

# CAL STATE NORTHRIDGE EVIDENCE REPORT

## TABLE OF CONTENTS

1. [Evidence and Change at CSUN](#) on page 2
2. [Quantitative Study—Pupil Learning Growth](#) on page 3  
[What We Have Learned So Far](#) on page 4  
[What We Plan To Do](#) on page 4
3. [Mixed-Methods, Longitudinal Study: Assessment of Candidate Knowledge, Skills, and Performance](#) on page 6  
[What We Have Learned So Far](#) on page 6  
[What We Plan To Do](#) on page 8
4. [Mixed-Methods, Longitudinal Study: Assessment of New Teacher Knowledge, Skills, and Performance](#) on page 12  
[What We Have Learned So Far](#) on page 12  
[What We Plan To Do](#) on page 12
5. [Program Improvement or Curriculum Redesign](#) on page 14  
[Contexts—funding, politics, culture, academic preparation](#) on page 14  
[The CSUN theory of change](#) on page 15  
[Course, Program, and Pedagogy Changes](#) on page 15
6. [Index](#) on page 17
7. [Glossary of Abbreviations](#) on page 18
8. [Credits](#) on page 19

# 1. Evidence and Change at CSUN:

CSUN views **evidence as the basis for all change** in our teacher preparation programs. To that end, we have constructed a **data warehouse** about our candidates at each stage of their development through induction and beyond. We have access to matching K-12 pupil and teacher data through collaboration with the LAUSD (Los Angeles Unified School District) and the CSU (California State University) Chancellor's office. We are conducting research studies in order to understand and ultimately improve our contribution to K-12 learning and achievement.

Thus, the **central research question** is: How does early and ongoing emphasis on clinical site experience<sup>1</sup>, subject matter, and pedagogical content knowledge affect teachers' instructional practices and impact on K-12 pupil learning (See the CSUN paper, [Schools as Clinics](#) and the Footnote below.) If we answer this question with convincing evidence, **we can legitimate—and make—change**.

**Our model of teacher change** acknowledges that policy, personal background, and social context influence teachers profoundly, over a lifetime. However, we make the most of the candidates' short exposure to a preparation program in two ways. First, we integrate subject matter, subject pedagogy, and professional knowledge in course work and field experiences. Thus, the novice's single-minded focus on teaching is enriched by strategic links to subject matter and professional skills. Second, the focus of the more experienced but not yet advanced teacher on day-to-day survival is expanded through induction and professional development. That way, this teacher has the chance to (re-) view subject and teaching skills from a perspective of experience that changes the ways that s/he sees and applies such content and pedagogy.

**Our model of institutional change** requires that evidence be valid, linked to CSUN's goals for [Effective Teaching](#); that methods and indices pertain to our domains of teacher effectiveness; that evidence be reliable, induced through consistent methods and interpretation; that results feed into cyclical assessment reports; that these reports go before faculty and institutional partners for program change; that budget requests be justified by review of these reports; that this process be supported by a stable data infrastructure and consistent training about methods of collection and analysis; and that sustainability also requires extensive partnerships with districts, transfer partners, and expert consultants. This process is cyclical, not linear.

**Only multivariate analyses** can triangulate the effects of programs on candidates, teachers, and K-12 pupils. Such analyses can disentangle the threads of the teacher preparation programs from the weave of policy, personal background, and social context. A quantitative, longitudinal study can expose the aggregate effects of programs and pathways; but a longitudinal study with a mix of methods is more likely to unravel the specific reasons for these effects.

There should be **consistent instrumentation, method, and samples across studies**. Separate analyses multiply these tools in a way that complicates findings, undermines sustainability, and taxes project personnel. Both of our **mixed-method**, longitudinal studies (Assessment of Candidate Knowledge, Skills and Performance; Assessment of New Teacher Knowledge, Skills and Performance) will use common instruments and common sample populations as much as possible.

---

<sup>1</sup> At our three clinical sites, cohorts of candidates take a sequence of courses together. They teach under the supervision of a K-12 coordinator and CSUN faculty. CSUN provides professional development on site and funds each partnership at 30K.

## 2. Quantitative Study—Pupil Learning Growth

**Evidence focus, rationale, team, and funding:** The **evidence group** has **focused** on a series of quantitative, longitudinal questions to theorize—and assess—the impact of teacher preparation programs and pathways, both undergraduate and post-baccalaureate, on K-12 learning.

The **rationale** has been to find evidence that either different programs or different pathways through programs have differential effects. The study is constructed to explore this **question**: do “program” and “pathway,” which are curricular concepts, have empirical validity in students’ experience.

**Funding** for the pilot and the CRESST (National Center for Research on Evaluation, Standards, and Student Testing) model have come from the TNE grant itself—75K in ’05-06 for the former and 125K for the latter in ’06-07. Substantial dollars in prior years have been devoted to a **data warehouse** and the training of IR (Institutional Research) and ITR (Information Technology Resources) staff to port local data on teacher candidates into a central database that then can be integrated with LAUSD teacher and student data. This capitalization of infrastructure—data bases and models—will reduce costs in the out years.

**Sources of Evidence:** The comprehensive data base from LAUSD and CSUN includes the following three categories of information: 1) student information including test scores, characteristics (e.g., free/reduced price lunch status, LEP (limited English proficiency), ethnicity, gender, etc); 2) teacher information including teacher preparation program, course grade, years of service, gender, age, etc; 3) school information including demographic information, school characteristics index, (% of students in free/reduced price lunch program, etc). Student information is linked to teacher and school information.

**Design:** These are the key questions behind the study; they specify, in sub-questions, the [central research question](#) that we emphasized at the top of this review.

- In what ways do CSUN teacher preparation programs impact K-12 pupil achievement, particularly in reading/writing, and math?
- Do pupils of teachers who graduated from different CSUN teacher preparation programs produce significantly different achievement patterns as measured by standardized NRT’s (norm-referenced tests) and district- developed tests? Are there pathways within these programs or weaving among them that have distinctive effects?
- Is there a relationship between teacher preparation factors (such as course grades and test scores) and pupil performance on standardized or district tests?
- Which set of variables have a greater impact on pupil test scores, those relating to teacher preparation programs, those relating to teacher in-service experience (number of years teaching, waivers for special education or emergency credentialing), or those relating to school characteristics (proportion of pupils in school lunch program, etc.)?
- Are there significant differences in K-12 pupil achievement patterns that are related to whether the teachers completed their credential program at CSUN, other CSU schools, or, non-CSU institutions?
- If there are differences, are these related primarily to school characteristics, pupil characteristics, or teacher in-service characteristics?

**What We Have Learned So Far:** A pilot study linking teacher preparation pathways of 200 teachers to K-12 learning and achievement indicated significant differences among the post-baccalaureate Traditional pathway, the ACT (Accelerated Collaborative Teacher) pathway (a cohort program offered at a school site), and the Multiple Subjects Internship pathway for teachers on emergency permits. These pathways were chosen because they were based upon different approaches to teacher preparation:

- Traditional--largely within a university setting,
- ACT/Cohort--largely within a supervised cohort at a school site, and
- Intern--largely within a dispersed, non-university setting.

**Method:** Multivariate analyses compared (grades 1 – 5) pupil learning from 1999 through 2002. The analyses controlled for the school's Academic Performance Indicator (Academic Yearly Progress on standardized tests combined with the proportion of pupils on the free-lunch program as an indicator of socioeconomic level). Learning was represented by scores in reading, math, language arts, and writing on standardized tests (Stanford 9, California Achievement Test), an augmented version of the standardized tests that included items aligned to state standards (Augmented Stanford or Augmented CAT), performance-based tests designed to align with state content standards (California Standards Test), and an English Language Assessment test.

**Findings:** These analyses yielded apparently significant differences between programs on six tests (SAT 9 Reading 2002, CST Reading 2002, CST ELA 2002, CST Language Performance Standard (English) 2002, SAT 9 Math 2000, and CST Augmented Math 2000). Students whose teachers graduated from one pathway consistently performed at a lower level than did students whose teacher graduated from either of the other two pathways. However, these results varied by grade level; for example, results were most consistent in grade levels 2 and 3. But that could have been a function of a larger number of teacher and pupil subjects at those grade levels. In addition, because the cell sizes significantly decrease as critical contextual variables are added (such as the pupil's English Language Level), the study needed an increased number of teacher and pupil subjects.

**What We Plan To Do:** Because of the need for a more complex model, we contracted with CRESST. CRESST determined that HLM/VAM (hierarchical linear/value-added) trials were appropriate because the data have both a hierarchical (HLM) and a cross-classified structure (VAM). Each student has a time-series of test scores. And students and teachers are nested within schools. However, students are very likely taught by a different teacher at each grade. For example, a student is taught by teacher A at grade 3, teacher B at grade 4, and teacher C at grade 5. In this situation, grade and teacher are crossed, not nested.

**Methods:** We delivered to CRESST an organized data matrix for elementary teachers. The **steps for assembling and integrating the fields** were:

- **Step 1:** Clean-up of the original Master File involved assigning a uniform format and width to fields that differ only by year.
- **Step 2:** Added information from the Employment Data File to the cleaned-up Master File. With the aid of the teacher IDs in both files, social security numbers for CSUN-trained teachers were added to the Modified Master File, along with employment data.
- **Step 3:** Created multiple year-specific files with the aid of the Modified Master File. In the process, similar variables or fields were grouped together.

- **Step 4:** Added CSUN ID data to each year-specific file. Using the Employment Data File and campus databases, relevant ID numbers were identified for the CSUN-trained teachers whose classes are included in the Master File.
- **Step 5:** Created a Combined File for each set of five year-specific files.

In, '06-07 CRESST will employ the **two different statistical models** for this study. (Data from '02-06 will augment the study.) The three-level hierarchical model will use K-12 students' scores. The effect of CSUN's teacher preparation program can be estimated within this framework. This approach is a less challenging statistical model than the value-added approach.

The other method uses a value-added model, which deploys the entire data base at a single time. The teacher effect is what remains of the year-to-year gain after removing the district mean gain and the factors that are idiosyncratic to students. Since students are taught by a different teacher each year, the data no longer have a nested structure. It is not possible to apply the three-level model in this way. The analysis plots a vector for each individual teacher effect. The vector registers the extent of the deflection to the trajectory of mean growth.

**Schedule:** The **evidence team** will receive and review a draft of the model by 10/06; the data runs will be completed and analyzed by 6/30/07.

**Expansion and Sustainability of Evidence-Gathering Activities:** Also, CSUN is partnering with the CSU Chancellor's office on a **large-scale** HLM/VAM analysis of the effects of CSU-trained teachers on K-12. We plan to share both the design of the data warehouse with other campuses in the CSU and to collaborate on modeling with the Chancellor's office. Ultimately, we intend to build a CSU infrastructure to **sustain** the gathering and modeling of program effects, while capitalizing on campus-based studies that are qualitative and quantitative but on a smaller scale. CSU survey reports—entered by teachers and supervisors—provide cross-checks on **reliability**. The system's TPAs and TPEs (teacher performance assessments and expectations) provide base lines for determining the alignment and **validity** of methods and instruments.

**Funding:** CSU has submitted a grant to Carnegie to support the project; CSUN-TNE's will use partnership funds--@120K--to support LA Basin campuses and the Chancellor's office to mount this system-wide effort. Additionally CSU has committed at least 150K plus in-kind in staff.

**Findings:** Preliminary work, based on a sample of 1,600 K-12 teachers and their elementary students, indicates that CSU-credentialed teachers account for 6-13% more learning, as measured by norm-ed and criterion-referenced tests, than those not so trained. Evidence so far suggests similar gains for SES-challenged students taught by CSU-trained teachers.

### 3. Mixed-Method, Longitudinal Study: Assessment of Candidate Knowledge, Skills, and Performance

**Evidence focus, rationale, team, and funding:** Ending the fourth year and entering the fifth, we are perched on the threshold of a significant change in our work. Initially, we undertook a number of activities—inputs—to affect candidate learning and affirm partnerships with schools. These were not evidence studies *per se*. But as TNE has matured into an endeavor to identify evidence that the inputs matter, we find ourselves with many projects on candidates, newly minted teachers, and veterans.

**What We Have Learned So Far:** Up to this time, projects have produced suggestive, not confirmative results, especially about pedagogical content knowledge. Still, we must remember that many projects were not conceived initially as designs for evidence, with controls, reliability, and attention to selection bias. Rather, they were activities to unite disparate partners in common endeavors—principally to align effort with the CSUN principles of [effective teaching](#). We have learned, too, that mixed-method observation, while hard to manage, provide multiple perspectives. We need many and different views to record the complex phenomena of learning and to avoid the reductionism that equates learning with tests, not with the visible but fleeting traces of minds as they encounter one another.

**Math:** The **central focus** has been whether knowledge about teaching a subject—the exemplary instance at CSUN is math—enriches a candidate’s depth of knowledge. The Deborah Ball MKT (mathematical knowledge for teaching) instrument was administered to students in multiple sections of arithmetic and geometry for elementary-level teachers.

Also, to **assess potential candidates’ readiness** to teach middle-school math, a Ball skills test was administered to cohorts of math majors, math-teaching candidates, and Liberal Studies students in a series of 200 and 300-level courses. The test was followed by an interest survey.

**Findings:** Results in the first study indicated a 0.3 SD (standard deviation) improvement in the arithmetic classes and a 0.6 SD improvement in the geometry classes in candidates’ understanding after interventions in teaching skills and tasks. The Ball instrument has been linked to—and validated against—K-12 learning as measured on the Terra Nova. Comparing ending content knowledge against initial ability in all sections, we concluded that the students in the treated sections did 5% better on average on the common final (the measure of content knowledge) compared to students of like ability in the untreated sections. We are now correlating content knowledge with MKT achievement to corroborate the claim that our intervention is more productive than the experience of the control population.

The second study yielded significant evidence of interest and ability—across pathways—to teach middle-school math and thus fill a desperate need.

**ITEP Science:** Sophomores ((Integrated Teacher Education Program) in the introductory Biology lab course and Liberal Studies seminar on integrating the sciences taught lessons in K-12, as part of their early field experience. The **governing question** was whether, with such concerted intervention, they would improve K-12 learning and improve their own mastery of content?

**Findings—Teacher Candidates:** Interviews revealed candidates’ exposure to teaching the Standards improved their disposition toward teaching science in an integrated fashion.—**Pupil Learning:** pre/post testing of 7<sup>th</sup> graders disclosed gains in learning. The work is suggestive. Further, professors who were actively involved in the science strand in TNE, prepared students who, in turn, taught K-12 students with the greatest gains.

**Writing:** The English Composition group **questioned** whether a series of self-reflective essays that required candidates to consider the utility of surface corrections, the role of collaborative process, and the application of scoring rubrics to K-12 writing—in middle and high school—would improve candidates’ own skills.

**Findings:** Scoring of the reflective essays, according to a holistic rubric, implied that such reflection improved organizational abilities and the understanding of the hierarchy of values in academic writing.

**Application of theoretical learning in the field:** Across the multiple-subjects program (LS), as well as in the mixed single-subject cohort at NAHS (Northridge Academy High School) and in the mixed-discipline cohorts at three [clinical sites](#), faculty and site supervisors assessed candidates in a family of studies. The **general motivation** was to see whether theoretical learning in the classroom rippled through field practice, where, in fact, it would deepen applied skills like pedagogical content knowledge and professional awareness.

- 50 supervisors’ field notes on 200 students were reviewed in Liberal Studies.
- A dozen reading tutors compiled field notes on their perceived effect on mentored students at the clinical sites.
- And 12 student teachers and teachers not trained at CSUN were interviewed at NAHS about their teacher training experience and understanding, on site, of how a school works.

**Findings: Student Teacher supervisors** recorded few instances of PCK and little attention to ELL (English language level) or to special needs. **SRI** (Stanford Research Institute) observers confirmed this; but they concluded that arranging candidates in cohorts who work together improves dispositions about teaching. At NAHS, teacher candidates valued the cohort and site experience, but they also seemed uncertain about professional roles in the school. Some teachers questioned whether indeed the CSUN cohort had adequate content and pedagogical knowledge.

**Use of All Current Findings:** The suggestion that there is a gap that separates theory about PCK, content mastery, and professional knowledge from execution in the field causes us to reflect and **change**. To address the methodological issue, **we have adopted** an approach that integrates instruments. This will facilitate training, make consistent the placement of items in domains, and align domains with CSUN TNE’s objectives for [effective teaching](#).

The gap between college classroom and practice the field appears in LAUSD’s qualitative studies of CSU-trained teachers. This issue reinforces the urgency of early, consistent, and possibly integrated experience with content, pedagogy, and professional knowledge. Already, instructors in math and English are **building activities** that include an understanding of how to teach subject matter into transition courses for candidates. Evidently, candidates’ self-reflection about how to explain subject matter to others engages professors; they view it as one means to encourage candidates to unpack the structure of knowledge in disciplines.

**What We Plan to Do:** To transform all these projects into evidence-focused efforts that assess candidates' content mastery, pedagogical knowledge, and professional knowledge of schools, as well as their dispositions toward teaching and learning, is not feasible. We need to **focus** on a limited set of projects, driven by our [central research question](#), to yield **evidence**. Parsimony increases the likelihood that the instruments are valid, their application reliable, the implementation feasible, the samples controlled, and the Ns sufficient and coordinated longitudinally. The instruments must be linked to the CSUN view of [effective teaching](#). The domains of the instruments must be both mutually consistent (albeit with different items) and mapped onto the CSU exit survey, the CSU survey of supervisors and teachers, the CSU/N TPEs, and the NCATE categories (National Council for Accreditation of Teacher Education). Our study, in sum, must be the **hub** that connects many spokes. Without such **rationale**, we will have dangling data.

So, we (**the executive committee and evidence team, in coordination with clinical site, induction, and arts and science teams**) commit to this **focus** on the following three studies:

- A multi-site **candidate learning study** that employs a family of instruments with complementary domains, scales, and protocols.
- The alignment of studies so that all questions must relate to the central, researchable question.
- A coherent approach to **longitudinal sequence** from candidates to new teachers to the more experienced.
- An emphasis on populations at **three types of sites**—where [clinical](#) dollars have been spent, where a historical relation has been nurtured with minimal investment, and where placements occur without significant CSUN intervention.

### **Study 1a: Focus on LS undergrad preparation on subject matter and PCK**

**Research Question:** How does intensive preparation (using coursework and fieldwork) of undergraduate LS (Liberal Studies) students in subject matter (math, science) and subject matter (math, science) pedagogy affect these students' practices and effectiveness in tutoring K-12 pupils as measured by pre/post tests or work-product of pupils? (**Funding: Evidence and Arts and Science Teams, @120k**)

**Design:** LS teacher candidates will receive a pre/post test of their math/science skills and knowledge of math/science methods prior to and after a lesson planning/observation/feedback cycle, modeled on the Ball protocol. After the intervention, potential candidates will tutor K-5 pupils. The pupils receiving the tutoring will be pre/post tested on math/science. Trained observers within each subject-matter will observe tutors and, using the same rubric across subject matter with some items being specific to the subject matter, rate the LS candidates on their practices.

**Study 1b:** Design is the same as above but instead of pre/post test of K-12 pupils, we will use pupils' portfolios to show whether improvement in writing/literacy skills, according to a rubric, can be associated with the knowledge and skills of candidates who have taken English 406, Writing for Teachers. This continues the [writing](#) study. (**Evidence, Culturally Relevant Teaching, and Arts and Science Teams. Funding: see above.**)

## Study 2: Focus on CSUN teacher preparation programs and pathways

**Research Question:** Are there differences in emphases and training approaches among CSUN teacher preparation pathways and programs with regard to [clinical site](#) experiences, PCK and professional knowledge? If so, how do these differences emerge in supervision seminars, methods courses and/or student teaching? (Also, see [Footnote](#) 1 on page 2 on clinical sites.)

**Design:** Document and describe differences in emphases and training for clinical site experiences and about PCK in math and literacy in Elementary and Secondary Education programs. Interview candidates at the beginning of their methods courses in math, reading, and/or language arts about their subject matter preparation and PCK. Use interviews and/or journaling, as well as ratings by university supervisor and cooperating teachers to track candidates' evolution of thinking and development in Subject matter and PCK during methods courses and student teaching. Use Clare's instrument for analyzing student work product to examine student teachers' impact on pupil learning in math or literacy. In Study 3, follow them through induction. (See Study 3, too: **Induction and NAHS, @120K—Funding.**)

## Study 3: Focus on CSUN teachers in two induction programs

**Research Questions:** (a) How do first and second year CSUN teacher graduates demonstrate their use of subject matter knowledge and PCK in math, reading, language arts? What is the impact on K-12 pupil learning in those subject areas? (b) How do the induction program and district policies affect teachers' implementation of PCK in their classrooms? Do they have the opportunity to implement PCK as learned in their preparation programs?

Identify first and second year CSUN teacher graduates and, using observations and interviews, describe their use of PCK in math and/or literacy, and professional knowledge. Measure impact on pupil learning through pupil work product using Lindsay Clare instrument across all studies PLUS whatever subject-specific existing measures exist (e.g., Ball) that can be used as a pre/post test and district CST tests.

**Design:** Identify first year CSUN teachers in the CSUN/LAUSD and the LAUSD only induction programs. Observe their classroom instructional practices as related to math, reading, and/or language arts. Interview them regarding the factors affecting this implementation. Examine impact on pupil learning using Lindsay Clare's model of analyzing pupil work product as well as district CST tests. Answer these questions: What are the salient characteristics of [clinical](#) TNE CSUN sites, other long-term CSUN clinical sites, and control sites? What is the impact of the approaches used at these sites on teacher candidates' PCK (pedagogical content knowledge) and professional knowledge? What is the impact on pupil learning?

**Expansion and Sustainability of Evidence-Gathering Activities:** Essentially, these itemized queries reduce to our **focus:** How does early and ongoing emphasis on clinical site experience, preparation in subject matter knowledge, and preparation in pedagogical content knowledge affect teachers' instructional practices and impact on K-12 pupil learning?

Such focused study will promote **validity and reliability** over time. To study effects on teachers and teachers' effects, we must apply similar instruments and indices, as well as treatments, in earlier phases of their development. And we must remain aligned with our domains of [effective](#)

[teaching](#), although over time teachers will fulfill these domains differently. We will satisfy validity by clustering the itemized queries with the appropriate domains in CSUN’s statement of effective teaching: subject matter (SM), subject matter pedagogy (PCK), dispositions (D), and professional knowledge (PK), including course/class management, and practical skills such as lesson planning. We will map these domains—against these CSU/N’s ongoing TPEs:

- Subject-Specific Pedagogical Skills for Single Subject Teaching Assignments
- Monitoring Student Learning During Instruction
- Interpretation and Use of Assessments
- Making Content Accessible
- Student Engagement
- Developmentally Appropriate Teaching Practices in Grades 4-8/9-12
- Teaching English Learners Learning about Students
- Instructional Planning
- Instructional Time
- Social Environment
- Professional, Legal, and Ethical Obligations
- Professional Growth

Domains must be carried through across the design of different kinds of instruments—observations, interviews, self-reflections, and rubrics for portfolio assessment. The same scale—or a map among scales—must be developed. And domains must reflect a similar hierarchy of knowledge/skills level when assessing candidates’ achievements. The chart below indicates which studies can capitalize on common instruments, how these instruments can be mapped against TPEs and CSUN TNE domains, which CSUN TNE domains are pertinent to the treatment, and what kind of items will populate the domains.

SUB-STUDIES	FORM	DOMAINS	SPECIFICS OF DOMAINS
LS STUDIES	TPE/TNE MAPPED	SM, PCK, DISP, PK	ALL—SM, PCK, DISP, PK
LS PORTF	WRITING RUBRIC	SM	SM—WRITING
WR PORT SAMPLE	WR RUBRIC	SM	SM—WRITING
WR PRE-POST	PRE-POST DESIGN	SM	SM—WRITING
WR OBSERVE	TPE/TNE MAPPED	SM, PCK, DISP	SM, PCK, DISP: WRITING
MATH OBSERVE	TPE/TNE MAPPED	SM, PCK, DISP	SM, PCK, DISP: MATH
MATH PRE/POST	PRE-POST-DESIGN	SM	SM: MATH
S-S PRE-POST	PRE-POST-DESIGN	SM	SM: S-S
S-S OBSERVE	TPE/TNE MAPPED	SM, PCK, DISP	SM, PCK, DISP: S-S
MATH S-S	TPE/TNE MAPPED	SM, PCK, DISP	SM, PCK, DISP
1-2 YR OB	TPE/TNE MAPPED	SM, PCK, DISP, PK	SM, PCK, DISP, PK
2> OB	TPE/TNE MAPPED	SM, PCK, DISP, PK	SM, PCK, DISP, PK
TC INTERVIEW	TPE/TNE MAPPED W/ CSU EXIT	DISP, PK	COLLAB, SUPPORT,
1-2 INT	TPE/TNE MAPPED W/ CSU EXIT	DISP, PK	COLLAB, SUPPORT, VALUE OF SM PREP AND TC EXPERIENCE
2>INT	TPE/TNE MAPPED W/ CSU EXIT	DISP, PK	COLLAB, SUPPORT, VALUE OF SM PREP AND TC EXPERIENCE
SUPERV INT	TPE/TNE MAPPED W/ CSU EXIT	SM, PCK, DISP, PK	COLLAB, SUPPORT, VALUE OF SM AND TC EXP
SITE EVAL	TPE/TNE MAPPED W/ CSU EXIT	SM, PCK, DISP, PK	COLLAB, SUPPORT, VALUE OF SM AND TC EXP

To the extent possible, we will carry out the field components of these studies across three sorts of sites—clinical sites, sites with historical/intellectual/program ties to CSUN, and more loosely

affiliated controls. That way, as Studies 2 and 3 imply, we can assess and then determine the role and definition of “sites.”

Site Type	School	Grade Level	Hi/Lo CSUN Presence	Programs
<b>A: Clinical \$\$ Investment</b>	Langdon	E	Hi	TNE
	Sepulveda	M	Hi	TNE
	Monroe	H	Lo	Trad
	NAHS	H	Hi	Trad
	CHIME	E	Hi	ITEP; others
<b>B: Historical Ties</b>	Canterbury	E	Hi	ITEP
	Fernangeles	E	Hi	ITEP
	Poly Family	E	Hi	ACT
		M	Hi	ACT
		H	Hi	ACT
<b>C: Control</b>	3- schools TBA	E	Lo	ACT/ITEP/TRAD/ OTHER
	3 schools TBA	M	Lo	“
	3 schools TBA	H	Lo	“

**Linkage to Pupil Learning:** Math has the advantage of using Ball instruments that have been linked to student gains in scores. However, the other disciplines only can adapt this link. Each sub-study also requires either linkage with a school-based test that, in turn, can be nested in mean performance for grade and school type. Or, in instances in which the candidates’ immersion in the K-12 class is brief, administration of a pre/post test, or use of a control, will be more useful. Also, evaluations of K-12 students’ prior knowledge provide an important context. In particular, social sciences, English, science, and Liberal Studies will adapt this latter approach, since their candidates’ immersion in K-12 is brief and episodic.

**Schedule:** The evidence and executive groups will conduct sessions to review instruments and construct training protocols in 8/06. Treatments will occur from 10/06 through 4/07. Field observations follow from 4/07 to 6/07. Results will be compiled and analyzed from 6/06-7/07.

## 4. Mixed-Methods, Longitudinal Study: Assessment of New Teacher Knowledge, Skills, and Performance

**Evidence focus, rationale, team, and funding:** As we said at the beginning, we have conceived 4 as an extension of 3. Together they **focus**, as a **longitudinal study**, on the ripple effects of CSUN programs and pathways on candidates, new teachers, and eventually experienced ones. This part of the longitudinal study will be coordinated by the NAHS and Induction teams, in consultation with the Arts and Science group and the evidence leadership in TNE.

**What We Have Learned So Far:** As we conclude AY, '06, several activities have focused on this **question**: To what degree do course work and field experience—especially in math, science, English, and social science—prepare early-stage teachers to do their jobs? Exemplary studies include:

**Math:** Single-subject faculty in math conducted observations and interviews with 50 1<sup>st</sup> and 2<sup>nd</sup> year CSUN-trained teachers to assess the continuity of their learning with their teaching.

**Social Science:** Social science faculty have struggled to develop a common understanding of pedagogy across their fields, both in K-12 and the University. Nevertheless, they roughed out an instrument with which they assessed ten CSUN-trained teachers and three board-certified ones. The teachers were interviewed with an open-ended form that explored their understanding of social science instruction in K-12. Once the data are analyzed this summer, the team can determine whether the candidates' views diverge from the intent of the program. If so, then the team can track back to reform the program.

**NAHS:** The team at Northridge Academy High School has conducted focus groups with 1<sup>st</sup> and 2<sup>nd</sup> year CSUN-trained teachers, to evaluate the adequacy of the program for the ensuing job. With the exception of the social science study, there have not been controls. And the studies have not yet dealt with inter-evaluator reliability through extensive training.

**Findings:** The math single-subject study indicates that graduates used PCK activities 5% of the time, asked high-level questions 25% of the time, and occasionally linked to prior experience to contextualize learning. The other studies have not completed the analysis of the data. Nevertheless, given what the studies of candidate learning has shown, the findings likely will encourage the exploration of the reasons for the apparent gulf between academic course work and experience, on the one hand, and practical wherewithal, on the other.

### What We Plan to Do—Expansion and Sustainability of Evidence-Gathering

**Activities:** Refer back to **Studies 2 and 3** for the complete context of the sites and sequence of the mixed-method, longitudinal study. In short, both induction and NAHS—interventions that can be evaluated against controls--provide a design for assessing three central questions, which merely parse the [central research question](#):

- Do the CSUN programs prepare candidates, now new teachers, adequately?
- What value-added does the Induction program add in these areas?

- Do K-12 teachers at a site linked closely to CSUN, as at NAHS, exhibit benefits in execution, disposition, and associated K-12 learning gains?

At **NAHS** the teams will evaluate whether CSUN-trained teachers perform better, with stronger student gains, than others. They will observe and interview eight CSUN-trained teachers, adapting questions from the CSU exit survey and supervisor/new teacher survey. This derivation allows for a comparison with mean score on both system-wide instruments. The team will explore whether selection bias can be compensated by gleaning like subjects from the larger survey pools. To the extent that questions and domains from these instruments are portable, the study can gain a control by focusing, too, on new teachers in the K-12 school control group. (See the previous [chart](#), as well as [Study 2](#).)

The **Induction group will focus** on three CSUN-trained new teachers in the CSUN induction program, three teachers without a CSUN credential but in the program, and three non-CSUN trained teachers now in the LAUSD stand-alone Induction. Adapting the CSU supervisor/new teacher survey domains on quality of preparation and support, the team will observe and interview the new teachers for three one-hour sessions. Again, the instruments will integrate, to the extent possible, itemized queries from other CSUN TNE teams to concentrate training and minimize reliability issues. Teachers will be matched demographically to reduce skewing. (See [Study 3](#).)

**Pupil Learning:** The quantitative longitudinal study, which will not be completed early enough to assist the qualitative study in '06-07, nests one-year gains in scores within tiers of environmental influences. These influences can deflect, from the mean, the vector that leads from teacher to K-12 learner. The data are on a large scale for programs and pathways—aggregated in mean values and in divergent tails. It takes unusual confidence to assign +/- to one teacher, based on data that are more meaningful in aggregates, although some school systems are doing so.

Because of the short time and the nature of the data, in '06-07 we will sort student gain scores by hand, not by a massive run, for the candidates and new teachers in this study. We will “explain,” in part, new teacher effectiveness by linking it with class scores on exams that have grade, school, and district values so that we can nest like within like. We will abbreviate the HLM, so to speak.

**Schedule:** As said before, the evidence and executive groups will conduct sessions to review instruments and construct training protocols in 8/06. Treatments will occur from 10/06 through 4/07. Field observations follow from 4/07 to 6/07. Results will be compiled and analyzed from 6/06-7/07.

## 5. Program Improvement or Curriculum Redesign

**Contexts—funding, politics, culture, academic preparation:** At first blush, program change and curriculum redesign seem to be easy. You find stuff out. You use that stuff to change other stuff. In turn, that stuff brightens the stuff inside pupil’s heads.

The **design** of the clinical site evaluation shows the logic—and the perplexity—of change. Candidates are paired with K-12 students who score below grade-level in reading. The candidates’ effects are measured by a rubric used to decode a self-reflection that they write. Their effects can be aggregated and hence compared with controls inside and outside the building. Candidates also complete a survey before they leave the site program, with items drawn from the CSU supervisor/new teacher questionnaire. The data from that effort provide a nuanced context. Site supervisors also assess their partnership with CSUN against NCATE standards. And of course, the candidates are observed by faculty and teachers who will use a form that shares domains with a family of other qualitative instrument so that alignment and comparability are possible. The evaluation is multivariate, with many opportunities to register the idiosyncrasies of the building and policies. Eventually, the HLM/VAM study will enable comparisons across like units. But next year, we will assemble a make-shift HLM by associating K-12 students’ test scores with candidates and new teachers’ interventions.

But what if the design of the study reveals that the sites add value significantly? Or what if the study discloses that LAUSD does not want outside parties in the schools? In either case, the finding would not lead to a remedy that program or curricular change can put in place. Resources and policy must change to allow the expansion of the sites; and change in policy suddenly can doom the sites. Politics and funding, in other words, are truly X factors.

So, program change must be put in context. Here are other nested contexts. VAM analysts suggest that a teacher accounts individually for no more than 10-15% of a pupil’s learning, although the effect compounds over time. Of that, it would seem improbable that the preparation program outweighs personality and previous experience in shaping the teacher. So, perhaps the program effect accounts for 33% of the teacher effect—or 3% to 6% of pupil learning. In turn, that program effect can be parsed by the proportion of time on task (CSUN’s TNE, in theory, devotes more time to applied components than traditional programs do):

	GE SKILLS	SUBJECT CONTENT	PCK AND FIELD	DISPOSITION, SKILLS, PROF. KNOWLEDGE
CSUN	25%	50%	10%	15%
TNE	25%	40%	15%	20%

Keeping in mind that the CSU transfer function limits the extent to which GE can be leveraged, reformers can change about 75% of the program effect on the teacher—and perhaps 2.5% to 5% of the indirect effect on pupils. But that assumes, contrary to academic practice, that the reformers control the curriculum that candidates take.

Further complicating change is the assumption that curriculum and program influence candidates more than professors’ practice. So, what is the effect if a professor **only lectures** about PCK, collaborative work, and learn-by-doing?

**The CSUN theory of change:** Researchers in education have developed and tested a variety of theories about change in teachers. In one of the earliest models, Fuller and Bown (1975) proposed that pre-service teachers were focused on their pedagogical skills, teachers in their first two years were focused on day-to-day survival, and only after that do teachers truly focus on pupil learning. Recent theories still start by looking at teacher development and rarely examine whether different programs of teacher preparation affect this evolution. None of these theories and little research link the effect of programs, through the evolution of teachers, to the impact on pupil learning.

At CSUN, we have developed two documents that begin to bridge this gap:

- a brief paper that sketches the morphology—over a career—of [effective K-12 teaching](#);
- and a conceptual model, in that paper, that shows the variables that affect that evolution.

Our model of teacher-change compensates for the limited time that candidates spend in a teacher preparation program. We integrate subject matter, relevant pedagogy, and professional knowledge in course work and field experiences to enrich novices' tendency to focus on teaching skills. We expand the focus of the first-to-second year teachers on day-to-day survival through induction and professional development. They can (re-)view subject and teaching skills from a perspective—s/he now is wholly on site—in which meaning acquires use value.

The conceptual model and the paper on [effective teaching](#) recommend multivariate measures of a teacher's effectiveness, as the teacher changes over time and as contexts change, too. We privilege K-12 learning as special evidence of such effectiveness. Of course, to the extent that we rely on inductive proof, our approach suffers the limitations of inductive reasoning—approaching probability but never achieving certainty. And to the extent that we equate learning only with results on tests, we strip from it the mental processes and human actions that, however fleeting and difficult to observe, that link knowledge to behavior.

We have constructed a sequence to implement the changes in programs and our behaviors that will effect changes in teacher candidates and teachers:

1. **DIRECT EVIDENCE:** The campus now requires direct evidence of learning to justify program change; and it understands the value of multivariate analyses.
  - a. CSUN's planning process requires that assessment precedes budget requests and explains curriculum change, beginning in '06-07.
2. **EVIDENCE REVIEW:** The College of Education regularly reviews evidence reports that are mentioned in this summary.
  - a. Liberal Studies and Math follow suit, leading the science and social science departments to review the effects of the interventions.
3. **INFRASTRUCTURE:** Large pieces of infrastructure, like a data warehouse and a highly skilled IR office are in place.
4. **SCALE AND SUSTAINABILITY:** The CSU Chancellor's Office and Education deans have accepted the logic of VAM/HLM and have paid into a partnership to retrieve data and analyze them.

**Course, program, and pedagogy changes:** Because of the relative youth of the project, most change has been in courses. There are more faculty who see TNE evidence as justifying change in courses than there are like-minded faculty in positions that route changes in programs.

Nonetheless, we have seen a few significant changes to programs—sequences of courses and/or services and relationships:

1. Attention to PCK has changed the content in sections of algebra, geometry, and math for educators. Preliminary evidence shows that an emphasis on MKT/PCK enriches content knowledge.
2. PCK has affected the delivery of the writing course, English 406, which teachers pass through. We have convincing evidence that self-reflection on writing as a teachable process improves candidates' writing and comprehension of priorities when teaching it.
3. And interest in PCK has motivated social scientists, who have surveyed candidates and teachers, to consider how best to prepare candidates to teach according to the new integrated Standard in California.
4. In Liberal Studies, science, and math professors, adjusting curriculum to the California Standards, working with candidates on subject-specific pedagogy, and observing the effect of these changes on field work, have developed new, blended courses.
5. Also, in Liberal Studies, surveys of candidates and teachers about their pathways through the program and their most valuable courses have allowed us to re-conceive advising and course scheduling.
6. Liberal Studies developed two seminars on teaching science based on our evidence that increased exposure improved their disposition toward the subjects.
7. Liberal Studies responded to deficits in candidates and teachers' ability to connect reading to writing problem in K-12 by developing a concentration in literacy.
8. We began three clinical sites where candidates take courses together and teach on site. This program integrates theory with practice and hinges on K-12/University partnerships, such as team-teaching.
9. We hired K-12 teachers-in-residence to instruct arts and sciences courses for future teachers. Thus, the candidates now get early access to a school professional while CSUN's host departments consult with the teachers about how to close the gap between University programs and K-12 school needs..
10. We cooperated with LAUSD on opening and developing NAHS. Functioning much like a clinical site—without cohorts—the site provides CSUN faculty with the equivalent of a teaching lab for observing candidates as tutors and beginning teachers and for field-testing curriculum and professional development..
11. In response to the documented gap between what we teach and what new teachers do, we have opened a joint Induction program with LAUSD to nurture beginning teachers.
12. We have created a system for observing candidates according to standard rubrics. Once entered into the data warehouse and integrated with background and performance data, this information has become a rich inventory for assessment and change. The TNE emphasis on evidence drove this change.

Also, we should not overlook the fact that, increasingly, evidence, not opinion, not fashion, drives change. The window is narrow, and usually the prevailing breeze blows against change. But the biggest change—in an institution's habit of mind—has occurred already

## 6. Index

activities, 6, 7, 12, 15  
arts and science teams, 8  
candidates, 2, 3, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16  
central research question, 2, 3, 8, 12  
Chancellor's office, 2, 5, 16  
change, 1, 2, 7, 14, 15, 16  
clinical site, 8, 14  
cohort, 4  
conceptual model, 15  
control, 11, 13, 14  
CRESST, 3, 4, 5  
CSU, 2, 3, 5, 7, 8, 10, 13, 14, 16  
data base, 3, 5  
data warehouse, 2, 3, 5, 15  
day-to-day survival, 2, 15  
design, 3, 8, 9  
domains, 2, 7, 8, 10, 13, 14  
effective teaching, 2, 6, 7, 8, 10, 15  
evidence team, 8  
executive committee, 8  
exit survey, 8, 10, 13  
findings, 4, 5, 6, 7, 12  
funding, 3, 5, 9, 14  
gap, 7, 15  
HLM, 4, 5, 13, 14, 16  
hypothesis, 6, 7  
induction, 2, 8, 12, 13, 15  
Intern, 4  
interviews, 7, 9, 10, 11, 12  
ITEP, 6, 11  
LAUSD, 3, 7, 13, 14  
LAUSD's qualitative studies, 7  
liberal studies, 6, 7, 11, 15, 16  
math, 4, 6, 11, 12, 15, 16  
mix of methods, 2  
mixed-method, 2, 6, 12  
model of institutional change, 2  
model of teacher change, 2, 15  
multi-site candidate learning study, 8  
multivariate analyses, 2, 4, 15  
NAHS, 7, 11, 12, 13  
NCATE, 8, 14  
nested, 4, 5, 11, 14  
observations, 9, 10, 11, 12, 13  
pathways, 2, 3, 4, 6, 9, 12, 13  
PCK, 7, 8, 9, 10, 12, 14, 15, 16  
pedagogical skills, 15  
pilot study, 4  
policy, 2, 14  
pre/post, 7, 8, 9, 11  
preparation, 1, 2, 3, 4, 5, 8, 9, 13, 14, 15  
preparation program, 2, 3, 5, 14, 15  
previous experience, 14  
pupil learning, 1, 3, 4, 7, 11, 13, 15  
qualitative study, 2, 13  
quantitative longitudinal study, 13  
rationale, 3, 6, 8, 12  
reading, 3, 4, 7, 14  
reliability, 5, 9, 12, 13  
research question, 8, 9  
science, 6, 12  
schedule, 5, 11, 14  
scores, 3, 4, 5, 11, 13, 14  
sites, 7, 8, 12, 14  
SRI, 7  
study 1a, 8  
study 1b, 8  
study 2, 9  
study 3, 9  
supervisors, 5, 7, 8, 14  
sustainability, 5, 9, 12, 16  
teachers, 2, 3, 4, 5, 6, 7, 8, 12, 13, 14, 15  
TPE, 10  
Traditional, 4  
validity, 2, 3, 5, 8, 9  
VAM, 4, 5, 14, 16  
What We Have Learned So Far, 1, 4, 6, 12  
What We Plan To Do, 1, 4  
writing, 7

## 7. Glossary of Abbreviations

ACT	(Accelerated Collaborative Teacher)
CRESST	(National Center for Research on Evaluation, Standards, and Student Testing)
CSU	(California State University)
DISP	(disposition)
ELL	(English language level)
HLM	(hierarchical linear model)
IR	(Institutional Research)
ITEP	((Integrated Teacher Education Program)
ITR	(Information Technology Resources)
LAUSD	(Los Angeles Unified School District)
LEP	(limited English proficiency)
LS	(Liberal Studies)
MKT	(mathematical knowledge for teaching)
NRT	(norm-referenced test)
NAHS	(Northridge Academy High School)
PCK	(pedagogical content knowledge)
PK	(professional knowledge)
SM	(subject matter)
SRI	(Stanford Research Institute)
TPAs	(teacher performance assessments)
TPEs	(teacher performance expectations)
VAM	(value-added model)

## 8. Credits

### Area

Theory of Change:

Longitudinal/Multi-site:

Data Warehouse:

HLM/VAM:

Studies 1, 2, 3:

NAHS:

Induction:

Clinical Sites:

CRRT:

Liberal Studies:

Social Sciences:

English:

Science:

Math:

Chancellor's Office:

Drafting, Editing:

### Individuals and Teams

Spagna, Cabello, Theodoulou, Humphrey, Rusche

Crowe, Spagna, White, Theodoulou

Cabello, Huber, Kay

Choi, Huber

Spagna, Cabello

Ericson/NAHS Team

Eaton, Chong/Team

Prosenjak/Team

Montano, Gudoski/Team

Cahn, Adams/Team

Graves, Donaldson, Andrews, Burke/Team

Barnard, Clark/Team

Herzog, Najy-Shadman/Team

Zeitlin, Gold, Mason/Team

Wright, Wilson

Bishop, Cabello, Eaton, Handler, Hellenbrand, Herzog, Humphrey,

Rusche, Spagna, Theodoulou, White