

Validation of a Leadership Scale in the Classrooms in an Indonesian University

Laurens Kaluge

Abstract: This article examines the Leader Behaviour Description Questionnaire Form XII (LBDQ XII) to determine whether it is applicable in Indonesia for teaching purposes. The instrument was translated into Indonesian with a slight modification in wording and distributed to the student teachers in State University of Surabaya. The sample, consisting of 300 subjects, was obtained by stratified random sampling from all of the faculties. By analysing the item characteristics (in terms of appropriateness, importance, and discriminating ability), confirmatory factor analysis, and calculating alpha and omega, the major results are: 59 items were considered applicable, 8 dimensions were confirmed by COSAN, and the sufficient high estimation of reliability.

Keywords: leadership behaviour, classroom teaching, test validation.

Leadership is a human phenomenon that occurs universally among all people regardless of culture (Bass, 1981). In spite of that, there are some differences in actual practices based on cultural values. Actually, leadership exists in all structures and levels of educational organisation that is defined as *educational leadership* for deeper study. Leadership has been regarded as an important factor in an organisation to empower people for certain purpose/s. Studies about and training in leadership are conducted for the desired purpose/s. But in fact, leadership in management has been more

Laurens Kaluge is a lecturer at the Faculty of Education, State University of Surabaya (Unesa).

developed than that in education because of such activities -the uses of behavioural science theory and research (Boles & Davenport, 1975; Bridges & Tierney, 1994).

Organisations in educational contexts (such as the classroom, school, board of education), of course need leadership. However, Boles and Davenport (1975) realised some difficulties in developing educational leadership because (1) all reading materials were aimed at persons who were aspiring to become superintendents or principals, (2) all available resources, seemingly, equated school administration with educational leadership, (3) none of the books gave the matter of how one leads more than passing attention. On the other hand, Fairman and Clark (1985) assert that apparently effective leaders are able to help faculties identify and clarify the meaning of basic skills within their schools and communities. Its effectiveness appears to be a prerequisite for providing a positive learning environment for all students within a school.

In the classroom setting, since a long time ago, teaching has been viewed as an act of leadership by some researchers such as Hemphill, Cantor, Rogers, Thelen (Bany & Johnson, 1964). They consider that teachers are leaders and that the teaching job can be studied as a series of leadership acts or roles. Bartky (1956), Amos and Orem (1967), Hoyle (1969), Kibble (1985), Harris (1986), Litchfield (1986) delineate classroom leadership as instructional leadership. However, there is no clear view about instructional leadership. Irwin (1985) points out that Instructional Leadership studies are generally classified under the major heading of School Effectiveness Research. He shows that these studies have some deficiencies: vagueness and lack of definition (key variables, e.g. 'leadership', 'instruction', 'student achievement', have not been clearly and/or operationally defined); standards (the use of absolute and arbitrary standards to assess effectiveness has not taken into consideration equivalencies for culturally disadvantaged or low socio-economic student learning. Extreme contrasts, and inappropriate comparisons have resulted); validity (there appears to be narrow focus, and the use of small samples, resulting in a lack of generalizability); accumulated research (in general, the bulk of instructional leadership research appears to have been done at the elementary, rather than at secondary school or tertiary levels).

The LBDQ XII is a research instrument developed in the United States and used for studying leadership. It can possibly be adopted for

Indonesian conditions, because, as stated before, leadership is a universal human phenomenon and because as Hoy and Miskel (1997) assert, the LBDQ has been used by students to describe teachers. However, Marsh (1985) points out the dangers that instruments in one setting will be used in new settings without first studying the suitability of the instruments to the new settings. In Indonesia, although there is no such valid and reliable instrument for research like this, Tan (1980) also warns social researchers to the same dangers of adopting an instrument from another country because of different context and culture. She suggests the necessity to reconsider its validity and reliability for Indonesian society and condition.

Based on such considerations, the main purpose of the study is to test the applicability of LBDQ XII as a reliable and valid instrument that can be used for identifying leadership behaviours of teachers in the classroom in Indonesia. In detail, the elaborated purposes were: to assess the appropriateness and the importance of items, to assess the discriminating power of items, to identify the dimensions of leadership behaviour, and to examine the reliability of the instrument.

METHODS

By request, the copyright owner, the Bureau of Business Research at Ohio State University sent the LBDQ XII and granted permission to use and to translate the instrument into Indonesian. Because this instrument was developed in America, it is highly (possible) that American expressions or slang were used - not easy to translate in Indonesian by simply using an English-Indonesian Dictionary. To overcome this problem, the writer took the following steps: (1) translating the instrument into Indonesian; (2) distributing two copies of the English and Indonesian version to two Indonesian Ph.Ds in English and Linguistics; (3) revising the translation based on the suggestions in step 2. In this step, some slight modifications had been made, i.e: "group" is changed into "class", "group member" is changed into "student", "work" is changed into "learn", "production" is changed into "learning outcomes"; (4) translating the Indonesian version into English. Then validate the translation by comparing the result (in English) with the original version of LBDQ XII; (5) revising the items based on step 4. A modification from 5 point scales was made to overcome the respondent doubt between the two close points.

To obtain the data for assessing the appropriateness, the importance, and the discriminating power of items which were stated as part of the research purposes, the instrument was designed to be rather different than the original: a blank column on the left of the page was available for the data about how frequently the behaviour was showed by the lecturer; two blank columns ("S" and "P") were added on the right side. The respondents were asked to tick under "S" column if they thought the item was appropriate, and under "P" column if it was important. On the contrary, if they thought the item was inappropriate or unimportant, they just put a cross; the top right of the first page was labelled, by using a pen, either A, B, or C. The meaning of these labels will be explained shortly.

The subjects were the students on the second and third years of undergraduate program. The first year students were not included because in this period they were adapting themselves to the campus life and they might not know enough lecturers. The final year students were also not included because most of them were busy with practical activities outside campus such as practice teaching in the isolated areas and Public Services (*KKN = Kuliah Kerja Nyata*). The present study used the proportional stratified random sampling procedure. Because no single stratum was small enough to be statistically insignificant, the researcher did not focus on any particular stratum, as Kerlinger (1990) indicates it has a high probability of being representative.

State University of Surabaya consists of six faculties which can be considered as strata: Faculty of Education, Faculty of Language and Arts, Faculty of Mathematics and Science, Faculty of Social Studies, Faculty of Engineering, and the Faculty of Physical Sciences. From the six faculties, 300 students were selected randomly and proportionally.

In the process of collecting data, the students were asked to describe their perceptions of the classroom leadership based on the LBDQ XII items. These questionnaires were labelled A, B, or C and mixed together and distributed to the students randomly. Students who got questionnaire A were requested to fill in their questionnaire based on their experience about a lecturer whom they regarded as a "good" leader. Students who got questionnaire B were asked to select lecturers whom they considered as an "average" leader for answering the questions. Those who got questionnaire C had to fill in the questionnaires based on their opinion on a lecturer whom they selected as a "poor" leader.

To achieve the purposes of the research, the analysis strategy was planned as follows: analysis of item characteristics. Each item was initially tested in terms of its appropriateness and importance by using the criterion of 80%; discriminant analysis was used to test the ability of each item to discriminate between the good and the poor lecturers as leaders; factor analysis was used to identify the dimensions of leadership behaviour as unities. The confirmatory factor analysis, in COSAN would be of benefit (Fraser, 1988); analysis of reliability. Alpha and omega coefficients were calculated to estimate the internal consistency of the instrument and the dimensions.

RESULTS

The characteristics of each item of LBDQ XII in terms of appropriateness and importance were judged by the percentage of students who considered each item to be appropriate or inappropriate, important or unimportant. The percentage was applied as a descriptive statistical technique. The criterion used for the judgment was 20%. This percentage was considered as extreme. If less than 20% of respondents judged an item as appropriate or important, it meant that the item could not be considered appropriate or important.

The discriminating ability of these items, to discriminate between good and poor teaching, has been highlighted by Norr and Crittenden (1975), dealing with evaluating college teaching as leadership. This procedure involved students in identifying and characterising their worst and best teachers. The procedure was fruitful for selecting effective items in a broad spectrum of teaching. The present study took the same procedure and analyses the data using *Discriminant Analysis*. By using discriminant analysis, the mean vectors of the leadership characteristics (good vs poor) in each item were evaluated by means of the Wilk's lambda test. Wilks' lambda is a statistic that has double advantages (Huck et al, 1974; Klecka, 1982; Youngman, 1979). First, it is useful in considering differences between groups. The lambda can be converted into an F statistic for testing the significance of differences. The larger the F value, the more likely it is to be significant. Second, it is useful in considering the cohesiveness or homogeneity within groups. The nearer the cases to their group centroid, the more cohesive they are and the smaller the lambda that will be produced.

Four original dimensions did not appear applicable: representation, tolerance of uncertainty, role assumption, and superior orientation. One dimension, initiating structure, loses 2 items and four dimensions (demand reconciliation, persuasion, tolerance of freedom, and production emphasis) loses one item in their construction. Three dimensions are considered completely applicable: consideration, predictive accuracy, and integration. In general, there are 8 applicable dimensions with 59 items. Four dimensions with 35 items and 6 item from other dimensions are not applicable.

As a result from this step, the inapplicable dimensions and items would be omitted from some of the further analyses. Those which were appropriate, important and had significant discriminating power, were used for identifying the dimensions and their reliability estimates.

For identifying the dimensions of instrument, the confirmatory factor analysis was used. Factor analysis was chosen because it provides a method of searching for a cluster of variables that correlate with one another and correlate less with members of other clusters and also for assessing the validity of empirical measures (Anastasi, 1997). Nunnally (1978) points out that factor analysis plays an important part with respect to three types of validity, i.e. predictive validity, content validity, and construct validity. Confirmatory factor analysis was used to confirm certain underlying dimensions considered as hypothesised dimensions.

The criteria used for examining the factor scales are: (1) the loading of equal or more than 0.3 is considered to be significant (Child, 1996; Youngman, 1979; Cohen & Manion, 1997); (2) a factor is considered as a useful dimension if it consists of at least 3 meaningful items (Youngman, 1979).

By using COSAN with the estimation of Least Squares (LS) and the 8 factor criteria, 59 applicable items were analysed. All items in that restricted factor are well confirmed with high loading. No factor has less than three significant and meaningful items. The factor intercorrelation matrix demonstrates high coefficients. The high coefficients mean that oblique solution as used in this analysis is the best way. The eight dimensions are demand reconciliation, persuasion, initiating structure, tolerance of freedom, consideration, production emphasis, predictive accuracy, and integration.

Reliability is the degree to which a test consistently measures whatever it measures. The more reliable a test is, the more confidence we can have that the scores obtained from the administration of the test are essentially the

same scores that would be obtained if the test were readministered. Reliability is expressed numerically, usually as a coefficient found by using one of some different methods. The present study chose to calculate omega because it provides the highest estimate of reliability that is closest estimate to the true reliability of the measure (Carmines & Zeller, 1994; McDonald, 1985). Comparing with Cronbach's alpha and theta, the position of omega is in the following order: alpha theta omega. McDonald asserts that alpha possibly has a very poor lower bound to reliability; however, this coefficient is still popular in many pieces of research publication. Due to such consideration, the omega method was carried out for cross checking the alpha approach.

Tabel 1 The Reliability of the Dimensions

Dimension	Alpha	Omega
1. Demand reconciliation	0.95	0.96
2. Persuasion	0.84	0.86
3. Initiating structure	0.84	0.88
4. Tolerance of freedom	0.78	0.82
5. Consideration	0.84	0.88
6. Production emphasis	0.83	0.85
7. Predictive accuracy	0.91	0.93
8. Integration	0.85	0.85

There is no absolute criteria for assessing reliability. Nunnally (1978) sets up the condition, if for example coefficient alpha is only 0.30 for a 40 item test the researcher should reconsider and begin the test construction anew. Mehrens and Lehman (1978) propose a reliability about 0.65 may suffice. The reliability coefficients as shows in Table 1 are sufficiently high. The internal consistency reliabilities of the eight dimensions satisfy the criteria set up by Nunnally, and Mehrens and Lehman as well.

DISCUSSION

In a classroom setting, the leadership dimensions of a teacher might be different from those of an educational administrator or any other leaders.

Ntgodill (1963) himself showed that in some settings certain subscales are not used or not so important. For example, integration and superior orientation are not used for describing community leaders, ministers, and senators as leaders. In this study, four subscales of the LBDQ XII were not fit to the classroom context. They are representation, tolerance of uncertainty, role assumption, and superior orientation.

The *representation* refers to the behaviour of a leader as a spokesperson who publicises and represents the group. But it is not a must for a lecturer to represent students. A student has the right to represent the class, so that the dependency of class members on the teacher can be reduced (Brownfield, 1973). A classroom has autonomy and the essential activities happen inside without the intervention of outsiders. The intervention can tentatively happen, not often as an essential part of lecturer's leadership in the classroom. Further more, the students can publicise their own activities or products.

Something can happen unexpectedly in certain conditions. For example sickness, or accidents to the members. As a result, the progress of the group is obstructed. The leader should be able to tolerate uncertainty and postponement without anxiety or being upset. The same situation might happen in the classroom. The role of a lecturer in teaching is fraught with uncertainty. The source of much of the uncertainty in education today can be attributed to the great social changes which have radically altered the traditional relationship between teachers and students (Watts, 1987). But such a situation can be tolerated under reasonable conditions. For teaching effectiveness, all the items in this dimension can be applied based on the students' good will in judging them. However the results show that *the tolerance uncertainty* subscale is not applicable. Perhaps this subscale may lead the leadership into a difficult situation in the classroom, especially relating to the class discipline.

A leader should actively exercise the leadership role rather than surrendering leadership to others. He should think and act properly to show clearly that he is not just a symbol of leadership. In the classroom, lecturers have a leadership role by the fact of their appointment. They have to teach the students in the class. They must accept their roles as class leaders with all their strengths and limitations (Watts, 1987). Because of such conditions, they should use various leadership tactics to overcome the problem and help to increase student capacity and creativity. They

are primarily coordinators of the groups and their ongoing activities. They appear to be leaders who mainly react to those behaviours that interfere with the flow of classroom events (Blumenfeld & Meece, 1985). In such situation, the *role assumption* subscale is not applicable.

Based on the hierarchy in the organisational structure, it is possible for a leader to have superiors in a superordinate level. The leader should maintain cordial relations with superiors, have influence with them, and strive for higher status. This is the situation where the *superior orientation* dimension is applied. However, the lecturers' superiors have influence in the outside of the classroom only. In the classroom, lecturers have their own privileges and the students do not need to know about their lecturers' relationship with superiors.

It is essential to recognise that in most university teaching situations the student and teacher must communicate with one another across the gap. Waugh indicates that the generation gap between teachers and students might create the communication problems. Hence, the teachers need to learn from, as well as with, their students in order to get mutual respect and to understand the manners and behaviours on both sides.

Besides those four dimensions, there are some inapplicable items in five dimensions of LBDQ XII. The dimensions are demand reconciliation, persuasion, initiating structure, tolerance of freedom, and production emphasis. Although *demand reconciliation* is an important dimension when there is a conflict between students in the classroom, one item fails to get a significant discriminating ability. The item is: getting swamped by details. This item appears as a non-effective one in solving the conflict among students.

Relating to *persuasion*, one item does not have significant discriminating power. The item concerns about 'persuading others that his/her ideas are to their advantage'. The expression of the item can be interpreted as a sly form of lecturer domination. Perhaps such a kind of domination produces a kind of reluctance in students, especially towards being involved in the practical politics of campus life.

Initiating structure is an important dimension of leadership. However, there are two ineffective items: (1) assigning students to particular tasks, and (2) maintaining definite standards of performance. The first item became ineffective due to the cultural influence which is called *Gotong Royong* (mutual cooperation) (Koentjaraningrat, 1995; 1985; Mintz, 1961; Peacock, 1973). According to Koentjaraningrat, this institutionalised con-

cept has positive and negative aspects. The *gotong royong* concept requires three main points: a harmonious life in the community, good relationships with other people, and being inconspicuous in a group or community. The last point is negative for developing individual capacity and recognising individual quality. If someone is going to be conspicuous, he will be mocked by the community: considered as having ambition to go ahead selfishly (Koentjaraningrat, 1995). Koentjaraningrat asserts that the *gotong royong* concept tends to treat individuals as equal, does not give a chance for them to develop individually, and does not appreciate individual work. The item of "assigning students to particular tasks" is not suitable for the Indonesian culture, especially in rural areas, which have *gotong royong* concept. The second ineffective item reveals the reaction to how the lecturers mark the students' work. In higher education, perhaps the students prefer using *norm-referenced* rather than *criterion-referenced* tests for marking their learning efforts.

Dealing with *tolerance of freedom*, one item has been discarded for non significant discriminating power. It is about reluctant to allow the students any freedom of action. This item expresses a dominant characteristics of leaders which is rather controversial to Indonesian culture. In the past, people were drilled by the ideas of *Pancasila Leadership* called *serasi, selaras, seimbang* (harmonious, compatible, and well-balanced). These terms indicate that a leader should not be too dominant and not too weak as well.

Learning outcomes as production are really one of the targets of educational leadership. It is expressed in the *production emphasis* subscale. However, one item can not perform well in discriminating ability. The item states driving hard when there is a job to be done. This item does not fit one of the Indonesian maxims *biar lambat asal selamat* (more haste, less speed). This maxim tends to be irrelevant to the modern style of life where everything is going fast, but it is still the reality in the Indonesian way of life.

CONCLUSIONS AND SUGGESTIONS

Conclusions

The leadership is a common phenomenon of social dynamics where people interact with each other. However, there are some similarities and differences in different places and contexts because of the nature of in-

teraction and cultural aspects. These were reflected by a great number of items which were considered respectively, applicable and inapplicable.

By the description of leader behaviours, good and poor leaders can be identified. The discriminating power of items showed significant differences between those characteristics.

There are similar underlying concepts of LBDQ XII in western and eastern (especially) countries. Sara (1981) executed a comparative study using LBDQ in four developing countries i.e Nigeria, Pakistan, Saudi Arabia, and Sudan and proclaimed it as a useful and universal instrument for leadership study. The LBDQ XII has not been used in such different countries, and now the present study reveals that demand reconciliation, persuasion, initiating structure, tolerance of freedom, consideration, production emphasis, predictive accuracy, and integration are still considered as dimensions of educational leadership in Indonesian higher institution classrooms.

A good instrument needs high internal consistency as measured by a reliability coefficient. Cronbach's alpha, which is considered appropriate in many pieces of research, is a poor lower bound of reliability as asserted by McDonald (1981; 1985), and Carmines and Zeller (1994). The results from this study show that the omega coefficients are better and higher than the alpha.

Suggestions

The results of the present study provide useful information for those who teach, especially in tertiary education classrooms, in order to improve their leadership. They can learn to assess and to improve their classroom environment through strategies that create an atmosphere conducive to teaching and learning.

In curriculum development, especially the courses related to the classroom leadership, the present findings give information about a set of behavioural descriptions to be relevant in Indonesia.

In teacher education, there is a need to know more about describing the desired behaviour in providing training to enhance the students' capacity for internalization and application (Cawelti, 1987). Referring to the results of this study, classroom leadership could be taken into consideration as a basic competency in teaching that can be described in terms of a number of dimensions and detailed variables.

Regarding the limitations of the study, further research could be performed by replicating the instrument in different classroom settings and reconsidering the possible dimensions that exist in Indonesian culture.

As the final remarks, the instrument produced by the present study can be used for further broad and in-depth investigation which relates leadership in the classroom to other aspects of the educational context, for instance quality of education, learning outcomes, class size, homogeneity, level of instruction and subject matter.

REFERENCES

- Amos, W.E. & Orem, R.C. 1967. *Managing Student Behavior*. St. Louis, Missouri: Warren H. Green.
- Anastasi, A. 1997. *Psychological Testing*. Upper Saddle River, NJ: Prentice Hall.
- Bany, M.A. & Johnson, L.V. 1964. *Classroom Group Behavior*. New York: The Macmillan.
- Bartky, J.A. 1956. *Administration as Educational Leadership*. Stanford, California: Stanford University.
- Bass, B.M. 1981. *Stogdill's Handbook of Leadership: A Survey of Theory and Research*. New York: The Free Press.
- Blumenfeld, P.C. & Meece, J.L. 1985. Life in Classroom Revisited. *Theory Into Practice*, 24 (1): 50-56.
- Boles, H.W. & Davenport, J.A. 1975. *Introduction to Educational Leadership*. New York: Harper & Row.
- Bridges, E.M. & Tierney, W.G. 1994. *The International Encyclopedia of Education*. Oxford: Pergamon Press.
- Brownfield, C.A. 1973. *Humanizing College Learning: A Taste of Hemlock*. New York: Exposition Press.
- Carmines, E.G. & Zeller, R.A. 1994. *Basic Measurement*. London: Sage Publication, Ltd.
- Child, D. 1996. *The Essentials of Factor Analysis*. London: Holt, Rinehart and Winston.
- Cohen, L. & Manion, L. 1997. *Research Methods in Education*. London: Croom Helm.
- Fairman, M. & Clark, E. 1985. Moving Toward Excellence: A Model to Increase Student Productivity. *NAASP Bulletin*, 69 (477): 6-11.
- Fraser, C. 1988. *COSAN User's Guide*. Armidale: Department of Behavioural Studies, University of New England-Australia.
- Harris, B.M. 1986. *Leadership for Quality Instruction: Looking toward the 21st Century*. Educational Resources Information Centre (ERIC), ED 270 885.
- Hoy, W.K. & Miskel, C.B. 1997. *Educational Administration: Theory, Research, and Practice*. New York: Random House.

- Hoyle, E. 1969. *The Role of the Teacher*. London: Routledge & Kegan Paul.
- Huck, S.W., Cormier, W.H. & Bounds, W.G. 1974. *Reading Statistics and Research*. New York: Harper and Row.
- Irwin, C.C. 1985. *What Research Tells the Teacher about Instructional Leadership*. Educational Resources Information Centre (ERIC), ED 257 173.
- Keeler, B.T. & Andrews, J.H.M. 1963. Leader Behavior of Principals, Staff Morale, and Productivity. *Alberta Journal of Educational Research*, 9: 179-191.
- Kerlinger, F.N. 1990. *Asas-Asas Penelitian Behavioral*. (Translated by L.R. Simatupang). Yogyakarta: Gadjah Mada University Press.
- Kerr, S. & Jermier, J.M. 1978. Substitutes for Leadership: Their Meaning and Measurement. *Organizational Behavior and Human Performance*, 22 (3): 375-403.
- Kibble, D. 1985. Leadership in the Classroom. *Education Today*, 34 (1): 25-31.
- Klecka, W.R. 1982. *Discriminant Analysis*. Beverly Hills, California: Sage.
- Koentjaraningrat. 1985. *Javanese Culture*. Singapore: Oxford University Press.
- Koentjaraningrat. 1995. *Manusia dan Kebudayaan di Indonesia*. Jakarta: Djambatan.
- Litchfield, D.J. 1986. If You Want Me to Be Instructional Leader, Just Tell What An Instructional Leader Does. *Peabody Journal of Education*, 63 (1): 202-205.
- Marsh, H.W. 1985. *The Applicability Paradigm: Student Evaluations of Teaching Effectiveness in Different Countries*. Sydney: University of Sydney.
- McDonald, R.P. 1981. The Dimensionality of Tests and Items. *The British Journal of Mathematical and Statistical Psychology*, 34: 100-117.
- McDonald, R.P. 1985. *Factor Analysis and Related Methods*. New Jersey: Lawrence Erlbaum Associate.
- Mehrens, W.A. & Lehman, I.J. 1978. *Measurement and Evaluation in Education and Psychology*. New York: Holt, Rinehart and Winston.
- Mintz, J.S. 1961. *Indonesia: A Profile*. Princeton, New Jersey: D. van Nostrand.
- Norr, J.L. & Crittenden, K.S. 1975. Evaluating College Teaching as Leadership. *Higher Education*, 4: 335-350.
- Nunnally, J.C. 1978. *Psychometric Theory*. New York: McGraw-Hill.
- Peacock, J.L. 1973. *Indonesia: An Anthropological Perspective*. Pacific Palisades, California: Goodyear Publishing.
- Stogdill, R.M. 1963. *Manual for the Leader Behavior Description Questionnaire - Form XII*. Columbus, Ohio: Bureau of Business Research, Ohio State University.
- Tan, M.G. 1980. Penggunaan Data Kuantitatif. In Koentjaraningrat (Ed.), *Metode-Metode Penelitian Masyarakat* (pp. 307-327). Jakarta: Gramedia.
- Watts, L. 1987. Classroom Management: A Group Affair. *The Practicing Administrator*, 9 (1): 42-46.
- Youngman, M.B. 1979. *Analysing Social and Educational Research Data*. London: McGraw-Hill Book Company.