

Using program evaluation for
understanding and improving
language education

John M. Norris

Georgetown University

September 19, 2014



US foreign language education: Status quo 2001

washingtonpost.com

The Washington Post

More Steps We Can Take Right Now

Beef Up the Country's Foreign Language Skills

“Americans need to be open to the world; we need to be able to see the world through the eyes of others if we are going to understand how to resolve the complex problems we face.”

Daniel Akaka, U.S. Senator from Hawaii
Daniel Akaka, U.S. Senator from Hawaii

The need for reform in language education



Modern
Language
Association

MLA

En
D
Ac

Resources

Job List

Publications

Bookstore

MLA Style

Convention

Governance

Membership

Home > MLA Report to the Teagle Foundation on the Undergraduate Major in Language and Literature

MLA Report to the Teagle Foundation on the Undergraduate Major in Language and Literature

In 2006, the MLA prepared white papers on language and literature concentrations in the United States. In 2007, the MLA brought together a group of experts, including college professors and industry professionals, to write the group's report to the Teagle Foundation on the undergraduate major in language and literature. The [report](#) also serves as a call to action for generations of students to a traditional core of liberal study: language, literature, and culture.

We know that today's students are less likely to choose language and literary study as majors than they were thirty-five or even fifteen years ago, and we wanted to explore ways to strengthen majors in our fields and attract new generations of students to what has been the traditional core of liberal study.

The need for reform in language education



Modern Language Association **MLA**

Resources Job List Publications Bookstore MLA Style Convention Governance Membership

Home > Professional Resources > Surveys, Reports, and Other Documents > Foreign Languages and Higher Education: New Structures for a Changed World

Download this report in PDF.

Foreign Languages and Higher Education: New Structures for a Changed World

Background

The Modern Language Association's Language and Culture Committee was formed in 2004 in response to the challenges of language education in today's world. The committee was established in response to the challenges of language education in today's world. Languages, characterized by deep translingual and transcultural competence, were a key focus of the committee's work. The committee was established in response to the challenges of language education in today's world. Languages, characterized by deep translingual and transcultural competence, were a key focus of the committee's work.

The language major should be structured to produce a specific outcome: **educated speakers who have deep translingual and transcultural competence.**

colleges and universities. It began working in 2004 and submitted its report to the Executive Council two years later. Committee members have made presentations at the MLA convention and at other public venues, including events sponsored by federal agencies, professional



US foreign language education: Status quo today

“The **lack of language skills and civic and global awareness** among American citizens increasingly jeopardizes their ability to interact with local and global peers or participate meaningfully in business, diplomatic, and military situations.

The United States is **not producing enough foreign-language speakers** to staff important posts in the U.S. Foreign Service, the intelligence community, and American companies.”

US Council on Foreign Relations, 2012



US foreign language education: Status quo today



HOME ABOUT SUBCOMMITTEES HEARINGS MEDIA LEGISLATION & NOMINATIONS LIBRARY CONTACT

Homeland Security & Governmental Affairs
OVERSIGHT of
GOVERNMENT MANAGEMENT
SUBCOMMITTEE

WEBSITE ARCHIVE

Search this Site **SEARCH**

OVERSIGHT OF GOVERNMENT MANAGEMENT HOME **MORE** ▼

Oversight of Government Management > Hearings

A National Security Crisis: Foreign Language Capabilities in the Federal Government

Oversight of Government Management, the Federal Workforce, and the District of Columbia

May 21, 2012 02:30PM
Location: Room 342, Dirksen Senate Office Building



US foreign language education: Status quo today

The screenshot shows the homepage of the American Association of University Professors (AAUP) website. The header includes the AAUP logo, navigation links (LOGIN, CONTACT, and social media icons), and a search bar. A yellow banner highlights the current issue: "SEPTEMBER-OCTOBER 2011". Below this, a blue breadcrumb trail reads "HOME » REPORTS & PUBLICATIONS » PUBLICATIONS » ACADEME » SEPTEMBER-OCTOBER 2011".

The main content area features a sidebar on the left with links for "Current Issue", "Previous Issues", "Contact", "Advertising", "Submissions", and "Subscriptions". Below these is a button that says "CHECK OUT THE CURRENT ISSUE".

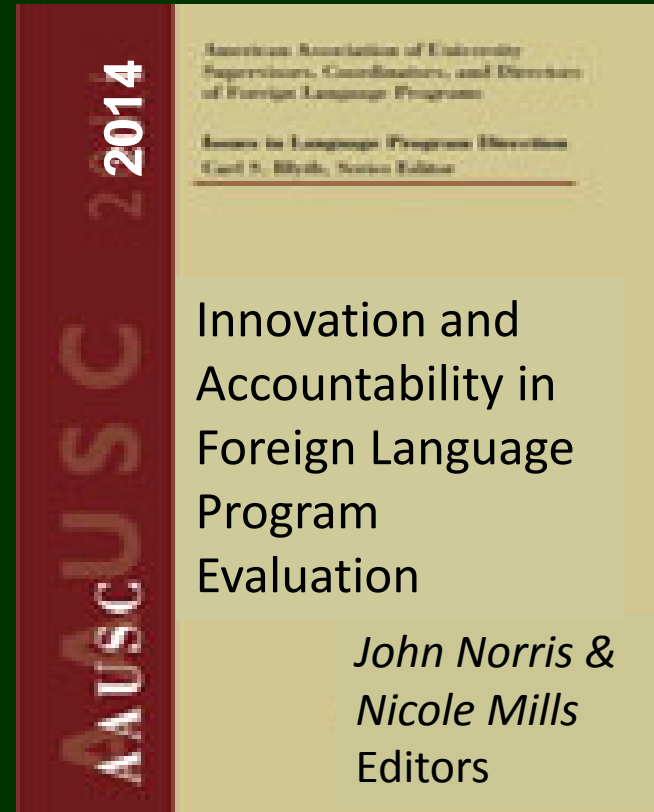
The central article is titled "The Real Language Crisis" and is circled in red. Above the title are social sharing buttons for Facebook (Like), Twitter (Tweet), and Pinterest (Pin it), along with a share count of 38. The article text begins with: "We are becoming a nation of second-language illiterates, and recent draconian cuts to language teaching in colleges and universities are exacerbating an already serious problem." It is attributed to "By Russell A. Berman". The first paragraph of the article reads: "We in America are confronting a stark alternative: either open ourselves to an appreciation of human plurality and the diversity of cultures around the globe or limit ourselves to a narrowly normative culture. Will American schools and colleges provide students with opportunities to learn to understand other voices, or will our educational system succumb to the temptations of isolationism and xenophobia?"



A reform role for evaluation?

“Innovation in FL education today is an absolute essential: we adapt, or we do not survive. Evaluation offers a systematic basis for language programs to pursue innovations of various kinds.”

Norris & Mills (2014, p. 11)





A reform role for evaluation?

But what kind of evaluation leading to what kind of reform?

"This emphasis compels us to justify our values and methods by translating them into the quantitative, quasi-scientific methods... We are not asked to identify what we want our students to know or understand or be prepared intellectually to grapple with. Rather, we are asked for the behaviors that our students will exhibit that will demonstrate their learning—and we are told that we must develop a quantitative instrument that will measure these behaviors."

Berger (2008)

"...[M]ore students are attracted to our program because the improvements in teaching and curriculum that grew out of the assessment process, which include increased emphasis on oral proficiency and culture for all languages, have enabled them to better move toward their goal of communicating fluently in the target language and acquiring an improved understanding not only of the target culture but also of their own."

Carstens-Wickham (2008)



A reform role for evaluation?

2 ways of seeing evaluation

Regulatory mechanism

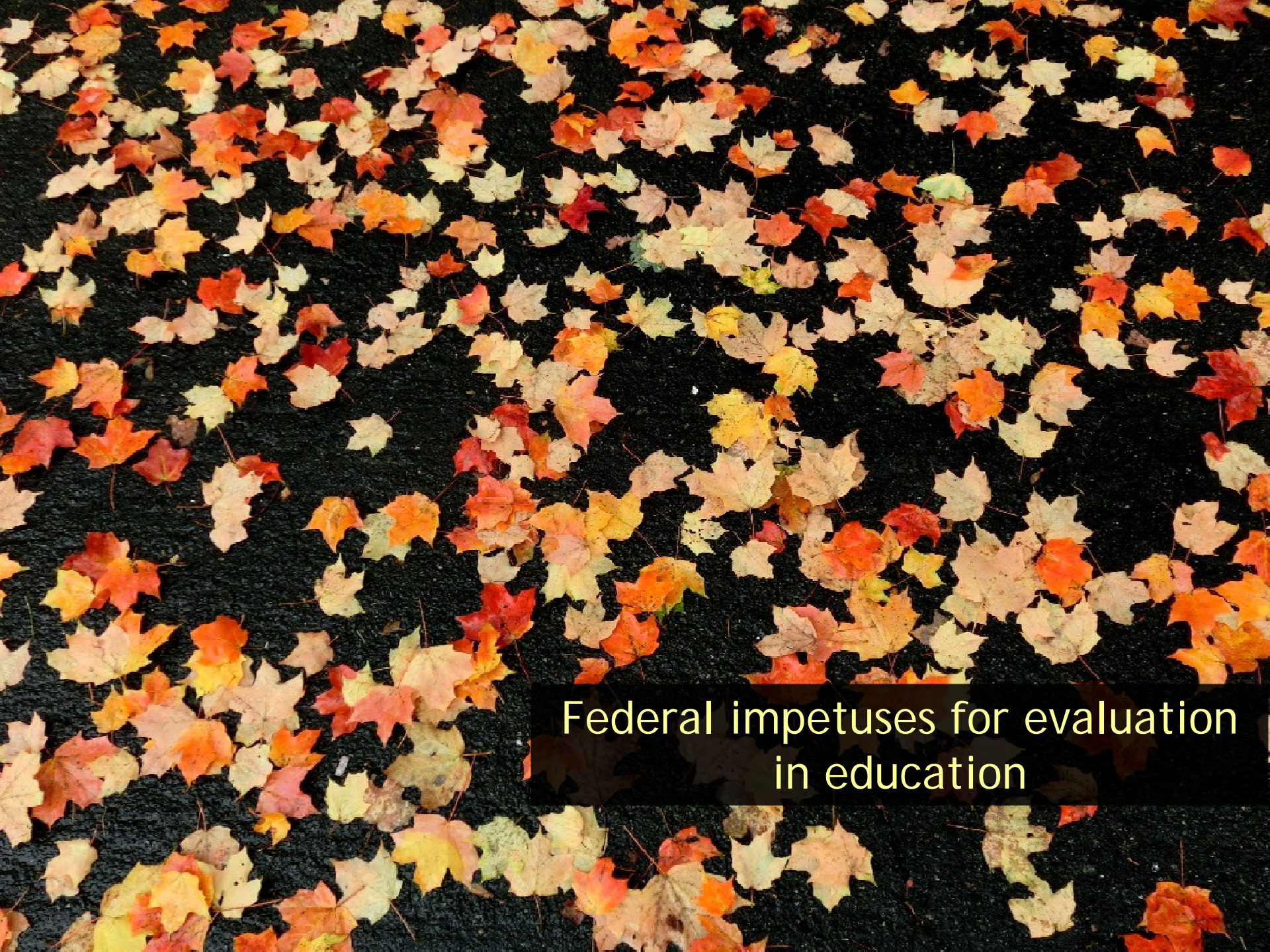
- agent of change
- external, mandated
- quasi-scientificist
- quantitative measurement
- comparing, monitoring, managing

Educative process

- capacity for dealing with change
- internal, proactive
- pragmatic practice
- aligned assessments
- inquiring, understanding, improving

How do we choose to see it?

What approach to evaluation enables reform?



Federal impetuses for evaluation
in education



Impetuses for evaluation: Accountability

Accountability movement: using standardized tests to hold teachers and students to performance expectations

"We're going to stand strong on accountability"
--Margaret Spellings, U.S. Secretary of Education



No Child Left Behind (K-12)

"...higher education institutions should measure student learning..."
--Spellings Commission (2007) on higher education



No *College Student* Left Behind???



Impetuses for evaluation: Accreditation





Impetuses for evaluation: Accreditation

Primary emphasis → Ensuring educational effectiveness

Primary indicator → Student learning outcomes

Primary mechanism → SLO assessment

Assessment mandate, e.g., WASC (2008):

The program has a fully-articulated, sustainable, multi-year assessment plan that describes when and how each outcome will be assessed and how improvements based on findings will be implemented. The plan is routinely examined and revised, as needed.



Impetuses for evaluation: GPRA

Government Performance and Results (Modernization) Act (2010)

To require quarterly performance assessments of Government programs for purposes of assessing agency performance and improvement...

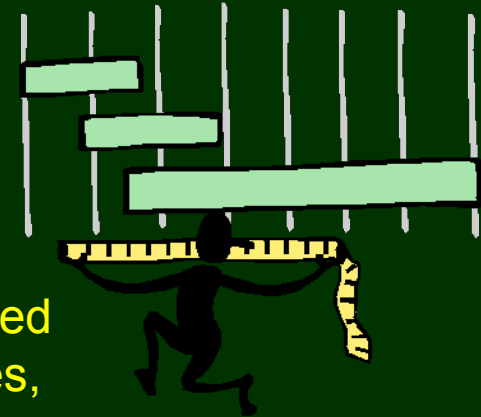
...include a description of the program evaluations used in establishing or revising general goals and objectives, with a schedule for future program evaluations to be conducted.

(1) establish performance goals to define the level of performance to be achieved during the year in which the plan is submitted and the next fiscal year;

(2) express such goals in an objective, quantifiable, and measurable form unless authorized to be in an alternative form...



Plan → measure → report → budgetary consequences...





Impetuses for evaluation: The White House

2014 Economic Report of the President Chapter 7: Evaluation as a Tool for **Improving** Federal Programs

The Administration continues to support the use of these tools, broadly and often, to facilitate continuous improvement in government programs as well as to identify best practices and effective new approaches that can be shared with organizations delivering services funded with Federal dollars. (p. 297)



What tools???



Impact Evaluation
(aka: Randomized Control Trials)



...evaluations measuring impact on outcomes using random assignment provide the most definitive evidence of program effectiveness. (p. 272)



Impetuses for evaluation: Key features

One approach to evaluation

- External, regulatory mechanism
- Focused on performance and outcomes measurement
- Valuing objectivity and experimental designs
- Encouraging quantification and efficiency

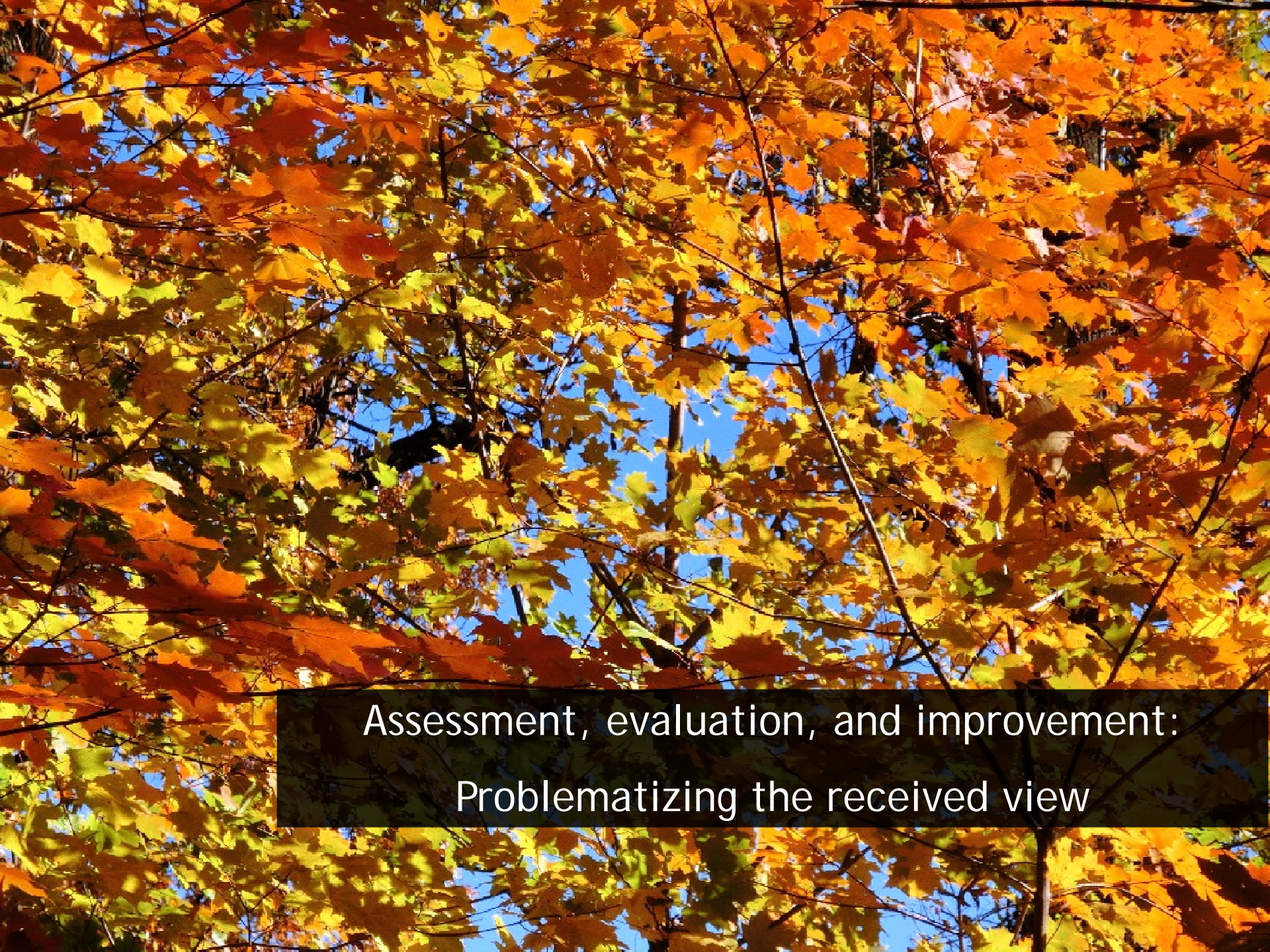


Does it happen?

If so, how?

- Is there any evidence to suggest that it works?
- Is it the best way to achieve program improvement?





Assessment, evaluation, and improvement:
Problematizing the received view



Encountering evaluation in practice

Jet-in-jet-out Expert
(JI JOE) review



Accountability testing

Managerial evaluation



Under these familiar approaches, evaluation gets done, but it generally meets only program-external bureaucratic or political needs; evaluation is done to programs (and teachers, and learners), not with or for programs. *Does it lead to improvement?*



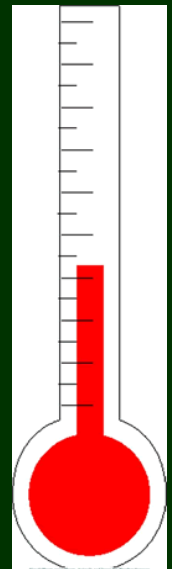
Encountering evaluation: Mis-representation

Spellings Commission (2007), on higher education accountability:

"...higher education institutions should measure student learning..."



Evaluation is measurement
(only)





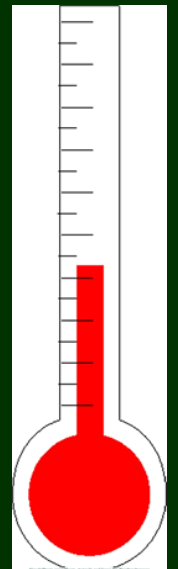
Encountering evaluation: Mis-interpretation

Falk (*Baltimore Sun*), Dean of Arts & Sciences, Johns Hopkins University:

“...the more we rely on standardized testing as our bellwether for the quality of education, the more we will value in education only those things that can be measured on standardized tests”.



Is measurement all that is
needed for improvement to
happen?





Encountering evaluation: Mis-guided practice

Do 'comparison shopping' and 'league tables' lead to improvement?

**INSIDE
HIGHER ED**

insidehighered.com

June 26

Assessment for 'Us' and Assessment for 'Them'

Spellings Commission on the Future of Higher Education

Assessment is done to show: “**how much** students learn in colleges and **whether they learn more** at one college than another...”

academic establishment, from whether a college has for-profit or nonprofit status to whether its classes are offered online or in brick-and-mortar buildings. Instead, they care — as we do — about results.”

This expectation for assessment as accountability has forced many faculty members and administrators to seek out ways to balance assessment for “us”, or assessment for “improvement,” with assessment for “them,” or assessment for “accountability.”



Encountering evaluation: Mis-guided practice

Testing/No Child Left Behind

Press Clips

Institute for Language and Education Policy

Putting Assessments to the Test

By Valerie Strauss
WASHINGTON POST

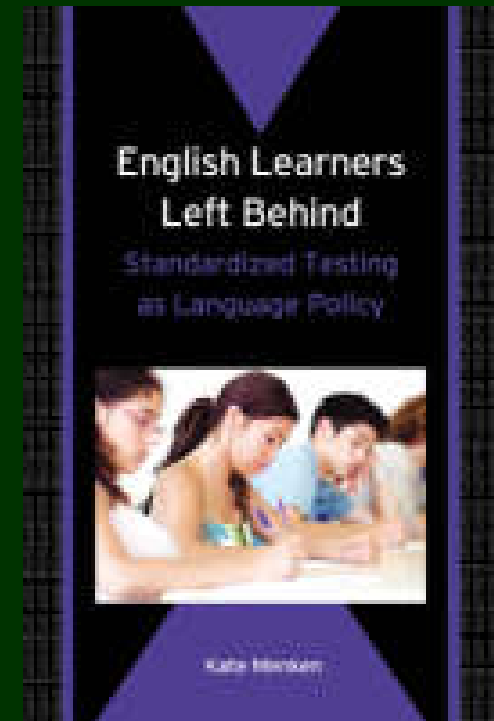
"There has been an explosion of mandates for more and more standardized tests with very little evidence to support their use"

--Walter Haney of Boston College's Center for the Study of Testing, Evaluation and Educational Policy.



Encountering evaluation: Mis-guided practice

Menken, K. & Solorza, C. (2014). No child left bilingual: Accountability and the elimination of bilingual education programs in New York City schools. *Educational Policy*, 8(1), 96-125.





Encountering evaluation: Compliance mentality

SLO Assessment for University Accreditation:

Lived realities

Standard Process

- State outcomes
- Measure behaviors
- Analyze the results
- Then what?
- (Let the chair/dean do it)

Why is this a typical kind of reaction...?



Encountering evaluation: Punitive orientation

- ~~LSU: German, Russian programs (and faculty)~~
- ~~George Washington: FL requirement~~
- ~~SUNY Albany: French, Italian, Russian programs (and faculty)~~
- ~~University of Iowa: MA and PHD in German and Linguistics~~
- ~~Brandeis University: BA in Hebrew and Yiddish~~
- ~~Indiana University of Pennsylvania: BA in French and German~~
- ETC.

“...no clear value to the institution...”

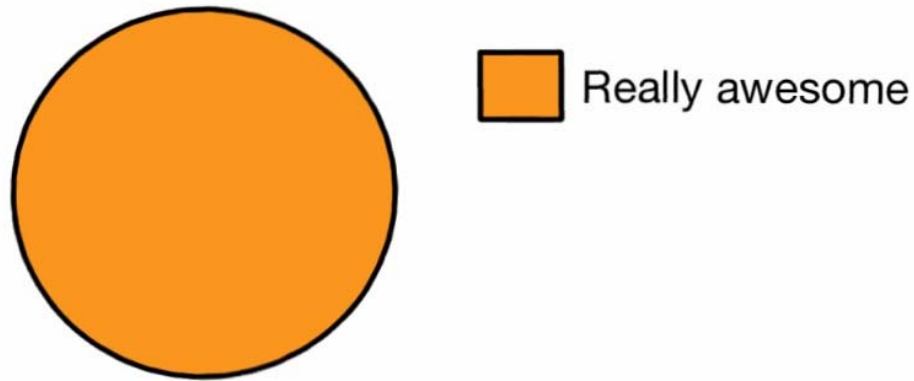


Encountering evaluation: Cynical survivalism

How to approach a self-study evaluation for program review:

Write "reports" using "data"

Figure 1. How awesome we are



SOURCE: Because we said so



Encountering evaluation: 'Just do it' mindset

Common ~~European~~ Framework of
Reference for Languages



One size fits all...?

A new measure to solve all of our
problems?

C2

C1

B2

B1

A2

A1



Encountering evaluation: 'Just do it' mindset

What is the appropriate ACTFL proficiency level for this student learning outcome?

Superior

Advanced High

Advanced

Intermediate High

Intermediate Mid

Intermediate Low

Novice High

Novice Mid

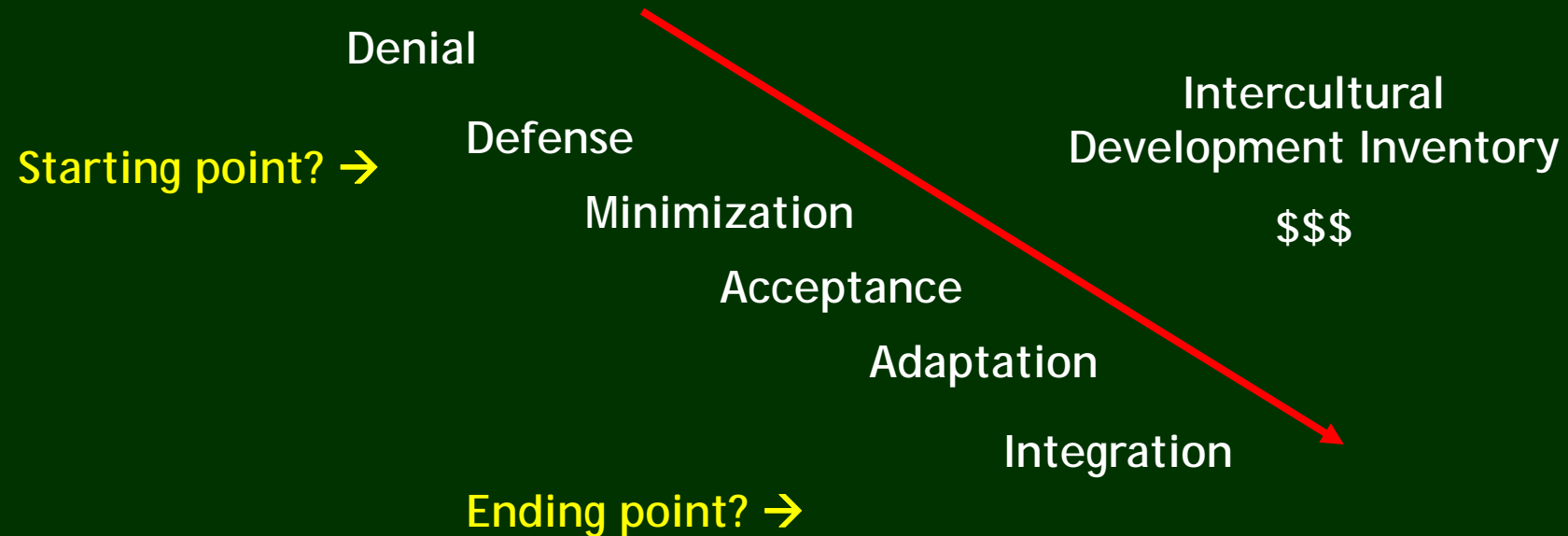


INTERMEDIATE - LOW ???



Encountering evaluation: 'Just do it' mindset

Assessing intercultural learning in study abroad:



Other intercultural outcomes not assessed?

Knowledge? Skills? Awareness? Etc.



Encountering evaluation: 'Just do it' mindset

Collegiate Learning Assessment (CLA)

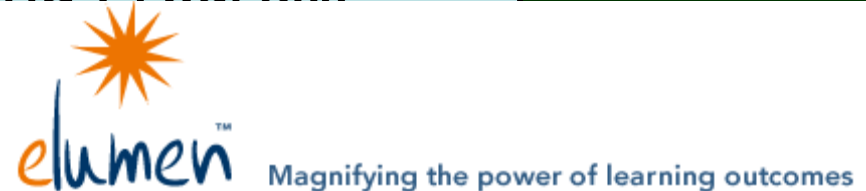


National Institute for Learning Outcomes Assessment
Making Learning Outcomes Usable & Transparent



Major Field Tests

MAPP - Measure of Academic Proficiency and Progress



Respectable efforts at assessments for common outcomes, but...

Too easily adopted without articulation to individual programs and their values

What Do College Graduates Know?

A Survey of Arizona Universities

24. Plato was a pupil of:

a. Aristophanes

18. What are the four basic forces in the

b. Socrates

universe?

c. Crito

d. Aristotle

a. Gravity, electromagnetism, strong

34. Renoir and Monet belong to which

nuclear force, weak nuclear force

school of art?

b. Gravity, electromagnetism, chemical

force, atomic force

a. Surrealism

c. Gravity, centrifugal force, centripetal

b. Identical Snoop Dogg.

force, Coriolis force

c. Impressionism

d. Gravity, electromagnetism,

d. Realist

thermodynamic force, quantum force

b. Cartoon by Charles Schulz

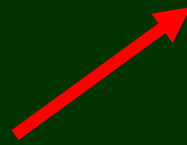
c. A mystery series

d. A jazz pianist

40-item, multiple-choice test...

"...an examination of how much those graduating from Arizona's three public universities—the University of Arizona, Arizona State University (ASU) and Northern Arizona University (NAU)—know about **history, science, math, literature, arts, civics and other subjects**. In other words, have they received a well-rounded, liberal education?" (p. 2)

"The vast majority of students surveyed earned a failing grade on our test of general knowledge and reasoning."





Encountering evaluation: 'Just do it' mindset

pick *a* prof™

**RateMy
Professors**

reviewum.com



RateMyTeachers.com
Teacher Ratings By Students and Parents

Quality Rating
Categories

- ☺ How easy?
- ☺ How fair?
- ☺ How good?

HOW HOT???





Encountering evaluation: Reactions

EVALUATION

Rate My Professor & Professors Strike Back

James Rhem, Executive Editor

Given that student evaluations of faculty remain one of the most written about and contentious areas of research in higher education, perhaps it's important to state at the outset that this article won't shed any new, scientifically valid light on the subject. This article looks at the popular website www.ratemyprofessor.com and its sister site "Professors Strike Back," a feature of www.mtvu.com.

After viewing the videos on "Professors Strike Back," we wondered what those teachers actually thought about "Rate My Professor" and the experience of "striking back." In reference to busy schedules

the Viacom Corporation. MTVu soon set up the "Professors Strike Back" feature where faculty can answer criticism, and at the same time contacted some faculty offering them the chance to be filmed responding to student comments on their teaching and to the idea of "RateMyProfessor" in general.

Frank Popper

In the highly edited videos that appear online, most faculty come across as quite emphatic and, at times, shockingly frank even when cloaked in an exaggerated persona. Frank Popper, who teaching land use planning at Rutgers and Princeton, clearly enjoyed making his over-the-top rebuttal to students. In response to a student comment that his course is a "rewarding experience if you have some

background in this field; if not you might get lost and

can't read. Certainly can't read at the college level."

To the comment that he "loves to hear himself talk," Popper replies:

"Well, what I'm doing there is anti-filibustering you students who don't know what you're talking about, and take too long to say it and get all the details wrong or upside down anyway. Of course I anti-filibuster you. Who would want to listen to you? You're like the worst senator in the world times ten and you're paying for this (you think). I have to listen to you? Not in my class."

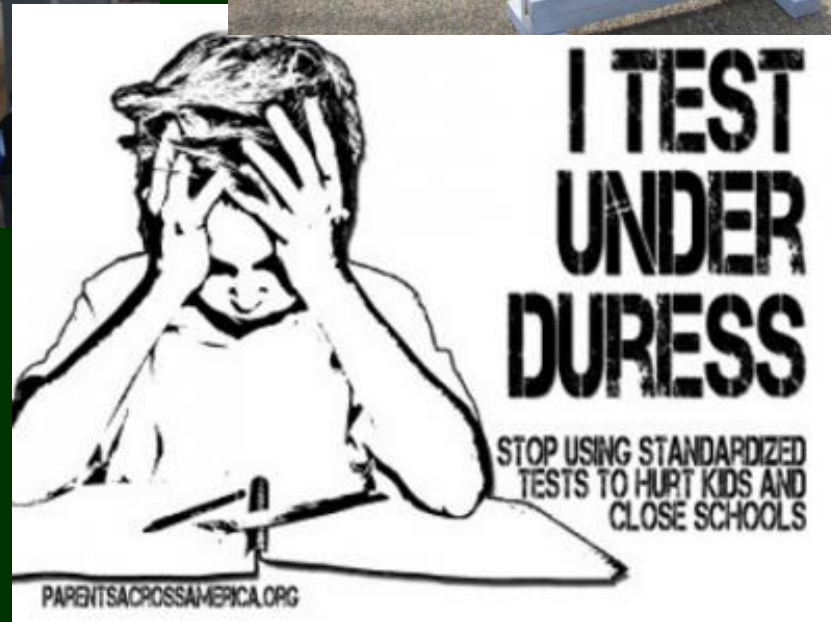
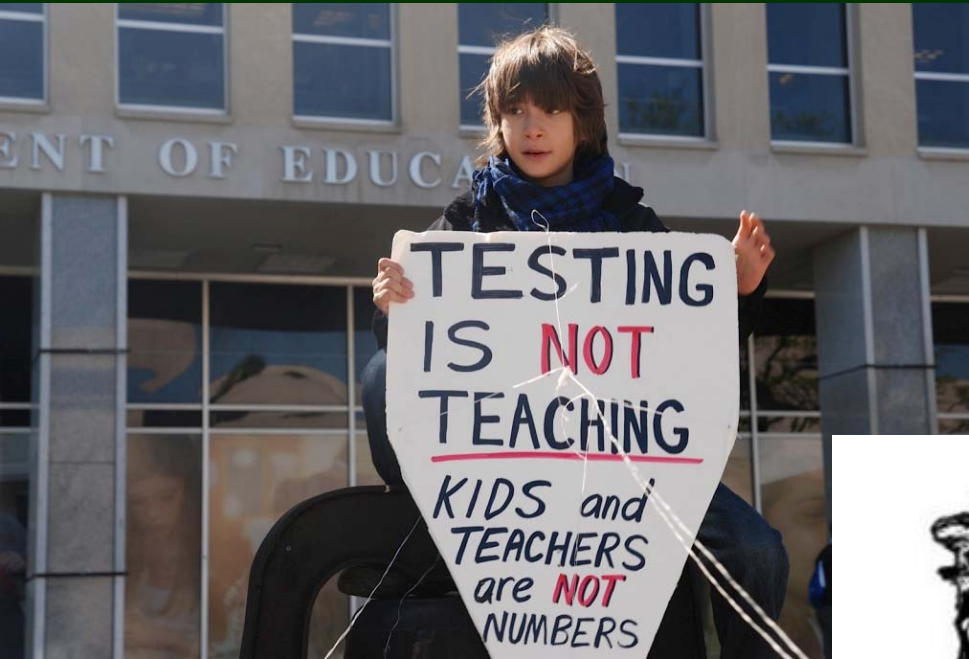
Popper, who drove over to MTV's studios in Greenwich Village to tape his reply, can't quite sustain his persona. What come across as a bit more seriousness begins to leak out in his comment on student ratings on the site in general:

"In general I think you are very young students who have problems—oh what would a psychiatrist say?—of impulse control. You may love me; you may hate me, but you tend to be inaccurate, illiterate, and inappropriate in your expres-





Encountering evaluation: Reactions



Increasing interpretation that evaluation is bad for education



Encountering evaluation: Reactions

Reduction of
evaluation to
testing





Encountering evaluation: Reactions

INSIDE
HIGHER ED

insidehighered.com

Nov. 13

Outcomes Assessment: No Gain, All Pain

By [Bernard Fryshman](#)

Something is certainly afoot. The public disclosure systems put forward by [the National Association of Independent Colleges and Universities](#) and the [Voluntary System of Accountability](#) from the country's two big public university groups are major national initiatives encompassing some of America's most impressive institutions. Miami Dade College's effort to [embed 10 desired learning outcomes into the curriculum](#), and a report of outcomes measurement by discipline, are two other accountability approaches that certainly bear watching.

Even homey old IPEDS has put on a fresh coat of paint, as the Education Department's [College Opportunities Online database](#) (COOL, née PEER) has become Navigator, with an even more attractive set of tools, in the hopes that this time a few more somebodies will use the hundreds of millions of data elements sitting patiently by.

All of this activity is more a function of the skill of Secretary of Education Margaret Spellings in moving her agenda than it is a recognition that there is merit to the numerical assessment of student outcomes.



**LITERARY STUDY, MEASUREMENT,
AND THE SUBLIME:
DISCIPLINARY ASSESSMENT**

Editors

Donna Heiland and Laura J. Rosenthal

With the assistance of
Cheryl Ching

Teagle Foundation (2011)



Encountering evaluation: Reactions

Holquist (2011), on dangers of outcomes assessment:

“...to sacrifice all other goals in the service of standardized outcomes across the board, no matter what the effect might be on different areas of scholarship, university systems, or individual students and professors—the humans who are the subject of the humanities.” (p. 79)

“But in the end, it is only in the scale of a whole lifetime that the worth of literary education may be measured, and that is a scale that cannot be ‘tuned.’” (p. 86)



Encountering evaluation: Reactions

Barrington (2003), on assessment in the liberal arts:

“To design and administer (intellectually honest) assessment plans that will measure such capabilities with a dozen or more standardized ‘learning objectives’ is next to impossible” leading to “pestilent repercussions” for the truly valued learning objectives that constitute the liberal arts, in that it “discourages teaching such skills because they are difficult to measure”.

MLA website blogger:

“What I would much rather see is a definitive statement from the MLA rejecting the assessment madness altogether. Let’s admit that, when all is said and done, what we do is not something that we can ‘know’, or that can be measured...”



Encountering evaluation: Reactions

Survey of US college FL educators, on evaluation:

*“frankly, a waste of time; it just causes us to jump through **meaningless hoops**. Good FL instructors already assess their students constantly both inside and outside of class and in a variety of ways. It's what we do. Much of the assessment craze seems to be a **waste of time for us.**”*

*My main concern is... “That it would not be a waste of everyone's time with **no concrete results**. That the people running it would be so afraid of stepping on toes that they **just babble about quality without looking at the details.**”*

*“Although required by our university and accrediting association, faculty see it as a burden that is essentially a waste of time. Some faculty refuse to participate. **Conclusions drawn from evaluations have little, if any, impact on decision making.**”*

*“Sometimes we think they are **just collecting dust on some administrator's shelf in the Dean's office**, cause we never hear anything from them...So, I'm not sure what those program evaluations are really accomplishing in our college or in our university.”*

“Evaluation is a colossal waste of everyone's time!”



Summarizing the received view

- Evaluation portrayed & perceived as a bureaucratic mechanism
- Associated with technocratic, external measurement tools
- Realized in perfunctory, compliance-oriented ways
- Often punitive rather than formative or transformative
- Misplaced, unscholarly reactions to evaluation
- Non-participation, no buy-in by FL educators
- Missed opportunities for follow-through on evaluation

So, what's the use?

And, what are the alternatives?



What are the alternatives?

How do we transcend these debilitating attitudes?
How can we transform evaluation into a useful process that leads to the improvement of FL education?

The Chronicle of Higher Education

Drake University language programs...


- Low enrollments, student dissatisfaction, poor external reviews
- Faculty refuse instructional development support
- Faculty refuse to create strategic plan for improvement

→ No acknowledgement of need to change

→ No engagement with evaluation findings

→ NO MORE LANGUAGE PROGRAMS, NO MORE FACULTY!

David Maxwell has a message for prospective foreign-language students: If you want to study French, take a college course. If you want to learn to speak it, take a plane to Paris. But steer clear of American classrooms, Mr. Maxwell



Re-envisioning evaluation:
A focus on utility and use



Learning from research on *useful* evaluation

When the American Evaluation Association put together its *Standards for Program Evaluation*, the foremost criterion they addressed was...

UTILITY: The Utility Standards are intended to ensure that an evaluation will serve the practical information needs of intended users.

Joint Committee on Educational Evaluation (1994)



Focus on who and why



Learning from research on *useful* evaluation

But what are the characteristics of evaluations that meet this utility criterion? Considerable research has gone into answering this question (see Patton, 2008), and the following characteristics have been associated strongly with evaluations that produce useful and used results...

evaluation findings

key stakeholders

Manageable & Feasible:

Adapted to available
time and resources

Democratic:

Negotiated
decision making

Educational & Transformative:

Users learn by participating

Responsive:

Evaluation responds to
primary intended users'
purposes

Clear & Understandable:

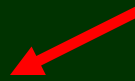
Transparent processes
and outcomes



Learning from research on *FL* evaluation

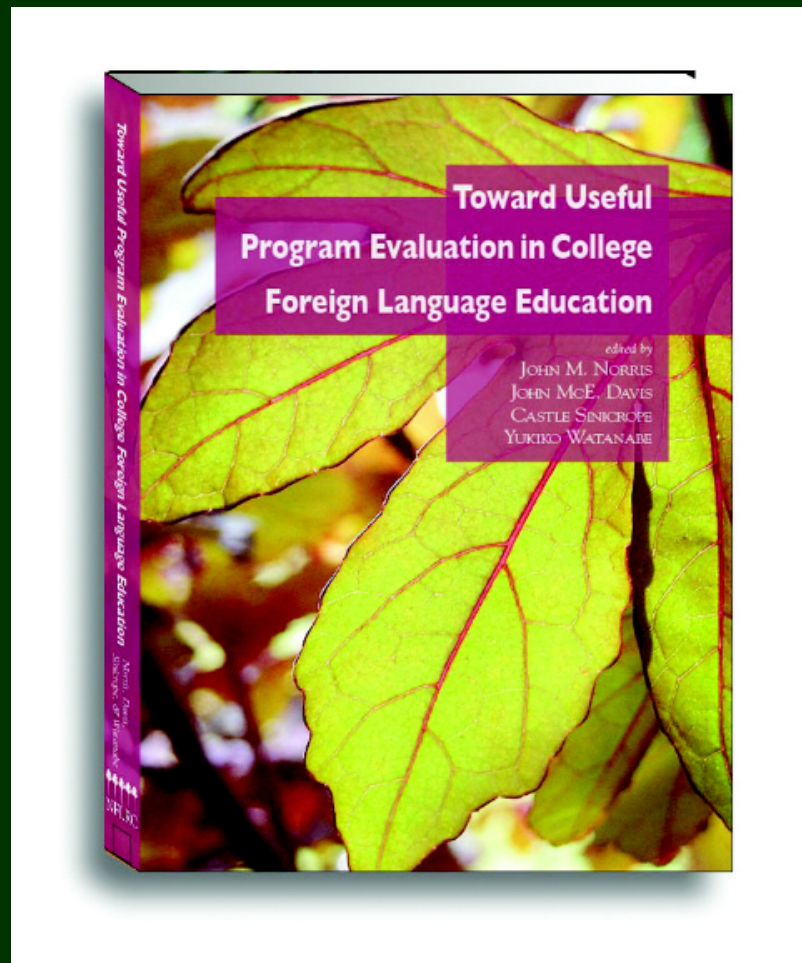
Observations from case studies:

- Leadership = Modeling, enabling
- Focus = Prioritization for feasibility
- Data = Real, empirical, cyclical, local
- Collaboration = Participation by multiple faculty, staff, students crucial for programmatic use/reform to ensue



New questions:

- Factors that predict useful evaluation?
- Contextual constraints?
- What learning/change happens?
- Which methods contribute most?





Learning from research on *FL* evaluation



Davis (2012): What factors contribute most to SLOA being used in college FL programs? (Survey of US FL departments)

Institutional support for assessment

Program leadership

Faculty collaboration

Institutional assessment infrastructure

Communication about assessment

Program resources for assessment

Program ethos to improve/innovate

Institutional policies for assessment

Collecting, using assessment data

Making changes and learning from process

1	2	3	4
Not at all			A lot
37%	27%	22%	14%



Learning from research on *FL* evaluation



Davis (2012): What factors contribute most to SLOA being used in college FL programs? (Survey of US FL departments)

Institutional support for assessment

Program leadership

Faculty collaboration

Institutional assessment infrastructure

Communication about assessment

Program resources for assessment

Program ethos to improve/innovate

Institutional policies for assessment

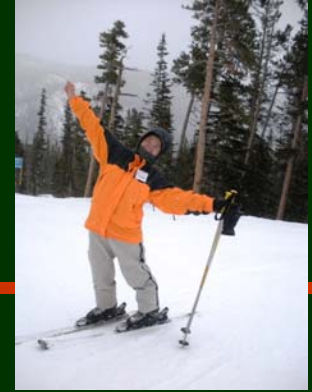
Collecting, using assessment data

Making changes and learning from process

1	2	3	4
Not at all			A lot
37%	27%	22%	14%



Learning from research on *FL* evaluation



Watanabe(2012): How/when/why do programs learn through SLOA?
(2-year case studies of 8 college FL programs)

Engaged,
participatory leaders

Clear sense of
program identity and
value(s)

Commitment to
collaborative
innovation

Capacity and
structure to make
decisions

“Proactive program ownership”

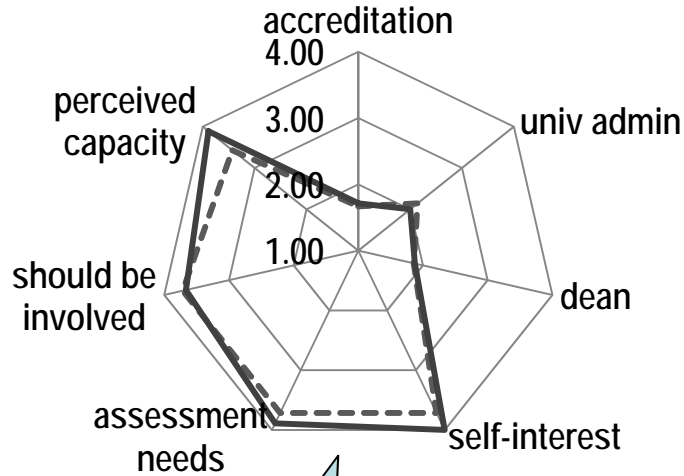
...the ability to take advantage of or suppress external accountability pressures and perceive the need for, self-interest in, and capacity to do outcomes assessment

Learning through evaluation, primarily...

- **changes in pedagogy**
- **changes in outcomes/assessment**
- **changes in curricular design**
- **changes in *understanding***

High assessment use for learning

Program B

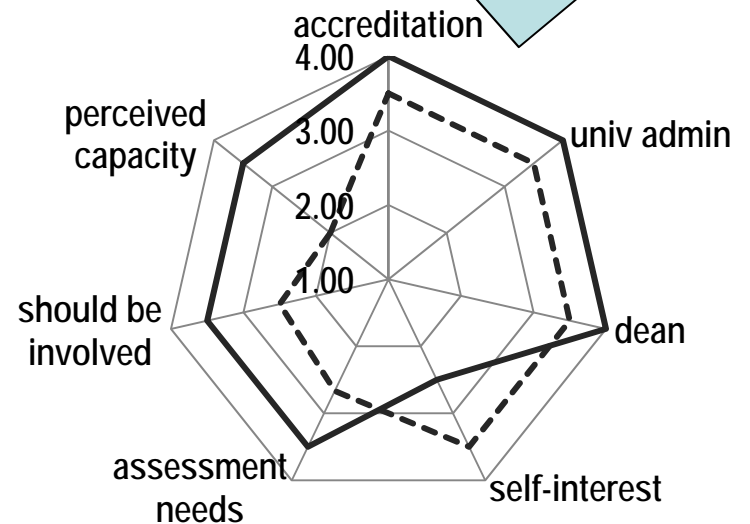


Internally motivated, not driven by program-external forces

----- = presurvey
———— = postsurvey

Program E

Primarily externally motivated



Low assessment use for learning



Useful evaluation in college FL programs: What have we learned?

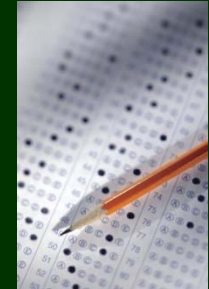


Context & conditions:
→ Localized program focus and identity, engaged leaders, collaborative faculty, commitment to innovation, structure for decision-making



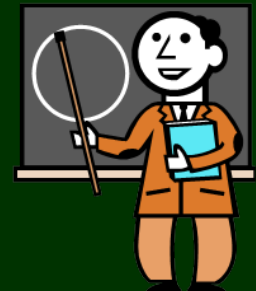
Process:

→ Participatory; focus on intended uses; prioritize important, feasible activities; collect locally meaningful data; follow through with joint decisions, actions; repeat the cycle



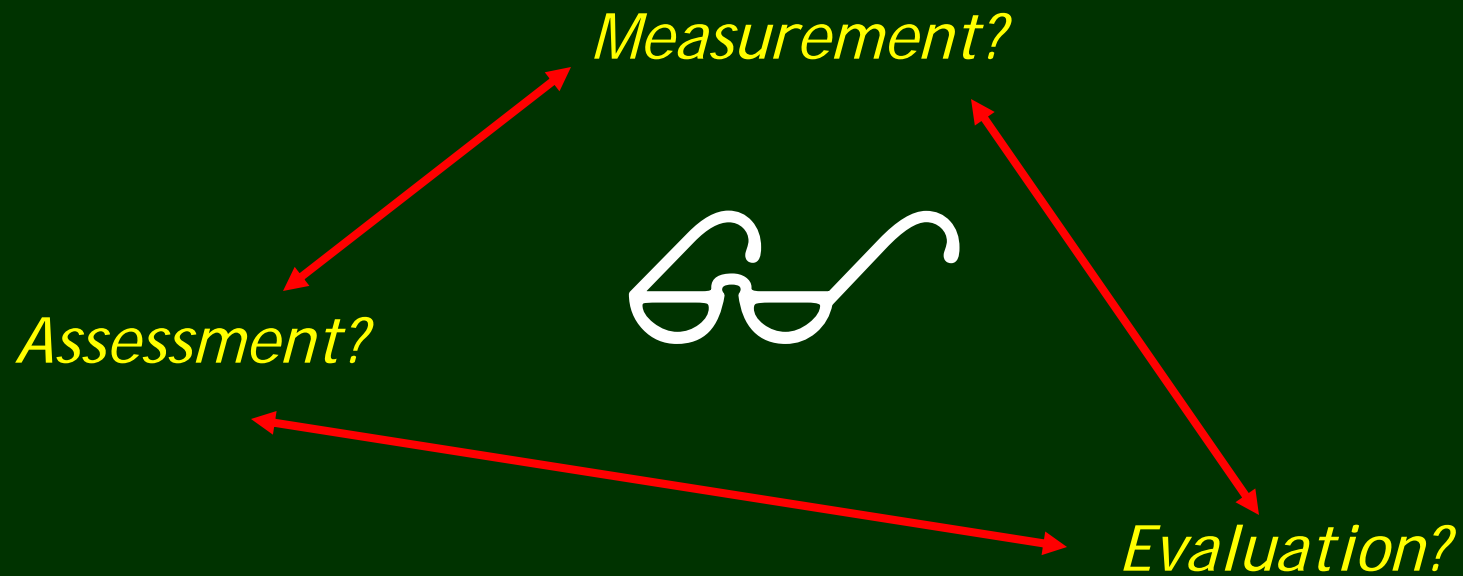
Methods:

→ Aligned with program and scholarly values; focused on substantial issues; designed, selected for specific uses; emphasize information yield





First steps towards useful evaluation: Vision correction





Resolving terminological confusion

Evaluation is the gathering of information about any of the variety of elements that constitute educational programs, for a variety of purposes that include primarily understanding, demonstrating, improving, and judging program value; evaluation brings evidence to bear on the problems of programs, but the nature of that evidence is not restricted to one particular methodology.

Norris (2006) *MLJ Perspectives*



Resolving terminological confusion

Outcomes assessment = the use of information about student learning for understanding and improving educational programs...

EVALUATION
of programs

ASSESSMENT
of learners

MEASUREMENT
of quantifiables



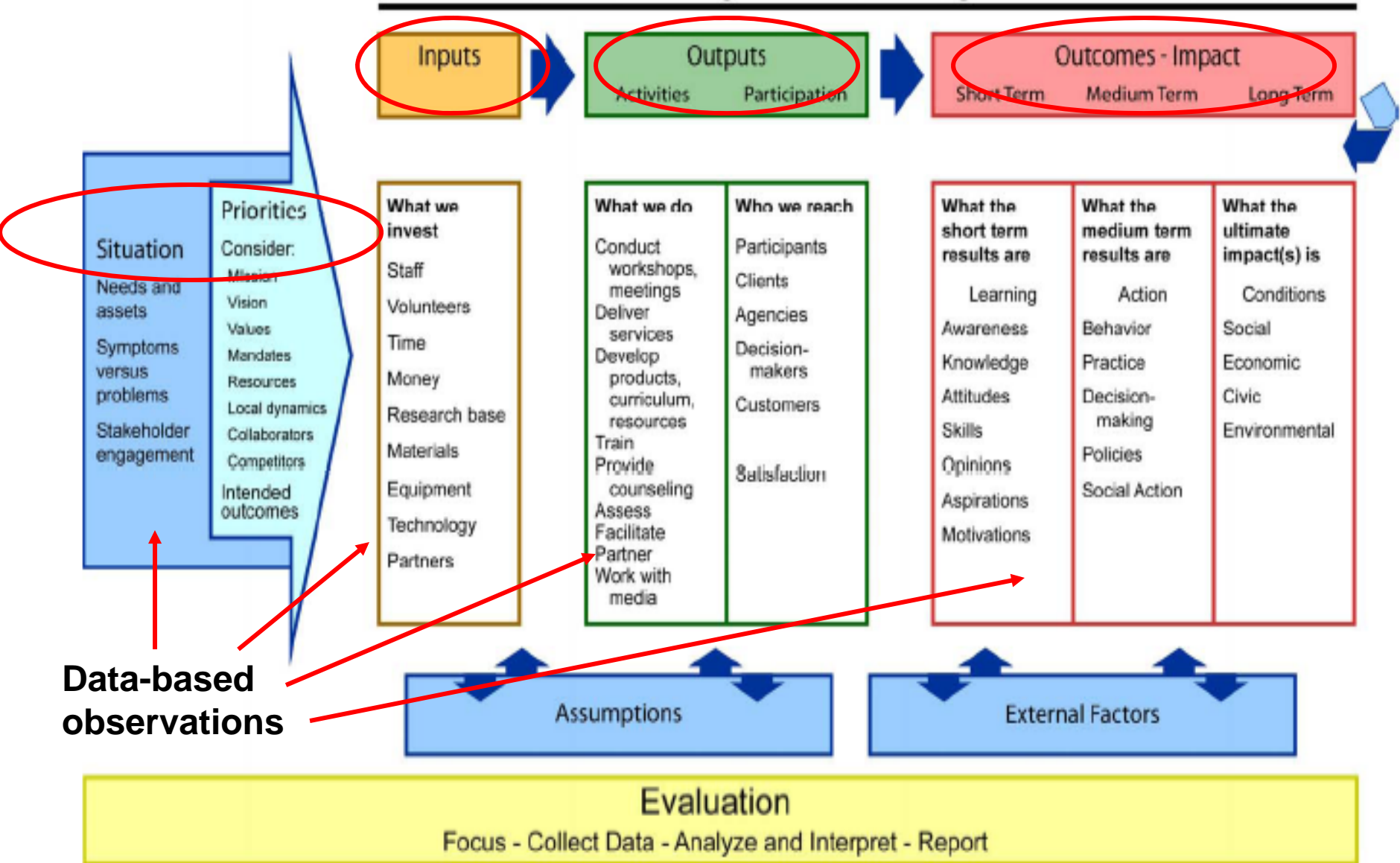
How does evaluation work?

PROGRAMMATIC

PRAGMATIC

PARTICIPATORY

Program Action - Logic Model





How does evaluation work?

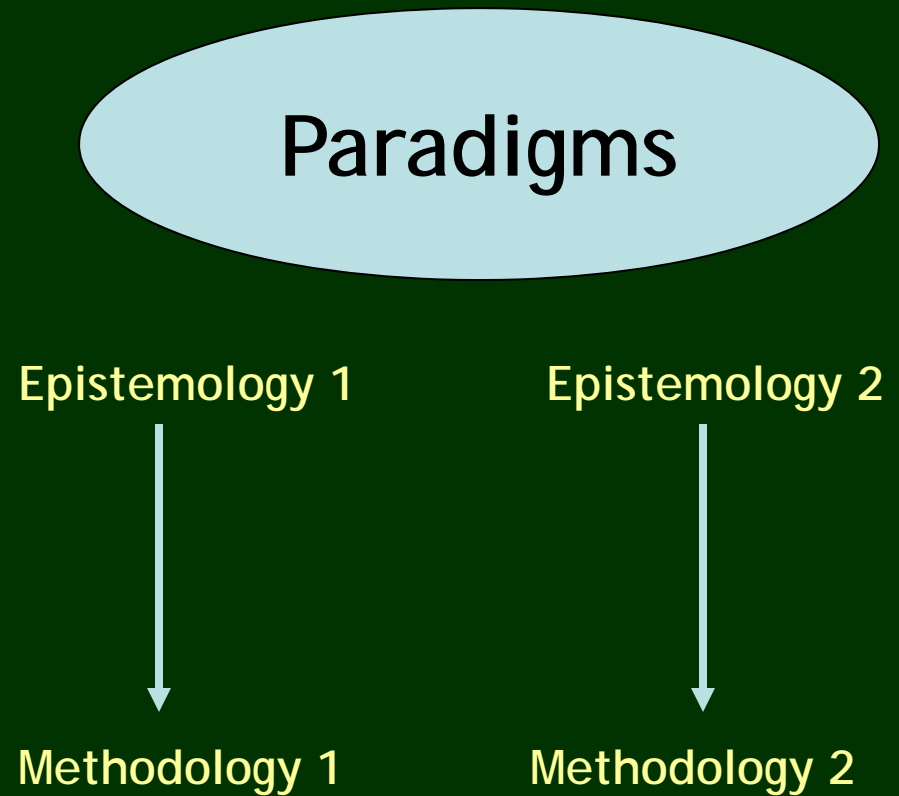




How does evaluation work?

Cronbach et al. (1980)

“The evaluator will be wise not to declare allegiance to either a quantitative-scientific-summative methodology or a qualitative-naturalistic-descriptive methodology.” (p. 7)

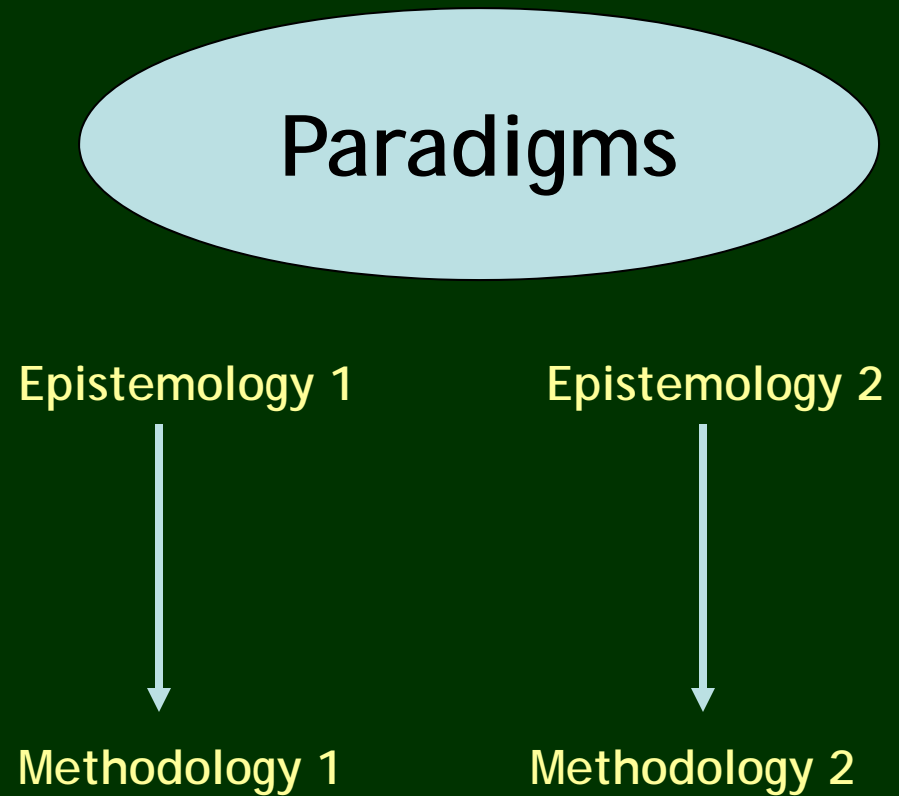




How does evaluation work?

Cronbach et al. (1980)

“The evaluator will be wise not to declare allegiance to either a quantitative-scientific-summative methodology or a qualitative-naturalistic-descriptive methodology.” (p. 7)



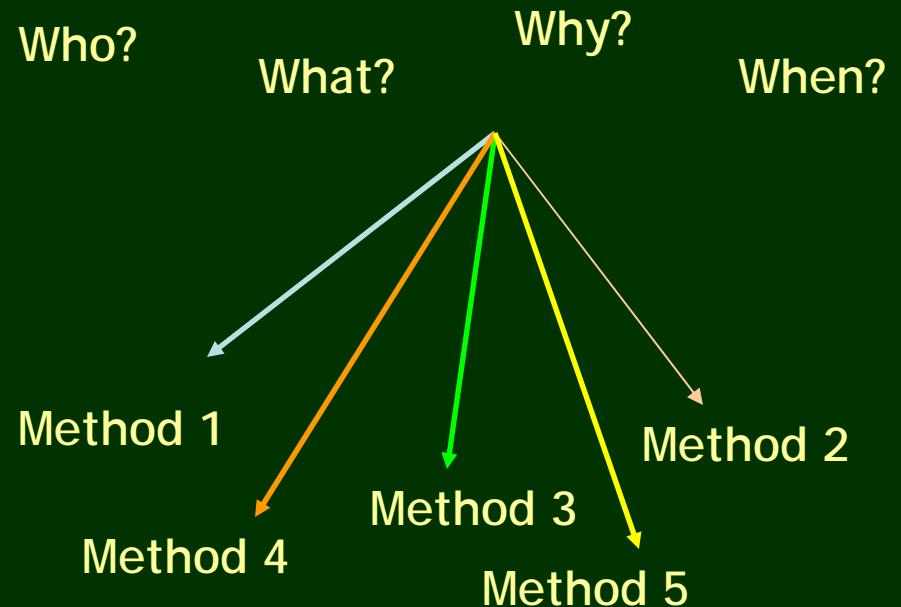


How does evaluation work?

Cronbach et al. (1980)

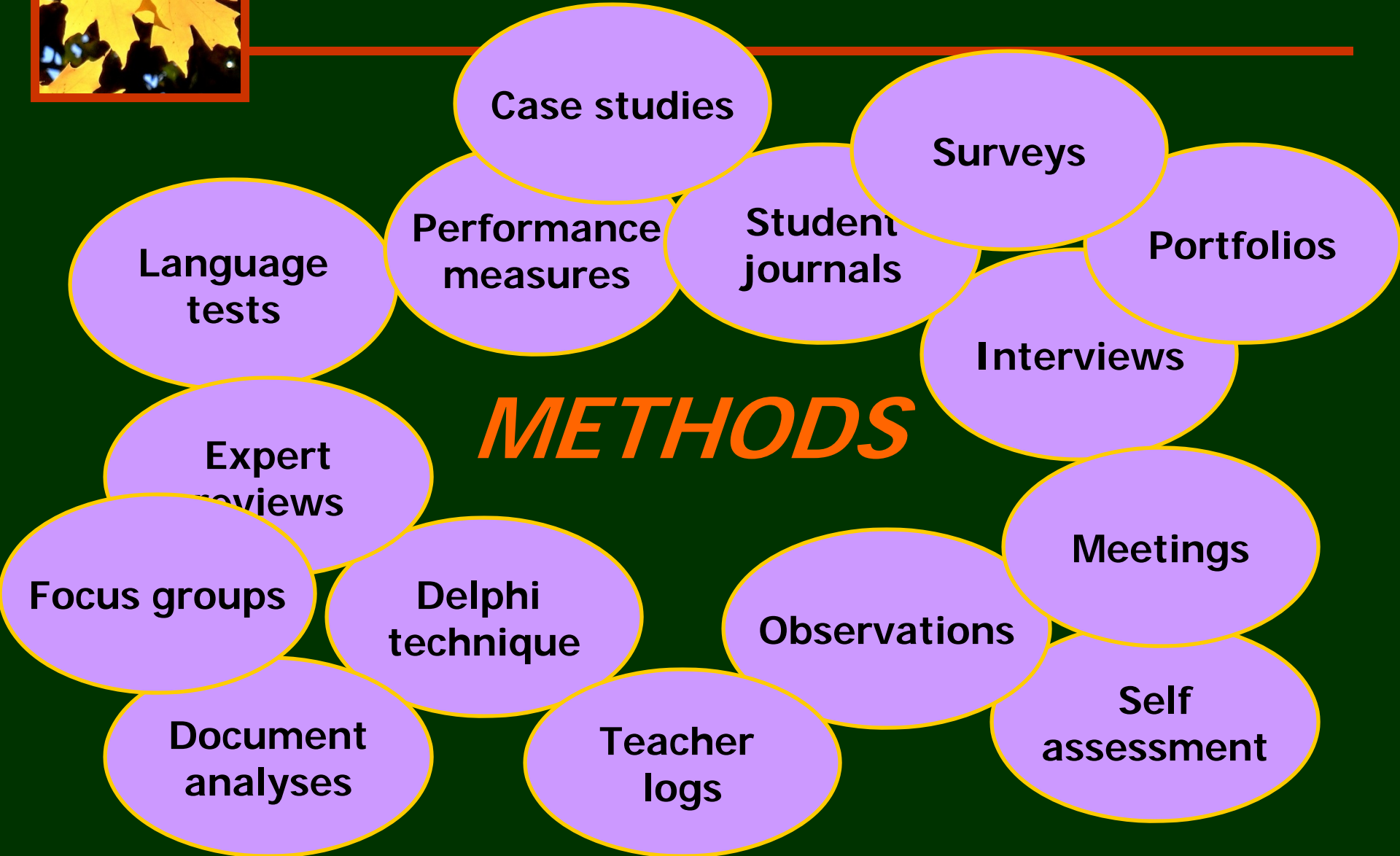
“The evaluator will be wise not to declare allegiance to either a quantitative-scientific-summative methodology or a qualitative-naturalistic-descriptive methodology.” (p. 7)

Pragmatism





How does evaluation work?





How does evaluation work?

1. **Participation** - stakeholders, representatives, primary intended users
2. **Prioritization** - challenges, questions in immediate need of answers
3. **Instrumentation** - what data will answer the questions?
4. **Collection** - how can we get data in available time/resources?
5. **Interpretation** - what do findings mean in context?
6. **Utilization** - what decisions & actions are taken?

Participation by language educators is essential throughout evaluation if contextual relevance and program improvement are sought.



Corrected vision...

What is the starting point for developing useful evaluations?

Who is asking for that information? Why?

Evaluative vision: Begin by asking....


- What are the positions utilized in the program?
- How can they be measured?
- What questions do they have about learners, teachers, courses, curriculum, etc.? Are they being met? What challenges do they face?

Who is doing the measuring and interpreting? Why?

Who is held responsible? Who can actually make changes in the program?



Useful evaluation



Evaluation in use for understanding and
improving language education



Example 1: Improving teacher induction practices

Rapid turnover
in GTAs

Variable teaching
experience

Flexible induction
practices

"Lack of
preparedness"

High student & uni
expectations for ELI

Feeling of 'sink or
swim' teaching



Context

- English Language Institute
- U.S. University
- Diverse International Students
- Required ESL coursework
- M.A. Graduate TA Instructors



Teacher-led
internal evaluation



ELI Administration =
Intended users

Teacher pre-service
needs?

Admin's induction
goals?



Improvements in
induction practices?

See Yang (2009)



Example 1: Improving teacher induction practices

Findings + Formative Uses

✓ Induction partially successful: **identified weaknesses**

✓ Induction outcomes not

Transformations



Teachers

supported
ed

ues

evaluation



Dedicated teacher
evaluator position
created

Interviews:

ELI Administrator: "I can now see how evaluation is connected to the evolution of the program"

teachers

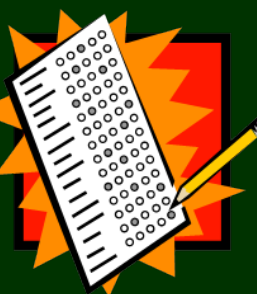
✓ Teachers value variety of practices: **enhance availability (online resources, observation, meeting with other teachers)**

✓ New teachers unaware: **early hiring, pre-work internships, annotated syllabi**

Surveys:

- Admin
- Former + new teachers

Methods



Course designers,
BPA supervisors,
Learners, Alumni all
involved in eval

2: Designing & validating a program

Context: Spanish language program

- BPA 8-week language training course
- Limited effectiveness of translation textbook approach
- "...students who had just finished BPA were not prepared to communicate"

Evaluation use:
Program
development and
design



Evaluation Step 1: Needs Analysis

- Interviews: Supervisors, trainers, agents
- Observations: Tactical training job tasks
- Analyses: Videos of agents doing tasks

7 target task types (e.g., vehicle stop, first aid, extracting suspect) + Language requirements of each task type



New TBLT Spanish Course

Evaluation use:
Judgment and
decision to continue
implementation

2: Designing & validating a program

Evaluation Step 2: Priority Questions

- Is the new program more effective in developing speaking task abilities?
- Does the Spanish oral proficiency of trainees change over the 8-week course?



Picture sequence narration

Versant Test of Spanish

Pre-post TBLT cohort

+sentence mastery

+fluency

+vocabulary

+pronunciation

External standardized measure

Final GT cohort

Fluency

Syntactic complexity

Lexical complexity

Grammatical accuracy

<

<

<

=

First TBLT cohort

Fluency

Syntactic complexity

Lexical complexity

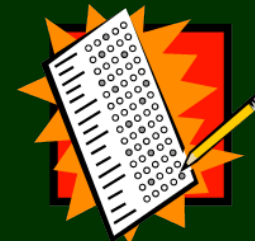
Grammatical accuracy

Evaluation use:
Focus on delivery
and outcomes for
improvement

2: Designing & validating a program

Evaluation Step 3: Formative Feedback

- How do pre-service trainees view key features of the course?
- How do in-service agents view the course?



See Gonzalez-Lloret & Nielson (2014)

Likert-scale and Open-ended Survey

Learner confirmation of effectiveness; course attributes to retain



Pre-service
Learning job through L2
Role-plays
Applicability

In-service
Job-related
Commands, etc.
NS interactions

Learner expression of additional needs; course adjustments to consider



+conversations?
+more role-plays?
+grammar/vocab

+non-job Spanish?
+more role-plays?
+grammar/vocab?



Example 3: Transforming a multi-lang



Contextual Challenges:

- Multi-language department
- Uncertain relation between languages
- Uncertain program value
- Questionable contribution to the institution
- Perceived minimal value by other programs
- Future of the department?

UNIVERSITY OF EVANSVILLE

[UE HOME](#) | [ABOUT UE](#) | [PROSPECTIVE STUDENTS](#) | [ACADEMICS](#) | [ALUMNI & FRIENDS](#) | [PARENTS](#) | [CONTACT UE](#) | [APPLY ONLINE](#)

Department of Foreign Languages

Vision

The Department of Foreign Languages at the University of Evansville will be recognized as one of the top language programs among universities in its category. The department embraces and enhances the University's global vision by

- emphasis on interdisciplinary programming
- internship/study abroad opportunities
- innovative language instruction
- experiential and service learning projects
- a wide range of language offerings
- outstanding faculty members interacting with a culturally diverse student body.

Mission

Foreign languages are an essential component of the great tradition of teaching and learning in the liberal arts. By teaching students to communicate in other languages, allowing them to gain knowledge and an understanding of other cultures, helping them to make connections with other disciplines, providing them with insight into the nature of language and culture and requiring them to participate in multilingual communities both at home and abroad, the programs in the Department of Foreign Languages play a critical role in preparing students for the personal and professional challenges of a multicultural society and a global marketplace.

Intended Uses → Understanding and enhancing program's contribution to student learning, enhancing institutional profile, survival



Example 3: Transforming a program

Step 1: Who are we and what do we do?

Actions:

- Full faculty discussion of learning outcomes
- Assessment committee convened, drafted SLOs
- Student focus groups vetted and suggested revisions
- Revised SLOs approved for majors across all FLs

Student Learning Outcomes

1. Students express themselves in the target languages in a variety of oral and written registers, keeping in mind the communicative context and conventions of the particular culture.
2. Students read and comprehend texts in the target languages tailored to a variety of communicative needs.
3. Students write documents in the target languages tailored to a variety of communicative needs, keeping in mind the conventions of the particular cultures.
4. Students understand native speech.
5. Students demonstrate a familiarity with the current events, the pop culture, and the social structures of the countries/cultures in which the target languages are spoken.
6. Students demonstrate understanding of language variation (social, dialectal, and contextual.)

ETC...

Realization that learning expectations were uneven across languages → came to consensus



Example 3: Transforming a program

Step 2: Where does learning occur

Realization that opportunity to learn was uneven across different FLs → added courses, modules

Actions:

- Reviewed course offerings, syllabi X outcomes
- Identified likely gaps
- Proposed revisions to courses, sequences
- Articulated core offerings across languages

I = Introduce
 R = Reinforced, practiced
 M = Mastery at target level
 A = Assessment collected

Courses, activities, and requirements	BA Major learning outcome 1	Outcome 2	Outcome 3
Lower-level language courses	I		
Upper-level language courses	R	I	
Electives	R	R	I, R
Seminars	M	M	M
Capstone course	M,A	M, A	A
Senior thesis	A	A	A



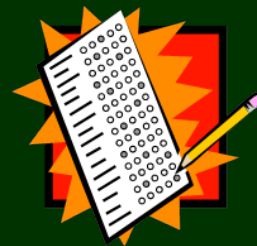
Example 3: Transforming a program

Step 3: What have students learned

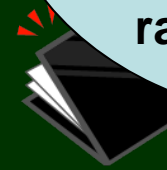
Realization that some outcomes not sufficient → added/adjusted course emphases, raised awareness

Actions:

- Specified uses: program improvement + public demonstration of value
- Identified major constructs: FL proficiency, content knowledge, program impact
- Developed data-collection methods: performance and reflection, objective and subjective
- Pilot-tested, revised, implemented plan



Anonymous survey of students' perceptions about learning



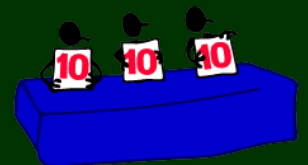
Capstone course



Portfolio presentation in FL to faculty committee



Online reading assessment, CEFR level rating (external indicator)



Committee Q&A, ratings on rubric

research senior essay, reflective narrative



Example 3: Transforming a multi-language program

Step 4: What happened? Follow through and consequences

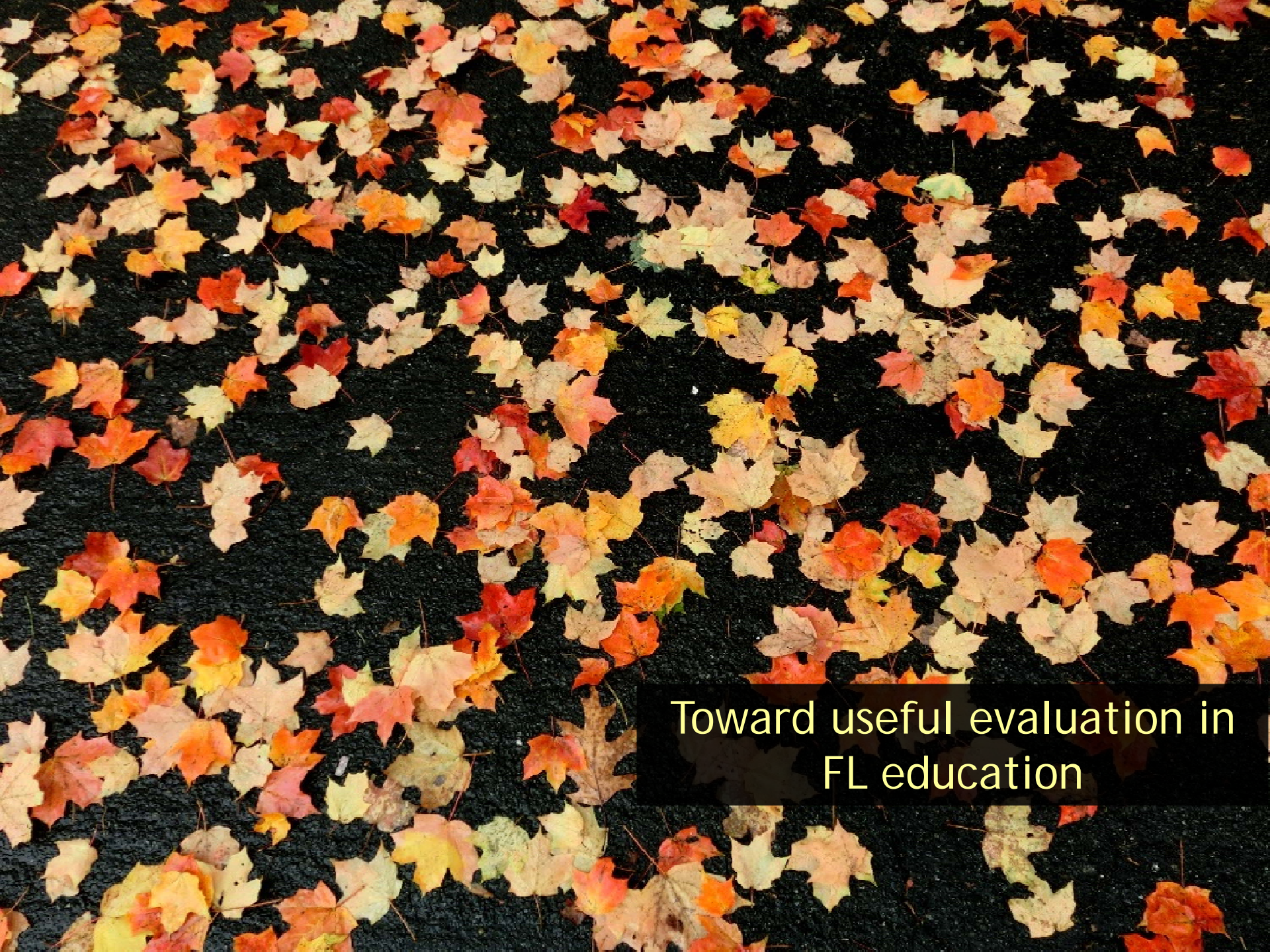
Actions:

- Posted SLOs, assessment plans to the website
- Developed program brochure based on SLOs (used with students and parents)
- Revised curriculum and courses based on findings
- Submitted unsolicited assessment report to dean, faculty senate
- Published assessment work

Positive changes:

- Enrollments increased
- Request for new faculty lines approved
- Chair named to newly formed university assessment committee
- Students' perceptions increasingly positive
- Faculty collegiality improved

See Grau-Sempere, Mohn, & Pieroni (2009)



Toward useful evaluation in
FL education



Useful evaluation: What have we learned?

Received View

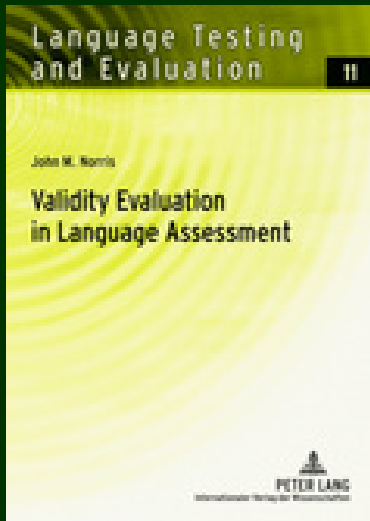
- Generic, one-size-fits all
- Accountability-driven
- Measurement-based
- One-shot judgments
- Problem-identifying
- Imposed, external

Useful Evaluations

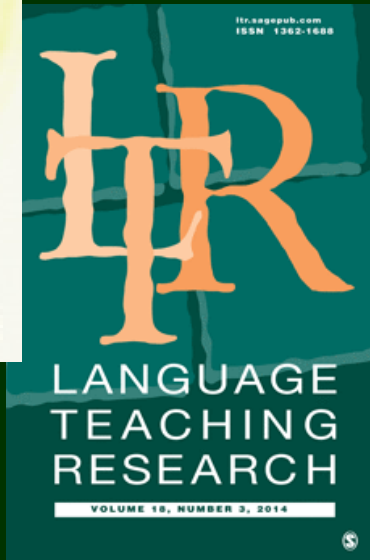
- Contextualized: specific language programs
- Intentional: multiple purposes and uses
- Diversified: methods articulated to uses
- Iterative reform: change takes time
- Problem-solving: improve via understanding
- Engaging, internal: stakeholders (especially educators) take interest in and act upon evaluation



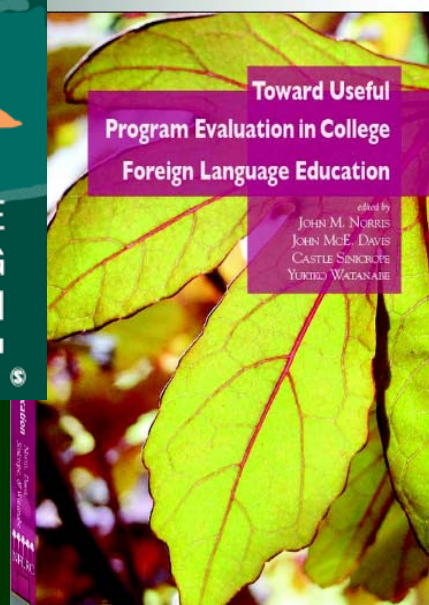
Useful evaluation: Examples in FL practice



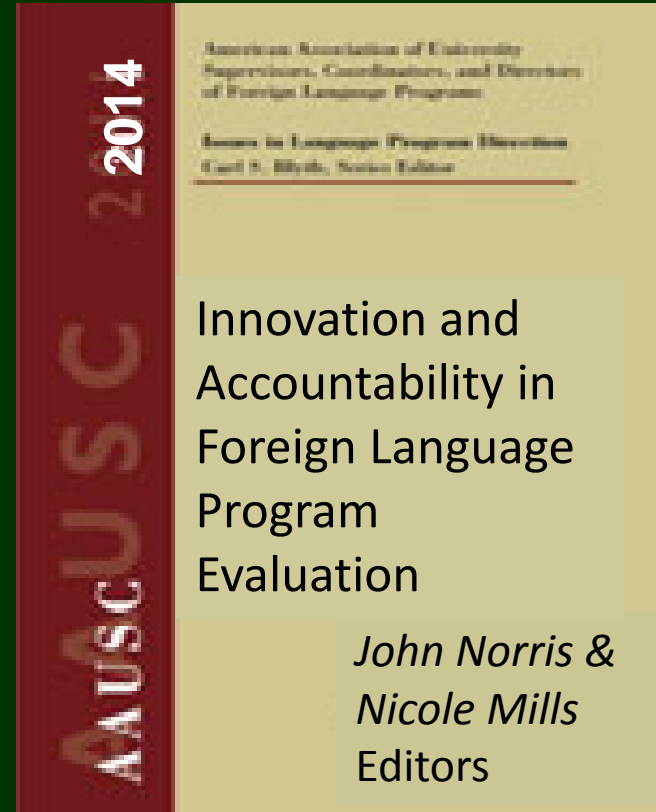
Norris (2008)



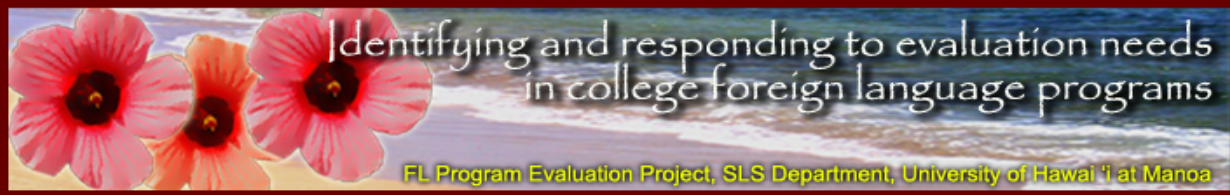
Norris (2009)



Norris et al. (2009)



Norris & Mills (2014)



HOME PROJECT PUBLICATIONS EVENTS RESOURCES EXAMPLES PERSONNEL CONTACT SEARCH

Welcome to the home page of the FL Program Evaluation Project at the University of Hawaii. Our mission is to enable useful evaluation in support of foreign language programs. We invite you to browse the web site, utilize the resources, and contact us with your ideas about program evaluation. This web site is updated frequently, so please check back soon for additional resources and project news.

Announcements

- May 28-June 6, 2007: Residential Summer Institute on FL Program Evaluation in Hawaii! Find out more at: <http://nflrc.hawaii.edu/prodev/si07d>
- June 6-10, 2007: Summer Seminar West of the Association of Departments of Foreign Languages, University of Hawaii, Honolulu; featured sessions on program evaluation and assessment. For more info, see: <http://www.adfl.org>





<http://www.nflrc.hawaii.edu/evaluation>

...more: <http://www.hawaii.edu/sls/lampe>



Association of Departments of Foreign Languages

University of Hawaii, National Foreign Language Resource Center
Summer Institute 2007

What is the value of evaluation in language education?

Provides a
framework for
discussion

Increases
awareness,
communication

Facilitates
solving of
problems

Encourages
heightened
commitment

Sheds light on
how programs
function

Makes student
learning more
efficient

Democratizes,
unifies,
engages...

“enables the field to articulate and demonstrate—internally and externally—the unique contributions of language studies in a pluralist and globalized world.”



Thank you!