also reported, in line with Schmidt and Hunter's (1996) recommendations.

The two indexes of construct validity presented by Westen and Rosenthal (2003) were used to test the consistency of the observed correlations against those expected on the basis of previous research. The first of Westen and Rosenthal's indexes, the alerting index ($r_{\text{alerting-CV}}$), is simply the correlation between the Fisher-transformed predicted and observed correlations. This provides a rough indication of the degree to which observed correlations are consistent with expectations. Unfortunately, $r_{\text{alerting-CV}}$ does not allow a straightforward test of statistical significance, for which purpose Westen and Rosenthal presented the contrast index ($r_{\text{contrast-CV}}$). The weakness of the contrast index is that it is strongly affected by the size of the expected correlations, so it does not provide a valid estimate of the degree of consistency between expected and observed statistics. Consequently, Westen and Rosenthal recommended that both indexes should be reported and interpreted together.

When the corrected correlations between the FFM and Citizenship Performance measures reported in Table 2 were compared with the Hertz and Donovan (2000) correlations cited in the introduction, the contrast index was significant ($r_{\text{contrast-CV}} = .20$; $p < .001$) and the alerting index was substantial ($r_{\text{alerting-CV}} = .78$). Consideration of these two indexes demonstrate that the Citizenship Performance scale reported in this study had correlations with the FFM measures that were strongly consistent with expectations, thus confirming Hypothesis 1.

 Insert Table 2 around here

Discussion

The results of Study 1 led to two conclusions. Firstly, the three Citizenship Performance scales ostensibly measured by the Borman items were in fact acting as indicators of a common latent factor. Thus, the results of Study 1 were consistent with Hoffman et al.'s (2007) arguments. Secondly, the resultant unifactorial Citizenship Performance scale showed good internal consistency, factor structure and external validity, suggesting that it has the potential to be a valid, reliable measure of Citizenship Performance. Study 2 was designed to further test these conclusions, as well as providing a test of Hypotheses 2 and 3.

STUDY 2: CONFIRMATION OF STRUCTURE AND CONVERGENT VALIDATION

Study 2 was undertaken using a different sample and different measures for validation. Whereas the participants in Study 1 were students, in Study 2 participants were staff employed in a range of occupations. The unifactorial Citizenship Performance scale was compared with effort, job knowledge, compliance and quality of work, variables that have clearly described theoretical associations with Citizenship Performance. Previously developed measures of Citizenship Performance were also used in Study 2, in order to further test the validity of the unifactorial scale. Although Study 1 provided evidence that the unifactorial scale had sound internal and inter-rater

reliability and simple structure as assessed by CFA, there are several other criteria to be satisfied to establish this as a valid measure. It was argued in the introduction that Citizenship Performance directly reflects levels of effort while it has only an indirect relationship with job knowledge. It was also argued that compliance with rules is an important component of Citizenship Performance, unlike quality of work, which should be directly associated with Task Performance. Hence, the unifactorial Citizenship Performance scale should be more strongly correlated with measures of effort and compliance than with measures of job knowledge and quality of work.

The final test of validity that was used in Study 2 was a comparison with previously developed measures of Citizenship Performance. For this purpose the unifactorial scale was compared with two commonly used Citizenship Performance scales, Interpersonal Facilitation and Job Dedication (Van Scotter & Motowidlo, 1996). Including these scales not only provided a test of convergent validity; it also allowed a direct comparison of the reliability and validity of these pre-existing scales with the unifactorial scale.

Method

Participants

Non-academic supervisors and employees of an Australian university were invited to participate in Study 2. Academic staff were not included because of the somewhat unusual nature of academic work and the relatively low level of supervision they receive, making ratings of their performance behavior problematic. Initially, a sample of supervisors was approached and, if they were willing to participate, their staff were also asked if they were willing. A total of 43 out of 46 supervisors and 238 out of 283 staff agreed to participate in the study. Participants were employed in a broad range of occupations, ranging from professional accountants to semi-skilled cleaners and caterers. The average age of participants was 38.5 years and 22.7% were male, which is broadly similar to the average age of all employees of this university (41.5 years) and the proportion of staff who are male (28%).

Measures

In addition to the unifactorial scale, two additional Citizenship Performance scales were used in Study 2. These correspond to two of the earliest components of Citizenship to be proposed, namely Conscientiousness (not to be confused with the similarly-named personality variable) and Altruism (Organ, 1988; Smith et al., 1983). Borman, Penner et al. (2001) claimed that these “are the two major or overarching dimensions” (p.67) of Citizenship Performance. The Interpersonal Facilitation and Job Dedication scales developed by Van Scotter and Motowidlo (1996) are among the most widely used for measuring these components of Citizenship Performance (Podsakoff et al., 2000), so these were used in Study 2. The Interpersonal Facilitation and Job Dedication scales have seven and eight items respectively, and were rated on the same scale as that used for the unifactorial scale developed in Study 1. Consistent with other measures of Citizenship Performance (Podsakoff et al., 2000), Job Dedication and Interpersonal Facilitation have been compared with personality (Van Scotter & Motowidlo, 1996), rewards (Van Scotter, Motowidlo, & Cross, 2000), exchange fairness (Aryee, Chen, & Budhwar, 2004), supervisor liking (Johnson, Erez, Kiker, & Motowidlo, 2002), and emotional intelligence (Law, Wong, & Song, 2004), among other predictor and criterion measures. Van Scotter and Motowidlo (1996) reported sound reliability for these scales, both in terms of inter-rater reliability (Job Dedication: .50; Interpersonal Facilitation: .69) and internal consistency (Cronbach’s alpha: .94 for both scales). Yet, as mentioned in the introduction there has been some uncertainty about the validity of the Job Dedication construct as an aspect of Citizenship and it appears that neither of these scales has previously been assessed using CFA.

Four single-item scales were used to test for validity of the Citizenship Performance scales, with supervisors asked to rate their staff on effort, job knowledge, compliance and quality of work. The specific wording used as the basis for these ratings was derived from the set of performance measures developed by Viswesvaran, Ones and Schmidt (2002) and read as follows — effort: how much work they expend in striving to do a good job (initiative; attention to duty; alertness; enthusiasm; industriousness; persistence); job knowledge: the extent to which they demonstrate expertise in job-related issues (keep up-to-date; know the job well); compliance with authority: their conformity with formal rules and procedures (positive perspective on rules & regulations; positive

attitude to supervision; conforms to organizational norms & culture); quality of their work: how well their tasks are done (accuracy; lack of errors; thoroughness). These scales were assessed using the relative percentile method (Goffin, Gellatly, Paunonen, Jackson, & Meyer, 1996), which requires raters to compare individuals with all other staff employed within the organization. This method appears to have greater reliability and validity than alternatives, such as behavioral observation scales (Goffin, et al., 1996).

Procedure

Study 2 formed part of a larger research project examining predictors of performance. After obtaining organizational and ethical approval, supervisors were invited to participate. If they agreed to do so meetings were arranged with their staff, at which time potential participants were informed of the requirements of the overall project. Participants were assured that no-one apart from the research team, including their supervisors or fellow staff, would have access to ratings that were collected in this research. Participants were also told that they could request at any time to have any material relating to them destroyed. No participants asked for this to happen. Supervisors were subsequently provided with lists of staff who had consented to participate, on the understanding that participation by staff and supervisors was voluntary and no positive nor negative consequence would result from a decision to participate or not participate. Supervisors were then asked to rate only participating staff on the three Citizenship Performance scales and the ratings of effort and job knowledge, according to how they believed these staff typically behaved at work. All staff who could be contacted after the completion of the research project were provided with a summary of the basic findings of the research.

Results

Descriptive statistics, reliabilities and inter-correlations for Study 2 measures are presented in Table 3. Previous research was identified to obtain estimates of the correlation between Citizenship Performance and effort (Entwhistle III, 2001: $r = .47$), job knowledge (Bettencourt, Gwinner, &

Meuter, 2001: $r = .36$), compliance (Greenslade & Jimmieson, 2007: $r = .46$), and quality (Podsakoff, Ahearne, & MacKenzie, 1997: $r = .29$) to allow a similar test of validity to that conducted in Study 1. The alerting and contrast indexes for the unifactorial Citizenship Performance measure ($r_{\text{alerting-CV}} = .77$; $r_{\text{contrast-CV}} = .34$; $p < .001$) showed that it had a good correspondence with expectations, as did Interpersonal Facilitation ($r_{\text{alerting-CV}} = .70$; $r_{\text{contrast-CV}} = .17$; $p < .05$). However, Job Dedication had relatively poor correspondence with expected correlations ($r_{\text{alerting-CV}} = .08$; $r_{\text{contrast-CV}} = .02$; $p = ns$). So it appears from this analysis that the unifactorial Citizenship Performance scale has similar or better validity than the other two scales.

Westen and Rosenthal's (2003) indexes are global indicators of validity relationships, somewhat analogous to an overall analysis of variance test. For a more detailed examination of the convergent and divergent validity of these scales Olkin's z test was used to test the significance of the hypothesized differences between the various correlations (May & Hittner, 1997). An examination of the correlations between the unifactorial scale and the measures of effort, job knowledge, compliance and quality of work showed that they formed a pattern that was consistent with predictions. Specifically, the unifactorial Citizenship Performance scale had a significantly higher correlation with effort than with job knowledge, thus confirming Hypothesis 2, and also had a significantly higher correlation with compliance than with quality of work, which was consistent with Hypothesis 3. As pointed out by an anonymous reviewer, this may reflect the fact that on face-value item 2 in this scale seems related to compliance. The corresponding correlations for the other Citizenship Performance scales were consistent with Hypothesis 2 but the correlations with Job Dedication were not consistent with Hypothesis 3. Although both Interpersonal Facilitation and Job Dedication had higher correlations with effort than with job knowledge, and the correlation of Interpersonal Facilitation with compliance was significantly higher than its correlation with quality of work, the correlation of Job Dedication with quality of work was actually higher than its correlation with compliance. These findings are consistent with indications from previous research that Job Dedication may reflect aspects of Task Performance. Taken together, the pattern of correlations presented in Table 3 show that the unifactorial scale had better convergent and divergent validity than the other two Citizenship scales.

Insert Table 3 around here

The inter-correlations between the three Citizenship Performance scales show associations that are equal to or greater than the highest uncorrected correlation between Citizenship scales (.67) reported by LePine et al. (2002). Thus, these results clearly support Hypothesis 4 that a scale based on a single latent Citizenship factor would be highly correlated with pre-existing measures of Citizenship Performance. Study 2 also allowed an additional test of the structure of the unifactorial scale using CFA. It appears that neither of the other two Citizenship Performance scales has been similarly assessed in previously published research, so Interpersonal Facilitation and Job Dedication were analyzed alongside the unifactorial scale to allow for comparison. These analyses, (see Table 4), revealed that the unifactorial Citizenship Performance scale had very good fit, with a non-significant Chi-Square and values for the three fit indices that were well within the relevant cutoff criteria. Both of the other scales had significant Chi-Square values, which some authors would argue renders their factor models invalid (e.g., Hayduk, Cummings, Boadu, Pazderka-Robinson, & Boulianne, 2007). However, when the other fit indices were considered, only RMSEA was unsatisfactory for Interpersonal Facilitation. but Job Dedication had unsatisfactory fit as assessed by both CFI and RMSEA, indicating that it did not have simple structure. This may seem to be inconsistent with the high alpha value for this scale, but this incongruity is apparent only — alpha does not directly reflect the factorial integrity of a scale (Miller, 1995). Examination of modification indices for the Job Dedication model showed that four pairs of error terms needed to be allowed to correlate before an acceptable level of fit was achieved, suggesting that the scale may be measuring more than one construct. The results of these CFAs show that the unifactorial Citizenship Performance scale had the best fit to the data gathered in Study 2 suggesting that it had a more valid internal structure.

Insert Table 4 around here

Discussion

Study 2 provided further evidence for both the validity and the reliability of the unifactorial Citizenship Performance scale. The unifactorial scale was strongly correlated with the other two measures of Citizenship Performance and was also correlated with other aspects of performance in a manner consistent with theoretical expectations, further confirming the validity of the scale. When compared with the other Citizenship Performance scales used in Study 2, it also had the best psychometric qualities having demonstrated the best factor structure fit and the best convergent and divergent validity.

GENERAL DISCUSSION

This article described the development and validation of a unifactorial scale for measuring Citizenship Performance. This scale appears to be the first unifactorial Citizenship Performance scale to have been successfully tested with CFA using contemporary criteria. The relationship of the unifactorial Citizenship Performance scale with other variables was consistent with theoretical expectations as well as previous empirical findings, providing both convergent and divergent validation. Finally, although the unifactorial Citizenship Performance scale was shown to be highly correlated with previously developed measures of Citizenship Performance, it had a better internal structure and external validity than these other scales. Therefore, the unifactorial Citizenship Performance scale appears to be a valid measure of the latent Citizenship Performance construct.

One of the major justifications for the idea that Citizenship is a single latent construct is the fact that most correlations with other variables do not vary between Citizenship Performance measures (Hoffman et al., 2007; LePine et al., 2002). An examination of the correlations in Study 2 might be taken as evidence that the three Citizenship scales are measuring different aspects of Citizenship because of their significantly different correlations with other performance measures. However, the fact that these differences are in a direction that is inconsistent with expectations for measures of Citizenship Performance can also be taken as evidence of limitations of the Job Dedication and Interpersonal Facilitation scales. Specifically, the original authors of the scale (Van Scotter & Motowidlo, 1996) found that Job Dedication overlapped with Task Performance. This

measurement overlap is consistent with the unsatisfactory CFA fit of the Job Dedication scale as well as with its higher than expected correlation with job knowledge. On the other hand, both Interpersonal Facilitation and the unifactorial Citizenship Performance scale had correlations with job knowledge that were significantly lower than the correlation with effort, which was consistent with expectations. Thus, the results of this research, when considered alongside previous findings, are consistent with the idea that Citizenship Performance is unifactorial.

One of the consequences of viewing Citizenship as a unifactorial latent construct is that focusing on different scales for Citizenship Performance “would be like interpreting relationships with individual items from a multi-item measure of a unidimensional construct” (LePine et al., 2002, p.61). In other words, without evidence to the contrary, researchers would do better to measure Citizenship Performance as a single scale rather than attempting to measure sub-scales so the scale presented here should prove to be a useful tool for organizational researchers. Nonetheless, various Citizenship Performance measures have previously been compared with a wide range of measures, including employee attitudes, task characteristics, leadership behaviors and organizational rewards (Podsakoff et al., 2000). Future research comparing the scale presented in this article with these constructs will assist in assessing this scale’s validity.

From a broader perspective, the research reported here is consistent with previous efforts intended to bring order to the broad literature on Citizenship Performance. Since Organ (1997) proposed that OCB and Contextual Performance were largely similar, there have been several major reviews that have helped to integrate the various Citizenship Performance literatures. Both Podsakoff et al. (2000) and Coleman and Borman (2000) provided integrative reviews, albeit from different perspectives and using disparate methodologies. Specifically, Podsakoff et al. (2000) approached their review from the OCB tradition and focused mainly on reviewing measurement methodologies and meta-analyzing correlations between Citizenship Performance and other variables. In contrast, Coleman and Borman (2000), the latter of whom was one of the initial proponents of the Contextual Performance construct, provided a content analysis of Citizenship Performance. LePine et al. (2002) and Hoffman et al. (2007) used multivariate modeling to advance upon the work of Podsakoff et al. (2000), leading to the recognition that for most purposes Citizenship Performance should be treated as

unifactorial. The research reported in this article also used multivariate modeling, but in this case to further the work of Coleman and Borman. The analyses of LePine et al. and Hoffman et al. and the research presented here all examined the structure of Citizenship Performance measures. However, the earlier research examined the relationship of Citizenship Performance with attitudinal variables, Task Performance and justice perceptions, while the two studies in this article considered links with personality and various independently measured precursors and components of performance. The fact that each of these articles reached similar conclusions about the structure of Citizenship Performance despite these differences should provide considerable confidence in their respective findings. It also confirms the wisdom of Organ's (1997) call for 'construct clean-up' for OCB or Citizenship Performance, however labeled.

REFERENCE LIST

- Allen, T. D., & Rush, M. C. (1998). The effects of organizational citizenship behavior on performance judgments: A field study and a laboratory experiment. *Journal of Applied Psychology, 83*(2), 247-260.
- Arbuckle, J. L. (2005). AMOS 6.0.0. Spring House, PA: Amos Development Corporation.
- Aryee, S., Chen, Z. X., & Budhwar, P. S. (2004). Exchange fairness and employee performance: An examination of the relationship between organizational politics and procedural justice. *Organizational Behavior and Human Decision Processes, 94*(1), 1-14.
- Barksdale, K., & Werner, J. M. (2001). Managerial ratings of in-role behaviors, organizational citizenship behaviors and overall performance: Testing different models of their relationship. *Journal of Business Research, 51*(2), 145-155.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57*, 289-300.
- Bettencourt, L. A., Gwinner, K. P., & Meuter, M. L. (2001). A comparison of attitude, personality, and knowledge predictors of service-oriented organizational citizenship behaviors. *Journal of Applied Psychology, 86*(1), 29-41.

- Blumberg, M., & Pringle, C. D. (1982). The missing opportunity in organizational research: Some implications for a theory of work performance. *Academy of Management Review*, 7(4), 560-569.
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71-98). New York: Jossey-Bass.
- Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9(1-2), 52-69.
- Bowles, S., Gintis, H., & Meyer, P. (1999). The long shadow of work: Education, the family, and the reproduction of the social division of labor. *Critical Sociology*, 25(2/3), 286-305.
- Brief, A. P., & Motowidlo, S. J. (1986). Prosocial organizational behaviors. *Academy of Management Review*, 11, 710-725.
- Coleman, V. I., & Borman, W. C. (2000). Investigating the underlying structure of the citizenship performance domain. *Human Resource Management Review*, 10, 25-44.
- Conway, J. M. (1999). Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology*, 84(1), 3-13.
- Cortina, J. M. (2002). Big things have small beginnings: An assortment of "minor" methodological misunderstandings. *Journal of Management*, 28(3), 339-362.
- Curran, P. J., Bollen, K., Chen, F., Paxton, P., & Kirby, J. B. (2003). Finite sampling properties of the point estimates and confidence intervals of the RMSEA. *Sociological Methods and Research*, 32(2), 208-252.
- Dwight, S. A., Cummings, K. M., & Glenar, J. L. (1998). Comparison of criterion-related validity coefficients for the Mini-Markers and Goldberg's Markers of the Big Five Personality Factors. *Journal of Personality Assessment*, 70, 541-550.
- Entwhistle III, G. H. (2001). *Measuring effort expended in the workplace: Discretionary effort and its relationship to established organizational commitment and attachment dimensions*. Unpublished Doctoral Dissertation, Boston University, Boston.

- Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling, 3*, 62-72.
- Goffin, R. D., Gellatly, I. R., Paunonen, S. V., Jackson, D. N., & Meyer, J. P. (1996). Criterion validation of two approaches to performance appraisal: The behavioral observation scale and the relative percentile method. *Journal of Business and Psychology, 11*(1), 23-33.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26-42.
- Gosling, S. D., Rentfrow, P. J., & Swann Jr, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*, 504-528.
- Greenslade, J. H., & Jimmieson, N. L. (2007). Distinguishing between task and contextual performance for nurses: Development of a job performance scale. *Journal of Advanced Nursing, 58*(6), 602-611.
- Hayduk, L. A., Cummings, G., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! testing! one, two, three - Testing the theory in structural equation models! *Personality and Individual Differences, 42*(5), 841-850.
- Hoffman, B. J., Blair, C. A., Meriac, J. P., & Woehr, D. J. (2007). Expanding the criterion domain? A quantitative review of the OCB Literature. *Journal of Applied Psychology, 92*(2), 555-566.
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology, 85*(6), 869-879.
- Johnson, D. E., Erez, A., Kiker, D. S., & Motowidlo, S. J. (2002). Liking and attributions of motives as mediators of the relationships between individuals' reputations, helpful behaviors, and raters' reward decisions. *Journal of Applied Psychology, 87*(4), 808-815.
- Koys, D. J. (2001). The effects of employee satisfaction, organizational citizenship behavior, and turnover on organizational effectiveness: A unit-level, longitudinal study. *Personnel Psychology, 54*(1), 101-114.

- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The source of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods, 9*(2), 202-220.
- Law, K. S., Wong, C.-S., & Song, L. J. (2004). The construct and criterion validity of emotional intelligence and its potential utility for management studies. *Journal of Applied Psychology, 89*(3), 483-496.
- LePine, J. A., Erez, A., & Johnson, D. E. (2002). The nature and dimensionality of Organizational Citizenship Behavior: A critical review and meta-analysis. *Journal of Applied Psychology, 87*(1), 52-65.
- LePine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: Evidence of differential relationships with Big Five personality characteristics and cognitive ability. *Journal of Applied Psychology, 86*(2), 326-336.
- Lievens, F., & Anseel, F. (2004). Confirmatory factor analysis and invariance of an organizational citizenship behaviour measure across samples in a Dutch-speaking context. *Journal of Occupational and Organizational Psychology, 77*, 299-306.
- Lounsbury, J. W., Gibson, L. W., Sundstrom, E., Wilburn, D., & Loveland, J. M. (2004). An empirical investigation of the proposition that 'School is Work': A comparison of personality-performance correlations in school and work settings. *Journal of Education and Work, 17*(1), 119-131.
- MacKenzie, S. B., Podsakoff, P. M., & Ahearne, M. (1998). Some possible antecedents and consequences of in-role and extra-role salesperson performance. *Journal of Marketing, 62*, 87-98.
- MacKenzie, S. B., Podsakoff, P. M., & Fetter, R. (1991). Organizational citizenship behavior and objective productivity as determinants of managerial evaluations of salespersons' performance. *Organizational Behavior and Human Decision Processes, 50*(1), 123-150.
- MacKenzie, S. B., Podsakoff, P. M., & Fetter, R. (1993). The impact of organizational citizenship behavior on evaluations of salesperson performance. *Journal of Marketing, 57*(1), 70-80.
- Marshall, H. H. (1988). Work of learning: Implications of classroom metaphors. *Educational Researcher, 17*(9), 9-16.

- May, K., & Hittner, J. B. (1997). Tests for comparing dependent correlations revisited: A Monte Carlo study. *The Journal of Experimental Education*, 65(3), 257-269.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- Miles, D., Borman, W. C., Spector, P. E., & Fox, S. (2002). Building an integrative model of extra role work behaviors: A comparison of counterproductive work behavior and organizational citizenship behavior. *International Journal of Selection and Assessment*, 10(1/2), 51-57.
- Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, 2(3), 255-273.
- Motowidlo, S. J. (2003). Job performance. In W. C. Borman, D. R. Ilgen & R. J. Klimoski (Eds.), *Handbook of Psychology: Industrial and organizational psychology* (Vol. 12, pp. 39-53). New York: John Wiley & Sons.
- Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, 10(2), 71-83.
- Munson, H. L., & Rubenstein, B. J. (1992). School IS work: work task learning in the classroom. *Journal of Career Development*, 18(4), 289-297.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402.
- Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA, US: Lexington Books/D C Heath and Company.
- Organ, D. W. (1997). Organizational citizenship behavior: It's construct clean-up time. *Human Performance*, 10(2), 85-97.
- Organ, D. W., & Paine, J. B. (1999). A new kind of performance for industrial and organizational psychology: Recent contributions to the study of organizational citizenship behavior. In C. L. Cooper (Ed.), *International review of industrial and organizational psychology* (Vol. 14, pp. 337-368). Chichester, England: John Wiley & Sons Ltd.

- Podsakoff, P. M., Ahearne, M., & MacKenzie, S. B. (1997). Organizational citizenship behavior and the quantity and quality of work group performance. *Journal of Applied Psychology, 82*(2), 262-270.
- Podsakoff, P. M., & MacKenzie, S. B. (1994). Organizational citizenship behaviors and sales unit effectiveness. *Journal of Marketing Research, 31*(3), 351-363.
- Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *Leadership Quarterly, 1*(2), 107-142.
- Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., & Bachrach, D. G. (2000). Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research. *Journal of Management, 26*(3), 513-563.
- Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills, 105*(3), 997-1014.
- Reise, S. P., Waller, N., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment, 12*(3), 287-297.
- Roberts, B. W., Bogg, T., Walton, K. E., Chernyshenko, O. S., & Stark, S. E. (2004). A lexical investigation of the lower-order structure of conscientiousness. *Journal of Research in Personality, 38*, 164-178.
- Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*, 506-516.
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1*, 199-223.
- Schnake, M., & Dumler, M. P. (2003). Levels of measurement and analysis issues in organizational citizenship behaviour research. *Journal of Occupational and Organizational Psychology, 76*(3), 283-301.
- Shevlin, M., Miles, J. N. V., Davies, M. N. O., & Walker, S. (2000). Coefficient alpha: A useful indicator of reliability? *Personality and Individual Differences, 28*, 229-237.

- Smith, C. A., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology, 68*(4), 653-663.
- Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*, 401-426.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. (4th ed.). Needham Heights, MA.: Allyn & Bacon.
- Van Scotter, J. R., & Motowidlo, S. J. (1996). Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology, 81*(5), 525-531.
- Van Scotter, J. R., Motowidlo, S. J., & Cross, T. C. (2000). Effects of task performance and contextual performance on systemic rewards. *Journal of Applied Psychology, 85*(4), 526-535.
- Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2002). The moderating influence of job performance dimensions on convergence of supervisory and peer ratings of job performance: Unconfounding construct-level convergence and rating difficulty. *Journal of Applied Psychology, 87*(2), 345-354.
- Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology, 84*(5), 608-618.
- Williams, L. J., & Anderson, S. E. (1991). Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management, 17*(3), 601-617.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806-838.

Appendix

*Unifactorial Citizenship Performance Scale*¹

1. Cooperates fully with others by willingly sacrificing own personal interests for the good of the team.
2. Knows and follows both the letter and the spirit of organizational rules and procedures, even when the rules seem personally inconvenient.
3. Consistently takes the initiative to pitch in and do anything that might be necessary to help accomplish team or organizational objectives, even if such actions are not normally part of own duties.
4. Avoids performing any tasks that are not normally a part of own duties by arguing that they are somebody else's responsibility (Reverse scored).
5. Goes out of his or her way to congratulate others for their achievements.
6. Looks for opportunities to learn new knowledge and skills from others at work and from new and challenging job assignments.

¹ The items in this scale were provided to the authors by Walter Borman, and the authors wish to express their gratitude to him for allowing these items to be published here.

Table 1

Descriptive Statistics, Reliabilities and Correlations of the Borman Citizenship Performance Scales.^a

		Mean	SD	Intra-class		Correlations ^{b*}	
				correlation		1	2
1	Personal Support	4.98	2.28	.45	.55	1.11 [†]	1.17 [†]
2	Organizational Support	3.81	2.28	.29	.57	.48	.95
3	Conscientious Initiative	4.65	2.74	.56	.71	.54	.67

^a N = 298 for all statistics, except for the inter-rater reliabilities for which N = 197.

^b Values for Alpha are reported in italics on the diagonal. Uncorrected correlations are reported below the diagonal. Correlations corrected for scale reliability (Cronbach's Alpha) are reported above the diagonal.

* All correlations are significant at $p < .001$.

[†] The corrected correlations that are greater than 1.0 appear to be the result of a combination of rounding error and random measurement error.

Table 2.

Descriptive Statistics and Correlations for the unifactorial Citizenship Performance scale with Mini-Markers.^{abc}

	Mean	SD	1	2	3	4	5	6	7
1. Extraversion	6.00	12.58	<i>.81</i>	0.15	0.17	0.21	0.16	.12	.12
2. Agreeableness	14.83	8.96	.12	<i>0.8</i>	0.39	0.39	0.19	.05	-.06
3. Conscientiousness	11.73	9.36	.14*	<i>.31***</i>	<i>0.81</i>	0.36	0.14	.27	.28
4. Emotional Stability	-18.67	10.30	.16*	<i>.30***</i>	<i>.28***</i>	<i>0.74</i>	-0.05	.10	-.08
5. Openness	30.49	9.15	.13*	.15*	0.11	-0.04	<i>0.78</i>	-.09	-.16
6. Citizenship Performance (1 st rater)	18.16	3.95	.10	.04	<i>.22**</i>	.08	-.07	<i>.81</i>	.67
7. Citizenship Performance (2 nd rater)	18.13	3.68	.10	-.05	<i>.23**</i>	-.06	-.13	<i>.54***</i>	<i>.81</i>

^a Correlations between Mini-Markers & Citizenship Performance (1st rater): $n = 266$; Correlations between Mini-Markers & Citizenship Performance (2nd rater): $n = 176$; Correlation between Citizenship Performance (1st rater) & Citizenship Performance (2nd rater): $n = 197$.

^b Benjamini & Hochberg (1995) False Discovery procedure used to adjust significance levels for multiple comparisons.

^c Values for Cronbach's Alpha are reported in italics on the diagonal. Uncorrected correlations are reported below the diagonal. Correlations corrected for scale reliability (Cronbach's Alpha) are reported above the diagonal.

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 3.

Descriptive Statistics, Reliabilities and Correlations for Study Two.^a

		Correlations ^{bc*}								
		Mean	SD	1	2	3	4	5	6	7
1	Unifactorial Citizenship Performance	22.63	4.42	<i>.84</i>	<i>.85</i>	<i>.87</i>	<i>.74</i>	<i>.34</i>	<i>.65</i>	<i>.50</i>
2	Interpersonal Facilitation	26.15	4.88	<i>.75</i>	<i>.87</i>	<i>.76</i>	<i>.65</i>	<i>.44</i>	<i>.57</i>	<i>.53</i>
3	Job Dedication	30.34	5.23	<i>.76</i>	<i>.67</i>	<i>.90</i>	<i>.72</i>	<i>.53</i>	<i>.48</i>	<i>.52</i>
4	Effort	75.28	16.10	<i>.68_a</i>	<i>.62_a</i>	<i>.67_a</i>	<i>na</i>	<i>.32</i>	<i>.56</i>	<i>.62</i>
5	Job Knowledge	75.18	15.47	<i>.31_b</i>	<i>.42_{bc}</i>	<i>.49_{bc}</i>	<i>.32</i>	<i>na</i>	<i>.32</i>	<i>.62</i>
6	Compliance	72.20	18.91	<i>.60_a</i>	<i>.53_{abc}</i>	<i>.46_b</i>	<i>.56</i>	<i>.32</i>	<i>na</i>	<i>.49</i>
7	Quality of Work	74.80	14.94	<i>.46_c</i>	<i>.49_{bc}</i>	<i>.59_{ac}</i>	<i>.62</i>	<i>.62</i>	<i>.49</i>	<i>na</i>

^a $n = 179$.

^b Where available, Cronbach's Alpha is reported in italics on the diagonal. Uncorrected correlations are reported below the diagonal. Correlations corrected for scale reliability (Cronbach's Alpha) are reported above the diagonal.

^c Correlations in the same column that do not share the same subscript are significantly different at $p < .05$.

* All correlations are significant at $p < .001$.

Table 4

Confirmatory Factor Analysis of Measures used in Study 2.^a

Measures	Items	χ^2	df	p	SRMR	CFI	RMSEA
Unifactorial Citizenship Performance	6	9.5	9	.392	.026	.999	.018
Interpersonal Facilitation	7	32.0	14	.004	.037	.974	.085
Job Dedication	8	101.0	20	.000	.065	.876	.151

^a $n = 179$.

Note. p = probability for χ^2 ; SRMR = Standardized Root Mean Residual; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.