This report summarizes the major activities and findings to date of the Career Ladders Research and Evaluation Project with respect to the five-year pilot-test currently underway in Arizona. Summary descriptions are given of each year's activities from 1985 to 1989: (1) development and planning for program evaluation; (2) application of research and evaluation instrumentation, methodology, reporting procedures and dissemination; (3) expansion and dissemination of data base information; (4) the emergence of a predictive model for effective career ladders program implementation; and (5) advancement in project focus, research design and methodology. Descriptive summaries of each year's progress are followed by bibliographic listings of related research reports, presentations and publications. Research and evaluation of the pilot-test program over a period of time has resulted in the following three major findings: (1) successful change and reform can be influenced by intervention programs such as career ladders, but it is primarily dependent upon well-developed and effective school systems; (2) the association between teacher performance and competency based on process and developmental evaluation can scientifically be related to reliable and valid student academic achievement measures; and (3) while the Career Ladders Pilot-Test Program does an excellent job of assuring teacher accountability, it has been far less successful with respect to assessing the corresponding accountability of other key factors of district operational effectiveness. (JD)
EDUCATIONAL REFORM IN ARIZONA:

1985 to 1990

A CHRONOLOGY OF RESEARCH & EVALUATION PROCEDURES
FOR ASSESSMENT OF THE PILOT-TEST CAREER LADDERS
TEACHER PERFORMANCE & INCENTIVE PROGRAMS:
1985 - 1990

by

Dr. Richard D. Packard, Manager
The Arizona Career Ladders Research & Evaluation Project

&

Dr. Mary I. Dereshiwsky, Research Associate
The Center for Excellence in Education / Research & Development Partnership

P. O. Box 5774, Northern Arizona University
Flagstaff, Arizona 86011

Presented to the Joint Legislative Committee on Career Ladders
State Capitol, House Wing

July 20, 1989
<table>
<thead>
<tr>
<th>Table of Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Overview</td>
<td>1</td>
</tr>
<tr>
<td>1985 - Development and Planning for Program Evaluation</td>
<td>3</td>
</tr>
<tr>
<td>Evaluation Design</td>
<td>3</td>
</tr>
<tr>
<td>Origination of Evaluative Components</td>
<td>3</td>
</tr>
<tr>
<td>Doctoral Dissertations and the Data Base</td>
<td>4</td>
</tr>
<tr>
<td>Documentation</td>
<td>4</td>
</tr>
<tr>
<td>1986 - Application of Research and Evaluation, Instrumentation, Methodology,</td>
<td>5</td>
</tr>
<tr>
<td>Reporting Procedures and Dissemination</td>
<td></td>
</tr>
<tr>
<td>Related Literature Research</td>
<td>5</td>
</tr>
<tr>
<td>Policy Research</td>
<td>5</td>
</tr>
<tr>
<td>The Perception Assessment Scale</td>
<td>5</td>
</tr>
<tr>
<td>Census Survey, Analysis and Reports of Findings</td>
<td>6</td>
</tr>
<tr>
<td>Documentation</td>
<td>6</td>
</tr>
<tr>
<td>1987 - Expansion and Dissemination of Data Base Information</td>
<td>7</td>
</tr>
<tr>
<td>Expansion of the Data Base</td>
<td>7</td>
</tr>
<tr>
<td>Information Dissemination and Reporting of Results</td>
<td>8</td>
</tr>
<tr>
<td>Documentation</td>
<td>8</td>
</tr>
<tr>
<td>1988 - The Emergence of a Predictive Model for Effective Career Ladders Program</td>
<td>9</td>
</tr>
<tr>
<td>Implementation</td>
<td></td>
</tr>
<tr>
<td>Directional Focus of the Research and Evaluation Project</td>
<td>9</td>
</tr>
<tr>
<td>The Career Ladders Pilot-Test Program Network</td>
<td>10</td>
</tr>
<tr>
<td>Individual School District Research and Dissemination</td>
<td>10</td>
</tr>
<tr>
<td>Instrumentation Reliability and Validity</td>
<td>10</td>
</tr>
<tr>
<td>Documentation</td>
<td>11</td>
</tr>
</tbody>
</table>
A CHRONOLOGY OF RESEARCH AND EVALUATION OF THE PILOT-TEST CAREER LADDERS TEACHER PERFORMANCE AND INCENTIVE PROGRAM
1985-1990

Introduction

Overview

In 1985 the Center for Excellence in Education at Northern Arizona University responded to the opportunity to evaluate the probability of program success of the career ladder pilot-test project initiated by the passage of S.B. 1336. In addition to yearly research reports on the current status of the program, the objectives of the Arizona Career Ladder Research and Evaluation Project have expanded to include making recommendations as to individual districts’ readiness to support major change and reform programs. The purpose of this report is to trace and summarize the activities of the research and evaluation process related to the five-year pilot-test project, from its legislated beginning through the present time, including forthcoming activities. Each summary description will be followed by a bibliographic listing of related research reports, presentations and publications.

Figure 1, page 2, presents a diagrammatic overview of the outputs of the research and evaluation project. Reading from left to right, this model depicts the process by which data collected by the center is systematically analyzed, summarized and formulated into specific policy recommendations to be made to the Joint Legislative Committee on Career Ladders (JLCCL) at the end of the five-year pilot-test. This process effectively serves as an overall framework for the more specific activities to be described in greater detail and in time order below.

Two major findings have emerged to date which have influenced the direction of research and evaluation activities. The first of these is a startling discrepancy in individual districts’ readiness to support change and reform movements such as career ladders, despite supposedly “uniform” legislative program requirements. As a result, evaluation has shifted from overall monitoring of compliance of these
I. DATA BASE DOCUMENTS INCLUDE A SERIES OF QUALITATIVE AND QUANTITATIVE PLANS, DISSERTATIONS, ANALYTIC STUDIES AND PUBLICATIONS

II. SUMMARY EVALUATION DOCUMENTS ARE DESCRIPTIONS AND LISTINGS OF EVALUATIVE JUDGMENTS BASED ON QUANTITATIVE AND QUALITATIVE RESEARCH METHODOLOGY

III. MODELS, DESIGNS AND STRUCTURES INCLUDES CONCEPTUAL MODELS WHICH DIAGRAMMATICALLY SUMMARIZE PROGRAM COMPONENTS AND PROCESSES

IV. POLICY DOCUMENTS INCLUDE ESSENTIAL REQUIREMENTS FOR LEGISLATION AND PROJECTED EFFECTS OF NON-LEGISLATION ON REFORM
guidelines, to include careful assessment of the current operating status of support and focus factors for change within districts on an individual basis.

The second conclusion apparent from the first five years of program research is an unrealistic expectation which is prevalent with respect to the change process in the social sciences. There has been a disturbing tendency for adopters of programs such as Career Ladders to simply “attach” them to existing organizational structures and expect both automatic integration and instantaneously successful outcomes. What these well-intentioned innovators fail to realize, however, is that new discoveries in the social and behavioral sciences are as evolutionary and developmental in nature as their counterparts in the medical, physical and natural sciences. Processes for change and improvement in these fields require a carefully planned start-up time, as well as continual monitoring, feedback and adaptation to changing conditions. Therefore, a careful understanding of the elements of the change process, as well as a long-run strategic plan for evaluating the resulting outcomes, are necessary for a complete and valid evaluation. Research efforts during the five-year pilot-test period have incorporated an identification and application of such a model to the change and reform process in education.

1985 - Development and Planning for Program Evaluation

Evaluation Design

Arizona legislation established the Center for Excellence in Education at Northern Arizona University as the designated evaluator of the "Career Ladders Pilot-Test Teacher Performance and Incentive Program." An evaluation design was formulated and approved, which consisted of an "improvement model" requiring districts to follow a yearly formative and cyclical planning, implementation and summative evaluation process. Along with evaluation, the Center was to provide monitoring of important legislated guidelines.

Origination of Related Evaluative Components

In 1984-85, the researcher culminated studies relating to basic human psychological needs and organizational functions which result in the greatest worker motivation and production. Several factors
were identified as being instrumental to successful change efforts. These included the quantity and quality of worker competencies, as well as their performance in meeting the objectives of public and private organizations. These research results were later incorporated into evaluative components of individual districts' career ladder plans. Their primary objective was to assess the districts' ability to provide adequate incentives for superior teacher performance. A second aspect of the evaluation process was districts' ability to motivate teachers to improve their skills and expand their responsibilities. This research later became a part of comparative studies which determine the relationship between a district's "organizational climate" or "psychological environment" and career ladders (CL) program success. A second objective was to evaluate trend data collected during the pilot-test period in order to determine the direction and extent of impact which the CL incentive plan has had on interpersonal communication within each organization.

**Doctoral Dissertations and the Data Base**

An important part of legislative policy intent was to involve university students more actively in various phases of the evaluation process. The first doctoral dissertation related to Career Ladders was conducted at Northern Arizona University. Recent technological advances with respect to the study variables, as well as the complex and broad-based nature of this evaluation process, have resulted in ongoing research efforts at all three state universities. The popular *effective schools* movement, for example, has been the focus of a number of current papers, theses, and dissertations. Specific areas being investigated include CL program goals of teacher improvement, administrative leadership, and enhanced accountability for student achievement.

**Documentation**


Related Literature Research

One doctoral study initiated continuing research and analysis of related literature. It involved an extensive review of such key areas as "differentiated staffing" and "incentive plans." In addition, it investigated the effects of other, more specific variables related to the many complex programmatic issues of concern. These included the following: teacher and administrator evaluation, inservice development, program input and ownership, change, reform, motivation, communication, leadership, organizational management, support systems, curriculum, student achievement, and testing.

An especially important objective of the literature review was identifying the reasons for the success or failure of previous educational reform movements. This focus was necessary, in order for the evaluators to identify those specific factors which would be most indicative of district "readiness" levels for successful program implementation.

Policy Research

The initial interest of outside agencies and research centers in the five-year pilot-test eventually led to an ongoing association with other social and political entities. Continued input from these agencies has extensively influenced research purpose and methodology of the "Arizona Project." Such active involvement has helped to determine the potential for legislative bodies to use "policy research" or "evaluation research" results in making legislative or policy decisions.

The Perception Assessment Scale

The primary career ladders survey instrumentation, entitled the Perception Assessment Scale, was designed in order to elicit information on attitudes and perceptions of organizational functions. This survey was developed by a research team composed of members from the three state universities and legislative
research analysts, and it was initially pilot-tested in two CL districts. The instrumentation included forced-choice and open-ended questions related to CL concepts and organizational climate, resulting in both quantitative and qualitative data. (Three years of data have accumulated from the scale which will allow for trend analyses.)

Census Survey, Analysis and Reports of Findings

The total population of "Phase I" school districts was surveyed, data were analyzed and initial findings were presented to the JLCCL. (Phase II and III districts were subsequently evaluated as they implemented their programs.) More specific and detailed data reports were provided to each school district for analysis and improvement of individual areas of weakness. The intent has been for each district to evaluate the findings and implement the initial improvement model of refinement and recycling, based on research evidence. As a result, change and reform would be based on input from teachers and other individuals who have a vested interest in improving themselves and the achievement of their students.

Documentation


1987 - Expansion and Dissemination of Data Base Information

Expansion of the Data Base

A dissertation related to performance-based teacher evaluation expanded the review of related literature and analyzed quantitative and qualitative data through basic descriptive, anecdotal and non-parametric statistical procedures. As a result of the relationship of success of reform movements and level of organizational operations, effective schools research was reviewed and more closely associated with CL research findings. In order to satisfy one of the most important concerns of legislative intent, plans were initiated which related to the scientific study of accounting for program effects on student achievement.
Information Dissemination and Reporting of Results

Considerable outside interest in the research results of the career ladders pilot-test program became evident during this time. This was a direct result of the focus on relevant issues such as teacher development and incentives, as well as their effect upon student achievement. Several state, national and international presentations related to topics such as the follow: g: a) teacher productivity, performance evaluation and rewards; b) program designs and structures; c) reform and change in education; and d) teacher productivity and instructional leadership roles.

A major research finding during this time concerned the extreme diversity of individual districts' readiness to implement a uniform set of legislative guidelines. This diversity in readiness levels existed despite the otherwise carefully matched similarities in CL program goals, designs and structures across participating districts. As a result, reporting of project results to the JLCCL became considerably more complex. Quantitative and qualitative data indicated that factors other than CL plans and goals were, in fact, instrumental to potential program success or failure.

Documentation


I. 1988 - The Emergence of a Predictive Model for Effective Career Ladders Program Implementation

Directional Focus of the Research and Evaluation Project

Research findings during the preceding year had revealed a dichotomy with respect to CL program implementation of districts which were otherwise similar in terms of goals and models for designs and structures. That is, some of these seemingly equivalent districts were successfully reaching their target
objectives, while others were failing to make observable progress. As a result, research efforts in 1989 were focused on identifying and assessing the primary reasons for these dichotomous results.

The extreme diversities in readiness found in the CL research, combined with private and public organizational and effective schools research, identified those operational factors within the districts which influenced performance and production. This led to the emergence and continual refinement of an integrated model for assessing readiness (see Figure 2, on page 19) of districts for successfully implementing the CL program.

The Career Ladders Pilot-Test Program Network

The Network is an association of the CL school districts which developed primarily as an important support for project research and evaluation. Extensive time and effort has been expended by this committee and its research sub-committee in assisting the researchers with evaluation content and collection of data.

Individual School District Research and Dissemination

Several career ladders teacher-leaders, administrators and doctoral students within districts initiated program research, writing and dissemination of findings in cooperation with the project evaluators at this time. A number of specific components of the model have been enhanced through in-depth study of topics such as the following: a) production and outcomes in student achievement; b) change theory; c) program designs and structures; d) development of district R & D; e) teacher development, leadership, mentoring, peer instructional coaching and clinical supervision; and f) cooperation with the development of a predictive achievement model. In addition, requests for information from non-career ladders districts throughout the state resulted in wider dissemination of CL findings to date.

Instrumentation Reliability and Validity

The Perception Assessment Scale survey items were evaluated with respect to response consistency through calculation of standard reliability indices. A factor analysis was also conducted within each subsection in order to identify clusters of questions which best defined each topic area.


A lack of knowledge and application of adequate procedures related to measurement science and technology became apparent as a result of the presumed legislative intent to connect teacher performance to student achievement. School districts across the country continue to make the assumption that when teachers are evaluated on their instructional processes, there is automatically a valid connection to student achievement. More advanced research designs and models using multivariate statistical analysis procedures are being applied within three dissertations at the three state universities. These quantitative methods are more realistic, since they enable researchers to assess the joint effects of variables such as CL participation, years of teaching experience, and grade level taught. In addition, multivariate methods allow for inclusion of multiple associated measures of student performance, such as a standardized test and a teacher-made test. Preliminary results indicate a potential breakthrough in support of initial legislative opinions that teachers
should be evaluated and rewarded based on levels of performance, rather than solely on years of experience and college credit hours.

**Individualized Assessment Through Case Study Methodology**

Data originally collected through the *Perception Assessment Scale* survey identified the extreme diversities of districts with respect to readiness levels of the support and focus factors of effective change processes (see Figure 2 on page 19 for a depiction of the model). Additional research and data-analysis procedures were necessary in order to provide a more comprehensive picture of the actual program implementation processes operating within these individual districts. Through *case study methodology* the researchers have piloted procedures which have provided valuable insights into the "operational functioning levels" of key support and focus components. *Focus-group interviews* conducted with all relevant sub-groups within a given system have provided an in-depth picture of participants' perceptions and feelings concerning their CL program experiences. *Matrix analysis procedures* have been used to summarize the open-ended responses to the *Perception Assessment Scale* survey, in order to identify the most commonly occurring areas of satisfaction and dissatisfaction with district- and school-level organizational climate. Such clustering and comparative analysis of qualitative (non-numeric) data represent the latest methodologic advances for understanding and reporting key participants' experiences in their own words. The result is a more valid, complete and rich data base, adding to existing knowledge in the social and behavioral sciences.

**Convergence and Profiling from Quantitative and Qualitative Data**

The foregoing variety of quantitative and qualitative data-analysis procedures, as applied to *policy* and *evaluation research* issues, represent the current state-of-the-art. The results of applying such multiple and diverse methods as *survey, case-study and model-validation techniques* are continually assessed for consistency in order to determine the extent of *methods convergence*. A high degree of methods convergence means that the application of a number of different analytic techniques and data collection procedures has resulted in similar answers to the research question under study. This effectively indicates cross-validation, allowing for a corresponding high level of confidence in the conclusions.
An especially promising avenue of current assessment research involves the development of a profile for each district, based on specific quantitative and qualitative data resulting from the evaluation process. (Figures 3 and 4, pages 20 and 21, contain examples of such profiles.) Individual components (support and focus factors) of the readiness model are graphically depicted in terms of their cumulative frequency of mention as being positive or negative. That is, the position of each factor relative to the baseline of zero indicates its direction and magnitude of impact upon district readiness for change. By examining the components of these profiles, one can identify at a glance those support and focus factors which are operating at satisfactory levels, and which are in need of remediation. The primary finding depicted in this profile is that those factors which should be supportive of teacher and student progress have instead turned out to have a negative impact.

A number of other research studies are currently underway; these are listed in the "research in progress" section. Topics being investigated include factors related to assessment of student achievement, measurement of teacher effectiveness, and identification of attitudes and opinions of educational policy leaders.

Documentation


Research in Progress

Planning Documentation into the 1990's


A Potential Breakthrough

Research and evaluation of the pilot-test program over a period of time has resulted in the following three major findings:

1. **Successful change and reform can be influenced by intervention programs such as career ladders, but it is primarily dependent upon well-developed and effective school systems.**

   Development of effective schools includes the following key variables: a) competent and supportive leadership; b) clear and positive communications and interpersonal relationships; c) reliable and valid evaluation of teachers and administrators; d) accurate application of advanced science and technology; e) curriculum and instruction which meets local student needs; and f) reliable and valid accountability for student achievement.

2. **The association between teacher performance and competency based on process and developmental evaluation can scientifically be related to reliable and valid student academic achievement measures.**

   Many organizational bureaucracies, including school systems, compensate personnel based solely on years of experience and formal course work. They are not predominantly...
rewarded based on competency and performance. Unfortunately, most districts are still operating on the assumption that mere observation of teachers' instructional processes which result in good teacher evaluations automatically imply a high rate of learning for their students. The preferred alternative, which has not been adopted to the extent warranted, is to examine the magnitude of student achievement gains. In most districts, more valid and reliable methods of evaluating teacher performance and student achievement are needed.

3. While the Career Ladders Pilot-Test Program does an excellent job of assuring teacher accountability, it has been far less successful with respect to assessing the corresponding accountability of other key factors of district operational effectiveness.

Teachers have effectively been held accountable in terms of time on task, expanded responsibilities focused on instructional improvement, good classroom teaching and communication skills and demonstrating student achievement gains. However, the program has not held districts accountable for the development of an adequate curriculum, for improvement in communications and governance operations (participatory management, shared responsibilities, etc.) or for the development of long-range plans for demonstrating student achievement gains in individual schools as well as across the total district. More effective indicators of district-wide progress would enhance the evaluative aspects of the program.

Review of Progress

From 1985 to the present, the career ladders pilot-test research and evaluation methodology has evolved from a general evaluation design, to identification and description of relevant concepts and processes and in-depth into a variety of data collection, analysis and reporting procedures. Since career ladders program success was found to be dependent primarily upon factors other than the legislative and district goals and district CL plans and structures, current research has centered on essential support and focus factors which must be operating effectively before a teacher incentive and development program can
progress. It has become apparent that an effective school organization and operation is necessary in order for all participating districts to be equally successful in carrying out the legislative intent of the CL policy.

Epilogue

Democracy and the Freedom to Pursue Individualities and Uniquenesses

Politically, in a democracy, individuals and groups, with their uniquenesses and diversities, are equally recognized in the pursuit of freedom and independence.

Socially and behaviorally, in a democracy, these diversities must be recognized and allowed the freedom to develop in the direction of their uniquenesses and at the rate of their own potentialities for growth.

Educationally, equal treatments or interventions placed on unequal (unique) individuals, socially and behaviorally, fail to recognize diversities and uniquenesses and the way in which human beings and organizations develop.

To make everyone developmentally equal would require genetic and environmental interventions and manipulation by powerful non-democratic forces; by doing so, everyone would have an externally mandated potential to become socially and behaviorally "equal."

Equal treatment in intervention and developmental time, applied to unequal and unique entities, does not democratically, socially or behaviorally recognize the uniqueness and diversity of individuals or groups within our free society.

There is no way to recover from being socially and emotionally unequal, and in a democracy it is definitely not necessary. All individuals must have the freedom, independence and support to develop their own potential.

Richard D. Packard, 1989
Figure 2

MAPPING THE DEVELOPMENTAL PROCESS FOR EFFECTIVE SCHOOL REFORM:

Designs and Structures

A Model of Interrelated Components of Program Support and Focus for Effecting Change and Reform in Education
DISTRICT READINESS PROFILE OF STRENGTHS AND WEAKNESSES

Figure 3

LOCAL FINANCE/FUNDING (SALARY SCHEDULE)

CHANGE AND IMPROVEMENT

NETWORK

GENERAL FINANCE/FUNDING

PROFESSIONAL INPUT (OWNERSHIP)

DISTRICT R & D (PROGRAM EVALUATION)

EVALUATION (TEACHER/ADMINISTRATOR)

LOCAL GOVERNING BOARD

MOTIVATION (INTRINSIC/EXTRINSIC)

ORGANIZATIONAL FACTORS (CLIMATE/COMMUNICATION)

Sufficiency of Readiness Baseline

Positive Readiness Factors

SUPPORT

RANK ORDER

WEIGHTED DIFFERENCES

20

24

25
DISTRICT READINESS PROFILE OF STRENGTHS AND WEAKNESSES

Figure 4

TEACHER SKILLS DEVELOPMENT/LEADERSHIP

STUDENT ACHIEVEMENT (PRODUCTION/OUTCOMES)

CURRICULUM/INSTRUCTION STUDENT ACHIEVEMENT MEASUREMENT

ADMINISTRATOR DEVELOPMENT/LEADERSHIP

Positive Readiness Factors

Sufficiency of Readiness Baseline

Negative Readiness Factors

WEIGHTED DIFFERENCES

FOCUS

RANK ORDER