From LBDQXII to LBDQ50: preferred leader behavior measurement across cultures

Gillian Warner-Soderholm

Department of Communication and Culture, Handelshoyskolen BI, Oslo, Norway

Inga Minelgaite

Department of Business Studies, University of Iceland, Reykjavik, Iceland, and

Romie Frederick Littrell

Department of Management,
National Research University Higher School of Economics, Moscow, Russia and
Centre for Cross Cultural Comparisons, Miami, Florida, USA

Abstract

Purpose – The purpose of this paper is to refine and validate the most widely used leader behavior measurement instrument, LBDQXII, into a more parsimonious instrument for assessing cognitive templates of preferred leader behavior across cultures.

Design/methodology/approach – The 100-item LBDQXII survey was administered to 6,451 participants from 14 countries; these data were used to refine the survey.

Findings – The shorter survey instrument is a valid and reliable tool for assessing preferred leader behavior. Four periods in the LBDQXII "evolution" are identified: emergence, expansion, stagnation and revival.

Research limitations/implications – The new LBDQ50 can be used to collect data across cultures, contributing to both global management development and scholarly studies.

Practical implications – This project corresponds to calls to shorten the well-established leader behavior instrument into a measurement tool that is reliable and valid across cultures and languages. This can be administered by both private and public organizations, contributing to greater effectiveness. Furthermore, it retains its scholarly scope encompassing follower-centric studies of leadership.

Social implications – Leadership processes are found in all aspects of life and can be better understood and improved within and across cultures using the shorter version.

Originality/value – An efficient instrument to measure preferred leadership behavior across and within cultures. The availability of the LBDQ50 will allow practitioners and researchers to advance understanding of preferred leadership behavior as a predictor of organizational effectiveness. Most such instruments are overly-long, which hinders data collection opportunities. This newly developed instrument can lead to better response rates and easier applicability in organizational settings.

Keywords Cross-cultural management, Leadership development, Surveys, Management development, Validation

Paper type Research paper

Introduction

An increasing body of evidence speaks of the importance of understanding culturally influenced leader behavior preferences (Littrell *et al.*, 2018; Mustafa and Lines, 2016). Hofstede (2001), House *et al.* (2004) and Littrell (2013) are among the seminal researchers who have found strong connections between leadership dimensions and cultural norms and values. The findings of many empirical studies, i.e. Dorfman *et al.* (2012) and Littrell *et al.* (2018), indicate that members of a society develop a cognitive template of preferred leader behaviors and that leaders hence tend to behave in a manner that is consistent with expectations of their respective societies in order to be more effective (House *et al.*, 2013;

Littrell *et al.*, 2018, p. 244). The most widely use instrument is the Leader Behavior Description Questionnaire XII (LBDQXII) (Northouse, 2013), which measure preferred leader behavior using 12 dimensions, and, per Littrell *et al.* (2018) can effectively describe desired leader behavior in particular cultures.

The reliability and validity of the LBDQXII has been well researched during its development and well-documented in the literature (Northouse, 2013; Schriesheim *et al.*, 1995). Yet, even though the value of the use of the LBDQXII in organizational learning in developing effective leadership is evident, the use of the instrument in today's research has become problematic, in part due to its length. The 100-item length of the LBDQXII gives rise to dangers of survey fatigue and over-surveying, and hence possible transient measurement errors in large-scale random sampling (e.g. Donnellan *et al.*, 2006; Schmidt *et al.*, 2003).

The challenge of survey fatigue and the impression of being over-surveyed (Rogelberg and Stanton, 2007; Stanton *et al.*, 2002) are seen when respondents in business organizations register careless completion of surveys, non-response or missing responses. These are growing problems in management research today. Research suggests that with overly-long surveys, such as the 100-item LBDQXII, respondents may respond carelessly due to frustration with the length of the survey and may then refuse to take part in management research in the future (Donnellan *et al.*, 2006). Calls have been made for a more concise and focused LBDQXII measurement instrument (also see Van Dick *et al.*, 2018). The complex, multivariate nature of modern organizational and leadership research is also challenged by a faster digitalized work pace and heightened performance expectations at work. Hence, there are pressing demands to make shortened, psychometrically sound measures available for both scholars and practitioners.

This study adds to the management and consultancy literature by developing and validating a more parsimonious survey instrument from the original LBDQXII in order to measure preferred leader behavior from studies of employed business people across 14 cultures.

In this paper, we first review and synthesize the extant literature underlying the LBDQXII model and related culture and leadership issues. Second, we detail our LBDQXII scale reduction work, reducing the LBDQII from 100 to 50 items. Third, we suggest implications for application of the shorter survey in assessing culturally endorsed effective leadership behavior.

Literature review
Culture and leadership
Dorfman *et al.* (1997) propose that the phenomenon of leadership is universal across cultures, but that the way in which it is operationalized is culturally specific, supported by Hofstede (2001) and Littrell *et al.* (2018). Project GLOBE (House *et al.*, 2004) provides compelling evidence that business people within cultural groups agree in their beliefs about leadership such that there are statistically significant differences between preferred leader behavior templates in societal cultures. In the same vein, Bass (1990) indicated that most people of the same culture hold a common set of beliefs about attributes of a leader that are culturally contingent. O'Connell *et al.* (1990) also posited that culture plays a strong role in influencing the content of leader behavior perceived as desirable by followers in that culture. We build upon this logic with data from 14 countries. For those new to culturally endorsed leader–follower research, we recommend further reading of the seminal studies by Project GLOBE (House *et al.*, 2004) and Hofstede (2001); the results from these research threads warrant complete reading. We have chosen to apply concepts from Hofstede (2001) in our later discussion of construct validity between the LBDQ and Hofstede's framework for two reasons. First, among others, our consortium has collected additional VSM08 data from the societies in our LBDQ study, so the two samples are

well matched. Furthermore, the data are freely available, so we can carry out the appropriate statistical tests for convergent and discriminant validity of the LBDQ50 survey vs VSM08 data from our results.

We employ Hofstede's seven-dimensional model of societal cultural values, based on the Values Survey Model 2008 (Hofstede *et al.*, 2008). This model includes the dimensions: individualism/collectivism, power distance, uncertainty avoidance, masculinity/femininity, long-term/short-term orientation, indulgence/restraint and monumentalism. Researchers can obtain the VSM08 English original text, the VSM08 Manual and 22 non-English translations on request from the authors.

Status of the LBDQXII today
Relevance. The LBDQXII is derived from the LBDQ developed at Ohio State University. The instrument requires responses to items describing the behavior of a person in a leadership or supervisory position of a working group or unit in which the subject is a member (Stogdill, 1963). The LBDQXII consists of 100 items which define 12 dimensions describing preferred leader behavior (see Table I).

Validity and reliability
The LBDQXII has a long history of application, development and testing, resulting in a large number of studies of the instrument (Glynn and DeJordy, 2010; Littrell *et al.*, 2018; Northouse, 2013; Schriesheim and Bird, 1979; Schriesheim *et al.*, 1995), confirming its validity and reliability in different organizational contexts (Boatwright *et al.*, 2010).

An extensive meta-analysis of the LBDQXII, carried out by Judge *et al.* (2004), triggered a revival of the instrument, as leadership researchers sought to find alternative measurements for leader-centric studies which had become a tradition (Chang and Lin, 2018). Following the prescriptions of Cronbach and Meehl (1955) and Littrell *et al.* (2018) reviewed the construct, content and criterion validity literature relating to the LBDQXII across cultures, finding some diverse effects attributable to culture.

Factor 1: representation measures to what degree the manager speaks as the representative of the group	Factor 7: role assumption measures to what degree the manager actively exercises the leadership role rather than surrendering leadership to others
Factor 2: demand Reconciliation reflects how well the manager reconciles conflicting demands and reduces disorder to system	Factor 8: consideration depicts to what extent the manager regards the comfort, well-being, status and contributions of followers
Factor 3: tolerance of uncertainty depicts to what extent the manager is able to tolerate uncertainty and postponement without anxiety or getting upset	Factor 9: production emphasis measures to what degree the manager applies pressure for productive output
Factor 4: persuasiveness measures to what extent the manager uses persuasion and argument effectively; exhibits strong convictions	Factor 10: predictive accuracy measures to what extent the manager exhibits foresight and ability to predict outcomes accurately
Factor 5: initiation of structure measures to what degree the manager clearly defines own role, and lets followers know what is expected	Factor 11: integration reflects to what degree the manager maintains a closely-knit organization; resolves inter-member conflicts
Factor 6: tolerance of freedom reflects to what extent the manager allows followers scope for initiative, decision and action	Factor 12: superior orientation measures to what extent the manager maintains cordial relations with superiors; has influence with them; is striving for higher status

Source: Summarized from Stogdill (1963)

Table I.
Preferred leader
behavior

Evolution of the LBDQ to the LBDQXII

The evolution of the LBDQXII instrument can be summarized by four time and research trend periods depicted in Figure 1: the first period took place when leadership trait theory was not giving satisfactory results and the need to look at explicit leader behavior in more complex ways was identified. This led to extensive research work in developing the LBDQ from 1,800 leader characteristics to question development and sorting of 150 questions, then assigning these to leader behavior with two subscales: consideration and initiation of structure (Hemphill and Coons, 1950). This resulted in the creation of the "Ideal LBDQ" with 40 items to measure these two subscales (Hemphill *et al.*, 1952).

The two-factor (consideration and initiation of structure) model was soon challenged as too limiting. Stogdill (1963) called for the development of additional factors adequately describing the complexities of leader behavior. Additional instruments were developed, based on the LBDQ, including the LBDQXII, with 12 dimensions, resulting in growing popularity of the LBDQXII.

The third phase, which we identify with the term, stagnation, can be characterized by diverse criticism. Nevertheless, it should be noted that during this period, which gave rise to alternative leadership theories, (e.g. transformational–transactional leadership), new studies still utilized the two-factor leadership paradigm and the LBDQXII.

The fourth phase was triggered by the meta-analysis of Judge *et al.* (2004), reviving interest in and application of the LBDQXII. Furthermore, recent paradigms in the leadership field, namely follower-centric leadership, servant leadership and leader–member exchange, brought attention to the LBDQXII, as this instrument enabled researchers to investigate follower-centric attitudes toward leaders. However, the revival phase of LBDQXII faced a major barrier – its lengthy set of 100 items.

In summary, as depicted below, the evolution of the LBDQXII focused on follower-centric measures and extensive research in identifying appropriate dimensions of preferred leader behavior across different types of organizations, cultures and situations since the 1960s.

Method

The LBDQXII item-reduction project was conducted by an international research consortium, organized by the consortium director (third author in this paper) in cooperation with country collaborators who collected national data. Researchers used systematic random sampling techniques to distribute the surveys to samples of employed business people of the same nationality as the respondent country.
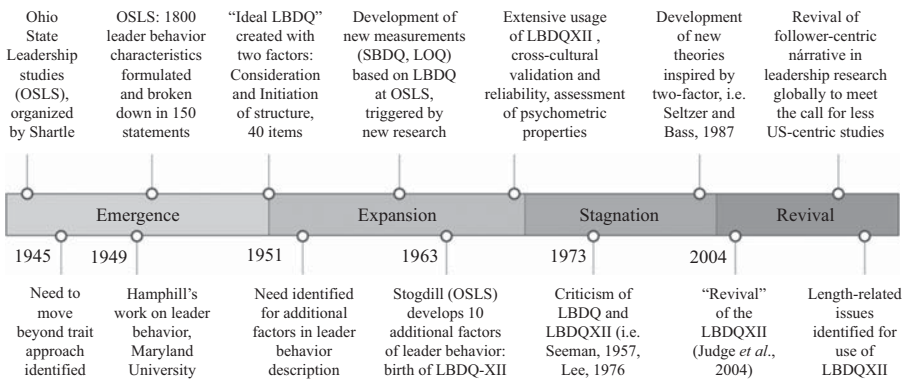

Figure 1.
Evolution of LBDQXII

Source: Created by authors, based on Bass (1990, pp. 511-534)

Survey item-reduction procedure
Clearly, in a time of globalization, theories of leader behavior need to be applicable across cultures; hence, we employed data from 14 countries representing all inhabited continents in order to validate a shorter version of the LBDQXII. Little guidance exists on how to reduce the length of a multi-item scale. The most common techniques include similar steps to those used in building and validating new models and measures, namely maximizing internal and external consistency and validity (Davila and Crawford, 2018; Ebrahimi Meharabani and Azmi Mohamad, 2015). In addition, the seminal study by Stanton *et al.* (2002) extends this method when reducing a scale by adding a third category for evaluating item removal for scale reduction, namely judgmental item qualities.

Such judgmental procedures refer to those issues that require subjective judgement and/or those that are difficult to assess in isolation from the context in which the scale is administered. This step is reminiscent of the Q-sorting step of the pool of items at an early stage of survey development. The expert panel has the knowledge to understand the cultural and multi-language settings of the survey and also the in-depth knowledge of the theories underpinning the items. They, at every stage, balance the trade-off of removing/keeping items that may only have moderate validating and reliability results, yet are essential to maintain the structure of the construct being measured. In our LBDQXII item-reduction techniques we applied all three stages recommended by the literature: validity and reliability tests and judgmental logic recommended by Stanton *et al.* (2002) to apply insights of the expert team in judging all items.

It is always a trade-off to refine/reduce a survey instrument while meeting all three criteria of validity, reliability and judgmental logic and not changing the actual findings in the data. Hence, we also validated the shortened scale by correlating mean score results in the 50-item scale with results in the 100-item scale within a test-re-test logic. Data were initially collected from 20 countries; raw data from countries which did not meet the stringent requirements for adequate sample size, back-to-back survey translation quality and appropriate respondent population were dropped from the final data set. Data from 14 countries, $n = 6,451$, remained and were employed in the reliability and validity tests described below. Sample descriptives are provided in Table II.

The data were analyzed in the three primary stages as recommended by the literature (Ebrahimi Meharabani and Azmi Mohamad, 2015). First, to test for reliability, Cronbach's α tests were conducted. Second, factor analysis tests were carried out for goodness-of-fit. Third, applying an iterative process, judgmental logic was applied at each stage. All 100 items were worked through manually to confirm "deletion sorting" with judgement logic which matched Cronbach's α results for best fit in a reduced scale: in first of three stages in sorting input, the four members of the expert country collaborators used Q-sorting logic applying the four criteria below. Findings were evaluated at the second stage by three expert panel members. Logic applied to deletions of items matched one or more of the following logic judgements for item purification:

(1) Items which were repeat questions in the same construct.

(2) Items which were culturally challenging to translate, i.e. LBDQXI item 28: "Needles the group": this is difficult to translate across languages and cultures.

(3) Items which have different meanings in a given society or culture, i.e. "a leader who encourages overtime." In Scandinavia, with all overtime paid, this would mean "encourage you to earn more while assisting the company." In many other western societies i.e. USA/UK, this could mean "encourages you to work long hours for the same basic pay," hence may have negative affect.

(4) Items which are not at the core of the construct, i.e. "publicizes the activities of the group." This is not seen as at the core of the factor "Representation," as it has a focus on public relations activities.

Countries	n	Gender	Language used in data collection
China	713	58% Male 42% Female	Han zi
Ghana	306	52% Male 48% Female	English or Swahili
Iceland	166	72% Male 28% Female	Icelandic
Iran	1,727	62% Male 38% Female	Persian
Japan	491	47%Male 53% Female	Japanese
Kenya	300	48% Male 52% Female	English Swahili
Lithuania	531	11% Male 89% Female	Lithuanian
New Zealand	221	75% Male 25% Female	English
Norway	801	51% Male 49% Female	Norwegian (Bokmål)
Russia	106	33% Male 67% Female	Russian
South Africa	231	70% Male 30% Female	English
South Korea	196	74% Male 26% Female	Korean
USA	362	71% Male 29% Female	English
Zambia	300	50% Male 50% Female	English or Swahili
Total	6,451	55% Male 45% Female	

Table II.
Sample characteristics

Table III below summarizes Cronbach's α and factor mean scores within the validity and reliability "test-re-test" logic by comparing results for LBDQXII analyses vs results from the reduced 50-item data set.

The 50 items in the scale were subjected to the same principal component factor analyses as the 100-item scale, using SPSS, v25. For the assessment of model fit of the LBDQ50, both absolute and comparative fit indices were used with AMOS structural

| LBDQXII Factors	1	2	3	4	5	6	7	8	9	10	11	12
Total (100 items)												
α	0.70	0.71	0.59	0.79	0.81	0.74	0.74	0.66	0.66	0.72	0.83	0.74
Mean	3.94	3.74	3.39	3.86	3.98	3.67	3.52	3.67	3.58	3.82	4.19	3.84
SD	0.64	0.82	0.53	0.62	0.58	0.57	0.69	0.59	0.53	0.64	0.72	0.55
Total with reduced items (50 items)												
α	0.69	0.79	0.42	0.72	0.71	0.77	0.79	0.68	0.70	0.69	0.74	0.73
Mean	3.98	3.53	3.47	3.83	4.04	3.86	3.48	3.86	3.75	3.82	4.19	3.88
SD	0.68	1.14	0.56	0.69	0.68	0.66	1.04	0.80	0.66	0.64	0.73	0.68
Total difference between full and reduced scale												
α	−0.01	0.08	−0.17	−0.07	−0.09	−0.03	0.05	0.02	0.04	−0.03	−0.09	−0.01
Mean	0.04	−0.21	0.07	−0.03	0.06	0.19	−0.03	0.19	0.17	0.00	0.00	0.04
SD	0.03	0.32	0.04	0.07	0.11	0.10	0.35	0.21	0.13	0.00	0.01	0.13

Table III.
LBDQXII and
LBDQ50 – comparison
of total scores and
validity results

equation modeling: confirmatory factor analysis. Following Ebrahimi Meharabani and Azmi Mohamad (2015) absolute fit indices to assess the predicted correlations/covariances of the model equal to their observed counterparts were carried out using RMSEA, GFI, CFI and CMIN/DF. Common interpretation in the literature confirms that RMSEA values are below 0.10, CFI and GFI values above 0.90 and CMIN/DF values below 3 (Jøreskog and Sørbom, 1993) indicate a good fit to the data. See Figures 2–13 for summaries of goodness-of-fit indices for each of the 12 dimensions. The full set of survey items in the 50-item field survey can be found at http://crossculturalcentre.homestead.com/ LeadershipReseach.html.

For the complete 100-item survey also see: http://crossculturalcentre.homestead.com/ LeadershipReseach.html.

Findings

For both the 100-item data set and the 50-item data set, inspection of the correlation matrixes revealed the presence of coefficients of 0.3 and above. Moreover, a Harman 1 factor analysis carried out on the LBDQ50 data indicate that this data set does not deviate from the common method bias issue, as only 24 percent of variance is explained by a single factor. In initial confirmatory factor analyses with the 100-item survey, Factors 1, 3, 4, 5, 6, 8, 9 and 11 showed more coherent item-factor loadings. In the factor analyses for the 50-item scale, a higher number of factors: (Factors 1, 2, 4, 5, 7, 8, 9, 10, 11 and 12), showed coherent item-factor loadings. The Kaiser–Meyer–Oklin values were 96, exceeding the recommended
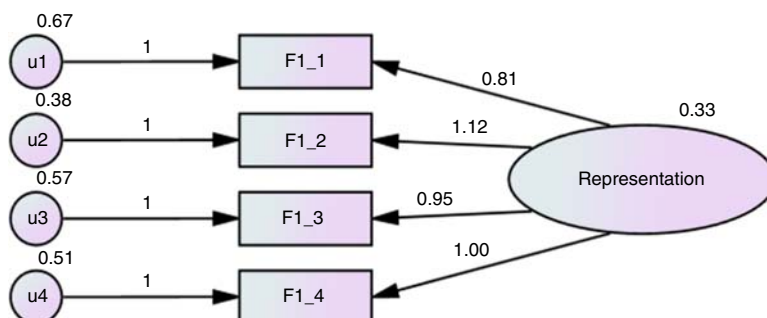
Notes: Goodness-of-Fit – CFI = 0.992; GFI = 0.997; RMSEA = 0.050;
CMIN/DF = 15.01; p(0.000)

Figure 2.
Measurement model for F1 representation
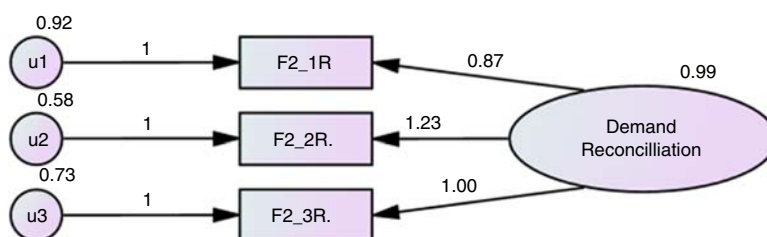
Notes: Goodness-of-Fit – CFI = 1.000; GFI = 1.000; RMSEA = 0.570;
CMIN/DF = 0.000; p(0.000)

Figure 3.
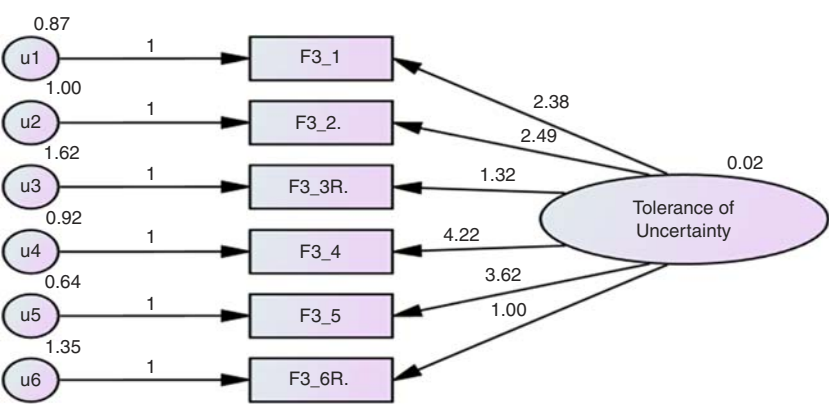Measurement model for F2 demand reconcilliation

Figure 4.
Measurement model
for F3 tolerance of
uncertainty

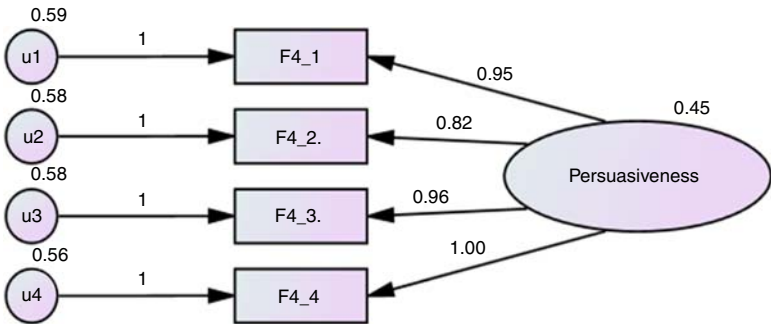**Notes:** Goodness-of-Fit – CFI = 0.636; GFI = 0.964; RMSEA = 0.112; CMIN/DF = 70.366; p(0.000)

Figure 5.
Measurement model
for F4 persuasiveness

Notes: Goodness-of-Fit – CFI = 0.954; GFI = 0.981; RMSEA = 0.134; CMIN/DF = 100.626; p(0.000)
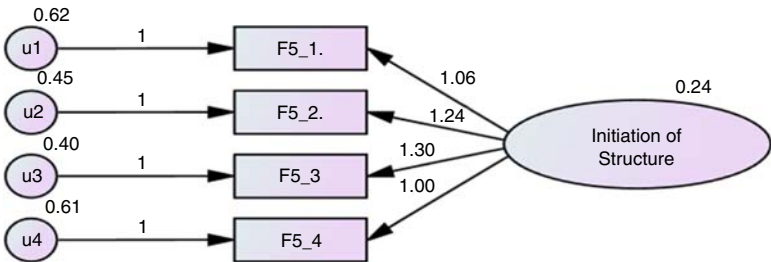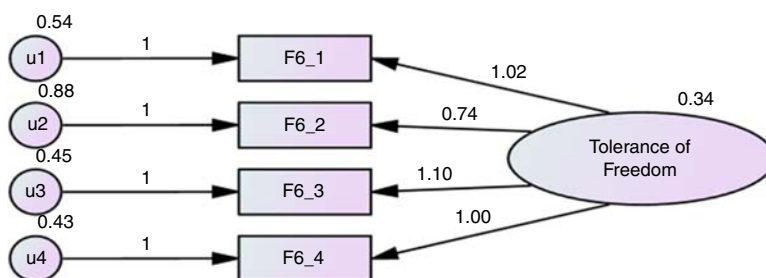

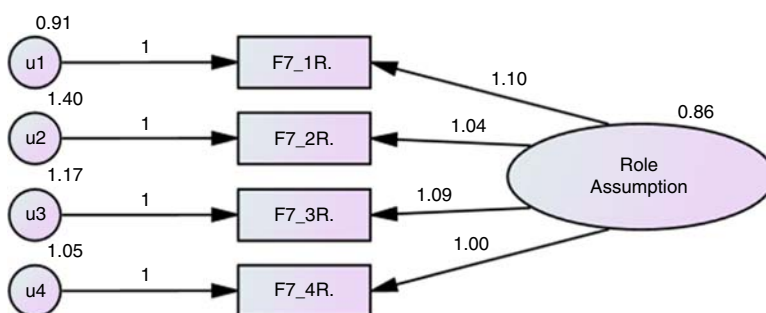**Figure 6.**
Measurement model
for F5 initiation of
structure

Notes: Goodness-of-Fit – CFI = 0.999; GFI = 0.999; RMSEA = 0.022; CMIN/DF = 3.774; p(0.000)
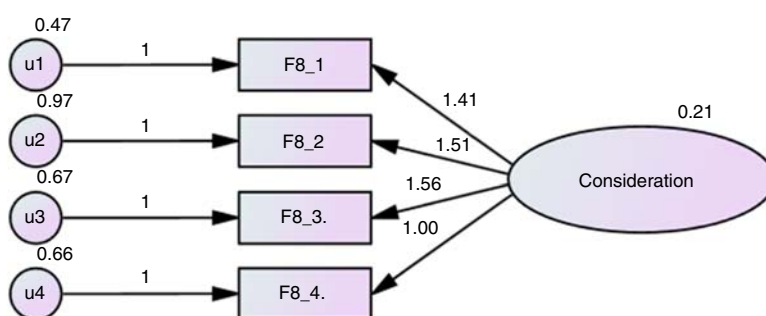
Notes: Goodness-of-Fit – CFI = 0.990; GFI = 0.997; RMSEA = 0.057;
CMIN/DF = 18.87; p(0.000)

Figure 7.
Measurement model
for F6 tolerance
of freedom

Notes: Goodness-of-Fit – CFI = 1.000; GFI = 1.000; RMSEA = 0.005;
CMIN/DF = 1.14; p(0.000)

Figure 8.
Measurement model
for F7 role assumption

Notes: Goodness-of-Fit – CFI = 0.990; GFI = 0.997; RMSEA = 0.056;
CMIN/DF = 18.50; p(0.000)

Figure 9.
Measurement model
for F8 consideration

value of 0.6 (Kaiser, 1970). Bartlett's tests of Sphericity (Nunnally and Bernstein, 1967)
reached statistical significance, supporting the factorability of the correlation matrixes.
Findings from structural equations modeling with AMOS showed the goodness-of-fit
indicated with RMSEA in our analyses with most variables have a good fit above the

Figure 10.
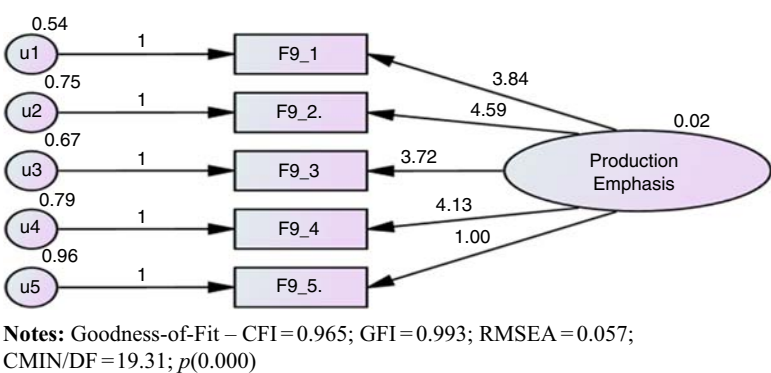Measurement model
for F9 production
emphasis

0.54 u1 — 1 → F9_1 — 3.84
0.75 u2 — 1 → F9_2. — 4.59
0.67 u3 — 1 → F9_3 — 3.72
0.79 u4 — 1 → F9_4 — 4.13
0.96 u5 — 1 → F9_5. — 1.00
Production Emphasis — 0.02

Notes: Goodness-of-Fit – CFI = 0.965; GFI = 0.993; RMSEA = 0.057;
CMIN/DF = 19.31; p(0.000)

Figure 11.
Measurement model
for F10 predictive
accuracy

0.59 u1 — 1 → F10_1. — 1.00
0.55 u2 — 1 → F10_2. — 0.92
0.37 u3 — 1 → F10_3. — 1.18
0.61 u4 — 1 → F10_4 — 1.00
Predictive Accuracy — 0.28

Notes: Goodness-of-Fit – CFI = 0.990; GFI = 0.993; RMSEA = 0.080;
CMIN/DF = 36.474; p(0.000)
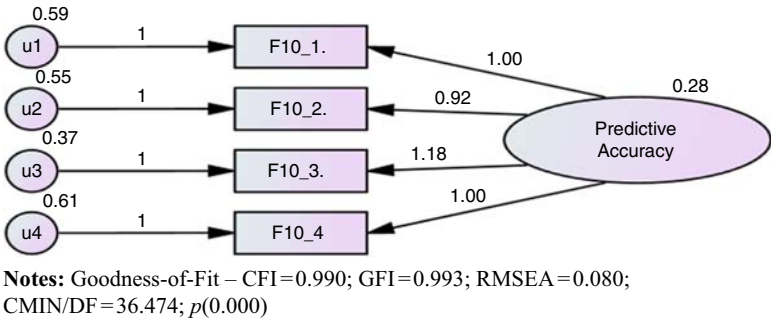
Figure 12.
Measurement model
for F11 integration

0.44 u1 — 1 → F11_1 — 0.94
0.35 u2 — 1 → F11_2. — 0.99
0.44 u3 — 1 → F11_3 — 1.00
Integration — 0.42

Notes: Goodness-of-Fit – CFI = 1.000; GFI = 1.000; RMSEA = 0.472;
CMIN/DF = 0.000; p(0.000)
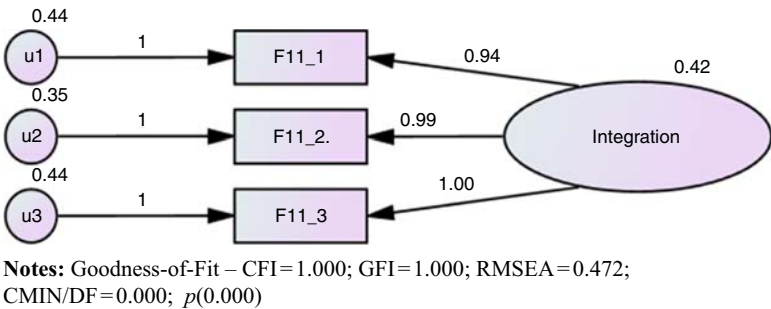
recommended level of less than 0.10; in addition, most goodness-of-fit values with GFI and CFI meet criteria with levels above 0.95. One possible explanation for why tests did not all produce acceptable results for all variables, especially in CMIN/DF tests, can be that the sample size is quite large, Anderson (1984, p. 156); these fit indexes are sensitive to sample size.

While dimension-to-dimension relationships are not a primary focus of this study, in order to show criterion-related validity, the 12 LBDQXII dimensions were correlated to the data from
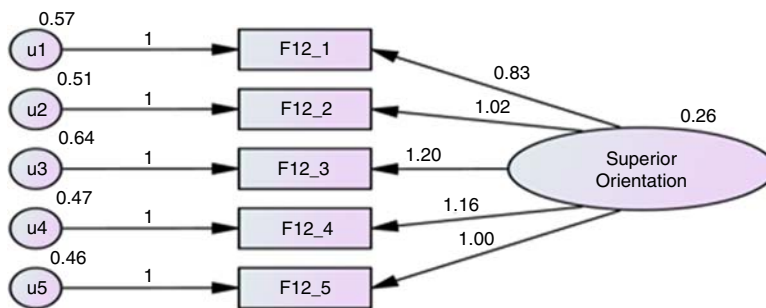
Notes: Goodness-of-Fit – CFI=0.992; GFI=0.997; RMSEA=0.50; CMIN/DF=15.0, *p*(0.000)

the respondents' cultural scores on the Hofstede 7D dimensions. Significant correlations were shown between four of the LBDQ dimensions and five of Hofstede's dimensions as follows: LBDQXII dimension 2: demand reconciliation and scores on Hofstede's dimension of Monumentalism. Also, between Dimension 3: tolerance of uncertainty and Hofstede scores for individualism, uncertainty avoidance, indulgence and monumentalism. Also, between Dimension 7: role assumption and Hofstede scores for power distance, indulgence and monumentalism. Furthermore, to support criterion validity, significant correlations were identified between Dimension 9, consideration and Hofstede's monumentalism.

Discussion

The results indicate that the 50-item short version of the LBDQXII is psychometrically acceptable. Indeed, for 5 of the 12 factors, there was an improved Cronbach's α with the 50-item scale compared to the 100-item scale. Comparative mean scores and comparative standard deviation scores indicate no significant differences between the 50-item scale means compared to the 100-item scale (Table II). We are confident of acceptable internal consistencies and comparable patterns of convergent, discriminant and criterion-related validity. Even more importantly, the mean factor score results of the LBDQXII 50-item instrument are comparable with those from the 100-item instrument with our data from 14 countries and $n = 6,451$, indicate that valid country comparison scores are also found in the shorter version. A valuable and novel finding with this "test-re-test approach" of running mean scores on the data, first with the 100 items, then with the 50, is that we see that we maintain very similar patterns of country comparisons and comparable scores in all 12 factors.

In summary, a practical shorter measure of the 12 factors of preferred leader behavior across cultures is developed and validated in this study. And as the role played by global managers in their immediate environment is culturally contingent, this underscores the need for managers to better understand the culturally specific follower-centric leadership profile in each society they work in. Individuals and organizations have different ways of achieving effective leadership, and the general conception of a leader whose effectiveness derives from his or her ability to engage in culturally sensitive leadership is clear.

On a methodological level, we contribute to the discussion of what statistical and judgmental logic procedures are needed when developing and validating shorter-item instruments in management research. We synthesized the validity, reliability and judgmental logic of Stanton *et al.* (2002) to protect against the danger of changing the underlying meaning of the 12-factor LBDQXII constructs when making a shorter scale. We followed the reliability approach of Donnellan *et al.* (2006) with a 20-item instrument to measure the five big personality traits.

Smith *et al.* (2000) caution that many well-intended researchers commit several "sins" in the process of developing shorter scales, such as not comparing reliability, validity and measurement results of the shorter measure to the original measure. We followed their advice and offer here clear guidelines regarding how to best compare psychometric properties of reduced scales to their "parent scales." Our theoretical contribution has been the mapping and synthesizing of the extant literature on the LBDQXII.

Results of the LBDQXII scale reduction project provide support for the quality and utility of the short version instrument to be used as the next generation of the LBDQXII instrument. Bass and Stogdill (1990) emphasized the need to maintain clarity regarding which instrument is used in empirical research when employing instruments like the LBDQXII, which originate from a different survey (LBDQ). Therefore, we suggest that a shortened, 50-item version of LBDQXII could be titled LBDQ50, contributing to clarity of the future empirical data and marking a new milestone in LBDQXII evolution.

Limitations
This study draws on data developed across cultures by an experienced research team studying and using the LBDQXII since 1996. While we are confident in the results of and interpretation of our analyses presented here, potential limitations of the study include:

(1) Salient leader preference dimensions particular to certain kinds of cultures have been overlooked. Future investigations replicating the original Ohio State study in the USA are planned.

(2) Additional research in Muslim-majority, Arabic-speaking countries is necessary, as well as studies of non-Arabic-speaking Muslim-majority countries.

(3) Our data set is limited in Southeast Asia and South Asia; research needs to be extended there.

Conclusions
We have addressed a continuous topic in management research, namely the pressing need for psychometrically sound yet shorter measurement scales to help us understand preferred leader behavior in global business. We detail a triangulation approach to scale reduction methodology, namely judgmental, validity and reliability methods. This can offer insights into both practitioners and scholars regarding quality and optimal length of any survey tool, both in and beyond management. We suggest that supplementing traditional reliability and validity methods for scale rationalization with judgmental logic is a must.

Leadership is clearly a set of complex, inter-related behaviors, and for global leaders to excel, understanding cultural expectations of what is "a good leader" for personnel and organizational management is a must. We believe that our study now offers "good fodder for theory development" (Caiazza *et al.*, 2018), both within scale validation and crafting the way forward for new leadership behavior development studies.

In closing, we note that shortening the length of established, yet overly-long instruments such as the LBDQXII may lead to subtle improvements in the experience and motivation of those participating in management research, one outcome that could yield big dividends for higher reliability and validity of the results and findings which can help global businesses manage cultural diversity better. We learned that it is possible to make very effective measures of leadership constructs with relatively few items. As such, we postulate that many other leadership instruments might be longer than necessary and therefore could be successfully shortened by taking an approach similar to ours – we invite other management scholars and consultants to take these steps.

References

Anderson, J.C. (1984), "The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis", *Psychometrika*, Vol. 49 No. 2, pp. 155-173.

Bass, B.M. (1990), *Bass and Stogdill's Handbook of Leadership: Theory, Research, and Managerial Applications*, 3rd ed., The Free Press, New York, NY.

Bass, B.M. and Stogdill, R.M. (1990), *Bass and Stogdill's Handbook of Leadership: Theory, Research, and Managerial Applications*, Simon and Schuster, Prentice Hall, Englewood Cliffs, NJ.

Boatwright, K.J., Lopez, F.G., Sauer, E.M., Van Der Wege, A. and Huber, D.M. (2010), "The influence of adult attachment styles on workers' preferences for relational leadership behaviors", *The Psychologist-Manager Journal*, Vol. 13 No. 2, pp. 1-14.

Caiazza, R., Cannella, A.A. Jr, Phan, P.H. and Simoni, M. (2018), "An institutional contingency perspective of interlocking directorates", *International Journal of Management Reviews*, Vol. 21 No. 3, pp. 277-293.

Chang, H.W. and Lin, G. (2018), "Effect of personal values transformation on leadership behavior", *Total Quality Management*, Vol. 19 No. 1, pp. 67-77.

Cronbach, L.J. and Meehl, P.E. (1955), "Construct validity in psychological tests", *Psychological Bulletin*, Vol. 52 No. 4, pp. 281-302.

Davila, A. and Crawford, M. (2018), "Including transcendental needs in a unified model of motivation", *Journal of Management Development*, Vol. 37 No. 5, pp. 385-396.

Donnellan, M.B., Oswald, F.L., Baird, B.M. and Lucas, R.E. (2006), "The mini-IPIP scales tiny-yet-effective measures of the Big Five factors of personality", *Psychological Assessment*, Vol. 18 No. 2, pp. 192-203.

Dorfman, P., Javidan, M., Hanges, P., Dastmalchian, A. and House, R. (2012), "GLOBE: a twenty-year journey into the intriguing world of culture and leadership", *Journal of World Business*, Vol. 47 No. 4, pp. 504-518.

Dorfman, P.W., Howell, J.P., Hibino, S., Lee, J.K., Tate, U. and Bautista, A. (1997), "Leadership in Western and Asian countries: commonalities and differences in effective leadership processes across cultures", *The Leadership Quarterly*, Vol. 8 No. 3, pp. 233-274.

Ebrahimi Mehrabani, S. and Azmi Mohamad, N. (2015), "New approach to leadership skills development (developing a model and measure)", *Journal of Management Development*, Vol. 34 No. 7, pp. 821-853.

Glynn, M.A. and DeJordy, R. (2010), *Leadership Through an Organization Behavior Lens. Handbook of Leadership Theory and Practice*, Harvard Publications, Boston, MA, pp. 119-157.

Hemphill, J., Siegel, A. and Westie, C. (1952), "An exploratory study of relations between perceptions of leader behavior, group characteristics, and expectations concerning the behavior of ideal leaders", unpublished staff report, Ohio State University, Personnel Research Board, Columbus, OH.

Hemphill, J.K. and Coons, A.E. (1950), *Leader Behavior Description Questionnaire*, Personnel Research Board, Ohio State University, Columbus, OH.

Hofstede, G. (2001), *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*, Sage Publications, Inc., Thousand Oaks, CA.

Hofstede, G., Hofstede, G.J., Minkov, M. and Vinken, H. (2008), *Values Survey Module 2008 Manual*, Institute for Research on Intercultural Cooperation, Maastricht.

House, R.J., Dorfman, P.W., Javidan, M., Hanges, P.J. and de Luque, M.F.S. (2013), *Strategic Leadership across Cultures: GLOBE Study of CEO Leadership Behavior and Effectiveness in 24 Countries*, Sage Publications, Thousand Oaks, CA.

House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W. and Gupta, V., Globe Associates (2004), *Leadership, Culture and Organizations: The GLOBE Study of 62 Societies*, Sage Publications, Inc., Thousand Oaks, CA.

Jøreskog, K. and Sorbom, D. (1993), *LISREL 8: User's Reference Guide*, Scientific Software, Chicago, IL.

Judge, T., Piccolo, R. and Ilies, R. (2004), "The forgotten ones? The validity of consideration and initiating structure in leadership research", *Journal of Applied Psychology*, Vol. 89 No. 1, pp. 36-51.

Kaiser, H.F. (1970), "A second generation little jiffy", *Psychometrika*, Vol. 35 No. 4, pp. 401-415.

Littrell, R.F. (2013), "Explicit leader behavior: a review of literature, theory development, and research project results", *The Journal of Management Development*, Vol. 32 No. 6, pp. 567-605.

Littrell, R.F., Warner-Søderholm, G., Minelgaite, I., Ahmadi, Y., Dalati, S., Bertsch, A. and Kuskova, V. (2018), "Explicit preferred leader behaviours across cultures: instrument development and validation", *Journal of Management Development*, Vol. 37 No. 3, pp. 243-257.

Mustafa, G. and Lines, R. (2016), "The emergence and effects of culturally congruent leadership: current status and future developments", *Entrepreneurial Business and Economics Review*, Vol. 4 No. 1, pp. 161-180.

Northouse, P.G. (2013), *Leadership. Theory and Practice*, 6th ed., Sage Publications, Thousand Oaks, CA.

Nunnally, J.C. and Bernstein, I.H. (1967), *Psychometric Theory*, Vol. 226, McGraw-Hill, New York, NY.

O'Connell, M.S., Lord, R.G. and O'Connell, M.K. (1990), "Differences in Japanese and American leadership prototypes: implications for cross-cultural training", Meeting of the Academy of Management, San Francisco, CA.

Rogelberg, S.G. and Stanton, J.M. (2007), "Introduction: understanding and dealing with organizational survey nonresponse", *Organizational Research Methods*, Vol. 10 No. 2, pp. 195-209.

Schmidt, F.L., Le, H. and Ilies, R. (2003), "Beyond alpha: an empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs", *Psychological Methods*, Vol. 8 No. 2, pp. 206-224.

Schriesheim, C.A. and Bird, B.J. (1979), "Contributions of the Ohio state studies to the field of leadership", *Journal of Management*, Vol. 5 No. 2, pp. 135-145.

Schriesheim, C.A., Cogliser, C.C. and Neider, L.L. (1995), "'Is it 'trustworthy'? A multiple-levels-of-analysis re-examination of an Ohio State leadership study, with implications for future research", *The Leadership Quarterly*, Vol. 6 No. 2, pp. 111-145.

Smith, G.T., McCarthy, D.M. and Anderson, K.G. (2000), "On the sins of short-form development", *Psychological Assessment*, Vol. 12 No. 1, pp. 102-111.

Stanton, J.M., Sinar, E.F., Balzer, W.K. and Smith, P.C. (2002), "Issues and strategies for reducing the length of self-report scales", *Personnel Psychology*, Vol. 55 No. 1, pp. 167-194.

Stogdill, R.M. (1963), *Manual for the Leader Behaviour Description Questionnaire-Form XII: An Experimental Revision*, Ohio State University, Bureau of Business Research, College of Commerce and Administration, Columbus, OH.

Van Dick, R., Lemoine, J.E., Steffens, N.K., Kerschreiter, R., Akfirat, S.A., Avanzi, L. and González, R. (2018), "Identity leadership going global: validation of the identity leadership inventory across 20 countries", *Journal of Occupational and Organizational Psychology*, Vol. 91 No. 4, pp. 697-728.

Corresponding author

Gillian Warner-Soderholm can be contacted at: gillian.warner.soderholm@bi.no