DOCUMENT RESUME

ED 376 120 SP 035 503

AUTHOR Tirri, Kirsi; Tirri, Henry
TITLE Applying Teacher Effectiveness Models in International, Cross-Cultural Contexts: A Case Study.
PUB DATE Apr 94
NOTE 28p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; *Comparative Education; *Cross Cultural Studies; Cultural Context; *Cultural Differences; Elementary Secondary Education; Foreign Countries; Higher Education; International Studies; Teacher Characteristics; Teacher Education; *Teacher Effectiveness; *Teaching Models
IDENTIFIERS *Developmental Teacher Evaluation Kit; *Finland; Indiana; Texas; United States

ABSTRACT

This study investigated the problem of applying a single model in international cross-cultural studies of effective classroom teaching. The context of the social-efficiency tradition was used as a framework for understanding how teacher education is reflected in actual teaching behaviors. The model chosen for study was the Developmental Teacher Evaluation Kit, developed by B. Harris and J. Hill, which is a six-component model based on American teaching effectiveness research. Its international applicability as a model of the basic skills used in effective teaching was conducted. U.S. teachers (N=167) from Indiana and Texas, and Finnish teachers (N=172) from 2 different areas were presented with 95 behavior statements in 6 performance areas (businesslike, friendly, verbally interactive, stimulating, individually oriented, and multimedia integrative) and were asked to evaluate their classroom behaviors against the described behaviors. U.S. data matched on five of the components, all except verbally interactive, while Finnish data matched on only two components--multimedia integrative and businesslike. Data show that the model is biased with respect to the different teacher education traditions, and it fits better the data from the tradition within which it was developed. (Contains 35 references.) (JDD)

Applying teacher effectiveness models in international, cross-cultural contexts: A case study

Kirsi Tirri

Department of Teacher Education

University of Helsinki

email: KTIRRI@cc.Helsinki.FI

Henry Tirri

Department of Computer Science

University of Helsinki

email: tirri@cs.Helsinki.FI

Running head: TEACHER EFFECTIVENESS IN INTERNATIONAL CONTEXTS

Objectives

The purpose of this study is to investigate the problems of applying a single model in international cross-cultural studies of effective classroom teaching. For such a model we have chosen the six component model used in the DeTEK system (Harris 1986, Harris & Hill 1982). This model is based on American teaching effectiveness research, mostly from 1960s and 1970s (Dunkin & Biddle 1974, Waxman & Walberg 1991). However, as we show in the theoretical part of our work, it still fits the modern context and reform traditions which affect teacher education in both of the countries we focus on, the United States and Finland.

In the empirical main part of our work we perform an extensive study on the international applicability of the model in question as a model of the basic skills used in effective teaching. The general framework for this study is an attempt to address several questions:

- Q1: What is the applicability of a case model (the Harris model) and the possible bias introduced by its origin in a particular educational tradition?
- Q2: What general problems can be identified in applying any single model in such a study?
- Q3: What type of a methodology should one use to perform comparisons of this cross-cultural nature?

In this paper we will focus on the first question (Q1) and only briefly mention issues related to the second item (Q2). More detailed accounts on some of the other aspects can be found in (Tirri 1993). Related to the question Q1, since the model used is based on American research tradition as well as empirical studies in the United States, one of our tasks here is to study the question whether or not the model is biased with respect to the different teacher education (teaching) traditions. We will show that indications for such a bias can be found in our data. Encouraged by this result we also investigate further the refined hypothesis that the model fits better the data from the tradition from within it was developed. It is obvious that due to the difficulties in

arranging the data gathering, empirical results of this type of an international cross-cultural comparison are scarce, which makes our observations even more valuable. As significant differences can be shown to exist in the straightforward attempt to fit the model structure to the data, we also perform a more detailed identification of the various abstract structures underlying the teacher classroom behavior, i.e., reveal structures underlying our data. This allows us to identify invariant components of such structures as well as more variable ones.

Theoretical Framework

Evidently the study is related to the studies in comparative education, more precisely to comparative pedagogy (Halls 1990). Since different teacher education traditions are reflected in the actual teaching behavior of the teachers from different countries, the reform traditions in teacher education identified by Liston & Zeichner (1991) are used. These reform traditions are: the Academic tradition, the Developmentalist tradition, the Social-reconstructionist tradition and the Social-efficiency tradition. The framework of the U.S. teacher education traditions as suggested by Zeichner (1992) has also been used as a heuristic in helping to identify the Finnish traditions in teacher education. In Finland teacher education is influenced by all of these traditions (Tirri 1993). However, here we restrict our discussion to the context of Social-efficiency tradition.

An effective teacher as a goal: The Social-efficiency tradition

The social-efficiency tradition in teacher education reform attempts to build a scientific base for the study of teaching. This knowledge base would provide the basis for building a teacher education curriculum. This reform tradition has much of the same content as the behavioristic paradigm in teacher education identified earlier by Zeichner (1983). Both these movements, whose goal is to produce effective teachers, draw from the teacher effectiveness research. Effective teaching is usually identified as teaching which causes most learning in the students (Dunkin & Biddle 1974, 13-14).

In his recent paper Zeichner contrasts two major traditions of practicum reform in the United States. They are the research-based practicum where teaching is seen as an applied science in one hand and as reflective practice on the other hand (Zeichner 1990). The key contributors to the applied science position are Gage and Berliner. In his book "The Scientific Basis of the Art of Teaching" (Gage 1978) Gage argues that there is a scientifically based knowledge about methods of teaching. According to Gage students should learn about the techniques of teaching whose efficacy has strong empirical support. Teachers need to have teaching methods as well as subject matter knowledge.

Gage advocates the broad version of the applied science perspective. In the broad version teachers would use the research as principles of procedure within a broader process of decision-making and problem-solving. As he states in his book "statistical results can help a teacher know the averages or trends around which individual cases will vary, and such knowledge can aid in understanding the individual" (Gage 1985, 4-5). According to Gage research results can be seen as helpful starting points, not prescriptions to be followed in all circumstances.

Berliner has argued that for the first time teacher education has a scientific foundation (Berliner 1984, 94). He thinks it is time to restructure teacher education programs including the practicum to provide for more systematic training in the knowledge, skills, and decision-making strategies that research has identified as being associated with desirable school outcomes. Berliner has called for the creation of pedagogical laboratories that provide experimental conditions for student teachers to try out the behaviors and teaching strategies to be learned (Berliner 1985).

In this study we join the broad version of applied science perspective. The self-evaluation instrument we have used is meant to be seen as a helpful starting point in the teacher's reflection on his own classroom behavior, not as a full prescription to be followed. The knowledge of teaching effective research and its findings should be one of the resources a reflective teacher uses (with critical mind) in guiding his own professional development.

Features of this teacher effectiveness trend can be found in Finnish teacher education (Puurula 1983). In the teacher training the student teachers are expected to master certain teaching skills. In the 1970's and 1980's these skills were based on S.C.T. Clarke's general teaching theory (Clarke 1970). Clarke's theory was widely used in assessment of student teachers in teacher training institutes. The modern trend has complemented this effectiveness approach by emphasizing teachers' pedagogical thinking (Kansanen 1991).

One should observe that in the current debates on teacher education reform the social-efficiency tradition has emerged again under the label of "research-based" teacher education. For example the proposals of the Holmes Group (1986) are very strongly influenced by this tradition. Throughout the century the common thread that ties the different approaches among this tradition together has been their reliance on the scientific study of teaching as the major source for determining the teacher education curriculum (Zeichner & Tabachnick 1991, 5-6).

The social-efficiency version of reflective teaching emphasizes the thoughtful application of particular teaching strategies that have been suggested by research on teaching. This version of reflective teaching falls into the conceptual orientation which Grimmett et al. (1990) have identified mediation of action where knowledge is used to direct practice as instrumental. This orientation can be seen as a technical definition of reflection. Examples of these perspectives are the reflective inquiry teacher education program (RITE) at the University of Houston (Freiberg & Waxman 1990), the PROTEACH program at the University of Florida (Ross & Kyle 1987) and Cruickshank's reflective teaching program at Ohio State University (Cruickshank 1987).

Although the programs differ in many details they all have the same emphasis on the intelligent use of "generic" teaching skills and strategies that have been suggested by research (Zeichner & Tabachnick 1991, 5-6).

DeTEK criteria for good teaching in the light of current research on effective teaching

The teaching effectiveness research used in developing DeTEK criteria is mostly dated from 1950 through 1975. Thus we want to adapt the DeTEK criteria to the modern context also and review the findings of the current teacher effectiveness research (1975 1990) with respect to the criteria chosen in DeTEK. The DeTEK criteria form a six component model. These components, which are called performance areas group together related criteria in order to form more abstract dimensions of effectiveness. These six components are: Businesslike, Friendly, Verbally Interactive, Stimulating, Individually Oriented and Multi-Media Integrative.

Behaviors identified in the DeTEK model are very much in accord with the current research on effective teaching (Waxman & Walberg 1991). The performance area "Businesslike" including teaching behaviors like time-on-task, teacher organization and classroom management skills, has especially strong support from the research. Similarly the verbal skills in the performance area "Verbally Interactive" are getting support from research especially in the area of clarity. The research approves a friendly, positive classroom climate with regular feedback from the teacher who avoids negative criticism. Harris's performance area "Friendly" has many teaching behaviors that describe a teacher who is capable of creating a warm atmosphere in the classroom.

There is evidence of the effectiveness of stimulating teaching behavior (Harris's performance area 4), but the appropriate level of teacher enthusiasm is under discussion. Individualized teaching can be justified in certain contexts but the evidence generally supports whole classroom teaching, where time-on-task can be maximized to all students. Harris's performance area Individually Oriented is defined more broadly than individualized teaching and emphasizes the need for every student to be treated as an unique learner. The use of multi-media in teaching can be supported by the need for variability, use of different teaching strategies and media to bring diversity to the classroom. More detailed discussion on these aspects can be found in (Tirri 1993, 62-71).

The theoretical framework for the methodology of our study is connected to the much debated notion of construct validation (Cronbach & Meehl 1955, Cronbach 1989). For our purposes the abstract discussion of general issues related to

construct validity reduces to the (in principle) simple question of finding positive or negative evidence in our data for the existence a theoretical artifact, the chosen model of effective teaching.

Methods

The instrument

The self-evaluation instrument used consisted of 95 behavior statements from the criteria list included in the DeTEK system (Harris and Hill 1982). These criteria were originally chosen based on the results of teacher effectiveness research, current research theory and professional wisdom. Like many others, Harris builds on Ryans's (1960) classic study of teacher characteristics. Choosing Ryans's characteristic patterns as the basis of the performance areas in the model is naturally debatable. For a more thorough discussion on choosing the criteria, the reader should consult the original text by Harris (Harris 1986, 72).

The self-evaluation instrument had two sections: background information and the main evaluation against the DeTEK criteria using a scale from 1-6. All the behavior statements for this part of the questionnaire were mixed together randomly including descriptions from all the levels of the original instrument: from the abstract main behaviors and from the indicators which describe the teacher behaviors in a very concrete way. For the gathering of Finnish data, the behaviors were translated directly into Finnish, though naturally in some of the cases the descriptions had to be slightly modified to fit the Finnish context. However, there are no essential differences in the descriptions from the original ones, except some divisions of disjunctive descriptions into several individual descriptions, together with the fact that some of the descriptions available were omitted to make the size of the instrument tractable.

The descriptions of teaching behaviors chosen represent all the six performance areas identified by Harris. In our instrument there are 14 (21)[1] behavior statements from performance area I (Businesslike), 22 (23) behavior descriptions from the performance area II (Friendly), 15 (14) from the performance area III (Verbally Interactive), 22(24) behavior statements from the area IV (Stimulating), 10 (21) from the area V (Individually Oriented) and 12 (19) behavior descriptions from performance area VI (Multi-Media Integrative). All the behaviors are presented in the context of classroom teaching and the teachers were asked to evaluate their classroom behaviors against the described statements using the given scale.

The analysis methods

The choice of suitable analysis methods are guided by our focus of interest; namely correlation between the different variables representing the behavior criteria of the underlying model, and grouping of these variables into more abstract concepts, performance areas, to form a model. Thus factor analysis is a natural choice for inclusion in the set of analysis methods to be used. Although the distributions of the measurements scores were far from normal, in the analysis phases the data was not normalized. Performing the normalization and checking the results revealed the results to produce a two-valued scale which would have collapsed the detailed variance information.

As stated earlier our empirical focus is twofold. In the first phase we concentrated on an attempt to fit the Harris six component model to the data gathered from the USA and Finland. Here we are facing a methodologically interesting situation. We have a reasonably well-defined model whose validity could in theory be tested with a confirmatory factor analysis (Jöreskog 1969). In such an analysis the space of possible loading matrices will be constrained by fixing some of the loadings to constants, and the nature of confirmatory factor analysis would satisfy our hypothesis testing requirement.

---

1 The original number of statements in the DeTEK model is indicated in the parenthesis.

Unfortunately in practice the high dimensionality of our data (95 variables, expected number of primary factors 18[2]) prohibits us from using methods such as LISREL (Jöreskog & Sörbom 1976) due to computational restrictions and the limited sample size. Consequently we performed a forced (six factor) factor analysis on the data sets, and relate the structures identified to the Harris model by comparing the corresponding variable sets. Although not as accurate as using "true" confirmatory models, such visual inspection gives us valuable information about the differences in the degree of fit between the US and Finnish data.

In our case the latent variables representing the various dimensions of effective teaching are naturally all highly correlated, and thus to achieve the simplest interpretation structure the use of direct oblimin rotation (Harman 1976, 334-341) is more appropriate than varimax solutions.

Although the visual inspection for the forced factor solutions did not result in a very good match, this does not give us much more information beyond the simple fact that a trivial one-to-one match is unlikely. Even if the underlying structures in the data are similar to the Harris model components, underfactoring tends to distort the results. Hence from the initial attempt to compare Harris's model and the forced factor solutions we proceeded by performing a unconstrained exploratory factor analysis for each of the data sets.

Any exploratory analysis has to be preceded by observations relating to the appropriateness of its use. Consequently we have tested both the hypothesis that the correlation matrices in question are identity matrices (Bartlett's test of sphericity) and the relationships of observed correlation coefficients to the magnitudes of the partial correlation coefficients (the Kaiser-Meyer-Olkin measure). In both of these respects the

---

2 The number of primary factors could be estimated from the previous study (Tirri 1991) where the same instrument was used for the subject-specific data of the current study. The assumption of the complex substructure of the Harris six component model was also supported by the properties of the method which Harris used to construct the model (see discussion in (Tirri 1991)).

correlation matrices seem to satisfy the requirements for factor analysis extremely well. The Bartlett sphericity test values ranged from about 10000 to 21000 with 0 significance level, and the KMO values from .81 to .95 for the whole of the data, values which Kaiser (1974) characterizes as meritorious or marvelous.

One of the central issues in exploratory analysis is the question of deciding the right number of factors; i.e., when to stop factoring. Using the typical criteria (KMO, scree-test) lead to the discovery of 18 primary factors. As Harris's model clearly is one with higher order composite factors, for comparison purposes we also performed second order factorization with factor scores as representatives of the primary factors. For calculating factor scores we chose the regression method over the Anderson-Rubin method as the latter always produces non-correlated scores even in the case where the original factors are estimated to be correlated. For these second order factors we performed a visual inspection against the Harris model components, but in this case the structures themselves also had a value of their own.

Data source

A group of teachers from Indiana and Texas (N=167) representing American teachers, and about the same number of Finnish teachers from two different areas (N=172) were chosen as the sample used for the study. A common self-evaluation instrument (for the Finnish group naturally translated into Finnish) was administered to both of the groups. The American teachers were selected using the expert opinion of the representativeness of the school districts chosen. The superintendents of the chosen districts and the principles of the schools assisted in the administration of the survey. In the case of Finnish teachers most of the questionnaires were distributed with the help of a research assistant who contacted the principals of the schools. The Finnish teachers consist of a sample of teachers both from the southern and a northern parts of Finland. A summary of the data statistics is presented in Table 1.

insert Table 1 approximately here

Results

How well did the Harris model fit the data - the results of the forced factor analysis

The easiest way to get an overall picture (to answer question Q1) is to check how the Harris model component variables are distributed among the factors in the forced 6 factor solution. In the ideal case of a perfect match with the Harris model one would expect to find a one-to-one match between the factors discovered and the components, and in the worst case a uniform distribution of the component variables into each of the factors. It is predictable that neither of these extremes is the case at hand.

Figures from 1 to 3 depict the Harris model component variable distributions for the complete set, and the US and Finnish data sets. From visual inspection of these histograms several interesting observations can be made.

---

insert Figure 1 approximately here

For the complete data one can clearly identify correspondence of two Harris components only: Businesslike (factor 4) and Friendly (factor 2). For the rest of the components the distribution has at least two significant peaks, although in the case of Stimulating one of the peaks is clearly more dominant. Even the Multi-Media Integrative component, which can be identified clearly both in the US and Finnish data sets has a two-peak distribution with peaks of almost equal in size.

The match for the US data is unquestionably the best. One can identify good matches for five of the Harris components: Individually Oriented (factor 1), Multi-Media Integrative (factor 2), Stimulating (factor 3), Friendly (factor 4) and Businesslike (factor

5). The only mismatched component is thus Verbally Interactive. This is a component which tends to be easily confused with components such as Stimulating and Businesslike as all of them involve teaching behavior with many verbal components. One should also observe that the match is nicely partitioned, none of the matched factors having high peaks for two Harris components.

---

insert Figure 2 approximately here

For the Finnish data the match is even less evident than for the complete data (which of course should be obvious as the clear match in the US data affects the match in the full set also). One can only identify two Harris components with a clear match: Multi-Media Integrative (factor 1) and Businesslike (factor 4). In addition the variable distribution for the rest of the components is closer to uniform than the corresponding distributions for the complete data.

Clearly one cannot find a complete fit (i.e., match all performance areas with unique factors) between the Harris 6 component model and our data. This was predictable already from the earlier preliminary study with the subject-specific data (Tirri 1993). However, in the preliminary study it was not clear to what extent the poorness of fit was due to the specific context of religious education. Interestingly in the current study the subject-specific data did not show significant variance from the other data sets, thus we can now with high confidence conclude that the previous observations are generalizable. However, even the worst case represented by the Finnish data clearly presents two of the Harris's components: "Multi-Media Integrative" and "Businesslike". This result is also supported by the variable level inspection – both of these components contain clear, unambiguous behavior descriptions which both the Finnish and American teaching traditions value as means for effective teaching (e.g., "Time-On-Task", "Use of audiovisual equipment" etc.). All this indicates that the underlying primary factors in the

used model are to a high degree invariant, but the grouping into the higher, second order structures is more debatable.

---

insert Figure 3 approximately here

The answer to the second major question of interest, "Is the Harris model construct culturally biased?" is already much more difficult to deduce from our analysis. By visual inspection the match of the six component model with the US data is unquestionably the best. One could identify counterparts for five of the Harris components: "Individually Oriented", "Multi-Media Integrative", "Stimulating", "Friendly" and "Businesslike". The only incompatible component is thus "Verbally Interactive", a component which tends to be easily confused with other components (e.g., "Stimulating" and "Businesslike"). On the other hand, for the Finnish data one can only identify two Harris components which have a clear counterpart. These observations seem to support the cultural dependency of Harris model, even with the possible distortions caused by the heterogeneity of the Finnish data set.

The invariant subcomponents in the Harris model - the results of the exploratory factor analysis

In the previous Section we describe the analysis of the data by constraining the factor structure space to solutions with six factors to be able to compare the results to the Harris six component model structure. Interesting structural differences across the data sets were identified. However, it is very likely that the Harris six component structure represents a very abstract model, and that the components themselves have an underlying substructure. Two indicators support this hypothesis. First, the method Harris used to construct the components is based on aggregates, for example by combining several dimensions empirically discovered by Ryans (Harris 1986, 71). Secondly a previous factor analytical study (Tirri 91) revealed a possible substructure for the components.

Consequently, one can reasonably assume that an unconstrained factor analysis would produce a vastly higher number of factors than six. To explore the unconstrained factor structures underlying the data we first performed a factor analysis for the data as a whole, and then on the US and Finnish data separately. Indeed much larger structures with 16 to 18 factors were extracted. These factors can be viewed as primary factors since in many cases they are specific subcomponents of Harris's components. The primary factor structures identified in both US and Finnish data are shown in the comparison presented in Figure 6.

Identification of such primary structures also allowed us to attempt to discover more abstract structures at the same abstraction level as the original Harris components without constraining the number of factors in the analysis, as such constraints can have a distorting effect. This was achieved by performing a second order factorization for the US and Finnish data using the factor scores of the previous analysis phases. It can be shown that the abstract structures discovered not only differ from each other, but also do not exhibit clear relationships with the original Harris components (even the number of factors in the second order factorizations differ from the six assumed in the Harris model).

---

insert Figure 4 approximately here

Since our interest is focused on cross-cultural differences we also compared the primary factor structure of the US data to the corresponding structure of the Finnish data. Since all the calculated matrices are available, we are able to perform the comparison based on the rotation method suggested by Kaiser, Hunka and Bianchini (Kaiser et al 1971) by using a slightly modified FACTREL software module. This gives us a more rigorous foundation for relating the primary factors in the two data sets than a purely visual inspection, or the use of simple statistical indicators such as the coefficient of congruence. In the light of the previous negative results by the forced factor analysis,

somewhat unexpectedly the primary structures exhibit a good structural match, in which 66% of the primary factors can be related to a unique mate. This interesting result is clearly illustrated in the summary Figure 6.

---

insert Figure 5 approximately here

Consequently a more detailed analysis of the data indicates that the apparent cultural bias is likely to be more a property of the second order structure; i.e., the particular grouping of primary factors performed by Harris, than for the elementary factor structure. Such an argument stems from the results of the explorative factor analysis performed for the data sets. The factor structure comparison revealed a remarkably high similarity between the two primary factor structures (the US and the Finnish data), while at the same time one could not identify a good match between the secondary structures for the same data sets (i.e., the second order factors). An obvious conclusion from this observation is that the elementary structures for effective teaching seem to be more invariant than the higher abstractions built on them. We conjecture that this is not a peculiarity of the particular model studied, but this phenomenom is generalizable also to other models of this nature.

---

insert Figure 6 approximately here

What general problems could be identified

When comparing measurements from different cultures the central problem always reduces to the same fundamental question of whether it is at all possible to scientifically compare measurements from different social and/or cultural systems and its units? To

illustrate the problems applicable to a cross-cultural study such as ours we briefly point out here some generic issues with concrete examples (thus providing only partial answer to question Q2). Needless to say, all of these aspects also affect the validity of our comparison task.

- Do the concepts being compared correspond? Are they similarly situated on the general-specific continuum? For example, we need to relate qualifying concepts such as "first level education" from different cultures. According to the Unesco Statistics yearbook (Anon. 1988) the term "first level" denotes grades 1-6 in Finland, but in US it also includes Kindergarten class (from K to 6th grade). Although in our case this does not have any substantial influence on the test situation, differences like this make drawing proper conclusions a very delicate task.

- How is the correspondence of measurements to be assessed? Do the indicators given to concepts correspond to each other? Or is it the case that although the definitions are identical, their semantics are different in different contexts? For example one of the instrument item assertions "I accept disagreements" is evidently understood differently among the Finnish teachers than their American colleagues.

- How easily can the concepts be identified? A theoretically perfectly clear concept may present operational and linguistic difficulties in its formulation. For example translating the variable "I am an exciting teacher" as "Olen jännittävä opettaja" is not a semantically equivalent translation, but cannot easily be improved either.

- Can the problem of how concepts are linguistically expressed be resolved? Different cultures verbalize different aspects of the same concepts. It has become customary to use bilingual or multilingual experts and repeated translation back and forth until unclear points are eliminated. In the case of an regular study, the cost of using experts is prohibitive.

Finally a fundamental issue related especially to our work applicable to most related work also. The study performed can be seen as a study of issues relating to the construct validity of the Harris model of effective teaching. To demonstrate construct validity one should focus both on convergent and discriminant validation (Anastasi 1988, Moss

1992, 233). For convergent validity one has to show that a particular test (behavior description in our case) correlates highly with variables with which it should theoretically correlate. For discriminating validity one attempts to show that the test does not correlate significantly with variables from which it should differ. If a test is both convergently and discriminantly valid, it is highly selective and measures one structure only. In our case we are able to test convergent validity; i.e., the intercorrelation of the variables within a Harris performance area. However, since it is already evident from the theoretical background that the different model components (performance areas) correlate with each other, we cannot hope to achieve discriminating behavior even in theory.

Importance of the Study

This study reports rare empirical case study of an important topic: cross-cultural issues related to applying models of effective teaching internationally. It raises interesting methodological questions and offers information about the behavior of a very typical model developed in American teaching tradition, when applied in another cultural context. The work at hand offers also a view to some of the very interesting and important aspects related to effective teaching, the various ways of understanding the concept and the issues related to the cross-cultural invariance of the notions involved.

References

Anastasi, A. 1988. Psychological testing (6th ed.). Macmillan. New York.

Anon. 1988. Statistical Yearbook 1988. Unesco.

Berliner, D. 1985. Laboratory settings and the study of teacher education. Journal of Teacher Education, 36, 2-8.

Berliner, D. 1984. The half-full glass: a review of research on teaching. In P. Hosford (Ed.) Using what we know about teaching. Alexandria, VA, Association of Supervision & Curriculum Development.

Clarke, S. 1970. General teaching theory. The Journal of Teacher Education, 21, 403-416.

Cronbach, L. (1989). Construct validation after thirty years. In Linn, R. (Ed.) Intelligence. Measurement, Theory, and Public Policy. University of Illinois Press.

Cronbach, L. & Meehl, P. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Cruickshank, D. 1987. Reflective teaching: The preparation of students of teaching. Reston, VA: Association of Teacher Educators.

Dunkin, M. & Biddle, B. (1974). The study of teaching. New York: Holt, Rinehart and Winston.

Freiberg, H. & Waxman, H. 1990. Reflection and the acquisition of technical teaching skills. In Clift, R., Houston, W. & Pugach, M. (Eds.) 1990. Encouraging reflective practice in education 119-139. Teachers College Press.

Gage, N. 1985. Hard gains in the soft sciences: The case of pedagogy. Bloomington, IN: Center on Evaluation, Development, and Research, Phi Delta Kappa.

Gage, N. 1978. The scientific basis of the art of teaching. New York: Teachers College Press.

Grimmett, P., MacKinnon, A., Erickson, G. & Riecken, T. 1990. Reflective practice in teacher education. In Clift, R., Houston, W. & Pugach, M. (Eds.) Encouraging reflective practice in education, 20-38. New York: Teachers College Press.

Halls, W.(Ed.) (1990). Comparative education. Jessica Kingsley Publishers/ Unesco.

Harman, H. (1976). Modern Factor Analysis (3rd Edition). Chigago: The University of Chigago Press.

Harris, B. (1986). Developmental Teacher Evaluation. Allyn and Bacon.

Harris, B. & Hill, J. (1982). Developmental Teacher Evaluation Kitt (DeTEK). Austin, TX: National Educational Laboratory Publishers.

Holmes Group 1986. Tomorrow's Teachers. East Lansing, MI: Author.

Jöreskog, K. (1969). A general approach to confirmatory factor analysis. Psychometrika, 34, 183-202.

Jöreskog, K. & Sörbom, D. (1976). LISREL III – Estimation of linear structural equation systems by maximum likelihood methods: A FORTRAN IV program. Chigago: International Educational Services.

Kaiser, H. (1974). An index of factorial simplicity. Psychometrika , 39, 31-36.

Kaiser, H., Hunka, S., & Bianchini, J. 1971. Relating factors between studies based upon different individuals. Multivariate Behavioral Research, 6, 409-422.

Kansanen, P. 1991. Pedagogical thinking: the basic problem of teacher education. European Journal of Education, 26, 251-260.

Liston, D. & Zeichner, K. (1991). Teacher education and the social conditions of schooling. Routledge.

Moss, P. 1992. Shifting conceptions of validity in educational measurement: implications for performance assessment. Review of Educational Research, 62, 229-258.

Puurula, A. 1983. The dilemma of humanistic teacher education and CBTE. In Kansanen, P. (Ed.) Current research on Finnish teacher education. University of Helsinki. Department of Teacher Education. Research report 13.

Ross, D. & Kyle, D. 1987. Helping Preservice Teachers Learn to Use Teacher Effectiveness Research. Journal of Teacher Education, 38,.40-45.

Ryans, D. (1960). Characteristics of teachers. Washington D.C.: American Council on Education.

Tirri, K. (1991). Teachers' professional development: exploring instruments for teacher self-evaluation. Licenciate Thesis in Education. Department of Teacher Education, University of Helsinki.

Tirri, K. (1993). Evaluating teacher effectiveness by self-assesment: a cross-cultural study. Research Report 122. Department of Teacher Education. University of Helsinki..

Waxmann, H. & Walberg, H. (Eds.) (1991). Effective teaching: Current research. McCutchan Publishing Corporation.

Zeichner, K. 1990. Changing directions in the practicum: looking ahead to the 1990s. Journal of education for teaching, 16, 103-132.

Zeichner, K. (1992). Conceptions of Reflective Practice in Teaching and Teacher Education. To appear in G. Harvard & P. Hodkinson (Eds.) Action and reflection in teacher education. Norwood, NJ:Ablex.

Zeichner, K. 1983. Alternative paradigms of teacher education. Journal of Teacher Education, 33, 3-9.

Zeichner, K. & Tabachnick, R. 1991. Reflections on reflective teaching. In Tabachnich, R. & Zeichner, K. (Eds.)1991. Issues and Practices in Inquiry-Oriented Teacher Education. The Falmer Press.

	Indiana N (%)	Texas N (%)	Kajaani N (%)	Helsinki N (%)	Subject specific N (%)	Total N (%)
Men	5 (6.5)	4 (4.4)	19 (30.6)	14 (28.6)	8 (13.1)	50 (14.7)
Women	72 (93.5)	86 (95.5)	43 (69.4)	35 (71.4)	51 (83.6)	287 (84.6)
Age <25	5 (6.5)	4 (4.4)	1 (1.6)	0 (0)	5 (8.2)	15 (4.4)
Age 25-34	15 (19.5)	27 (30)	22 (35.5)	18 (36.7)	15 (24.6)	97 (28.6)
Age 35-45	32 (41.6)	31 (34.4)	25 (40.3)	18 (36.7)	17 (27.9)	123 (36.3)
Age 46-55	19 (24.7)	20 (22.2)	11 (17.7)	10 (20.4)	14 (23.0)	74 (21.8)
Age >55	6 (7.8)	8 (8.9)	3 (4.8)	3 (6.1)	10 (16.4)	30 (8.8)
Exp. 1-5	20 (26)	20 (22.2)	13 (21)	15 (30.6)	14 (23)	82 (24.2)
Exp. 6-10	11 (14.3)	19 (21.1)	12 (19.4)	10 (20.4)	11 (18)	63 (18.6)
Exp 11-20	29 (37.7)	27 (30)	22 (35.5)	16 (32.7)	15 (24.6)	109 (32.2)
Exp >20	16 (20.8)	21 (23.3)	15 (24.2)	8 (16.3)	21 (34.4)	81 (23.9)

Table 1. Summary for the statistics of the data used.

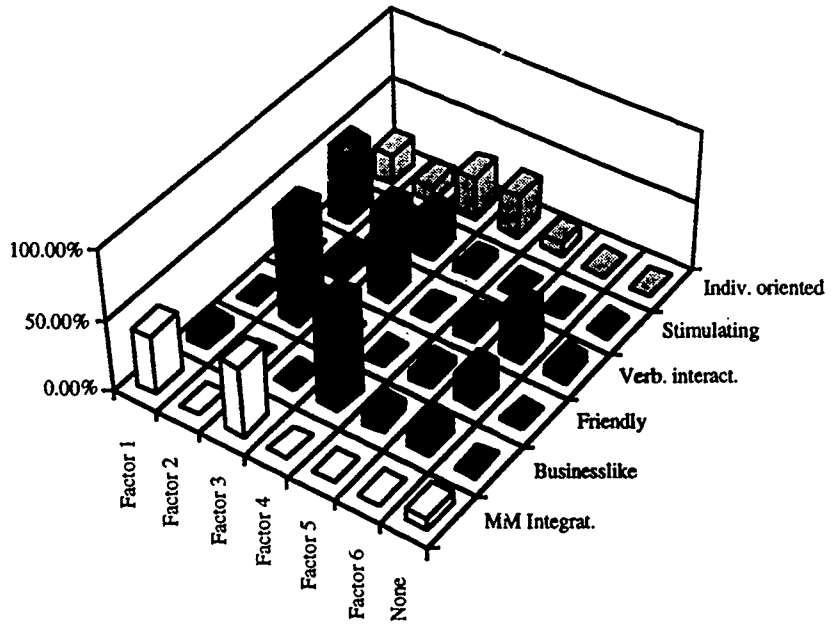

Figure 1. Distribution of the Harris model component variables into the forced 6 factor solution for the complete data.

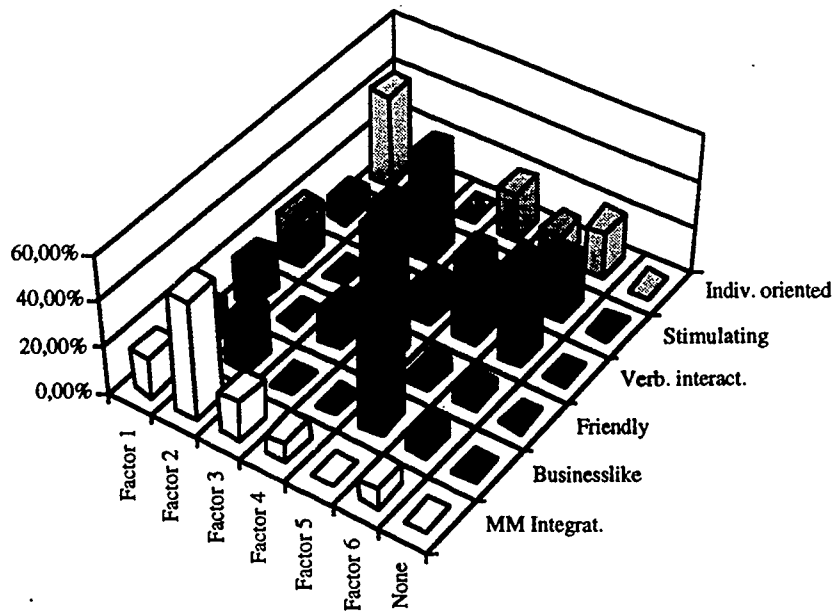

Figure 2. Distribution of the Harris model component variables into the forced 6 factor solution for the US data.

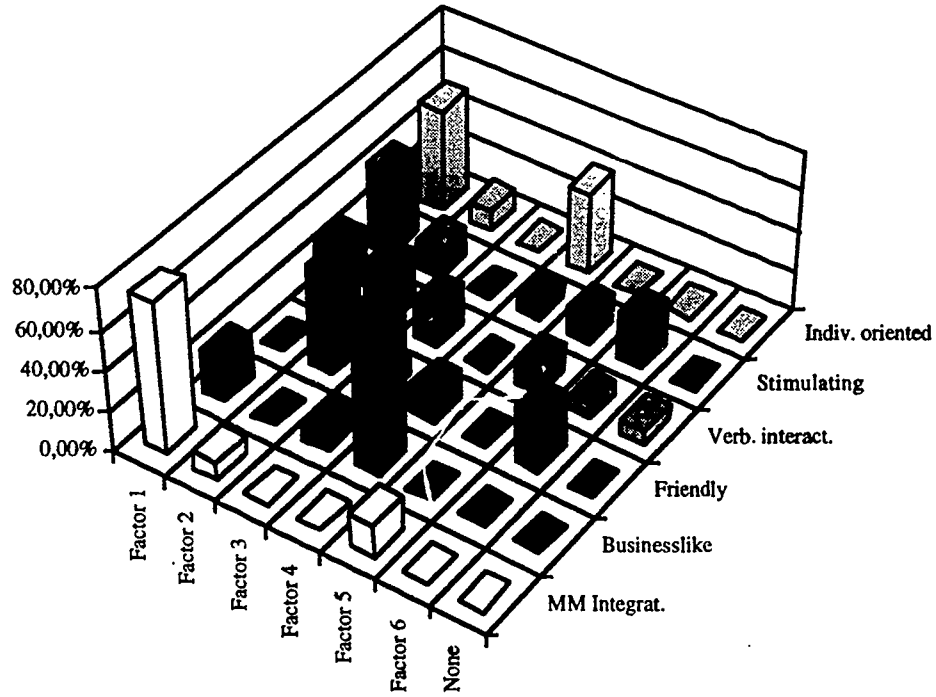

Figure 3. Distribution of the Harris model component variables into the forced 6 factor solution for Finnish data.

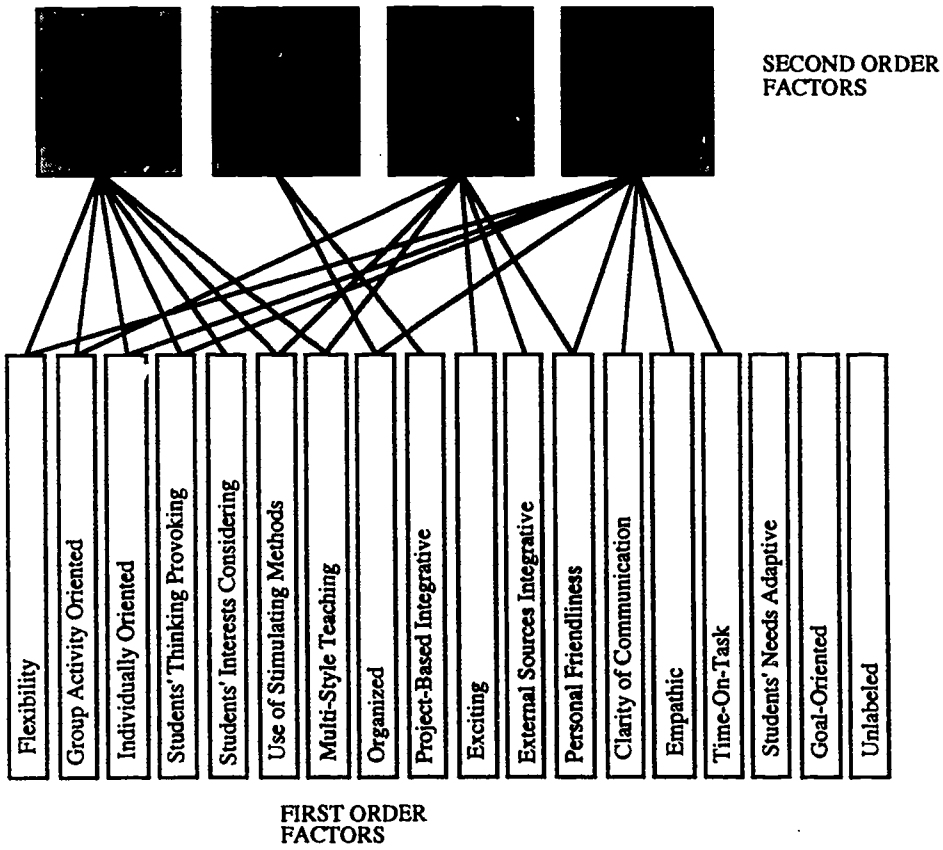

Figure 4. The second order factor structure for the US data. Grouping of the primary factors is denoted by lines.

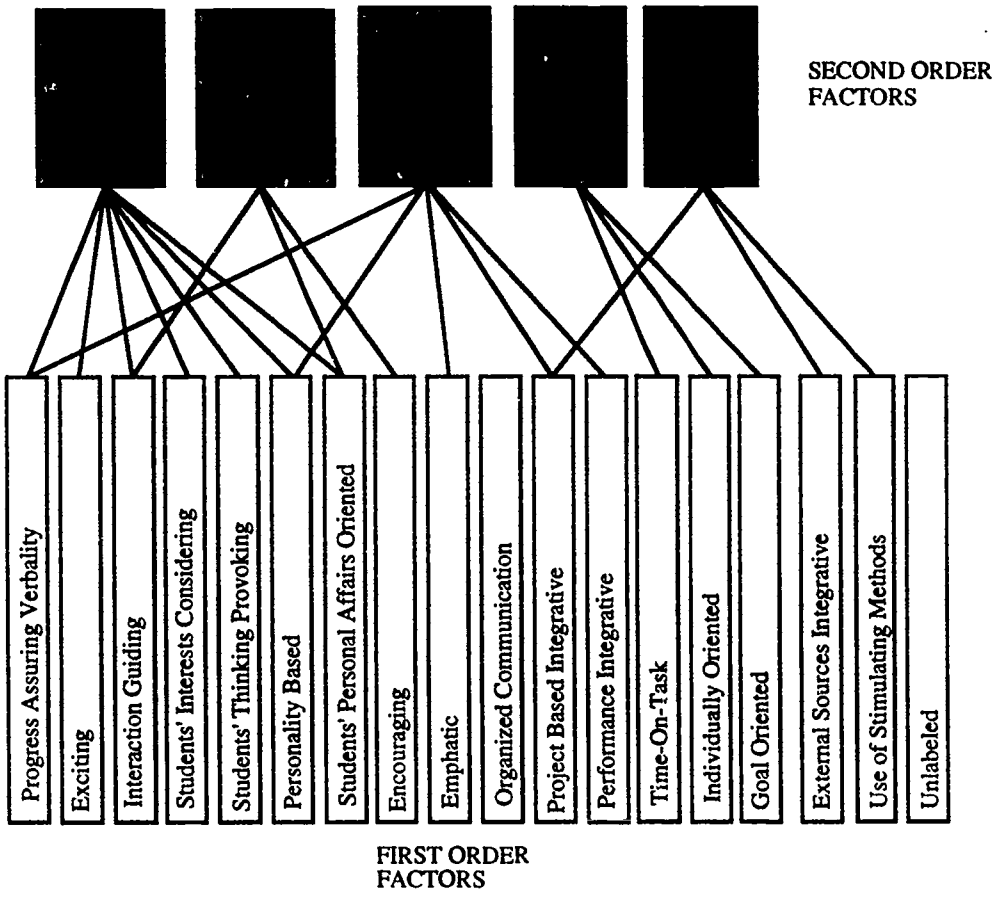

Figure 5. The second order factor structure for the Finnish data.

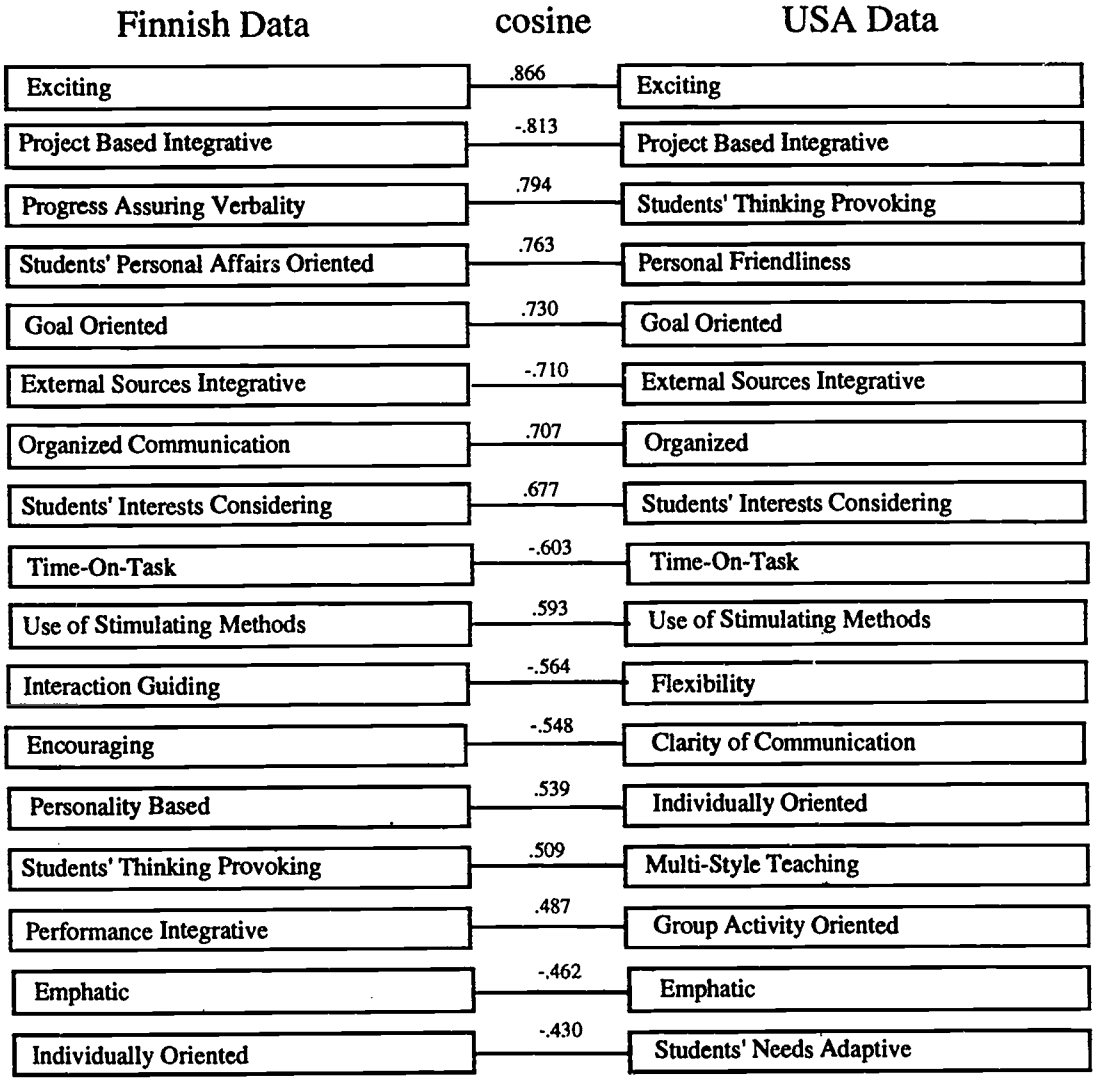

Figure 6. The pairing of related factors in the US and Finnish data. The values in the middle are the factor cosines for the pairs as calculated by the Kaiser, Hunka and Bianchini method.