Organizational citizenship behavior in Chinese society: a reexamination

Hao Zhaoa*, Jihong Wub, Jian-Min Sun and Charles Weizheng Chend

aLally School of Management and Technology, Rensselaer Polytechnic Institute, Troy, NY, USA; bSchool of Management and Economics, University of Electronic Science and Technology of China, Chengdu, P.R. China; cDepartment of Psychology and School of Labor and Human Resources, Renmin University, Beijing, P.R. China; dBusiness School, Sichuan University, Chengdu, P.R. China

Organizations in emerging economies such as China have attracted an increasing amount of research attention during the past few decades, and it is time to scrutinize findings from earlier studies. In this paper, we test the psychometric soundness of a widely used indigenous measure of organizational citizenship behavior in the Chinese context developed by Farh, Earley, and Lin ((1997), 'Impetus for Action: A Cultural Analysis of Justice and Organizational Citizenship Behavior in Chinese Society,' Administrative Science Quarterly, 42, 421–444). Using multiple samples from Mainland China, we performed systematic tests of the content validity, discriminant validity, factorial validity and nomological validity of this five-dimension measure. We found that their two culture-specific dimensions, namely (harm to interpersonal) harmony and (personal use of organizational) resources, in fact measure the construct of deviant behavior, and cannot prove the existence of indigenous contents. We discuss our results in light of the ongoing convergence–divergence debate, and provide suggestions to further improve the quality of indigenous research.

Keywords: China; convergence; deviant behavior; divergence; organizational citizenship behavior

Introduction

There has been a convergence–divergence debate regarding the role of national culture on employee values and behaviors (Dunphy 1987; Brislin 1993; Rotundo and Xie 2008). Proponents of convergence perspective emphasize the similarities across societies, believing cultural influence is secondary to the impacts of industrialization and globalization on employee values and behaviors. Modern societies share common characteristics that are not dictated by national cultures when they strive to meet economic, technological and administrative demands of industrialization. This view was predominant before 1970s, but has since then faced growing doubts when researchers became interested in the management of Japanese (Dunphy 1987) and Chinese enterprises (Ralston, Holt, Terpstra and Yu 1997). Divergence theorists highlight the cultural differences among nations. They believe that work values and behaviors are deeply rooted in a nation’s culture despite industrialization, and management practices must be adjusted to the local cultural context. People in different cultures may even define organizational constructs differently, so the first step of indigenous research is to explore the construct domain and develop indigenous instruments, if necessary, that are meaningful in the local

*Corresponding author. Email: zhaoh@rpi.edu

ISSN 0958-5192 print/ISSN 1466-4399 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/09585192.2012.669786
http://www.tandfonline.com
context (Tsui 2007). Ralston et al. (1997) proposed a third, ‘in-between’ perspective (i.e.,
crossvergence), which is ‘a continuum between the polar extremes of convergence and
divergence’ (p. 183), but a researcher’s relative position on this continuum still matters.

The divergence perspective is also challenged from time to time, because empirical
evidence is generally sparse (Tsui 2007) and sometimes counterintuitive. For example,
Chinese are often characterized by its collectivistic culture and Americans by its
individualistic culture, thus Chen (1995) found it a surprise that Chinese employees
preferred differential reward allocation rules that would enlarge in-group inequity while
Americans preferred egalitarian rules. He speculated that while China is achieving
industrialization, Chinese enterprises’ strong economic orientations are more powerful in
shaping employee preferences than traditional cultural values. Similarly, Rotundo and Xie
(2008) reported that Chinese managers tie with or score higher than North Americans in
individualism, and they argued that early models of national culture may not have captured
the major social, political and economic changes that define contemporary China. They
also found that Chinese managers and North American managers define the construct
domain of deviant behavior similarly.

In the area of organizational citizenship behavior (OCB), however, Farh et al. (1997)
developed a 20-item indigenous measure of OCB for Chinese society, claiming that Chinese
employees define the construct domain of OCB differently from western employees. They
attributed the newly identified contents to the collectivistic Chinese culture. This was
groundbreaking research upon publication due to its articulation of culture-specific contents.
It immediately became one of the most frequently cited articles on management and
organizations in Greater China (Li and Tsui 2002), and is often acknowledged as an
exemplar indigenous study in narrative reviews (e.g., Tsui 2004; Zhang 2010). An increasing
number of empirical studies have used this indigenous instrument to measure OCB and study
its relationship to other important organizational constructs in Greater China, including
Hong Kong (e.g., Leung 2008), Mainland China (e.g., Chen, Aryee and Lee 2005; Farh,
Hackett and Liang 2007) and Taiwan (e.g., Yen and Niehoff 2004; Lin and Peng 2010).

Despite the popularity and influence of the OCB measure developed by Farh et al.
(1997), to our knowledge, virtually no subsequent study has been conducted to
systematically evaluate the psychometric properties of this important measure, which
becomes a concern given the less-than-optimal evidence from the original study. Is there
really indigenous content of OCB in Greater China? Given the growing interest in
developing indigenous measures in emerging economies, what procedures can help
minimize the contamination from research artifacts and enhance the validity of the
findings? Such concerns prompted us to conduct the current research to systematically
evaluate the psychometric soundness of this measure, and reflect how we can enhance the
quality of indigenous research in general.

OCB in Chinese context
OCB research originated and developed in western societies (e.g., Katz 1964; Smith, Organ
and Near 1983; Organ 1988; Motowidlo, Borman and Schmit 1997). Scholars use OCB to
describe discretionary employee behaviors that can benefit the work organization and its
members, such as helping colleagues, working late and protecting organizational reputation.

Does OCB have different meanings and internal structure in the Chinese culture? Farh
et al. (1997) took an inductive approach to generate, screen and categorize OCB items in
Taiwan. They identified five dimensions of Chinese OCB: (1) identification with the
company, such as promoting and protecting company reputation, and making suggestions
for improvement; (2) altruism toward colleagues, such as helping others in need in the organization; (3) conscientiousness, such as going beyond the minimum role requirement in one’s work; (4) interpersonal harmony, which is negatively worded and includes examples such as speaking ill of the supervisor or colleagues behind their back; and (5) protecting company resources, which is also negatively worded and includes examples such as abusing company resources. Farh et al. (1997) compared the five dimensions with western OCB measures and concluded that the last two dimensions (i.e., harmony and resources) were unique in Chinese society, and called them ‘emic’ dimensions. The other three dimensions that are also identified in western OCB research are called ‘etic’ dimensions. Farh et al. (1997) attributed the emic dimensions to a Chinese cultural root: familistic collectivism.

Later, Farh, Zhong and Organ (2004) examined OCBs in Mainland China. Instead of building on the original model of Farh et al. (1997), they repeated the inductive process to have subjects generate a large pool of OCB items. The authors and some doctoral students then subjectively categorized the items into 27 categories, which were then grouped through another wave of subjective evaluations into 10 broader and conceptually meaningful OCB dimensions, with a total of 595 items. They did not list their items or report essential psychometric properties of the dimensions, nor did they go further to create a more parsimonious and practically meaningful measure from this study. It is thus not a surprise that researchers, including the authors (e.g., Farh et al. 2007), frequently adopt the original Farh et al. (1997) OCB measure in studies of Mainland China.

Although the major contribution of Farh et al. (1997) was the identification of the emic contents of OCB, we find it surprising that many subsequent studies using this measure to study Chinese employees’ OCB chose to exclude the two emic dimensions (e.g., Chen and Francesco 2003; Cheng, Jiang and Riley 2003; Chen et al. 2005; Ding and Lin 2007; Farh et al. 2007; Lin and Peng 2010). Do such choices imply a weakness of this measure? A close look at the original study of Farh et al. (1997) reveals less-than-optimal psychometric properties of this five-dimension measure. Specifically, there was an unusual pattern of inter-correlations among the five dimensions: the three etic dimensions had relatively strong inter-correlations with each other ($r = 0.50$ to $0.57$) and the two emic dimensions were strongly correlated ($r = 0.58$) with each other, but the correlations across etic and emic dimensions were rather weak. For instance, the harmony and the resources dimension had correlation coefficients of only 0.19 and 0.23, respectively, with the identification dimension. Yen and Niehoff (2004), which also used samples from Taiwan, found very similar results: the correlations across emic and etic dimensions seemed much smaller in magnitude than inter-correlations within emic or etic dimensions.

Are the unusual correlations caused by the negative directions of the emic items? Farh et al. (1997) reasoned that the negative orientation of their emic items was not a research artifact, because otherwise these two dimensions would have loaded on a common factor reflecting the ‘negative’ orientation in their exploratory factor analysis (EFA). However, EFA is exploratory in nature, and the emergence of five separate first-order factors does not necessarily mean they represent the same construct. The negatively worded items may unintentionally tap into such a different domain of content that they should not have been combined with other items into a single higher-order factor. Unfortunately, Farh et al. (1997) did not address this plausible cause. Farh et al. performed a confirmatory factor analysis (CFA) of their five-factor model using a different sample and acknowledged that fit indices were ‘not optimal’ (1997, p. 427; e.g., GFI = 0.87), but they did not perform a higher-order CFA to investigate whether the five dimensions represent two distinct higher-order factors.
To our knowledge, few subsequent studies have explicitly tested the proposed dimensionality of the Farh et al. (1997) OCB measure. Research by Hui, Law and Chen (1999) is an exception. They shortened the 20-item scale to 15 items to use it in Mainland China, and performed CFAs to conclude that the five-factor model is better than a null model and a single-factor model, yet the fit indices of the five-factor model remained modest. They did not further test whether these five factors represent a single higher-order factor.

**Negative OCB or deviant behavior?**

There are theoretical grounds for us to expect that the two emic OCB dimensions provided by Farh et al. (1997) may, in fact, not measure OCB, but rather a different construct: deviant behavior. There is an increasing level of agreement among researchers that there are three broad individual performance domains (e.g., Viswesvaran and Ones 2000; Rotundo and Sackett 2002; Dalal 2005; Dalal, Lam, Weiss, Welch and Hulin 2009): task performance, OCB and deviant behavior. Task performance reflects the effectiveness with which job incumbents perform activities that contribute to the organization’s technical core, either directly by implementing a part of its technological process or indirectly by providing it with needed materials or services (Borman and Motowidlo 1993). OCB can be defined as intentional but discretionary employee behavior that improves the functioning of the organization (Organ 1988; Farh et al. 1997). Deviant behavior, or sometimes referred to as counterproductive work behavior, is defined as intentional employee behavior that harms the legitimate interests of an organization (Gruys and Sackett 2003). While the distinction between task performance and OCB has been largely documented and accepted (e.g., Bateman and Organ 1983; Borman and Motowidlo 1993, 1997), the relationship between deviant behavior and OCB remains relatively elusive, and the conceptual differences are so vague in most prior research that OCB measures often overlap with deviant behavior measures, an artifact that may bias the observed correlations between these constructs in empirical research (Dalal 2005; Spector, Bauer and Fox 2010).

While Farh et al.’s (1997) labels of the two emic dimensions, namely ‘resources’ and ‘harmony’, are in positive directions like the other three dimensions, all items on these two dimensions were worded negatively, which was inconsistent with the wording for items of the other three dimensions. These negatively worded behavior items appear to read very similar to typical examples of deviant behaviors that explicitly harm the organization’s welfare. For instance, one item of Farh et al.’s (1997) resources dimension reads ‘[This employee] conducts personal business on company time (e.g., trading stocks, shopping, going to barber shops)’, while an item from Bennett and Robinson’s (2000) widely used deviant behavior measure reads ‘[This employee] takes an additional or longer break than is acceptable at his/her workplace’ and an item from Rotundo and Xie’s (2008) deviant behavior measure developed in China reads ‘Use work time to do things for self’. Farh et al. (1997) referred to Chinese language to defend the inclusion of such items as negative OCB: ‘it is common in the Chinese language for positive attributes to be expressed using a negation of a negative’ (p. 430). They argued that if their scores are reversed, those items can be combined with the etic items to represent positive OCBS.

Dalal (2005) explicitly challenged such rationale by arguing that ‘negation (the addition of *not*) does not necessarily imply the lexical opposite’. In reality, what is ‘not beautiful’ is not necessarily ‘ugly’, and ‘not exceeding job requirements’ does not necessarily mean ‘failing to meet job requirements’. We concur with this observation, and following this logic, we feel employees who do not demonstrate discretionary behavior to help the organization (low OCB) do not necessarily intentionally harm the organization, and
employees who do not intentionally harm the organization (low deviant behaviors) may not
go beyond the job requirements to benefit the organization. Spector and Fox (2010)
proposed an emotion-centered model to explain how some employees simultaneously or
sequentially engage in both OCBs and deviant behaviors. In either case, Spector and Fox
argued that OCBs and deviant behaviors are not the opposite form of each other, and the
correlations between them are expected to be moderate in magnitude at best. Thus, it is
worthwhile to scrutinize the psychometric properties of these negatively worded items of
Farh et al.’s (1997) OCB measure to see whether they accidentally tap into the construct
domain of deviant behavior. Given the indigenous nature of this measure, it is of course best
to test such possibilities in China.

In the current study, we performed such tests using multiple samples from different
locations in Mainland China. First, we examined the content validity of the measure
through subject matter experts’ opinions. Second, we conducted an EFA to see if the emic
OCB items can be differentiated from deviant behavior in the applied setting (discriminant
validity). Then, we tested the factorial validity of the Farh et al.’s (1997) OCB measure
using a second-order CFA. Lastly, we tested the nomological validity of this measure by
investigating the correlation patterns of the five dimensions to two theoretically relevant
antecedent constructs: affective organizational commitment and leader–member
exchange (LMX).

Study 1: content validity

Following an approach described by Schriesheim, Powers, Scandura, Gardiner and
Lankau (1993), we used subject matter experts’ judgment to evaluate the content of the
20 items of Farh et al. (1997). It can provide critical insight into whether the items
effectively sample the exact content that they are intended to capture without measuring
another domain. For comparison purposes, we also included 19 items from the deviant
behavior measure developed by Bennett and Robinson (2000).

We sent invitation e-mails to 108 doctoral students majoring in management and
organization studies in two research universities in the Sichuan province. Doctoral students
are appropriate for this purpose because they have relevant research knowledge and
experience, and because past research has used doctoral students for similar purposes
(e.g., Tinsley 1998; Cable and Yu 2006). Participation was voluntary and we offered
10 RMB (approximately 1.5 USD) for their time. Twenty-four students agreed to
participate, representing a response rate of 22%.

These experts were asked to sort the 39 items as either OCB or deviant behavior, or
leave a blank if neither was viewed as a fit. We explained that it was not a test, and that
they only needed to refer to the definitions we provided, without checking the literature for
the original instrument structures. We presented them the broad definition of OCB from
Organ (1988), which was also presented by Farh et al. (1997, 2004) to their subjects for
item generation purposes. We also presented the definition of deviant behavior from
Bennett and Robinson (2000). For half of the experts we shuffled the items in a random
order to avoid potential memory bias from rating a block of items. We followed the
standardized translation and back-translation procedure (Brislin 1970) to ensure that the
translations were faithful to the original meanings.

Results of the content validity analysis are shown in Table 1. The majority of these
experts (83–100%) believed the 7 items of the harmony and resources dimensions
measure deviant behavior, while all items of the other three dimensions were ‘correctly’
categorized as OCBs (88–100%). The agreement indices were above the threshold value
Table 1. Results of content validity and discriminant validity (Studies 1 and 2).

<table>
<thead>
<tr>
<th>Item</th>
<th>Study 1: experts</th>
<th>Study 2: EFA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OCB (%)</td>
<td>DB (%)</td>
</tr>
<tr>
<td>Identification 1</td>
<td>Willing to stand up to protect the reputation of the company.</td>
<td>100</td>
</tr>
<tr>
<td>Identification 2</td>
<td>Eager to tell outsiders good news about the company and clarify their misunderstandings.</td>
<td>100</td>
</tr>
<tr>
<td>Identification 3</td>
<td>Makes constructive suggestions that can improve the operation of the company.</td>
<td>100</td>
</tr>
<tr>
<td>Identification 4</td>
<td>Actively attends company meetings.</td>
<td>92</td>
</tr>
<tr>
<td>Altruism 1</td>
<td>Willing to assist new colleagues to adjust to the work environment.</td>
<td>96</td>
</tr>
<tr>
<td>Altruism 2</td>
<td>Willing to help colleagues solve work-related problems.</td>
<td>100</td>
</tr>
<tr>
<td>Altruism 3</td>
<td>Willing to cover work assignments for colleagues when needed.</td>
<td>96</td>
</tr>
<tr>
<td>Altruism 4</td>
<td>Willing to coordinate and communicate with colleagues.</td>
<td>96</td>
</tr>
<tr>
<td>Conscientiousness 1</td>
<td>Complies with company rules and procedures even when nobody watches and no evidence can be traced.</td>
<td>96</td>
</tr>
<tr>
<td>Conscientiousness 2</td>
<td>Takes one’s job seriously and rarely makes mistakes.</td>
<td>92</td>
</tr>
<tr>
<td>Conscientiousness 3</td>
<td>Tries hard to self-study to increase the quality of work outputs.</td>
<td>88</td>
</tr>
<tr>
<td>Conscientiousness 4</td>
<td>Does not mind taking on new or challenging assignments.</td>
<td>100</td>
</tr>
<tr>
<td>Conscientiousness 5</td>
<td>Often arrives early and starts to work immediately.</td>
<td>92</td>
</tr>
<tr>
<td>Harmony 1</td>
<td>Uses illicit tactics to seek personal influence and gain with harmful effect on interpersonal harmony in the organization.</td>
<td>94</td>
</tr>
<tr>
<td>Harmony 2</td>
<td>Uses position power to pursue selfish personal gain.</td>
<td>92</td>
</tr>
<tr>
<td>Harmony 3</td>
<td>Takes credits, avoids blames, and fights fiercely for personal gain.</td>
<td>92</td>
</tr>
<tr>
<td>Harmony 4</td>
<td>Often speaks ill of the supervisor or colleagues behind their backs.</td>
<td>100</td>
</tr>
<tr>
<td>Resources 1</td>
<td>Conducts personal business on company time (e.g., trading stocks, shopping, going to barber shops).</td>
<td>100</td>
</tr>
<tr>
<td>Resources 2</td>
<td>Uses company resources to do personal business (e.g., company phone, copy machines, computers, and cars).</td>
<td>100</td>
</tr>
<tr>
<td>Resources 3</td>
<td>Views sick leave as benefit and makes excuses for taking sick leave.</td>
<td>100</td>
</tr>
<tr>
<td>DB 1</td>
<td>Makes fun of someone at work.</td>
<td>96</td>
</tr>
<tr>
<td>DB 2</td>
<td>Says something hurtful to someone at work.</td>
<td>96</td>
</tr>
<tr>
<td>DB 3</td>
<td>Makes an ethnic, religious, or racial remark at work.</td>
<td>100</td>
</tr>
<tr>
<td>DB 4</td>
<td>Curses at someone at work.</td>
<td>96</td>
</tr>
<tr>
<td>DB 5</td>
<td>Plays a mean prank on someone at work.</td>
<td>96</td>
</tr>
<tr>
<td>DB 6</td>
<td>Acts rudely toward someone at work.</td>
<td>96</td>
</tr>
<tr>
<td>DB 7</td>
<td>Publicly embarrasses someone at work.</td>
<td>88</td>
</tr>
<tr>
<td>DB 8</td>
<td>Takes property from work without permission.</td>
<td>96</td>
</tr>
<tr>
<td>DB 9</td>
<td>Spends too much time fantasizing or daydreaming instead of working.</td>
<td>92</td>
</tr>
<tr>
<td>DB 10</td>
<td>Falsifies a receipt to get reimbursed for more money than he/she spent on business expenses.</td>
<td>100</td>
</tr>
<tr>
<td>DB 11</td>
<td>Takes an additional or longer break than is acceptable at his/her workplace.</td>
<td>83</td>
</tr>
<tr>
<td>DB 12</td>
<td>Comes in late to work without permission.</td>
<td>92</td>
</tr>
<tr>
<td>DB 13</td>
<td>Litters his/her work environment.</td>
<td>96</td>
</tr>
<tr>
<td>DB 14</td>
<td>Neglects to follow his/her boss’s instructions.</td>
<td>100</td>
</tr>
<tr>
<td>DB 15</td>
<td>Intentionally works slower than he/she could have worked.</td>
<td>100</td>
</tr>
<tr>
<td>DB 16</td>
<td>Discusses confidential company information with an unauthorized person.</td>
<td>100</td>
</tr>
<tr>
<td>DB 17</td>
<td>Uses an illegal drug or consumes alcohol on the job.</td>
<td>100</td>
</tr>
<tr>
<td>DB 18</td>
<td>Puts little effort into his/her work.</td>
<td>96</td>
</tr>
<tr>
<td>DB 19</td>
<td>Drags out work in order to get overtime.</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: \( n = 24 \) for Study 1 and \( n = 545 \) for Study 2. Numbers reported for Study 1 are agreement ratio and for Study 2 are factor weights. OCB, organizational citizenship behavior; DB, deviance behavior.

*This item was not included in the survey due to clerical error.*
of 75% recommended by Hinkin (1998). As expected, experts categorized all 19 items from Bennett and Robinson (2000) as deviant behavior.

**Study 2: discriminant validity**

Although experts’ categorization is useful in identifying the content difference of the items, it is also necessary to test the discriminant validity of the five OCB dimensions against deviant behavior in an applied setting. If the harmony and resources items are part of OCB as assumed by Farh et al. (1997), they should show stronger relationships to other dimensions of the same construct (i.e., OCB) than to a different construct (i.e., deviant behavior). While dimension-level correlations are acceptable to test discriminant validity, we decided to be more rigorous and perform an item-level EFA to test if any particular item from this OCB measure cannot be differentiated from deviant behavior.

In doing so, we drew our Sample 2 from a large retail company located in Beijing. With endorsement from the top management, we asked all 67 supervisors in the organization to rate their subordinates. Participation was voluntary and no monetary incentives were offered. Data were directly collected by the research team instead of any representatives of the company. We instructed supervisors to rate each subordinate’s frequency of specified behaviors on a seven-point Likert scale, ranging from 1 = never to 7 = always (at least once a week). We dropped responses from five supervisors due to large percentage of missing data, leaving a total of 62 supervisors rating 545 subordinates. The number of subordinates each supervisor rated varied from 1 to 38, with a median of 6.

To estimate whether the clustered data structure affects results (thus multilevel analysis is needed), we calculated the design effect of our data, which indicates how much the standard errors are underestimated in a complex sample compared with a simple random sample, using the formula from Muthén (1997). Our design effects of Farh et al.’s (1997) items varied from 1.01 to 1.87, all below the cutoff value of 2, suggesting that the clustered ratings are relatively independent and it is appropriate to conduct individual-level analysis. The subordinates’ average age was 31 and 60% of them were female. They had an average work experience of 7.3 years and had worked in their current position for 4.9 years. About 27% had bachelor’s degrees, 41% had diplomas from junior colleges and 27% had high school diplomas.

The 20 items comprising the Farh et al.’s (1997) OCB measure and 18 items comprising the Bennett and Robinson’s (2000) deviant behavior measure were presented in a random order. The 19th item from the Bennett and Robinson’s deviant behavior measure, ‘Makes an ethnic, religious, or racial remark at work’, was omitted from the questionnaire due to clerical error.

We performed an EFA on the 38 items using principle component estimation followed by varimax rotation. Inspection of the scree plot indicated that four factors would be optimal (Cattell 1966). So we retained four factors, which accounted for 65% of the variance. Following Ford, MacCallum and Tait’s (1986) suggestion, we considered an item loaded on a specific factor if its loading was 0.4 or larger and it did not cross-load with any other factor. The results of the EFA are also shown in Table 1.

All seven emic OCB items, together with all 18 items of deviance behavior, loaded on the first factor, and 10 of the 13 etic OCB items loaded on the second factor. One etic OCB item under the conscientiousness dimension loaded on the third factor, and another etic item under the identification dimension loaded on the fourth factor. The third and fourth factors had only one item each and were thus dropped, leaving the first and second factor representing deviant behavior and OCBs, respectively. Through this EFA, it is apparent
that Farh et al.’s (1997) emic items of personal use of organizational resources and harm to interpersonal harmony are essentially regarded the same as deviant behavior by Chinese supervisors in an applied setting.

Study 3: factorial validity

Because previous results suggested the potential existence of two higher-order factors for this 20-item measure, we conducted a second-order CFA using AMOS 7 (Arbuckle 2006) on a new sample to confirm such a structure.

Our Sample 3 consists of 244 employees rated by 118 supervisors. We contacted top executives or HR managers of 18 companies, of which 12 companies agreed to participate and provided their employee rosters. We randomly selected two to three employees from each unit of the company, and administered the paper-and-pencil questionnaires to these pre-selected employees and their supervisors in scheduled on-site appointments, and collected all the completed questionnaires on the same day. The response rate was 92%. Most subjects in Sample 3 were drawn from Sichuan province (67%) and Ningxia province (28%). The participating companies represent various industry sectors such as manufacturing, financial services, restaurant, real estate and telecom. Each participant was given a gift of 20 RMB in value (approximately 3 USD) for his or her time.

The percentage of males was 41% among the subordinates, and 55% among the supervisors. About 43% of supervisors were in their thirties, and 27% were in their forties. The subordinates were relatively younger, with 57% in their twenties and 29% in their thirties. About 57% of supervisors had worked in their organizations for more than 3 years and 42% of the subordinates had worked at their current employer for more than 3 years. Of the subordinates, 25% were employed at state-owned companies and 75% were in private companies.

The number of subordinates each supervisor rated varied from two to three. Similar to Study 2, we calculated the design effects of our data, which were all below 2 (varying from 1.18 to 1.62), suggesting that the clustered ratings are relatively independent and it is appropriate to conduct individual-level analysis.

Our hypothesized model includes two second-order factors, representing OCB and deviant behavior, respectively. The OCB factor has three first-order factors: identification, altruism and conscientiousness, while the deviant behavior factor has two first-order factors: harmony and resources. We compare this model to an alternative model with a single second-order factor of OCB that was assumed by Farh et al. (1997) and to an independence model using chi-square difference tests. To provide a metric for the latent constructs and to identify the measurement model, in each analysis the first construct loading for each latent construct was set to 1. To be stringent, we did not allow covariance between any error terms of observed indicators. We evaluated model fit with multiple fit indices, including the comparative fit index (CFI), normed fit index (NFI), standard root-mean-square residuals (SRMR) and root-mean-square error of approximation (RMSEA). Fit indices for all models are based on the maximum likelihood method. A model is often deemed as acceptable if CFI and NFI are more than 0.90 (Byrne 1998), SRMR is less than 0.08 and RMSEA is less than 0.10 (Kenny 2010).

Results of the CFA are shown in Table 2. The hypothesized model with dual second-order factors has a good fit, \( \chi^2 = 371.95, \quad df = 164, \quad p < 0.01; \quad CFI = 0.94, \quad NFI = 0.90, \quad SRMR = 0.05, \quad RMSEA = 0.07, \) which is a significant improvement over the single second-order factor model, \( \chi^2 = 581.40, \quad df = 165; \quad \Delta \chi^2 = 209.45, \quad \Delta df = 1, \quad p < 0.01. \) The hypothesized model is also better than the independence model. We conclude that
Table 2. CFA fit indices (Study 3).

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$ (df)</th>
<th>$\Delta\chi^2$ (Δdf)</th>
<th>CFI</th>
<th>NFI</th>
<th>SRMR</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual second-order factors</td>
<td>371.95 (164)**</td>
<td>–</td>
<td>0.94</td>
<td>0.90</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>Single second-order factor</td>
<td>581.40 (165)**</td>
<td>209.45 (1)**</td>
<td>0.88</td>
<td>0.84</td>
<td>0.12</td>
<td>0.10</td>
</tr>
<tr>
<td>Independence</td>
<td>3603.13 (190)**</td>
<td>3231.18 (26)**</td>
<td>0</td>
<td>0</td>
<td>0.62</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Notes: $\chi^2$, chi-square; df, degrees of freedom; GFI, goodness-of-fit index; CFI, comparative fit index; NFI, normed fit index; SRMR, standardized root-mean-squared residual; RMSEA, root-mean-square error of approximation; $n = 244$. **p < 0.01.
the hypothesized dual-factor model fits the data best, which supports the factor structure identified through the EFA from Study 2. Figure 1 shows the path coefficients of the hypothesized model, and Figure 2 shows the path coefficients of the alternative single second-order factor model. All path coefficients are statistically significant at the 0.01 level. Disturbance and error terms are omitted from the figure for clarity.

Study 4: nomological validity

In the taxonomy of Campbell (1960), content validity, discriminant validity, convergent validity and factorial validity are aspects of trait validity of a measure, and because they are in the theoretical vacuum, they are necessary but not sufficient evidence for establishing construct validity. For complete construct validity, measures must also demonstrate nomological validity. Nomological validity concerns whether the operationalization of a construct relates to other constructs as anticipated in theoretically driven hypotheses (Campbell 1960). So, to test the nomological validity of Farh et al.’s (1997) measure, from the nomological network surrounding OCB we carefully picked two concepts that are salient and frequently studied: affective organizational commitment and LMX. The former represents employees’ affective attachment to the employer organization, while the latter represents the quality of the relationship between an employee and his/her direct supervisor. OCB and deviant behavior are distinct concepts and would exhibit somewhat different patterns of relationships with antecedents, so we anticipate Farh et al.’s (1997) etic
dimensions (which measure OCB) and emic dimensions (which measure deviant behavior) to exhibit different correlation patterns with affective commitment and LMX.

According to commitment theory (Mowday, Porter and Steers 1982; O’Reilly and Chatman 1986), affectively committed employees are willing to aid the organization in all possible ways, even if direct reward is not contingent on the aid. Social exchange theory (Blau 1964) posits that employees who are in good relationships with their direct supervisor are more likely to reciprocate and help the supervisor and the organization in general. Such theoretical assumptions are empirically supported in both North America (e.g., O’Reilly and Chatman 1986; Podsakoff and Mackenzie 1993; Wayne, Shore and Liden 1997) and China (e.g., Hui et al. 1999; Chen and Francesco 2003; Wang, Law, Hackett, Wang and Chen 2005). Thus, our Hypothesis 1 is that OCB (as measured by Farh et al.’s etic dimensions) is positively related to affective commitment and LMX.

Affective commitment and LMX do not necessarily reduce the occurrence of deviant behaviors. Tang and Chiu (2003, p. 25) reasoned that in modern China where living expenses are rising high and traditional work ethics is eroded, ‘organizational commitment … will not “deter” or “undermine” [employees’] motives to engage in unethical behaviors’ such as personal use of organizational resources. As for LMX, subordinates who maintain high quality exchange relationships with their supervisor can expect the supervisor to ‘bail’ them out if they get into trouble (Scandura and Graen 1984). Some employees may be ‘spoiled’ and engage in deviant behaviors, believing they will not be caught or punished. Those forces will offset the positive influences of affective

Figure 2. CFA results for the single second-order factor model (Study 3). OCB, organizational citizenship behavior. Scores for harmony and resources were reversed before calculation. Standardized path coefficients are shown. All path coefficients are statistically significant at 0.01 level. Disturbance and error terms are omitted from the figure for clarity.
commitment and LMX in regulating employees’ deviant behaviors, so our Hypothesis 2 is that deviant behavior (as measured by Farh et al.’s emic dimensions) is in general unrelated to affective commitment or LMX in China.

We tested the nomological validity of Farh et al.’s OCB measure using Sample 3. To reduce potential method variances from the common source that can inflate the correlations, we correlated supervisor-reported scores of OCB to subordinate-reported scores of commitment and LMX. We used 5 items adapted from an indigenous scale developed by Ling, Zhang and Fang (2001) to measure affective organizational commitment. To assess the convergent validity of this indigenous affective commitment measure, we performed a CFA of its items and items from a shortened version of the Organizational Commitment Questionnaire developed by Mowday, Steers and Porter (1979) among a subset of this sample (n = 169). We found that they loaded on a single factor as expected, and at the scale level they were highly correlated (r = 0.90). We measured LMX using the 7-item measure from Scandura and Graen (1984), and used the same back-translation procedure described in Study 1.

Table 3 shows the means, standard deviations and internal reliabilities of the seven variables as well as the correlation coefficients among them. Internal reliabilities were assessed by Cronbach’s alphas, and each of the five dimensions of Farh et al. (1997) measure had acceptable internal reliabilities, varying from 0.85 to 0.91. Affective commitment and LMX also had good internal reliabilities of 0.91 and 0.87, respectively. Similar to inter-correlation patterns of Farh et al. (1997) and Yen and Niehoff (2004), the correlations coefficients across etic and emic dimensions (r = 0.26 to 0.36) seem much lower than inter-correlations within etic dimensions (r = 0.75 to 0.78) and within emic dimensions (r = 0.79).

Consistent with theoretical assumptions, affective commitment is significantly correlated with the three etic dimensions of OCB in the current study, with effect sizes very close to that reported in a previous study conducted in China (i.e., Chen and Francesco 2003). Similarly, LMX is significantly and positively related to the three etic dimensions of OCB as anticipated, with effect sizes comparable to that reported in previous studies conducted in China (Wang et al. 2005; Xu, Huang, Lam and Miao 2012). Our Hypothesis 1 is supported.

Like Tang and Chiu (2003), who found no relationship between affective commitment and deviant behavior in Hong Kong, in our study affective commitment is not correlated to harm to interpersonal harmony or to personal use of organizational resources. Like Liu (2005) who reported non-significant correlations between perceived supervisor support and deviant behaviors in Taiwan and Ferris, Brown and Heller (2009) who reported non-significant correlations between LMX and deviant behaviors in the USA, we found LMX is not correlated to harm to interpersonal harmony or to personal use of organizational resources. Our Hypothesis 2 is thus supported. In summary, Farh et al.’s (1997) etic OCB dimensions relate to other constructs as theoretically predicted, but the harmony and resources dimensions operate the same way as deviant behavior.

Discussion
Attempts to design indigenous measures are relatively new, so additional scrutiny is critical if these instruments are to be considered effective tools for assessing and analyzing organizational phenomena in a new cultural context. OCB is a very frequently studied employee behavioral outcome in management literature, so understanding its construct domain in the global setting has important implications for both researchers and practitioners. To our knowledge, despite the wide adoption of Farh et al.’s (1997) OCB measure, the present study is the first to
Table 3. Correlation results (Study 4).

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identification</td>
<td>M</td>
<td>5.55</td>
<td>1.08</td>
<td>(0.87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Altruism</td>
<td>M</td>
<td>5.78</td>
<td>1.03</td>
<td>0.78**</td>
<td>(0.91)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Conscientiousness</td>
<td>M</td>
<td>5.44</td>
<td>1.05</td>
<td>0.75**</td>
<td>0.75**</td>
<td>(0.88)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Harmony</td>
<td>M</td>
<td>6.43</td>
<td>0.90</td>
<td>0.34**</td>
<td>0.27**</td>
<td>0.34**</td>
<td>(0.86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Resources</td>
<td>M</td>
<td>6.46</td>
<td>0.92</td>
<td>0.33**</td>
<td>0.26**</td>
<td>0.36**</td>
<td>0.79**</td>
<td>(0.85)</td>
<td></td>
</tr>
<tr>
<td>6. Affective commitment</td>
<td>S</td>
<td>5.28</td>
<td>1.37</td>
<td>0.13*</td>
<td>13*</td>
<td>0.18**</td>
<td>0.05</td>
<td>0.01</td>
<td>(0.91)</td>
</tr>
<tr>
<td>7. LMX</td>
<td>S</td>
<td>5.14</td>
<td>1.22</td>
<td>0.16*</td>
<td>0.24**</td>
<td>0.22**</td>
<td>0.06</td>
<td>0.03</td>
<td>0.55**</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are Cronbach’s alphas. Scores for harmony and resource were reversed before calculation. LMX, leader–member exchange; M, manager; S, subordinate; \( n = 244 \). * \( p < 0.05 \); ** \( p < 0.01 \).
systematically examine the measure’s psychometric properties. We found Farh et al.’s harmony and resources dimension did not reflect the intended definition of OCB, but instead tapped into the content of deviant behaviors. Thus, the emic contents of OCB reported in that study was unfortunately a false positive.

This study fuels ‘a fair debate on the value of universal versus local theories’ (Tsui 2007, p. 1354). For now, our results cast doubt on previous indigenous findings, and put demands on divergence proponents to provide more scientific evidence. We caution that one unsuccessful empirical attempt does not necessarily mean that the divergence perspective is wrong; rather, this study shows how intriguing and challenging indigenous research is. National culture is ‘an elusive concept’ (Ralston et al. 1997, p. 179), and it may take decades, if not longer, to theorize, document and conclude on its influences on employee values, behaviors and management practices. It is highly important for researchers to stay objective in this process: being blind to contextual difference will impede the generation of global knowledge, but overestimating contextual influences will make the conversation between the East and the West unnecessarily difficult. A scientist remains neutral and open to new facts until a convincing body of evidence is amassed demanding a change. Within-culture variances in large societies like China and the USA can be so huge (Vandello and Cohen 1999) that it is difficult to estimate whether the cross-culture variances outweigh within-culture variances, or vice versa. Thus, the real challenges for scholars are theoretical reasoning and empirical assessment of the magnitude of cultural impact on a construct’s domain, and to determine whether the cross-culture difference is large enough to justify a separate model for a specific organizational phenomenon.

In the area of OCB, some researchers have clearly argued that the construct domain of OCB can transcend cultures. Lieveen, Van Hoye and Anseel (2004) tested the validities of a popular American OCB measure in Dutch-speaking Belgium and concluded that ‘the forms of OCB that are predominantly studied in the US seem to hold relatively well’ (p. 299) in this international context. Euwema, Wendt and Van Emmerik (2007) studied employees from 33 societies, and found the measure of work group OCB has a stable structure across countries, and OCBs are not directly affected by society-level collectivism and power distance. Cohen (2006) found that Arab employees and Jewish employees do not differ in their demonstrated altruism, a type of OCB, and collectivism is only related to altruism at the individual level (instead of society-level). In China, while Farh et al.’s (1997) OCB measure gained legitimacy and has been widely adopted, a small group of researchers kept using western OCB measures, primarily the one developed by Podsakoff, MacKenzie, Moorman and Fetter (1990), to study Chinese employees’ OCBs (e.g., Chen, Hui and Sego 1998; Lam, Hui and Law 1999; Hui, Lee and Rousseau 2004; Wang et al. 2005; Law, Wang and Hui 2010). They reported satisfactory psychometric properties of western measures among Chinese employees. For example, Lam et al. (1999) administered the Podsakoff et al.’s (1990) scale to samples from the USA, Hong Kong, Japan and Australia, and concluded the scale has ‘conceptual equivalence across all subsamples’ (p. 594). Hui et al. (2004) justified their choice of the western measure by arguing that ‘there appears to be reasonable evidence that this [Podsakoff et al. (1990)] OCB measure is acceptable in Chinese contexts and can facilitate comparative OCB research’ (p. 316). These studies did not intend to systematically explore the construct domain of OCB in China and cannot completely preclude the potential existence of indigenous OCB contents. But until more scientific studies are done, from which a more definitive answer to the above question can be drawn, researchers wishing to study OCB in Chinese societies are not obligated to use a particular measure because of concerns over culture-specific contents.
This study lends support to the conceptual distinction between OCB and deviant behavior (e.g., Viswesvaran and Ones 2000; Rotundo and Sackett 2002; Dalal et al. 2009; Spector et al. 2010). It would be fair to state that the weakness of Farh et al. (1997) reflected a limitation of that time, as the distinction between OCB and deviant behavior started to attract attention after this measure was published. But just like the original OCB model, recent advancements in the conceptual distinction between OCB and deviant behavior are limited to studies conducted in the West and have unfortunately neglected OCB measures developed and used in the East. For example, Dalal’s (2005) meta-analysis of this topic did not include any primary study from eastern societies, which is surprising because at least Farh et al. (1997) was published in the intended study period. A possible explanation is that scholars become hesitant about the dimensionality of OCB when an unfamiliar culture is involved. Our study is thus still important and in time, through its articulation of the conceptual distinction between OCB and deviant behavior in China. We find affective commitment and LMX are positively related to OCB, but they are unrelated to deviant behavior. Consistent with Spector et al. (2010), the results suggest that in China OCB and deviant behaviors are not the opposite ends of each other, and some employees may engage in both types of behaviors. Future studies can further explore the relationships between OCB and deviant behavior in China. For example, if OCBs are never recognized by the organization, will Chinese employees be more or less likely to engage in deviant behaviors simultaneously to ‘make it even’ than westerners? Such examinations will help demonstrate cultural influences.

This paper helps improve future indigenous research in the area of OCB and beyond in terms of research design. Our findings suggest that the novel findings in some indigenous research may be due to research artifacts instead of true variance from the focal population. Efforts to conduct indigenous research in emerging economies are relatively new, and researchers face additional challenges and complexities such as a lack of solid theoretical grounds, unknown research domains, little historical data to build on, unavailability of existing instruments, translation inaccuracy in test items, confused subjects due to unfamiliarity with behavioral research and inexperienced research assistants or site coordinators who may have unintentionally affected responses (Tsui 2004). Some challenges are culture-related, others are not, and unfortunately different influences are often intertwined and may produce false positives when researchers intend to study culture-unique phenomena. Because indigenous research is highly vulnerable to research artifacts, researchers need to exercise due caution in research design, implementation and results interpretation to rule out alternative explanations.

In terms of research design, Farh et al. (1997, 2004) used full induction approach to explore indigenous contents of OCB. Is full induction the best choice for indigenous scale development? Tsui (2004, p. 506) indicated that full induction is appropriate only when there is no a priori framework at all, and if there is extensive research on the topic in another context, researchers can adopt a partial inductive approach by beginning their research with a preliminary framework. It means researchers do not have to stay passive throughout the process. In indigenous scale development, providing subjects with only an open definition may seem to free item-generation process from researchers’ bias, but this practice often produces items that assess extraneous content domains (Hinkin 1998) and it creates unwanted difficulties in subsequent item-screening and interpretation, when researchers still need to rely on their own judgments to help make meaningful solutions. Given the extensive OCB-related studies in western societies, they can potentially inform the design of inductive research in some ways. For example, when asking subjects to generate examples of OCB, researchers need to thoroughly investigate the nomological network in
advance to identify adjacent constructs that may confuse subjects, and then either highlight the conceptual difference in the definition presented to subjects or collect additional data to test discriminant validity against adjacent constructs. Such preliminary framework may also come from the local culture, and the definitions should be broad enough to allow reasonable latitude for respondents to generate examples that they are familiar with.

In terms of implementation and interpretation, for research robustness, we advocate different research teams replicating preliminary indigenous findings under multiple settings before they can be accepted. Open-mindedness and constructive criticism from peers are especially valuable to improve the quality of indigenous research in China, where the pressure for harmonious social relationship can be high. Replication is not only necessary in the ‘new’ context; researchers should also replicate the study in the ‘old’ context (e.g., USA), so that one can prove that the emic content is really something new, not something overlooked when the original model was proposed. Such replications are necessary to help avoid false positives.

**Limitations and directions for future research**

Although the results of this study are instructive in various ways, its implications should be considered in light of the study’s limitations and directions for future research. First, we conducted this study with the purpose of assessing the Farh et al. (1997) OCB measure. As such, we did not generate our own items, or propose possible alternative factors of relevance to OCB. We start with this measure because it is a highly representative and influential indigenous study. Pointing out the weakness of this measure will reignite interest in this seemingly closed area and pave the road for future studies. Although we recommend dropping the two emic dimensions in their current formats, it is still possible that protecting organizational resources and interpersonal harmony are part of Chinese OCB, if rigorously designed future studies revealed so. To qualify as OCB, we expect that employees need to go as far as to be highly frugal when using organization resources or dedicate personal resources for business use. Similarly, interpersonal harmony may be part of Chinese OCB, if employees demonstrate harmony-maintaining behaviors beyond formal role expectations, such as continued support toward colleagues after being offended.

The second limitation is that our samples were all from Mainland China while Farh et al.’s (1997) samples were from Taiwan. Given the potential existence of subcultural differences between Mainland China and other geographic areas of Greater China, future replications in Taiwan and Hong Kong will help establish the external validity of the present study. Nevertheless, we believe our study is still meaningful for the following two reasons. First, Farh et al. (1997) intended to propose a model for ‘Chinese society’ and did not restrict its use in Taiwan. Mainland China has the largest population in the world and a profound impact on global economy and politics, making any indigenous model intended for ‘Chinese society’ misleading if it does not apply to Mainland China. Second, Farh et al. (1997) and other studies from Taiwan (e.g., Yen and Niehoff 2004) and Hong Kong (e.g., Leung 2008) that used this measure also reported problematic inter-correlations across etic and emic dimensions, a pattern replicated in our study, suggesting the measure’s weakness is not specific to Mainland China. There is no doubt that Chinese culture is complicated and subcultures are operating, but generating global knowledge does not mean developing a new model for each geographic location. Farh et al. (1997) investigated the construct domain of OCB in Taiwan and in Mainland China (Farh et al.2004) separately, but it was a pity that they did not compare or integrate the two studies toward a universal model for Greater China. To attract interest from other parts of the world and increase the
credibility of indigenous research, at the current stage we feel it is ideal for researchers in Greater China to collaborate and develop a universal model of Chinese OCB, if employees across this region agree on OCB’s contents and dimensionality. Such efforts would facilitate connecting context-specific research to the existing knowledge base in the West. Once the universal model is established, modifications can be made for each specific subculture in the next step of indigenous research.

**Conclusion**

It is time to further question, examine and verify the validity of existing indigenous research. The present study represents an important step toward the goal of increased quality of indigenous measures. The search for unique contents of Chinese OCB should continue, and we encourage scholars to develop a universal model for Chinese society.

**Acknowledgement**

This research was partially funded by the National Natural Science Foundation of China (Grant No. 70802010).

**References**


Liu, H. (2005), ‘Exploring the Predictors of Employee Deviant Behaviors,’ Master’s thesis, National Taiwan University of Science and Technology, Dept. of Business Administration.


