

**Diagnosing Organizational Cultures: A Conceptual and Empirical Review of Culture
Effectiveness Surveys**

Short title: Diagnosing Organizational Cultures

Daniel Denison*

International Institute for Management Development
Chemin de Bellerive 23
Lausanne Switzerland 100
+41.21.618.0311
denison@imd.ch

Levi R. G. Nieminen

Wayne State University
Department of Psychology
5057 Woodward Ave.
Detroit, MI 48202
906.367.2744
at2440@wayne.edu

Lindsey Kotrba

Denison Consulting
121 W. Washington, Suite 201
Ann Arbor, MI 48104
734.302.4002

The authors wish to thank the International Institute for Management Development and the University of Michigan Business School for the support that they have provided for this research. We also want to thank Jay Janovics, Joana Young, Hee Jae Cho, Ashley Guidroz, Paul Curran, and Ryan Smerek for their contributions.

NOTE: Please note that this manuscript is currently under peer review. Do not share this manuscript or cite without the authors' permission.

Abstract

Methodological reviews within the organizational culture domain have repeatedly called for additional research evaluating the reliability and validity of survey measures of organizational culture. In this review, we first trace the development of survey research within the organizational culture tradition, focusing specifically on the category of instruments that assess aspects of culture related to organizational effectiveness. Although surveys of this kind are the most direct diagnostic assessments of organizational culture, our review suggests that research support is generally inadequate to establish the reliability and validity of the majority of instruments in this category. Next, our review identifies several considerations that are unique to the development and validation of culture effectiveness surveys, and thus warrant special attention. We identify three key challenges for future culture researchers to address: *testing nested models*, *aggregating data to the organizational level*, and *establishing criterion-related validity*. Finally, using archival data collected with the Denison Organizational Culture Survey, we present an empirical illustration of the three challenges identified above and conclude by considering limitations and opportunities for future research.

Keywords: organizational culture
 cross-level analysis
 organizational effectiveness
 survey
 scale validation

Diagnosing Organizational Cultures: A Conceptual and Empirical Review of Culture Effectiveness Surveys

This review paper is organized around three main sections. In *Section I*, we present a historical overview describing the use of survey instruments within the organizational culture research tradition, and more specifically, the role of surveys in research investigating linkages between organizational culture and performance. Borrowing from existing taxonomies, various survey approaches are differentiated according to their theoretical diversity, the research purpose that is served, and the format of information that is produced. After describing the landscape of culture surveys, we present an overview of key methodological reviews by Ashkanasy, Broadfoot, and Falkus (2000) and Jung, Scott, Davies, et al. (2009), focusing on these authors' conclusions about the availability of various types of reliability and validity evidence. In *Section II*, we update these authors' work by focusing specifically on the category of measures referred to by Ashkanasy *et al.* as *effectiveness profiling instruments*. Although surveys of this type are the most direct diagnostic assessments of organizational culture, relatively little attention has been paid to their systematic evaluation. Based on our update, we identify and describe several key considerations that are under-examined "gaps," and therefore, important challenges for future researchers to address. These include the approach taken to testing nested models, the issues regarding the aggregation of individual responses on culture surveys to the organization level, and the considerations for linking survey responses to performance criteria. Finally, *Section III* provides an empirical illustration of our approach to each of these challenges using archival data based on the Denison Organizational Culture Survey.

SECTION I: Measurement of Organizational Culture

The importance of *organizational* culture was first described by Elliott Jaques in his 1951 book titled, *The Changing Culture of a Factory*. Jacques invoked *culture* – described as informal social structures – as a way to explain the failure of formal policies and procedures to resolve the un-productive dynamic between managers and employees at the Glacier Metal Company. Later, the concept was re-introduced to the field by Andrew Pettigrew (1979), whose work pointed to culture as the “social tissue” that contributes to collective sense-making in organizations (p. 574). Informal social structures and collective sense-making are still reflected in modern definitions of organizational culture, although new concepts have been integrated as well. Although there is no widely agreed upon definition, most organizational scholars concur that the core definitional content includes the values, beliefs, and assumptions that are held by the members of an organization and which facilitate shared meaning and guide behavior at varying levels of awareness (Alvesson, 2011; Denison, 1996; Schein, 1992; Smircich, 1983). Also, the potential for multiple cultures (or *sub-cultures*) within a single organization is generally acknowledged in definitions (Martin, 1992; Martin & Meyerson, 1988).

Alongside conceptual developments, measurement perspectives on organizational culture have evolved greatly over time. Early scholarship reflected the sociological and anthropological origins of the culture concept and therefore emphasized qualitative, ethnographic research methods (Rousseau, 1990). Similarly, culture was conceptualized mainly from an *emic* perspective, in which cultures are viewed as idiosyncratic, rather than *etic* perspective, in which cultures are viewed as comparable (Denison, 1996). Hence, the historical and epistemological forces guiding early scholarship mainly discounted the possibility that organizational cultures could be studied within a nomothetic framework using standardized survey instruments (Schein, 1992; Trice & Beyer, 1993).

Despite these challenges, the earlier “culture wars” waged by advocates of different epistemological and methodological views have generally subsided (Martin & Frost, 1996; Martin, Frost, & O’Neill, 2006). Instead, the strengths and limitations of various methodologies, including the survey, are recognized within a multi-method framework (Hofstede, Neuijen, Ohayv, & Sanders, 1990; Ostroff, Kinicki, & Tamkins, 2003; Rousseau, 1990; Sackmann, 2006). For comparative organizational researchers, surveys provide the foundation for quantitative assessment and cross-organization comparison (Xenikou & Furnham 1996). Additionally, surveys are less expensive and time consuming than clinical or ethnographic methods, provide normative information about an organization’s culture, facilitate the benchmarking and organizational change process, and allow for direct replication (Ashkanasy et al., 2000; Cooke & Rousseau, 1988; Tucker, McCoy, & Evans, 1990). In balance, researchers have generally acknowledged two main limitations of survey methodologies, specifically their inability to access “deeper” cultural elements such as symbolic meaning, semiotics, and fundamental assumptions (e.g., Schein, 1992; Smircich, 1983; Rousseau, 1990; Van Maanen, 1988) and their use of a priori content (e.g., pre-defined questions) which may fail to capture all relevant aspects of culture. Two additional assumptions are that survey respondents’ perceptions of the culture are accurate and meaningful when aggregated to the group level (Sackmann, 2006). As a result, culture surveys are most appropriate when the focus of investigation is at the level of “observable and measurable manifestations of culture,” such as values and behavioral norms, and when the research purpose calls for making comparisons across organizations using the same set of culture concepts (p. 132; Ashkanasy et al., 2000).

The Culture-Performance Link

Limitations notwithstanding, the number of standardized quantitative instruments has grown rapidly since the mid-1980s, and new approaches continue to be developed (Jung et al., 2009). Importantly, culture surveys have opened up several areas of inquiry previously difficult to investigate systematically. One area of particular interest involves the effect of organizational culture on performance. As noted by Sackmann (2011), the notion of a culture-performance link was implicit in the concept's first introduction to the organizational sciences and was reinforced by managers seeking yet another performance management tool. However, research support has been slow to accumulate, with several prior reviews failing to find compelling evidence of a culture-performance link (Lim 1995; Siehl & Martin, 1990; Wilderom, Glunk, & Maslowski, 2000).

More recently, a review by Sackmann (2011) in the second volume of the *Handbook of Organizational Culture and Climate* produced somewhat more optimistic conclusions. First, the number of empirical studies reviewed was more than five times the number included in the prior reviews (e.g., Wilderom et al., 2000 reviewed a total of 10 studies). Other macro trends include the increasing globalization (e.g., cross-cultural studies) and specialization (e.g., industry-specific studies) of culture-performance research, as well as the emergence of sub-disciplines. Most importantly, the sum effect of these trends is a richer understanding of the culture-performance relationship.

The majority of empirical studies reviewed demonstrate *direct effects* of culture on performance. For example, Sackmann (2011) concludes that:

a combination of external and internal orientation is an ideal combination for a direct positive relationship with performance, even though a stronger external focus seems to be more important – regardless of industry and country (p. 211).

Evidence of a more complicated relationship is also beginning to emerge, with a smaller number of studies demonstrating *indirect* (or mediated) effects, *interactive* (or moderated) effects, *non-linear* effects, and *reciprocal* effects. Together, these studies demonstrate the potential of a contingency-based perspective to augment the literature on direct effects and contribute to a more dynamic and contextualized understanding of the culture-performance link. Clearly, survey methodologies have played, and will continue to play, a pivotal role in the advancement of knowledge in this domain (Ashkanasy et al., 2000).

Important Methodological Considerations

Although Sackmann's (2011) review underscores significant progress in understanding the culture-performance link, it also highlights the persistence of many of the methodological problems described in previous reviews, some of which focus on the challenges associated with measuring culture (e.g., Wilderom et al., 2000). For example, Sackmann described how the wide variety of survey instruments used makes it difficult to establish clear patterns across studies, instead creating "a rather broad and colorful picture of the link between different culture dimensions and performance measures" (p. 196). In contrast, some have cast this diversity as a healthy form of pluralism and focused on the different research purposes served by the different approaches (e.g., Scott, Mannion, Davies, & Marshall, 2003; Jung et al., 2009). Additional problems include the lack of clear theoretical grounding of many culture surveys and the failure to provide sufficient evidence of scale reliability and validity (Ashkanasy et al., 2000). Below, we describe research progress toward each of these challenges.

Theoretical Diversity. The majority of existing culture surveys assess specific behavioral norms and values that are thought to stem from an organization's culture (Ashkanasy et al., 2000). An underlying model specifies the way in which specific norms and values are

grouped into meaningful themes or *dimensions* (e.g., factor structures), as well as how those dimensions ought to relate to one another (e.g., higher-order factors). Searches of the literature have revealed anywhere from 74 (Ott, 1989) to 114 unique dimension labels (van der Post, de Coning, & Smit, 1997), indicating that surveys differ extensively in terms of their *nominal* categorizations of the content of culture. Beyond clear differences in labeling, a handful of studies have sought to determine the conceptual overlap of culture dimensions across surveys (Delobbe, Haccoun, & Vandenberghe, 2002; Detert, Schroeder, & Mauriel, 2000; Ginevičius & Vaitkūnaitė, 2006; Xenikou & Furnham, 1996).

Xenikou and Furnham (1996) used a quantitative approach to test the convergent validity of similar dimensions between the *Organizational Culture Inventory* (Cooke & Lafferty, 1989), the *Culture Gap Survey* (Kilman & Saxton, 1983), the *Organizational Beliefs Questionnaire* (Sashkin, 1984), and the *Corporate Culture Survey* (Glaser, 1983). First, a content mapping procedure was used to identify pairs of similar dimensions across the instruments. Then all four instruments were administered to 157 employees from two organizations. Results indicated that 16 out of 30 dimensions demonstrated moderate convergence with correlations ranging from .40 to .57. Next, the authors performed a principal components analysis to determine whether scale scores across the four instruments could be represented within a common factor structure. Findings indicated that the 30 dimensions generally clustered into six correlated factors.

Detert et al. (2000) expanded on the study by Xenikou and Furnham (1996) by examining the dimensional overlap among 25 multi-concept frameworks of culture. These authors began with a comprehensive literature review to identify models and measures that describe culture along multiple aspects or dimensions. Then, using a qualitative content mapping procedure, these authors undertook an iterative process of grouping dimensions on the basis of conceptual

similarity until further grouping could not be justified. The resulting model included eight broadly defined themes: the basis of truth and rationality, the nature of time and time horizon, motivation, stability versus change-innovation-personal growth, orientation to work- task-coworkers, isolation versus collaboration-cooperation, control-coordination-responsibility, and internal-external orientation and focus.

Together, these studies provide initial evidence that the dimensions assessed by different culture surveys can be described in terms of a set of higher-order culture dimensions. Higher-order frameworks seem particularly useful in light of the difficulty of accumulating research findings based on different survey instruments (Sackman, 2011). On the other hand, Xenikou and Furnham's (1996) suggestion that the broad themes extracted in their research might also serve as a useful basis for developing new scales, awaits further evaluation. For example, it may prove difficult to write items with the level of specificity required to yield valid and reliable information but which also map clearly onto broadly defined culture concepts. Along similar lines, Ashkanasy et al. (2000) described the tradeoffs between model parsimony from a factor analytic standpoint and practical utility, such as when a reduced factor structure (e.g., a 2-factor model versus a 10-factor model) results in greater ambiguity about the meaning of dimensions.

A nested factor structure, in which survey results are interpretable at more than one level of specificity (i.e., in terms of first-order dimensions nested within higher-order factors), provides one possible solution. That is, first-order dimensions are specific enough to facilitate clear statements about behavioral norms and values, whereas the higher-order factors are broad enough to allow conceptual linkages to other instruments and theoretical models. Examples of culture surveys that propose a nested structure include Cooke and Lafferty's (1986) *Organizational Culture Inventory* (see also, Cooke & Rousseau, 1988), Woodcock and Francis's

(1989) *Organizational Values Questionnaire*, and the DOCS (Denison & Neale, 1996). An associated disadvantage is the increased complexity of factor analytic support required. Indeed, we are aware of no published evidence to support these models' nested structures; the nested model underlying the DOCS is tested at the conclusion of this manuscript.

Finally, it should be noted that although the inductive approaches used by Xenikou and Furnham (1996) and Detert et al. (2000) identified a fair degree of commonality across surveys, not all culture concepts were redundant, nor were they all operationalized in a consistent manner. Instead, differences in the content and format of instruments reflect the specific purposes and research agendas for which they were developed (Jung et al., 2009; Rousseau, 1990). Specific considerations related to research purpose are described below.

Research Purpose. Ashkanasy et al. (2000) developed a taxonomy based on the purpose of culture surveys and the type of information that is produced. First, instruments were classified as *typing* or *profiling*. Typing instruments categorize organizations into one of several mutually exclusive culture types. For example, the competing values framework differentiates four types of cultures – *clans*, *adhocracies*, *hierarchies*, or *markets* – based on a distinct configuration of values and behaviors (Quinn & Rohrbaugh, 1983). According to Ashkanasy and colleagues, typing instruments are guided by several questionable assumptions and impose important limitations. For example, typing approaches conceptualize culture as a discontinuous variable and assert that same-type organizations possess highly similar cultures. The associated risk is an overly simplistic or even stereotypical view of an organization's culture (Jung et al., 2009). Furthermore, the proposition that culture types are orthogonal has not received empirical support. Instead, research generally demonstrates moderate-to-strong positive correlations among descriptors of culture (Hartnell, Ou, & Kinicki, 2011).

In contrast, *profiling* instruments describe culture in terms of several non-orthogonal dimensions within a profile. No assumptions are made regarding the interrelationships among dimensions, such that organizations can be high or low on each dimension assessed. Mean scores on each dimension provide a nuanced representation of an organization's culture. In addition, some instruments allow researchers to quantify the extent of disagreement around dimension mean scores by computing the between-person standard deviation (Chan, 1998). Here, a parallel can be drawn to the concept of *weak cultures* (i.e., where disagreement about cultural values is high) versus *strong cultures* (i.e., where disagreement about cultural values is low). In this way, it is possible that profile representations of culture are meaningful both in terms of mean levels of endorsement and dissensus among organizational members (e.g., Gordon & DiTomaso, 1992; Martin, 1992; Trice & Beyer, 1993).

Ashkanasy et al. (2000) further grouped profiling instruments according to *descriptive*, *effectiveness*, and *person-culture fit* research purposes. Our focus here is on the former two categories.¹ According to Ashkanasy and colleagues, descriptive measures provide a method for understanding how organizational cultures differ without making any direct statement about how potential differences affect external criteria such as organizational performance. In other words, these instruments “measure values but do not attempt to attach any significance to those results” (p. 138). In contrast, effectiveness instruments were developed to facilitate an understanding of organizational culture that could help to explain differences in the performance of organizations (Sparrow, 2001). Jung et al. (2009) offered a similar distinction between *formative* and *diagnostic* culture assessments. In describing the difference, these authors stated that formative instruments are a mode of “cultural exploration as an end in itself,” whereas the purpose of diagnostic instruments is to “realign existing cultures to characteristics associated with high-

performance organizations” (p. 1090). In other words, diagnostic instruments are concerned with the cultural characteristics of maximally effective organizations and building those characteristics into a prescriptive model with generalized applicability. Alternatively, designers of formative measures either do not make an explicit connection to organizational performance or view the culture-performance relationship as dependent on a host of contextual factors.

The priorities underlying scale construction and validation are also somewhat different for formative and diagnostic instruments. For formative instruments, the primary considerations include the *internal* reliability and validity of the survey (e.g., does the data support the proposed factor structure?) and the adequacy with which the survey covers meaningful descriptive content (e.g., are any key culture concepts omitted?). Diagnostic instruments must also satisfy concerns about internal validity but have the additional priority of demonstrating that the dimensions of culture assessed indeed relate to organizational effectiveness outcomes. As a result, effectiveness measures are generally more focused than descriptive measures, retaining only those dimensions with a strong theoretical or empirical linkage to performance (Ginevičius & Vaitkūnaitė, 2006; van der Post, 1997). A more subtle difference involves the evaluative component underlying some items from diagnostic instruments. Whereas purely descriptive instruments may follow more closely from an anthropological view of cultures as value-neutral (i.e., there is no *good* or *bad* culture), concealing the valence of different responses may be difficult or of less concern when the focus is on performance-relevant aspects of culture (Jung et al., 2009). Although this may appear somewhat tautological, it is important to point out that respondents to diagnostic instruments are asked to indicate the intensity of specific behavioral norms and values within their organization and not whether those norms and values lead to higher effectiveness.

Consequently, these surveys do not necessarily conflate the predictor-criterion space for research that examines the culture-performance link.

Reliability and Validity. A small number of review articles and book chapters have reviewed the reliability and validity evidence available for culture survey instruments (e.g., Rousseau, 1990; Sackmann, 2006; Scott et al., 2003; Walker, Symon, & Davies, 1996). Our purpose here is not to offer a full description of these authors' findings but rather to highlight some general conclusions from the two most comprehensive reviews to date (i.e., Ashkanasy et al., 2000 and Jung et al., 2009). Consistent with our focus on the culture-performance link, we then offer a more detailed review of the reliability and validity evidence for nine surveys that follow an effectiveness profiling (or diagnostic) approach to culture assessment.

Ashkanasy et al. (2000) reviewed a sample of 18 culture surveys using the following evaluative criteria: scale reliability, consensual validity (i.e., evidence of agreement/aggregation), construct validity, and criterion-related validity. Overall, these authors found that reliability and validity evidence were generally lacking for the majority of instruments reviewed. More specifically, no supporting research was located for 10 of the 18 instruments reviewed, and two others possessed minimal evidence of reliability or validity (e.g., consensual validity only). The remaining six instruments possessed two or more types of evidence. Two instruments – the *Organizational Culture Profile* (O'Reilly et al., 1991) and the *Organizational Culture Inventory* (Cooke & Lafferty, 1989) – had research support for all four types of evidence. Among the three effectiveness surveys reviewed, two possessed no evidence of reliability or validity and the third possessed evidence of consensual validity alone. Finally, across all surveys, the type of evidence most commonly reported was criterion-related validity (available for 33% of surveys), and the least commonly reported was consensual validity (available for 22% of surveys).

More recently, Jung et al. (2009) presented a comprehensive review of 70 culture instruments, 48 of which were quantitative survey measures.² These authors' primary focus was to evaluate instruments for use in healthcare contexts (see also, Scott et al., 2003). However, the instruments reviewed ultimately covered a broad range of industry-specific and generalized applications. Reliability and validity evidence was evaluated using the framework by Fitzpatrick, Davey, Buxton, and Jones (1998). At least two study authors judged the adequacy of evidence along each of the following evaluative criteria: internal consistency reliability, test-retest reliability, inter-rater reliability (or agreement), content validity, criterion validity, predictive validity, convergent/ discriminant validity, cross-cultural equivalence, and dimensional validity (i.e., factor analytic support). As an example, evidence of inter-rater reliability was judged *adequate* if inter-rater reliability statistics indicated sufficient levels of agreement, *unclear* if analyses were reported but failed to provide strong support, and *no assessment* if no inter-rater reliability analyses were located.

Overall, results demonstrated that acceptable levels of evidence are available for only a minority of the key reliability and validity criteria by which culture survey are evaluated (Jung et al., 2009). For example, 13% of all judgments (i.e., across all surveys reviewed and all evaluative criteria considered) indicated an *adequate* level of evidence. 60% of judgments indicated that no statistical analyses could be located, and the remaining 27% of judgments fell into the marginal or mixed support category. A subset of instruments was evaluated somewhat more favorably, but these too appear to be in varying stages of preliminary analysis. For example, among the 20 instruments with the strongest supporting evidence, the average number of evidence types judged as adequate was fewer than two of ten ($M = 1.74$, $SD = 1.07$). In comparison, the remaining instruments received less than one adequate judgment on average ($M = 0.59$, $SD = 0.62$). Finally,

across all surveys, the types of evidence most commonly reported were predictive validity (available for 54% of surveys) and internal consistency reliability (available for 46% of surveys). The least commonly reported evidence included test-retest reliability and evidence of convergent/discriminant validity (both available for 10% of surveys).

In summary, prior reviews highlight the need for further methodological research focused on the development and validation of culture surveys in general, as well as more specific attention toward the class of instruments that seek to diagnose the effectiveness of organizational cultures. In particular, Ashkanasy et al.'s (2000) review underscored the dearth of research conducted on effectiveness profiling surveys. However, their review did not include all effectiveness instruments, and several years have since passed providing the opportunity for new studies to accumulate. Jung et al. (2009) also reviewed several effectiveness instruments but offered no conclusions about these surveys as a unique category. Therefore, we provide a brief update of this literature below.

SECTION II: Effectiveness Profiling Instruments – An Update

First, we scanned the literature and conducted electronic database searches to identify effectiveness instruments not included in the reviews by Ashkanasy et al. (2000) and Jung et al. (2009). We also contacted several well-known researchers from this domain and posted inquiries to relevant research databases. In total, we identified nine effectiveness profiling instruments including the six instruments referenced by the prior reviews and three additional instruments located by our searches. Several instruments were ultimately excluded from our review for one or more of the following reasons: (a) the survey authors did not refer to cultural values or beliefs as the target of assessment – for example, the *Voice Climate Survey* by Langford (2009) focused on employee opinions of work practices, (b) the survey authors chose to assess a subset of

performance relevant dimensions but clearly advanced a descriptive purpose for the instrument (e.g., the *Organizational Culture Survey* by Glaser and Zamanou, 1987), or (c) the instrument was broadly concerned with organizational effectiveness but not as the primary purpose of culture assessment – for example, the *GLOBE* measures by House and colleagues were primarily concerned with the interplay of leadership and culture across societal and organizational levels of analysis (see House, Javidan, & Dorfman, 2001). We next searched for and retrieved any published or unpublished studies that reported reliability and validity information for the nine instruments that were retained. Instruments were reviewed according to the evidence types described by Jung et al. (2009).

Detailed findings from our review are summarized in tables. Table 1 describes the structure of the nine instruments and corresponding evidence for reliability. Table 2 describes the evidence for each survey's validity. While our review points to *more* research evidence for effectiveness instruments than was obtained by Ashkanasy et al. (2000), it also points to several problematic trends and remaining gaps. One trend observed for several instruments was the sharp decline (or non-existence) of research following initial publication. Five of nine instruments fit this general pattern, including the three reviewed by Ashkanasy et al. (i.e., *Organizational Beliefs Questionnaire*, *Organizational Values Questionnaire*, and *Organizational Culture Survey*) and two others – the *OASIS Culture Questionnaire* (Cowherd & Luchs, 1988) and the *Organization Assessment Survey* (Usala, 1996a; 1996b). Aside from the two studies we located – one by Muldrow et al. 2002 reporting use of the *Organization Assessment Survey* as part of a culture change intervention with two government agencies and the previously described study of the *Organizational Beliefs Questionnaire*'s convergence with other culture surveys (Xenikou & Furnham, 1996) – research interest in these five instruments seems to have halted altogether.

[Table 1 here]

[Table 2 here]

Research interest appears to have been somewhat stronger for the *Organizational Culture Survey* or the OCS (van der Post et al., 1997). The OCS was developed through an extensive literature review and synthesis of 114 dimensions of culture. A preliminary version of the survey, which included 169 items assessing culture along 15 synthesized dimensions, was administered to 408 employees from eight organizations. Item reliability analyses were used to reduce the total number of items to 97. Retained items were submitted to a factor analysis, which supported the presence of 15 correlated factors (van der Post et al., 1997). A second study by van der Post, de Coning, and Smit (1998) provided evidence of criterion-related validity by demonstrating significant correlations between the OCS dimensions and a dichotomous composite of financial performance among the study sample of 49 organizations. Unfortunately, several important details were excluded from these studies such as statistical details of the factor analysis, information about the sampling methods used and the number and background of survey respondents, and evidence regarding aggregation of culture scores to the organization level in the second study. Furthermore, we question whether the two or three managers sampled per organization in the second study provided a representative assessment of each organization's culture.

We located four studies that have used the OCS since its development, although only one had a methodological focus. Erwee et al. (2001) examined the cross-cultural equivalence of the OCS (developed in South Africa) with a sample of 326 managers from the Australian Institute of Management. Although these authors concluded that the OCS was portable to the Australian context on the basis of reliability and item analyses, an exploratory factor analysis supported a

single-factor solution rather than the 15-factor solution proposed by van der Post et al. (1997). Therefore, the measurement invariance of the OCS was not supported by this research (Vandenberg & Lance, 2000), and future studies need to establish the number and meaning of the dimensions assessed. More recent studies by Strydom and Roodt (2006) and Liebenberg (2007) used the OCS to test how employee satisfaction and personality affect perceptions of organizational culture. These studies provide some evidence of the discriminant validity of the OCS, although both conducted analyses at the individual level instead of aggregating to the organization level. Also at the individual level of analysis, a study by Rieker (2006) linked perceptions of organizational culture to the quality of formal mentorship relationships experienced within two United States Airforce organizations.

Clearly, additional research is needed to establish the construct validity of the OCS and determine the number and meaning of the latent factors assessed. The high internal consistency ($\alpha = .99$) and single-factor solution reported by Erwee et al. (2001) call into question whether multiple concepts are indeed measured. Additional research is also needed to replicate and extend the model's predictive validity with a larger and more representative sample than that reported by van der Post et al. (1998). Future research also needs to establish the instrument's measurement properties and test convergent/discriminant validity at the aggregate level. With the study by van der Post et al. (1998) as the only exception, the studies reviewed here did not assess *organizational* culture but rather individuals' perceptions of organizational culture. Finally, our review indicates that while the OCS has potential as a practically useful assessment of performance relevant aspects of culture, we find little evidence linking the instrument to a particular theoretical perspective or paradigm within the culture literature. Thus, the measure's contribution to the advancement of knowledge in this domain is unclear at this time.

Two instruments reviewed were in early stages of development and validation. The *Questionnaire of Dimensions of Organizational Culture* was developed inductively by Ginevičius and Vaitkūnaitė (2006) on the basis of a comprehensive review of other instruments and the dimensions that correlated with effectiveness outcomes in prior studies. The authors then content analyzed the dimensions extracted from their review and produced a final instrument with 12 dimensions and 48 items. A preliminary factor analysis based on individual respondents from 23 organizations supported a four-factor model, and correlational analyses provided mixed support for the four factors predicting subjective performance and employee satisfaction outcomes. Subsequent studies by Aydin and Ceyla (2008, 2009) reported significant positive correlations between overall culture and employee satisfaction, and between dimensions of culture and perceived organizational effectiveness. The other recently developed instrument, the *Value Performance Index* (Schönborn, 2010), was constructed to assess the three levels of culture specified by Schein (1992). The initial survey with 135 items was administered to 2,873 managers from 46 companies in three European countries. Based on an exploratory factor analysis, 13 dimensions were identified and subsequently labeled. Correlational analyses demonstrated significant predictive relationships with a dichotomous composite index of financial performance for 7 of the 13 dimensions.

Although additional studies are certainly needed before these instruments can be fully evaluated, our review highlights their potential as predictive tools and identifies several key challenges that warrant further attention. For example, while both instruments present preliminary evidence of criterion-related validity, issues with the levels-of-measurement (e.g., individual-level rather than organization-level analysis), criteria (e.g., satisfaction instead of organizational performance outcomes), and sampling methods that were used (e.g., manager-

only samples) need to be addressed in future research. Future studies should also seek to clarify the theoretical basis for the proposed measures and their internal dimensional structures. This is particularly evident for the instrument by Ginevičius and Vaitkūnaitė (2006), which was developed inductively and has produced inconsistent factor structures across studies.

The final instrument in our update is the *Denison Organizational Culture Survey* or DOCS (Denison & Neale, 1996). Our review indicates that the DOCS is perhaps the only effectiveness instrument that has advanced beyond the initial stages of scale development. This assertion is based on the amount of research interest that the DOCS has generated, as well as the comparatively higher number of methodological investigations. Due to the high volume of unpublished dissertations and technical reports that have used the DOCS – we count over 30 dissertations alone – our review could not be comprehensive within the space limitations of this manuscript, and we have therefore focused on being exhaustive of the published research while selectively incorporating a small number of key unpublished studies (e.g., Boyce, 2010). Furthermore, we feel that the increased onus on published research is consistent with the current state of research on the DOCS and the natural progression within any research literature toward a focus on those studies that have survived the crucible of peer review.

In the pages that follow, we introduce the DOCS in somewhat greater detail by describing its development and what differentiates it from other effectiveness instruments and theoretical perspectives on organizational effectiveness. In line with the other instruments reviewed, we then summarize available evidence for reliability and validity and describe the key limitations of prior studies and remaining methodological challenges. The final section of our manuscript describes three specific hurdles that are widely applicable to culture surveys in the

effectiveness profiling category, and provides an illustration of how research can address these challenges based on archival data that was collected using the DOCS.

The Denison Organizational Culture Survey (DOCS)

The development of the DOCS reflects a slightly different approach than the other effectiveness instruments reviewed here and a stronger grounding in organizational culture theory. The survey instrument and corresponding measurement model was constructed in tandem with the development of a theory of cultural effectiveness that concentrates on four key traits as drivers of organizational performance: *involvement*, *consistency*, *adaptability*, and *mission* (Denison & Mishra, 1995 - see Table 3 for trait definitions). The traits were extracted from a line of research by Denison and colleagues combining qualitative and quantitative methods to examine the cultural characteristics of high and low performing organizations (Denison, 1984; Denison, 1990; Denison, Haaland, & Goelzer, 2003; Denison & Mishra, 1995; Fey & Denison, 2003). Together, these studies indicated that in general, the highest performing organizations find ways to empower and engage their people (involvement), facilitate coordinated actions and promote consistency of behaviors with core business values (consistency), translate the demands of the organizational environment into action (adaptability), and provide a clear sense of purpose and direction (mission).

[Table 3 here]

The meaning and importance of similarly defined concepts have been described by a number of organizational scholars interested in the characteristics of high performance organizations (Gordon & DiTomaso, 1992; Katz & Kahn, 1978; Kotter & Heskett, 1992; Lawler, 1986; Martin, 1992; Mintzberg, 1989; Saffold, 1988; Schein, 1992; Selznick, 1957; Spreitzer, 1995, 1996). However, the underlying model is somewhat more complex than the

straightforward amalgamation of popular concepts. The traits are organized into a framework that draws on contemporary theories of the dynamic tensions underlying organizational functioning and effectiveness (Denison, Hooijberg, & Quinn, 1995; Denison & Spreitzer, 1991; Quinn & Cameron, 1988). These tensions reflect the fundamental pushes and pulls experienced by the organization in response to its operating environment (Lawrence & Lorsch, 1967). The primary tensions represented by the model involve the competing demands of an external versus internal focus and between stability and flexibility. The traits assessed by the survey provide a framework for understanding how (and to what extent) organizational cultures balance these seemingly contradictory demands. For example, *mission* and *consistency* provide cultural support for stability, whereas *adaptability* and *involvement* provide cultural support for flexibility (see Denison, 2001; Denison & Mishra, 1995). The framework leans on similar conceptual origins as the competing values framework advanced by Quinn and colleagues (Quinn & Cameron, 1988; Quinn & Rohrbaugh, 1981), but maintains a few important differences. As already mentioned, one important difference is that the competing values framework has spawned a range of typological culture assessments, whereas the DOCS uses a profile approach. Beyond the measurement implications of this difference, the choice to use a profile approach also reflects an important theoretical distinction. Unlike the competing values framework, a focus of the work by Denison and colleagues is on the *balance* or mixture of cultural elements. Their model proposes that it is not only possible for an organization to display strong internal *and* external values (as well as stability *and* flexibility), but that it is of critical importance to the organization's long-term effectiveness. Thus, the model proposes that the most effective organizations are those that display "full" profiles as indicated by high levels of all four traits (Denison, 1990; Gillespie et al., 2008).

Another important difference is the second-order measurement model. Each trait is assessed by three indexes that operationalize a specific facet of the trait at the level of observable and measurable behavioral norms and values. The structuring of indexes within traits and corresponding definitions are shown in Table 3. The survey consists of 60 items or 5 items per index. As a result of the second-order structuring, information about culture is provided at two levels of abstraction. The indexes are designed to present culture information around 12 clearly understandable and actionable content areas (e.g., team orientation, customer focus, goals and objectives), whereas the traits organize these concepts into broader principles that are portable across organizational contexts and support the theoretical grounding and relevance of the model and instrument (Denison & Mishra, 1995).

Demonstrating empirically the linkages between the DOCS and organizational effectiveness outcomes (i.e., predictive validity) was central in the DOCS early development as well as subsequent studies that added or modified the concepts included in the survey. The earliest of machinations focused mainly on the participative or *bottom-up* aspects of culture and management and establishing their connection to bottom-line financial performance metrics such as organizations' return on assets (Denison, 1984). These concepts evolved into the *involvement* and *consistency* traits in subsequent iterations. Denison (1990) described the counterbalancing role of top-down and externally focused aspects of highly effective cultures, and these were subsequently incorporated in the shortened 8-item version of the DOCS reported by Denison and Mishra (1995). This study laid much of the theoretical foundation for the traits and dynamic tensions model and presented an empirical analysis based on data from 764 organizations. In addition to demonstrating the predictive validity of the four traits for a variety of performance indicators, this study provided initial evidence that the culture traits affect different aspects of

organizational effectiveness. Specifically, profitability outcomes were best predicted by the culture traits that support stability (i.e., mission and consistency), whereas growth outcomes were best predicted by the culture traits that support flexibility (i.e., involvement and adaptability). This pattern is consistent with the idea that stability contributes to efficient and productive systems while flexibility allows the organization to change and grow in tandem with the demands of the external marketplace. That the traits contribute to the organization's survival in different ways is an important finding not only from an understanding perspective, but also as further justification for retaining all four traits in the measurement model. No other effectiveness profiles that we are aware of have established differential prediction of effectiveness criteria within a multi-concept framework.

Subsequent studies have examined the generalizability of the DOCS predictive validity to other effectiveness outcomes and also across industry and national-culture boundaries. For example, studies by Gillespie et al. (2008) and Boyce (2010) demonstrated the positive linkages between scores on the DOCS and levels of customer satisfaction and sales among franchise car dealerships and home construction firms. The former study reported a cross-sectional analysis in which culture trait scores accounted for between 11 and 28% of the variance in customer satisfaction, and the latter study provided evidence of time-lagged positive effects of overall culture scores on customer satisfaction 1 year afterwards and the number of cars sold 2 years afterwards. Studies by Fey and Denison (2003), Denison et al. (2003, 2004), and Bonavia et al. (2009) examined the survey's psychometric characteristics and predictive validity with organizational samples from nine non-US countries. For example, comparisons between Asian organizations and "the rest of the world" indicated similar mean levels and predictive patterns between the indexes and effectiveness outcomes, although the authors also provided examples of

how the expression of specific values and behavior can vary somewhat across contexts (Denison et al., 2003). A second study comparing US- to Russian-based organizations, demonstrated the importance of all four traits in both contexts but also indicated that flexibility and involvement were more closely associated with overall perceptions of effectiveness (and mission less closely associated) among Russian organizations (Fey & Denison, 2003). Together, these studies provide initial evidence that the DOCS can be translated to other languages (e.g., Spanish in Bonavia, Gasco, & Tomás, 2009) and applied cross-culturally with similar support for reliability and validity of the model and survey items. However, further research that directly tests the measurement and structural invariance of the DOCS across cultures would be useful to provide a more complete understanding of these similarities and differences (e.g., see Steenkamp & Baumgartner, 1998).

While it is apparent that there is a stronger empirical foundation for inferring the predictive validity of the DOCS than the other instruments in our review, we also identified several limitations of this research as well as remaining “gaps” in the other types of evidence considered (e.g., test-retest reliability). First, as already noted, several of the studies reviewed above did not apply the current 60-item DOCS. Thus, the culture-to-performance linkages demonstrated by Denison and Mishra (1995) and Fey and Denison (2003) using shorter versions of the survey establish the importance of the concepts in the model and help support the evidence for the instrument’s predictive validity. Similarly, the reliability and factor analytic evidence provided by Fey and Denison should be qualified as indirect forms of support. Gillespie et al. (2008) and Kotrba et al. (in press) have presented the first evidence of the second-order factor solution, providing good evidence that the second-order model is the most appropriate fit to the data.

The study by Gillespie et al. (2008) also shares the second main limitation identified by our review – some studies’ focus on single-industry samples or on analyzing predictive relationships in a way that segments the overall sample by industry. First, it is important to clarify the context in which we view this as a limitation. From a research perspective, comparative analyses within a particular industry or organizational setting are attractive due to the potential elimination of industry-level confounds and an improved understanding of the research in context. From another perspective, the diagnostic approach discussed here is founded on the generalizability of the traits assessed and their relative universality as drivers of performance. Thus, demonstrating that the predictive relationships generally hold across industry, organizational setting, and national-cultural boundaries is an important piece of validity evidence, not only for the DOCS, but also for the diagnostic approach to organizational culture assessment in general. By the same means, identifying the bounds of these relationships would help to clarify when a diagnostic versus a contingency-based approach (e.g., descriptive profiles) to assessment is more appropriate.

A final noteworthy limitation involves the use of single-respondent samples or manager-only samples used in some of the studies conducted on the DOCS (e.g., Denison & Mishra, 1995; Fey & Denison, 2003). Although a scan of the literature reveals that reliance on one or a few “trusted” respondents (e.g., managers with extensive organizational experience) is not an infrequent practice (e.g., Birkinshaw, Hood, & Jonsson, 1998; Delaney & Huselid, 1996; Delery & Doty, 1996; Geringer & Hebert, 1989; Lee & Beamish, 1995; Shaw et al., 1998), it raises obvious questions about the representativeness of these samples as well as disallowing consideration of agreement and the possibility of cultural fragmentations.

Based on the preceding discussion, we identified three key challenges for which diagnostic instruments in general have demonstrated weak or incomplete supporting evidence. Instruments must first pass a psychometric test to make certain that individual respondents can discern the underlying structure proposed by the theory. Second, the respondents within each organization must show a high level of agreement in order to claim that organizational characteristics are being measured. And third, the organizational level characteristics must show a close link to the organizational level outcomes suggested by the model or in the case of diagnostic instruments, organizational effectiveness. Although these challenges do not cover the full scope of factors that must be satisfied, they are in our view foundational considerations and a good starting point for further methodological scrutiny. In the section below, we evaluate the DOCS with respect to each of these considerations, and in so doing, offer closure to some of the limitations and gaps that were identified above.

SECTION III: Empirical Illustration of Three Challenges

In this section of the paper, we present an empirical illustration, addressing the challenges identified in *Section II*. The analyses presented here are based on archival data from 160 privately held companies from a variety of industries and geographic locations. These organizations voluntarily completed the DOCS between 1997 and 2001. The organizations in the sample were generally large, employing on average nearly 60,000 employees and generating two billion U.S. dollars in sales annually, although several smaller private firms were also included. In total, 35,474 individuals completed the DOCS, with at least 25 respondents sampled per organization. Response rates ranged from 48% to 100%, with an overall mean across organizations equal to 60%, which is well within the range recommended in the management literature (e.g., Baruch, 1999). The specific samples drawn from each of these organizations

varied to a certain extent, with some organizations surveying all members and others surveying within and across specific divisions, geographic locations, and levels of the organization. In all cases, consultants worked with the organizations to ensure that the survey methodology produced a sample that was representative of the organizations' employee populations. Table 4 summarizes the organizational characteristics and demographics for individuals in the final sample.

[Table 4 here]

Surveys with missing data on any of the sixty items comprising the DOCS were excluded from this analysis.³ Responses to all items were measured on a 5-point Likert-type scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Respondents also rated the organization on the following six dimensions of performance relative to similar companies: sales/revenue growth, market share, profitability/ROA, quality of goods and services, new product development, and employee satisfaction. These items were rated on a 5-point Likert-type scale ranging from 1 = *low performer* to 5 = *high performer*. Although perhaps less attractive than objective indicators of performance, past researchers have demonstrated that subjective measures of organizational effectiveness can be useful proxies for objective sales or profitability data (e.g., Baer & Frese, 2003; Guthrie, 2001; Wall et al., 2004).

Challenge 1: Evidence of Internal Reliability and Validity. Two key pieces of evidence were examined to determine the extent to which the parameters implied by the theoretical model are supported by the obtained covariance matrix. First, we examined the internal consistency reliability of the twelve indexes to determine if the 5-item subsets hang together as internally reliable scales. Second, we conducted a series of confirmatory factor

analyses to determine if the pattern of relationships between the observed variables and latent traits support the hierarchical structure of the proposed model.

[Table 5 here]

Table 5 presents the results for the first step in the analysis. Included in this table are the item means, standard deviations, and item-total correlations, as well as the internal consistency estimates for the indexes. Coefficient alphas for the twelve indexes ranged from .70 to .85 indicating an acceptable level of internal consistency for all indexes (Nunnally, 1978). Item-total correlations exceeded .50 for over two-thirds of the 60 items in the survey. Item 15 from the *capability development* index (“Problems often arise because we do not have the skills necessary to do the job”) showed an unusually low item-to-total correlation of .23. This negatively-worded item was retained because (a) the alpha coefficient for the index itself still reaches an acceptable level of .70, and (b) the item was judged to have adequate content validity based on its fit with the definition provide for this index. Nevertheless, researchers using this scale in the future may wish to exclude this item. Table 6 presents the correlations between indexes. Values ranged from .45 to .74 with an overall mean correlation of .59. Together, these results support the internal consistency reliability of the indexes as 5-item subscales and also indicate that individuals perceive moderate to strong relationships between aspects of culture.

[Table 6 here]

[Figure 1 here]

Next a second-order confirmatory factor analysis was tested using the 60 items from the DOCS as observed variables, the 12 indexes as first-order factors, and the 4 higher order traits as second-order factors. Figure 1 shows a visual depiction of the second-order model that ultimately provided the best fit to the data. These results are described first in terms of the pattern of factor

loadings and inter-trait correlations and then in terms of accompanying indices of fit for the models tested. Table 7 displays the item loadings on their respective factors. Loadings generally fell in the .60 to .75 range, indicating considerable shared variance within those items intended to measure the same underlying concepts. Tables 8 and 9 present index loadings on their respective traits and between-factor correlations among the second order trait factors. These values range from the low .70s to the mid-.90s, indicating overlap in the variance explained by first-order factors and strong relationships between second-order factors.

[Table 7 here]

[Table 8 here]

[Table 9 here]

Model fit was evaluated using several fit indices, including the root mean squares error of approximation or RMSEA (Hu & Bentler, 1998), the goodness-of-fit index or GFI (Jöreskog & Sörbom, 1989), the normed fit index or NFI (Bentler & Bonnett, 1980), and the comparative fit index or CFI (Bentler, 1990). The RMSEA is recognized as the most sensitive index to models with mis-specified solutions and is indicative of close fit at values lower than .05 (Browne & Cudeck, 1993; Hu & Bentler, 1998; MacCallum, Browne, & Sugawara, 1996). The GFI is analogous to a squared multiple correlation, assessing the shared variance between the covariance matrix implied by the specified model and the observed covariance matrix. The NFI compares the proposed model to a model where no interrelationships between the variables are allowed. And finally, the CFI takes sample size into account when comparing the hypothesized model with the independence model while avoiding some of the problems found with other incremental indices (i.e., indices for comparing multiple models; Bentler, 1992). For the GFI,

NFI, and CFI values greater than .90 have generally been recommended as indicative of acceptable fit (Bentler & Bonett, 1980; Hu & Bentler, 1998; Raykov & Marcoulides, 2000).

The chi-square and fit indices for the second-order model were: $\chi^2(1692) = 122,715.83$, $p < .01$, GFI = .88, NFI = .98, CFI = .98, and RMSEA = .04. In general, these values indicate good fit for the second-order model, with RMSEA, NFI, and CFI values meeting recommended guidelines. Although GFI was lower than the recommended cut-off, the collection of fit indices as a set suggest that the specified model closely approximates the observed pattern of relationships between the items making up the culture assessment. However, in order to determine whether the second-order model provides a better fit than alternative representations of the data, several competing models were tested and compared. The first alternative model excluded the 12 intermediate first-order factors (i.e., the culture indexes), such that the 60 items were forced to load directly onto the four latent traits. The second alternative model was more restrictive still, forcing all 60 items to load directly onto a single latent factor, thus eliminating the four culture traits proposed by the model. The chi-square and fit indices for the first alternative model specifying four first-order factors were: $\chi^2(1704) = 157,276.98$, $p < .01$, GFI = .85, NFI = .98, CFI = .98, and RMSEA = .05. Comparison to the second-order model indicated significantly worse fit as evidenced by a significant change in chi-square ($\Delta\chi^2(12) = 34,561.15$, $p < .001$), higher RMSEA (.05 versus .04) and lower GFI (.85 versus .88). The second alternative model resulted in a further decline in model fit and indicated poor fit overall: ($\chi^2(1710) = 173,663.25$, $p < .01$, GFI = .78, NFI = .78, CFI = .79, and RMSEA = .06). These results indicate that the second-order hierarchical model yields a significantly better reproduction of the observed covariance matrix than the alternative model specifications, and therefore lends support

to the conceptualization of the 60 items as representing twelve indexes that group into four higher-order culture traits.

Challenge 2: Evidence for Aggregation to the Organizational Level. Before aggregating individual ratings of culture into an organizational-level variable, it is first necessary to show that those ratings are sufficiently homogeneous (Dansereau & Alutto, 1990; Klein et al., 2000). There are a number of statistical methods for assessing the homogeneity of responses within groups, including indices of agreement such as r_{wg} for single item measures or $r_{wg(j)}$ for multi-item measures (James, Demaree, & Wolf, 1984) and indices of reliability such as ICC(1) and ICC(2) (Shrout & Fleiss, 1979). Agreement was evaluated by computing $r_{wg(j)}$ for each organization as a function of the five items comprising each index of the DOCS and based on deviation from the uniform response distribution. Values greater than .70 have generally been recognized as sufficient response consistency to justify aggregating individual responses to the group level (e.g., Klein, Griffin, Bliese, et al., 2000). ICC(1) and ICC(2) were computed as omnibus indexes of intra-organizational reliability also at the cultural index level. ICC(1) indicates the proportion of total variance attributable to organization membership, and ICC(2) indicates the extent to which organizations are reliably differentiated by the measure (Bryk & Raudenbush, 1992). *F*-values from random effects oneway ANOVAs provide a statistical significance test for the ICC(1) values.

[Table 10 here]

The agreement and reliability indices for each index of the DOCS are shown in Table 10. Mean $r_{wg(j)}$ across organizations and culture indexes ranged from .85 to .89. As shown in Figure 2, the $r_{wg(j)}$ values observed for individual organizations ranged from the mid-.70s to the mid-.90s, well within the range of previous guidelines for establishing sufficient evidence of

agreement. ICC(1) ranged from .06 to .10 across culture indexes indicating that between 6 and 10% of the variance in culture ratings can be accounted for by organization membership. Corresponding *F*-values demonstrated that this proportion of variance was statistically significant in all cases ($p < .001$). ICC(2) ranged from .93 to .96 demonstrating high reliability for the organization-level means on each index. As a whole, these results support the aggregation of individual ratings of culture to the organization level and demonstrate that organizations are reliably differentiated by the DOCS.

[Figure 2 here]

These results also suggest that the positioning of inter-rater agreement in the literature as a *threshold* that must be reached in order to justify aggregation is somewhat misguided. Our results suggest that nearly all of the organizations that we have surveyed meet the minimal criteria to justify aggregation, so it doesn't really discriminate very well. But despite the fact that all of the organizations made it over the hurdle, there are still significant variations between the organizations. Thus it may be quite possible that internal consistency is more important to consider as a *variable* rather than as a *threshold*. In addition, as Kotrba et al. (in press) have shown, internal consistency can be an important measure of culture strength that is quite closely linked to performance.

Challenge 3: Evidence of Criterion-related Validity. As reinforced throughout this manuscript, demonstration of criterion-related validity holds special importance for the category of measures described as diagnostic culture instruments. In this section, we evaluate the criterion-related validity of the indexes from the DOCS as organization-level predictors of several subjectively rated indicators of organizational performance. More complete analyses of objective criterion-related validity have been presented in a number of studies over the years. A

complete review of these studies is beyond the scope of this paper, as is the introduction of a new criterion-based validity study based upon objective indicators of performance.

Descriptive statistics for the aggregated culture index scores are shown in Table 10. Correlations between the culture indexes and ratings of sales growth, market share, profitability, quality of products and services, new product development, and employee satisfaction are presented in Table 11. As the table shows, most of these validity coefficients were statistically significant at the .01 level and had magnitudes of at least .30. The strongest relationships were observed between aspects of culture and employee satisfaction, with correlations ranging from .42 to .79 (average $r = .63$). Slightly weaker correlations were observed for organizational ratings of new product development (average $r = .37$), quality (.36), sales growth (.26), and profitability (.25). The weakest relationships were observed for culture predicting ratings of market share, with correlations ranging from .04 to .26 (average $r = .13$). When the six performance measures were combined into a unit-weighted composite, correlations between the culture indexes/traits and performance ratings ranged from .44 to .68 (average $r = .58$). Overall, these results support the positive link between the cultural indexes measured by the DOCS and aspects of organizational effectiveness measures.

[Table 11 here]

As in past research, these results also show that some features of organizational culture are better predictors of specific effectiveness criteria than others (Denison & Mishra, 1995; Gillespie et al., 2007). The pattern of correlations observed here indicates that the internally focused traits *involvement* and *consistency* are generally better predictors of operating performance measures such as quality and profitability, whereas the externally focused traits *mission* and *adaptability* are generally better predictors of sales growth. Similarly, *mission* - and

particularly, *strategic direction and intent* - was the only significant predictor of market share. Other noteworthy trends were that new product development was least strongly correlated with the *consistency* trait, and *involvement* was clearly the strongest predictor of employee satisfaction. Together, these findings indicate that the aspects of culture assessed within the DOCS likely contribute to overall organizational effectiveness in complementary and slightly different ways.

Discussion

Perspectives on how best to measure the cultures of work organizations have shifted over time (Martin et al., 2006). With certain limitations acknowledged, researchers and practitioners alike have adopted surveys as a quantitative tool for understanding the normative behaviors and values that characterize an organization's culture. Provided growing evidence of the role that culture can play in bottom-line performance (Sackmann, 2011), culture surveys also complement a range of management purposes within the broad domains of organizational development and change by serving as a means of diagnosis, feedback, and benchmarking. Effectiveness profiling instruments represent the category of culture surveys most directly aligned with these applications (Ashkanasy et al., 2000). Below, we briefly revisit some of the key findings from our review and highlight several directions for future research.

One purpose of our review was to determine what progress has been made in the development and validation of instruments in this category. Perhaps, the most troubling finding was the lack of continued research interest documented for five of the nine instruments reviewed, including all three surveys previously reviewed by Ashkanasy et al. (2000). Despite that these particular instruments have not been reported extensively in the literature, many of their concepts were borrowed or adapted in the construction of the four remaining instruments. Two in

particular have followed the approach of synthesizing the dimensions of other culture surveys (i.e., Ginevičius & Vaitkūnaitė, 2006; van der Post et al., 1997). This too – that is, the development of new instruments on a purely inductive basis – is indicative of a somewhat problematic trend. Although re-organizing old concepts into new frameworks can and should lead to improved prediction over multiple iterations (e.g., by dropping the concepts that are either unrelated or only weakly related to organizational performance), we question whether such instruments can “pass” tests of construct validity without a clear guiding theory or contribute in a meaningful way to scholarly *understanding* in this domain. In contrast, the approaches taken in the remaining two instruments represent a blend of inductive and theory-driven components (i.e., Denison & Neale, 1996; Schönborn, 2010). Consequently, it is possible to demonstrate whether the response data as gathered vis-à-vis the items in the survey support the internal structuring of concepts as specified by the theory. If so, higher confidence can be lent to the results and validity of the model and data. Thus, survey developers should seek to make explicit the connections between their preferred measurement models and the theories from which they follow, particularly as dominant theoretical perspectives emerge over time.

Although attention to theory was not always apparent in the construction of diagnostic instruments, criterion-related validity on the other hand was always a central concern, despite that evidence was in some cases unavailable and in some others, weak. One issue that is paramount is the sort of research evidence that is needed to demonstrate a strong linkage between the aspects of culture assessed and organizational performance outcomes. Along these lines, several of the limitations described herein reflect the usual criticisms that accompany developing research areas such as the need for longitudinal evidence, more and better effectiveness criteria, larger and more representative samples of organizations, and so on. Of

course, researchers should be encouraged to address these issues in future studies, and doing so would go a long ways to strengthen the evidence presented for observed relationships between survey ratings of culture and performance outcomes. At the same time, a handful of the other limitations identified reflect specific and perhaps underappreciated considerations in the validation of effectiveness instruments. We focus on two below.

Generalizability Considerations

One example involves the generalizability of predictive relationships. Unlike descriptive approaches to measuring culture, the diagnostic approach generally leads to an inference about cultural effectiveness without necessary consideration of possible contingency factors. Consequently, evidence for effectiveness instruments ought to demonstrate that culture-performance relationships are robust across the range of settings for which the instrument is intended. Support can be demonstrated through a variety of strategies including use of diverse validation samples, as illustrated in our empirical evaluation of the DOCS, or by applying more sophisticated techniques that test measurement and structural equivalence across specific settings such as two or more national-cultures or industries.

The stream of research cited for the DOCS illustrates some of the complexity involved in cross-cultural comparative work. For example, these studies have demonstrated that while the culture concepts assessed retain similar meanings across national settings, the specific manifestations of these concepts can differ (e.g., Denison et al., 2003). This suggests that effectiveness instruments may need to be versatile enough to accommodate information about culture at varying levels of specificity. The second-order framework underlying the DOCS provides one possible solution. The same line of research also indicates that while all four traits contribute to organizational performance across national-cultures, the rank-ordering of traits in

terms of the magnitude of predictive relationships also varies somewhat. Thus, generalizability is clearly a multifaceted issue and one which requires a programmatic research effort in order to fully elucidate the boundaries.

Multilevel Considerations

Another set of issues has to do with the shift from individuals to organizations as the primary unit of analysis. Several studies cited in our review committed an atomistic fallacy by inferring organization-level relationships on the basis of regressions or correlations with individuals (Diez-Roux, 1998). Unfortunately, this type of evidence does little to substantiate criterion-related validity for an organizational assessment because culture-performance relationships among organizations may not reflect the individual-level relationships. Instead, an appropriate test involves examining the relationships between aggregated culture ratings and matched effectiveness criteria. There are a number of methods available for handling multilevel data including the approach illustrated here using the DOCS (see Bliese, Halverson, & Schriesheim, 2002). Whichever analysis strategy researchers adopt, two main points of interest involve first demonstrating that individual ratings can be used to represent the overall culture of organizations in a valid and reliable manner, and second testing the culture-performance linkages at the organization level.

Conclusion

In conclusion, our review identified a total of nine survey instruments whose objective is to diagnose organizational cultures by assessing those values and behavioral norms most directly related to organizational performance. The review indicated a number of problematic trends and remaining gaps in the types of reliability and validity evidence that support these instruments, underscoring the need for additional methodological research. Each of the “active” instruments

reviewed appear to be in varying stages of development and evidence gathering, while research on several others appears to have fallen off. Our review also identified the DOCS as the most well-researched effectiveness instrument to date. We therefore provided a more in-depth discussion of the background, strengths, and limitations that differentiate this particular instrument from other culture effectiveness surveys. And finally, our empirical illustration helps to clarify several key challenges extracted from our review, while simultaneously “closing” some of the remaining gaps found for the DOCS.

Endnotes

¹ Readers interested in fit applications of culture surveys are referred to Chatman and O'Reilly's work on the *Organizational Culture Profile* (Chatman, 1991; O'Reilly et al., 1991).

² A full report of these authors' search methods and findings are presented in a publically available compendium (<http://www.scothub.org/culture/instruments.html>).

³ The full set of items comprising the DOCS has been published elsewhere (see Denison et al., 2003).

References

- Alvesson, M. (2011). Organizational culture: Meaning, discourse, and identity. In Ashkanasy, N., Wilderom, C., & M. Peterson (Eds.), *The handbook of organizational culture and climate*, 2nd edn., p. 11-28. Thousand Oaks, CA: Sage Publications.
- Ashkanasy, N., Broadfoot, L., & Falkus, S. (2000). Questionnaire measures of organizational culture. In Ashkanasy, N., Wilderom, C., & M. Peterson (Eds.), *The handbook of organizational culture and climate*, p. 131-162. Thousand Oaks, CA: Sage Publications.
- Aydin, B., & Ceylan, A. (2009). The role of organizational culture on effectiveness. *Ekonomika A Management*, 3, 33-49.
- Aydin, B., & Ceylan, A. (2008). The employee satisfaction in metalworking manufacturing: How do organizational culture and organizational learning capacity jointly affect it? *Journal of Industrial Engineering and Management*, 1, 143-168.
- Baer, M. & Frese, M. (2003). Innovation is not enough: Climates for initiative and psychological safety, process innovations and firm performance. *Journal of Organizational Behavior*, 24, 45-68.
- Baker, W. (2000). *Achieving Success through Social Capital*. San Francisco, CA: Jossey-Bass.
- Barney, J. (1986). Organizational culture: Can it be a sustained source of competitive advantage? *Academy of Management Review*, 11, 656-665.
- Baruch, Y. (1999). Response rate in academic studies - A comparative analysis. *Human Relations*, 52, 421-438.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the *Bulletin*. *Psychological Bulletin*, 112, 400-404.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Birkinshaw, J., Hood, N., & Jonsson, S. (1998). Building firm-specific advantages in multinational corporations: The role of subsidiary initiative. *Strategic Management Journal*, 19, 221-241.
- Bliese, P. D., Halverson, R. R., & Schriesheim, C. A. (2002). Benchmarking multilevel methods: Comparing HLM, WABA, SEM, and RGR. *Leadership Quarterly*, 13, 3-14.

- Block, L. (2003). The leadership-culture connection: An exploratory investigation. *Leadership and Organization Development Journal*, 24, 318-334.
- Bonavia, T., Gasco, V. J., & Tomás, D. B. (2009). Spanish adaptation and factor structure of the Denison Organizational Culture Survey. *Psicothema*, 21, 633-638.
- Boyce, A. S. (2010). *Organizational climate and performance: An examination of causal priority*. Unpublished dissertation, Michigan State University, Lansing, MI.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. Bollen & J. Long (Ed.), *Testing structural equation models*, p. 136-162. Newbury Park, CA: Sage.
- Bryk, A. S., & Raudenbush, S. W. (1989). Methodology for cross-level organizational research. In S. Bacharach (Ed.), *Research in the sociology of organizations*, p. 233-273. Greenwich, CT: JAI Press.
- Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis. *Journal of Applied Psychology*, 83, 234-246.
- Chatman, J. (1991). Matching people and organizations: Selection and socialization in public accounting firms. *Administrative Science Quarterly*, 36, 459-485.
- Cooke, R. A., & Lafferty, J. C. (1989). *Organizational culture inventory*. Plymouth, MI: Human Synergistics.
- Cooke, R., & Rousseau, D. (1988). Behavioral norms and expectations: A quantitative approach to the assessment of organizational culture. *Group and Organizational Studies*, 13, 245-273.
- Cowherd, D. M., & Luchs, R. H. (1988). Linking organization structures and processes to business strategy. *Long Range Planning*, 21, 47-53.
- Dansereau, F., and Alutto, J. (1990). Levels of analysis issues in climate and culture research. In B. Schneider (Ed.), *Organizational Climate and Culture*, p. 193-236. San Francisco: Jossey-Bass.
- Delaney, J. T., & Huselid, M. A. (1996). The impact of human resource management practices on perceptions of organizational performance. *Academy of Management Journal*, 39, 949-969.
- Delery, J., & Doty, D. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of Management Journal*, 39, 802-835.

- Delobbe, N., Haccoun, R. R., & Vandenberghe, C. (2002). *Measuring core dimensions of organizational culture: A review of research and development of a new instrument*. Unpublished manuscript, Universite catholique de Louvain, Belgium.
- Denison, D. R. (1984). Bringing corporate culture to the bottom line. *Organizational Dynamics*, 13, 4-22.
- Denison, D. (1990). *Corporate culture and organizational effectiveness*. New York: John Wiley & Sons.
- Denison, D. (1996). What IS the difference between organizational culture and organizational climate? A native's point of view on a decade of paradigm wars. *Academy of Management Review*, 21, 619-654.
- Denison, D. R. (2001). Organizational culture: Can it be a key lever for driving organizational change? In C. Cooper, Cartwright, S., & Earley, P. (Eds.), *International handbook of organizational culture and climate*. Chichester: Wiley, 347-372.
- Denison, D. R., Haaland, S., & Goelzer, P. (2003). Corporate culture and organizational effectiveness: Is there a similar pattern around the world? *Advances in Global Leadership*, 3, 205-227.
- Denison, D., Hooijberg, R., & Quinn, R. (1995). Paradox and performance: Toward a theory of behavioral complexity in managerial leadership. *Organization Science*, 6, 524-540.
- Denison, D. R., Lief, C., & Ward, J. L. (2004). Culture in family-owned enterprises: Recognizing and leveraging unique strengths. *Family Business Review*, 17, 61-70.
- Denison, D., & Mishra, A. (1995). Toward a theory of organizational culture and effectiveness. *Organizational Science*, 6, 204-223.
- Denison, D. R., & Neale, W. S. (1996). *Denison organizational culture survey*. Aviat, Ann Arbor, MI.
- Denison, D. R., & Spreitzer, G. (1991). Organizational culture and organizational development: A competing values approach. *Research in Organizational Change and Development*, 5, 1-21.
- Detert, J. R., Schroeder, R. G., & Mauriel, J. J. (2000). A framework for linking culture and improvement initiatives in organizations. *Academy of Management Review*, 25, 850-863.
- Diez-Roux, A. V. (1998). Bringing context back into epidemiology: Variables and fallacies in multilevel analyses. *American Journal of Public Health*, 88, 216-222.
- Erwee, R., Lynch, B., Millett, B., Smith, D., & Roodt, G. (2001). Cross-cultural equivalence of the organisational culture survey in Australia. *Journal of Industrial Psychology*, 27, 7-12.

- Fey, C., & Denison, D. R. (2003). Organizational culture and effectiveness: Can an American theory be applied in Russia? *Organization Science, 14*, 686-706.
- Fitzpatrick, R., Davey C., Buxton, M., & Jones, D. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment, 2*, 19-43.
- Geringer, M., & Hebert, L. (1989) Control and performance of international joint ventures. *Journal of International Business Studies, 20*, 235-254.
- Gillespie, M., Denison, D., Haaland, S., Smerek, R., & Neale, W. (2008). Linking organizational culture and customer satisfaction: Results from two companies in different industries. *European Journal of Work and Organizational Psychology, 17*, 112-132.
- Ginevičius, R., & Vaitkūnaitė, V. (2006). Analysis of organizational culture dimensions impacting performance. *Journal of Business Economics and Management, 7*, 201-211.
- Glaser, R. (1983). *The corporate culture survey*. Bryn Mawr, PA: Organizational Design and Development.
- Glaser, R., Zamanou, S., & Hacker, K. (1983). Measuring and interpreting organizational culture. *Management Communication Quarterly, 1*, 173-198.
- Gordon, G., & DiTomaso, N. (1992). Predicting corporate performance from organizational culture. *Journal of Management Studies, 29*, 783-798.
- Guthrie, J. P. (2001). High-involvement work practices, turnover, and productivity: Evidence from New Zealand. *Academy of Management Journal, 44*, 180-190.
- Harris, P. R., & Moran, R. T. (1984). *Managing cultural differences*. Houston, TX: Gulf.
- Hartnell, C. A., Ou, A. Y., & Kinicki, A. (2011). Organizational culture and organizational effectiveness: A meta-analytic investigation of the competing values framework's theoretical suppositions. *Journal of Applied Psychology, 96*, 677-694.
- Hofstede, G., Neuijen, B., Ohayv, D., & Sanders, G. (1990). Measuring organizational cultures: A qualitative and quantitative study across twenty cases. *Administrative Science Quarterly, 35*, 286-316.
- House, R., Javidan, M., & Dorfman, P. (2001). Project GLOBE: An introduction. *Applied Psychology: An International Review, 50*, 489-505.
- Hu, L.T., & P. M. Bentler (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453.
- Jaques, E. (1951). *The changing culture of a factory*. London: Tavistock.

- James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*, 85-98.
- Jöreskog, K.G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications*. Chicago: SPSS Publications.
- Jung, T., Scott, T., Davies, H. T., Bower, P. Whalley, D., McNally, R., & Mannion, R. (2009). Instruments for exploring organizational culture: A review of the literature. *Public Administration Review, 69*, 1087-1096.
- Katz, D., & Kahn, R. (1978). *The social psychology of organizations*. New York: John Wiley & Sons.
- Kilman, R. H., & Saxton, M. J. (1983). *The Kilman-Saxton culture-gap survey*. Pittsburgh, PA: Organizational Design Consultants.
- Klein, K.J., Griffin, M.A., Bliese, P.D., Hofmann, D.A., Kozlowski, S.W.J., James, L.R., Dansereau, F., Yammarino, F.J., Gavin, M.B., & Bligh, M.C. (2000). Multilevel analytical techniques: Commonalities, differences, and continuing questions. In K. Klein & S. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations*, 512-553. San Francisco, CA: Jossey-Bass.
- Kotrba, L. M., Gillespie, M. A., Schmidt, A. M., Smerek, R. E., Ritchie, S. A., & Denison, D. R. (in press). The effects of cultural consistency on business performance. *Human Relations*.
- Kotter, J. & Heskett, J. (1992). *Corporate culture and performance*. New York: Free Press.
- Langford, P. H. (2009). Measuring organisational climate and employee engagement: Evidence for a 7 Ps model of work practices and outcomes. *Australian Journal of Psychology, 61*, 185-198.
- Lawler, E. (1986). *High involvement management*. San Francisco, CA: Jossey-Bass.
- Lawrence, P., & Lorsch, J. (1967). *Organization and environment: Managing differentiation and integration*. Boston, MA: Harvard University Division of Research.
- Lee, C., & Beamish, P. W. (1995). The characteristics of performance of Korean joint ventures in LDCs. *Journal of International Business Studies, 26*, 637-654.
- Liebenberg, J. S. (2007). *Factors influencing a customer service culture in a higher education environment*. Unpublished dissertation. Rand Afrikaans University.
- Lim, B. (1995). Examining the organizational culture and organizational performance link. *Leadership and Organization Development Journal, 16*, 16-21.

- Lindell, M. K., & Brandt, C. J. (1999). Assessing interrater agreement on the job relevance of a test: A comparison of the CVI, T, $r_{wg(j)}$, and $r^*_{wg(j)}$ indexes. *Journal of Applied Psychology, 84*, 640-647.
- Lindell, M. K., Brandt, C. J., & Whitney, D. J. (1999). A revised index of interrater agreement for multi-item ratings of a single target. *Applied Psychological Measurement, 23*, 127-135.
- MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.
- Martin, J. (1992). *Cultures in organizations: Three perspectives*. New York: Oxford University Press.
- Martin, J., & Frost, P. (1996). Organizational culture war games: A struggle for intellectual dominance. In S. Clegg, C. Hardy, & W. Nord (Eds.), *Handbook of organization studies*, p. 599-621. Thousand Oaks: Sage.
- Martin, J., Frost, P., & O'Neill, O. A. (2006). Organizational culture: Beyond struggles for intellectual dominance. In S. Clegg, C. Hardy, T. Lawrence, & W. Nord (Eds.), *The Sage handbook of organization studies*, 2nd edn, p. 725-753. Thousand Oaks: Sage.
- Martin, J., & Meyerson, D. (1988). Organizational cultures and the denial, channeling and acknowledgment of ambiguity. In L. Pondy (Ed.), *Managing ambiguity and change*, p. 93-125. New York: Wiley.
- Mintzberg, H. (1989). *Mintzberg on management*. New York: Free Press.
- Nunnally, J. (1978). *Psychometric theory*. New York, McGraw-Hill.
- O'Reilly, C., Chatman, J., & Caldwell, D. (1991). People and organization culture: A profile comparison approach. *Academy of Management Journal, 34*, 487-516.
- Ostroff, C., Kinicki, A. J., & Tamkins, M. M. (2003). Organizational culture and climate. In *Handbook of psychology*. p. 565-593. John Wiley & Sons.
- Ott, J. S. (1989). *The organizational culture perspective*. Pacific Grove: Brooks-Cole Publishing Company.
- Pettigrew, A. M. (1979). On studying organizational cultures. *Administrative Science Quarterly, 24*, 570-581.
- Quinn, R., & Cameron, K. (1988). *Paradox and transformation: Toward a theory of change in organization and management*. Cambridge, MA: Ballinger.
- Quinn, R., & Rohrbaugh, J. (1981). A competing values approach to organizational effectiveness. *Public Productivity Review, 5*, 122-140.

- Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Rieker, D. J. (2006). *An evaluation of how organizational culture can perpetuate a formal mentoring relationship*. Unpublished thesis. Air Force Institute of Technology, Wright-Patterson Air Force Base, OH.
- Rousseau, D. M. (1990). Assessing organizational culture: The case for multiple methods. In B. Schneider (Ed.), *Organizational culture and climate*, p. 153-192. San Francisco, CA: Jossey-Bass.
- Sackmann, S. A. (2006). *Assessment, evaluation, improvement: Success through corporate culture*. Gütersloh: Bertelsmann Stiftung.
- Sackmann, S. A. (2011). Culture and performance. In N. Ashkanasy, C. Wilderom, & M. Peterson (Eds.), *The handbook of organizational culture and climate*, 2nd edn, p. 188-224. Thousand Oaks, CA: Sage Publications.
- Saffold, G. (1988). Culture traits, strength, and organizational performance: Moving beyond 'strong' culture. *Academy of Management Review*, 13, 546-558.
- Sashkin, M. (1984). *Pillars of excellence: Organizational beliefs questionnaire*. Bryn Mawr, PA: Organizational Design and Development.
- Sashkin, M. & Fulmer, R. (1985). *Measuring organizational excellence culture with a validated questionnaire*. Paper presented at the Academy of Management, San Diego, California.
- Schein, E. (1992). *Organizational culture and leadership*, 2nd edn. San Francisco, CA: Jossey-Bass.
- Schönborn, G. (2010). Value performance: On the relation between corporate culture and corporate success. *Journal of Psychology*, 218, 234-242.
- Scott, T., Mannion, R., Davies, H., & Marshall, M. (2003). The quantitative measurement of organizational culture in health care: A review of the available instruments. *Health Services Research*, 38, 923-945.
- Selznick, P. (1957). *Leadership in administration*. Evanston, IL: Row & Peterson.
- Shaw, J., Delery, J., Jenkins, D., & Gupta, N. (1998). An organizational-level analysis of voluntary and involuntary turnover. *Academy of Management Journal*, 41, 511-525.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.

- Siehl, C., & Martin, J. (1990). Organizational culture: A key to financial performance? In B. Schneider (Ed.), *Organizational climate and culture*, p. 241-281. San Francisco, CA: Jossey-Bass.
- Smircich, L. (1983). Concepts of culture and organizational analysis. *Administrative Science Quarterly*, 28, 339-358.
- Sparrow, P. (2001). Developing diagnostics for high performance cultures. In C. Cooper, S. Cartwright, & C. Earley (Ed.), *The international handbook of organizational culture and climate*. New York: John Wiley & Sons, 85-106.
- Spreitzer, G. (1995). Psychological empowerment in the workplace: Dimensions, measurement, validation. *Academy of Management Journal*, 38, 1442-1466.
- Spreitzer, G. (1996). Social structural characteristics of psychological empowerment. *Academy of Management Journal*, 39, 483-504.
- Steenkamp, J. B., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78-107.
- Strydom, A., & Roodt, G. (2006). Developing a predictive model of subjective organisational culture. *Journal of Industrial Psychology*, 32, 15-25.
- Taylor, S., Levy, O., Boyacigiller, N., & Beechler, S. (2008). Employee commitment in MNCs: Impacts of organizational culture, HRM and top management orientations. *The International Journal of Human Resource Management*, 19, 501-527.
- Trice, H. M., & Beyer, J. M. (1993). *The culture of work organizations*. Englewood Cliffs, NJ: Prentice Hall.
- Tucker, R. W., W. J. McCoy, and L. C. Evans (1990). Can questionnaires objectively assess organizational culture? *Journal of Managerial Psychology*, 5, pp. 4-11.
- Usala, P. (1996a). *Clerical/technical study technical report: Psychometric analysis of the organizational assessment items: Work unit experiences and organizational experiences scales*. Unpublished technical report by the U.S. Office of Personnel Management, Personnel Resources and Development Center, Washington D.C.
- Usala, P. (1996b). *Psychometric analysis of the organizational assessment items: Work unit experiences and organizational experience scales*. Unpublished technical report by the U.S. Office of Personnel Management, Personnel Resources and Development Center, Washington D.C.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-70.

- van der Post, W. Z., de Coning, T. J., & Smit, E. M. (1997). An instrument to measure organizational culture. *South African Journal of Business Management*, 28, 147-168.
- van der Post, W. Z., de Coning, T. J., & Smit, E. M. (1998). The relationship between organizational culture and financial performance: Some South African evidence. *South African Journal of Business Management*, 29, 30-41.
- Van Maanen, J. (1988). *Tales of the field*. Chicago, IL: University of Chicago Press.
- Wall, T.D., Michie, J., Patterson, M., Wood, S.J., Sheehan, M., Clegg, C.W., & West, M. (2004). On the validity of subjective measures of company performance. *Personnel Psychology*, 57, 95-118.
- Walker, H., Symon, G., & Davies, B. (1996). Assessing organizational culture: A comparison of methods. *International Journal of Selection and Assessment*, 4, 96-105.
- Wilderom, C. P., Glunk, U., & Maslowski, R. (2000). Organizational culture as a predictor of organizational performance. In N. Ashkanasy, C. Wilderom, & M. Peterson (Eds.), *The handbook of organizational culture and climate*, p. 193-209. Thousand Oaks, CA: Sage Publications.
- Woodcock, M. & Francis, D. (1989). *Clarifying organizational values*. Aldershot, UK: Gower.
- Muldrow, T. W., Buckley, T., & Schay, B. W. (2002). Creating high-performance organizations in the public sector. *Human Resource Management*, 41, 341-354.
- Xenikou, A. & Furnham, A. (1996). A correlational and factor analytic study of four questionnaire measures of organizational culture. *Human Relations*, 49, 349-371.
- Zheng, W., Yang, B., & McLean, G. (2010). Linking organizational culture, structure, strategy, and organizational effectiveness: Mediating role of knowledge management. *Journal of Business Research*, 63, 763-771.

Table 1

Summary of Reliability Evidence for Culture Effectiveness Profiling Instruments

Instrument	Structure	Internal consistency ^a	Test-retest	Aggregation
Denison Organizational Culture Survey (Denison & Neale, 1996)	60 items, 12 dimensions, 4 traits	> .70 (Fey & Denison, 2003); .88 to .97 (Gillespie et al., 2008)	n/a	Adequate r_{wg} , ICC(1), and ICC(2) (Gillespie et al., 2008)
OASIS Culture Questionnaire (Cowherd & Luchs, 1988)	33 items, 5 dimensions	n/a	n/a	n/a
Organizational Assessment Survey (<i>Usala, 1996</i>)	100 items, 17 dimensions	n/a	n/a	n/a
Organizational Beliefs Questionnaire (Sashkin, 1984)	50 items, 10 dimensions	.35 to .78 (Xenikou & Furnham, 1996)	n/a	Low within-organization variance (<i>Sashkin and Fulmer, 1985</i>)
Organizational Culture Survey (van der Post et al., 1997)	97 items, 15 dimensions	.79 to .93 (van der Post et al., 1997)	n/a	n/a
Organizational Culture Survey Instrument (Harris & Moran, 1983)	99 items, 7 dimensions	n/a	n/a	n/a
Organizational Values Questionnaire (Woodcock & Francis, 1989)	60 items, 12 values	n/a	n/a	n/a
Questionnaire of Dimensions of Organizational Culture (Ginevičius & Vaitkūnaitė, 2006)	48 items, 12 dimensions	n/a	n/a	n/a
Value Performance Index (Schönborn, 2010)	105 items, 13 dimensions	.71 to .94 (Schönborn, 2010)	n/a	n/a

Notes: References shown in italics are unpublished sources. ^a Values shown indicate lower and upper bounds of alphas reported across dimensions or factors.

Table 2

Summary of Validity Evidence for Culture Effectiveness Profiling Instruments

Instrument	Dimensionality	Convergent / discriminant validity	Cross-cultural application	Predictive validity	Sensitivity to change
Denison Organizational Culture Survey (Denison & Neale, 1996)	Factor analytic support for indexes (Bonavia et al., 2009; Fey & Denison, 2003; Taylor et al., 2008); Factor analytic support for second-order model (Gillespie et al., 2008)	Leadership (Block, 2003); commitment (Taylor et al., 2008); knowledge management, org. structure, strategy (Zheng et al., 2010)	Asia, Australia, Brazil, Japan, Jamaica, and South Africa (Denison et al., 2003); Russia (Fey & Denison, 2003); Spain (Bonavia et al., 2009)	Longitudinal evidence linking culture to sales and customer satisfaction (Boyce, 2010); cross-sectional with “hard” performance metrics (Denison, 1984; Denison & Mishra, 1995; Gillespie et al., 2008); cross-sectional with perceived effectiveness outcomes (Denison et al., 2003; Fey & Denison, 2003)	Longitudinal study of 95 car dealerships (Boyce, 2010)
OASIS Culture Questionnaire (Cowherd & Luchs, 1988)	n/a	n/a	n/a	Case study demonstrating link between culture gap scores and profitability (Cowherd & Luchs, 1988)	n/a
Organizational Assessment Survey (Usala, 1996)	Factor analytic support (Usala, 1996)	n/a	n/a	n/a	Two case studies demonstrating change over time (Muldrow et al., 2002)
Organizational Beliefs Questionnaire (Sashkin, 1984)	n/a	Other culture questionnaires (Xenikou & Furnham, 1996)	n/a	n/a	n/a
Organizational Culture Survey (van der Post et al., 1997)	Factor and content analysis (van der Post et al., 1997)	Job satisfaction, personality (Liebenberg, 2007; Strydom & Roodt, 2006); mentoring (Rieker, 2006)	Australia (Erwee et al., 2001)	15/15 dimensions correlated with financial performance composite (van der Post et al., 1998)	n/a

Table 2 continued...

Title and author-year	Dimensionality	Convergent / discriminant validity	Cross-cultural application	Predictive validity	Sensitivity to change
Organizational Culture Survey Instrument (Harris & Moran, 1984)	n/a	n/a	n/a	n/a	n/a
Organizational Values Questionnaire (Woodcock & Francis, 1989)	n/a	n/a	n/a	n/a	n/a
Questionnaire of Dimensions of Organizational Culture (Ginevičius & Vaitkūnaitė, 2006)	EFA with little support for dimensional structure (Aydin & Ceyla, 2008, 2009; Ginevičius & Vaitkūnaitė, 2006)	Employee satisfaction (Aydin & Ceyla, 2008; Ginevičius & Vaitkūnaitė, 2006)	n/a	2/4 factors correlate with overall performance index (Ginevičius & Vaitkūnaitė, 2006); 10/10 dimensions correlate with perceived performance composite (Aydin & Ceyla, 2009)	n/a
Value Performance Index (Schönborn, 2010)	EFA to define dimension structure (Schönborn, 2010)	n/a	n/a	7/13 dimensions correlated with dichotomous performance composite (Schönborn, 2010)	n/a

Notes: References shown in italics are un-published.

Table 3

Definitions of Culture Traits and Indexes from the DOCS

Effectiveness traits and corresponding index definitions

Involvement concerns the personal engagement of individuals within the organization and reflects a focus on the internal dynamics of the organization and on flexibility.

Empowerment – Individuals have the authority, initiative, and ability to manage their own work. This creates a sense of ownership and responsibility toward the organization.

Team Orientation – Value is placed on working cooperatively toward common goals for which all employees feel mutually accountable. The organization relies on team effort to get work done.

Capability Development – The organization continually invests in the development of employee's skills in order to stay competitive and meet on-going business needs.

Consistency refers to shared values, and efficient systems and processes and reflects an internal and stable focus.

Core Values – Members of the organization share a set of values which create a sense of identity and a clear set of expectations.

Agreement – Members of the organization are able to reach agreement on critical issues. This includes both the underlying level of agreement and the ability to reconcile differences when they occur.

Coordination and Integration – Different functions and units of the organization are able to work together well to achieve common goals. Organizational boundaries do not interfere with getting work done.

Adaptability refers to employees' ability to understand what the customer wants, to learn new skills, and to change in response to demand. The focus of adaptability is external and flexible.

Creating Change – The organization is able to create adaptive ways to meet changing needs. It is able to read the business environment, react quickly to current trends, and anticipate future changes.

Customer Focus – The organization understands and reacts to their customers and anticipates their future needs. It reflects the degree to which the organization is driven by a concern to satisfy their customers.

Organizational Learning – The organization receives, translates, and interprets signals from the environment into opportunities for encouraging innovation, gaining knowledge, and developing capabilities.

Mission refers to an organization's purpose and direction, and reflects a focus external to the organization and on stability.

Strategic Direction and Intent – Clear strategic intentions convey the organization's purpose and make it clear how everyone can contribute and "make their mark" on the industry.

Goals and Objectives – A clear set of goals and objectives can be linked to the mission, vision, and strategy, and provide everyone with a clear direction in their work.

Vision – The organization has a shared view of a desired future state. It embodies core values and captures the hearts and minds of the organization's people, while providing guidance and direction.

Table 4

Demographic Characteristics of Organizational and Respondent Sample

Organizational Category	<i>n</i>	% of sample	Demographic Category	<i>n</i>	% of sample
<i>Country</i>			<i>Age</i>		
Australia	3	1.9	<20	22	0.1
Canada	5	3.1	20-29	3,006	8.5
France	2	1.3	30-39	8,034	22.6
Germany	4	2.5	40-49	7,680	21.6
Great Britain	8	5.0	50-59	3,650	10.3
India	2	1.3	>60	283	0.8
Japan	5	3.1	No response	12,799	36.1
Netherlands	2	1.3	<i>Gender</i>		
Norway	1	0.6	Male	14,104	39.8
Sweden	1	0.6	Female	8,369	23.6
Switzerland	8	5.0	No response	13,001	36.6
United States	119	74.4	<i>Educational level</i>		
<i>Industry</i>			High school	2,059	5.8
Basic Materials	23	14.4	Some college	3,983	11.2
Consumer Cyclical	19	11.9	Associate degree	1,910	5.4
Consumer Staples	22	13.8	Bachelor's degree	7,231	20.4
Health Care	17	10.6	Some graduate work	1,894	5.3
Energy	1	0.6	Master's degree	4,115	11.6
Financials	17	10.6	Doctoral degree	710	2.0
Capital Goods	17	10.6	Other	266	0.7
Technology	25	15.6	No response	13,306	37.5
Pharmaceuticals	1	0.6	<i>Function</i>		
Communication Services	10	6.3	Finance and Accounting	2,033	5.7
Utilities	7	4.4	Engineering	1,863	5.3
Transportation	1	0.6	Manufacturing and Production	1,928	5.4
<i>Employee Population^a</i>			Research and Development	1,548	4.4
Fewer than 1,000	11	7.2	Sales and Marketing	5,083	14.3
1,000 to 5,000	26	17.0	Purchasing	864	2.4
5,001 to 10,000	12	7.8	Human Resources	917	2.6
10,001 to 20,000	16	10.5	Administration	1,031	2.9
20,001 to 50,000	30	19.6	Support staff	1,973	5.6
50,001 to 100,000	28	18.3	Professional staff	1,820	5.1
100,001 to 200,000	20	13.1	No response	16,414	46.3
More than 200,000	10	6.5	<i>Organizational level</i>		
<i>Organizational Revenue^b</i>			Non-management	9,018	25.4
Under \$100 million	7	5.3	Line management	4,960	14.0
\$100 million — \$1 billion	17	13.0	Middle management	4,765	13.4
\$1 billion — \$5 billion	35	26.7	Senior management	1,031	2.9
\$5 billion — \$10 billion	14	10.7	Executive/Senior Vice President	280	0.8
\$10 billion — \$20 billion	15	11.5	CEO/President	71	0.2
\$20 billion — \$30 billion	18	13.7	Owner	12	0.0
\$30 billion — \$50 billion	14	10.7	No response	15,337	43.2
More than \$50 billion	11	8.4	<i>Years with organization</i>		
			Less than 6 months	1,042	2.9
			6 months to 1 year	1,432	4.0
			1 to 2 years	2,315	6.5
			2 to 4 years	3,093	8.7
			4 to 6 years	2,017	5.7
			6 to 10 years	2,952	8.3
			10 to 15 years	2,998	8.5
			More than 15 years	5,989	16.9
			No response	13,636	38.4

Note. ^a Information on employee population was unavailable for 7 organizations. ^b Information on organizational revenue was unavailable for 29 organizations.

Table 5

Alpha Coefficients and Descriptive Statistics for the Culture Survey (N=35,474)

Dimension	Index	Item	Item-total correlation	Mean	S.D
Involvement	Empowerment $\alpha = .76$	1	.43	3.94	.81
		2	.59	3.13	1.01
		3	.57	3.11	1.07
		4	.56	3.24	.98
		5	.51	3.13	1.04
	Team Orientation $\alpha = .82$	6	.56	3.53	1.00
		7	.70	3.47	1.02
		8	.61	3.31	1.06
		9	.63	3.46	1.01
		10	.54	3.24	.98
	Capability Development $\alpha = .70$	11	.43	3.39	1.03
		12	.54	3.31	.95
		13	.56	3.45	1.05
		14	.56	3.62	.98
		15	.23	3.30	1.08
Consistency	Core Values $\alpha = .71$	16	.47	3.13	1.03
		17	.39	3.34	.94
		18	.61	3.47	1.01
		19	.36	3.74	.94
		20	.51	3.84	.92
	Agreement $\alpha = .74$	21	.54	3.42	.94
		22	.41	3.50	.94
		23	.60	2.94	.91
		24	.47	3.09	.96
		25	.50	3.15	.97
	Coordination & Integration $\alpha = .78$	26	.43	3.22	1.00
		27	.60	3.03	1.00
		28	.62	2.70	.98
		29	.53	3.01	1.08
		30	.59	3.20	.93
Adaptability	Creating Change $\alpha = .76$	31	.56	2.82	1.04
		32	.53	3.29	.99
		33	.61	3.37	.96
		34	.46	2.82	.99
		35	.48	3.21	.87
	Customer Focus $\alpha = .74$	36	.57	3.34	.91
		37	.60	3.48	.93
		38	.49	3.01	1.03
		39	.53	3.44	1.01
		40	.36	3.57	1.00
	Organizational Learning $\alpha = .74$	41	.52	3.34	.98
		42	.52	3.04	1.04
		43	.46	2.79	1.08
		44	.46	3.73	.93
		45	.56	2.76	1.02
Mission	Strategic Direction & Intent $\alpha = .86$	46	.70	3.63	.99
		47	.51	3.24	.96
		48	.75	3.48	.96
		49	.80	3.44	1.00
		50	.67	3.29	1.15
	Goals & Objectives $\alpha = .80$	51	.60	3.24	.92
		52	.56	3.38	.97
		53	.58	3.70	.86
		54	.56	3.67	.91
		55	.60	3.37	.97
	Vision $\alpha = .79$	56	.63	3.05	.98
		57	.65	3.32	1.00
		58	.41	2.59	.99
		59	.60	3.02	.99
		60	.60	3.10	.93

Table 6

Correlation Matrix for the 12 Indexes of the Culture Survey (N = 35,474)

Traits	Indexes	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
Involvement	1. Empowerment	3.31	.71											
	2. Team orientation	3.40	.77	.74										
	3. Capability development	3.41	.69	.64	.66									
Consistency	4. Core values	3.50	.66	.61	.61	.57								
	5. Agreement	3.22	.66	.63	.65	.61	.64							
	6. Coordination & Integration	3.03	.73	.61	.63	.55	.57	.65						
Adaptability	7. Creating change	3.10	.69	.57	.58	.57	.47	.58	.60					
	8. Customer focus	3.37	.69	.49	.50	.48	.45	.49	.48	.54				
	9. Organizational learning	3.13	.71	.65	.66	.65	.58	.66	.63	.65	.54			
Mission	10. Strategic direction & Intent	3.41	.82	.58	.58	.58	.58	.57	.58	.56	.50	.61		
	11. Goals & objectives	3.47	.69	.61	.61	.59	.60	.60	.61	.57	.52	.63	.74	
	12. Vision	3.02	.73	.60	.60	.60	.57	.61	.62	.61	.52	.68	.73	.71

Table 7

First-Order Factor Loadings

Item	<u>Factor</u>					
	Empowerment	Team Orientation	Capability Development	Core Values	Agreement	Coordination & Integration
1	.49					
2	.69					
3	.66					
4	.65					
5	.62					
6		.65				
7		.78				
8		.68				
9		.68				
10		.66				
11			.62			
12			.67			
13			.64			
14			.68			
15			.30			
16				.69		
17				.45		
18				.74		
19				.37		
20				.56		
21					.67	
22					.54	
23					.66	
24					.55	
25					.63	
26						.55
27						.67
28						.67
29						.59
30						.73

Note. All loadings are significant at the $p < .01$ level.

Table 7 continued...

Item	Factor					Vision
	Creating Change	Customer Focus	Organizational Learning	Strategic Direction & Intent	Goals & Objectives	
31	.62					
32	.63					
33	.74					
34	.53					
35	.60					
36		.71				
37		.73				
38		.58				
39		.65				
40		.43				
41			.58			
42			.62			
43			.60			
44			.54			
45			.68			
46				.77		
47				.56		
48				.84		
49				.86		
50				.75		
51					.72	
52					.66	
53					.60	
54					.59	
55					.74	
56						.77
57						.73
58						.43
59						.74
60						.64

Note. All loadings are significant at the $p < .01$ level.

Table 8

Second-Order Factor Loadings

Index Factor	<u>Trait Factor</u>			
	Involvement	Consistency	Adaptability	Mission
1. Empowerment	.96			
2. Team Orientation	.94			
3. Capability Development	.92			
4. Core Values		.91		
5. Agreement		.94		
6. Coordination & Integration		.89		
7. Creating Change			.88	
8. Customer Focus			.70	
9. Organizational Learning			.99	
10. Strategic Direction & Intent				.91
11. Goals & Objectives				.97
12. Vision				.97

Note. All loadings are significant at the $p < .01$ level.

Table 9

Second-Order Factor Correlations

Trait Factor	Involvement	Consistency	Adaptability
Involvement	--		
Consistency	.94	--	
Adaptability	.92	.93	--
Mission	.84	.88	.89

Note. All correlations are significant at the $p < .01$ level.

Table 10

Descriptive Statistics and Aggregation Evidence for the 12 Culture Indexes

Index	Mean	SD	Mean $r_{wg(j)}$	Min $r_{wg(j)}$	Max $r_{wg(j)}$	ICC(1)	ICC(2)	<i>F</i> -value
Empowerment	3.31	0.71	.87	.74	.94	.10	.96	25.32
Team Orientation	3.40	0.77	.86	.73	.95	.08	.95	19.99
Capability Development	3.41	0.69	.86	.75	.94	.08	.95	18.85
Core Values	3.50	0.66	.88	.73	.94	.08	.95	21.31
Agreement	3.22	0.66	.88	.81	.94	.07	.94	17.96
Coordination & Integration	3.03	0.73	.86	.78	.95	.09	.95	21.62
Creating Change	3.10	0.69	.87	.75	.95	.06	.94	16.18
Customer Focus	3.37	0.69	.87	.76	.95	.06	.93	15.33
Organizational Learning	3.13	0.71	.86	.74	.96	.06	.94	15.89
Strategic Direction & Intent	3.41	0.82	.85	.67	.95	.08	.95	20.91
Goals & Objectives	3.47	0.69	.89	.77	.96	.08	.95	20.10
Mission	3.30	0.67	.87	.74	.94	.08	.95	19.51

Note. $N = 35,474$. All *F*-values are statistically significant at the .001 level.

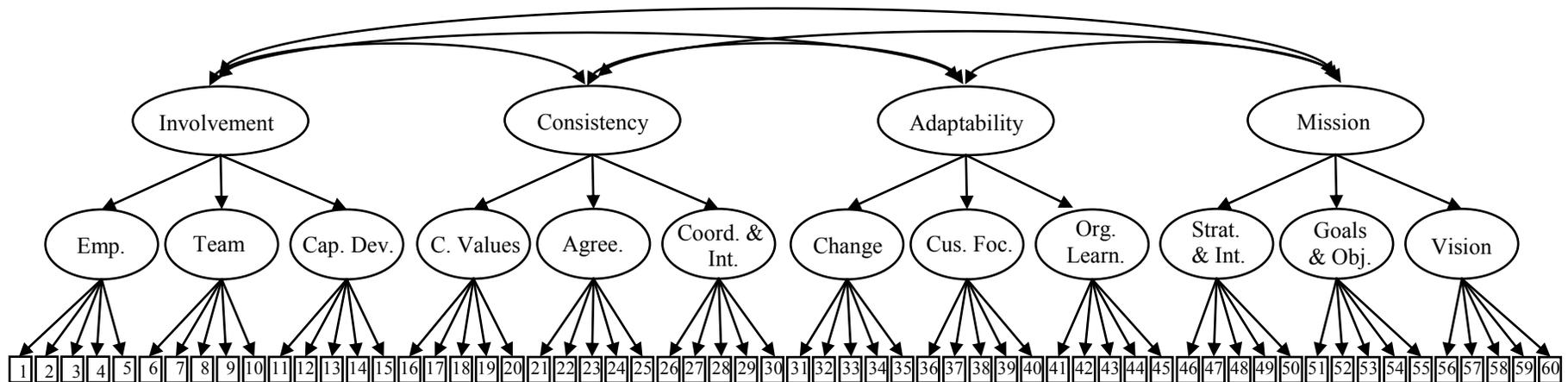
Table 11

Correlations Between Culture Indexes and Indicators of Organizational Performance (N = 155)

<i>Trait</i>							
Index	Sales Growth	Market Share	Profit	Quality	New Product	Employee Satisfaction	Overall Performance
<i>Involvement</i>	.24**	.13	.23**	.39**	.41**	.79**	.61**
Empowerment	.20*	.11	.21**	.37**	.36**	.74**	.57**
Team Orientation	.17*	.11	.20*	.32**	.36**	.70**	.51**
Capability Development	.33**	.16	.26**	.41**	.43**	.77**	.65**
<i>Consistency</i>	.20**	.12	.28**	.42**	.26**	.62**	.58**
Core Values	.20**	.15	.27**	.36**	.21**	.52**	.53**
Agreement	.26**	.13	.29**	.43**	.32**	.66**	.60**
Coordination & Integration	.11	.07	.21**	.36**	.17*	.53**	.48**
<i>Adaptability</i>	.29**	.10	.24**	.34**	.45**	.66**	.60**
Creating Change	.35**	.13	.24**	.31**	.49**	.63**	.57**
Customer Focus	.21**	.08	.16*	.31**	.27**	.42**	.44**
Organizational Learning	.20*	.04	.21**	.27**	.39**	.65**	.54**
<i>Mission</i>	.36**	.19*	.31**	.38**	.47**	.62**	.68**
Strategic Direction & Intent	.40**	.26**	.32**	.38**	.53**	.55**	.66**
Goal Orientation	.26**	.15	.27**	.35**	.39**	.57**	.60**
Vision	.34**	.10	.29**	.34**	.41**	.66**	.65**

Note. * $p < .05$. ** $p < .01$.

Figure 1. Depiction of the second-order factor model underlying the DOCS.



Notes. Item loadings, factor loadings, and trait intercorrelations are shown in Tables 7, 8, and 9, respectively. The twelve culture indexes (from left to right above) are: *empowerment, team orientation, capability development, core values, agreement, coordination and integration, creating change, customer focus, organizational learning, strategic direction and intent, goals and objectives, and vision.*

Figure 2. Within-group agreement (r_{wg}) values for the 160 organizations

