Judging Instructionally Sensitive Assessment: Teachers’ and Students’ Perspectives

Kellie Wills¹, Deanna Sands², Min Li¹ & Maria Araceli Ruiz-Primo²

¹University of Washington at Seattle
²University of Colorado Denver

New Orleans, LA.

The work reported herein was supported by National Science Foundation (DRL-0816123). The findings and opinions expressed in the paper do not reflect the position or policies of the National Science Foundation, but the authors.
Instructionally sensitive assessments have been defined based on three dimensions: (1) represent the intended and enacted curriculum (assessment material has been taught); (2) reflect the quality of the enacted curriculum (instruction); and (3) have formative value (teachers must be able to use the assessment information to adjust instruction). This study provides evidence on the first and third characteristics of the definition. It examines: (1) teachers’ and students’ judgments of test items regarding how well the test items represented the intended and enacted curriculum, and (2) teachers’ judgments with respect to the formative values of the test items. The research questions pursued are: What is the agreement between item developers, teachers, and students about the level of instructional sensitivity of the items? What factors influence the proximity judgments according to teachers and students? Do teachers find the information provided by their students’ results useful for planning adjustments to their teaching?

Study Context

Instructionally sensitive assessments have been in the spotlight lately. In defining instructional sensitivity the focus has been on the student performance on the assessments and how it can accurately reflect the quality of the instruction they received (Burstein et al., 1990; McClure as cited by Linn, 1983; Popham, 2007; Schmidt et al., 1983; Yoon & Resnick, 1998). This study described here is part of a larger project focusing on developing and evaluating an approach to constructing instructionally sensitive assessments. In the Developing and Evaluating
Instructionally Sensitive Assessments (DEISA) project, we (self-citation, 2008) have expanded the focus of instructional sensitivity to three characteristics mentioned above.

The DEISA project is based on an approach proposed by Ruiz-Primo, Shavelson, Hamilton, and Klein (2002) that involved two dimensions: a multi-level and a multi-faceted. The former rests on the idea of collecting evidence at different distances from the enactment of a curriculum or a program (See Figure 1): immediate – assessment either in the form of embedded assessment or part of an instructional activity that are exposed to students; close – assessment close to what students experience in the classroom; proximal – assessment focused on the same topic and concepts, but not as close to what students experienced; and, distal assessments developed based on state or national standards.

![Distance of Assessment Diagram](https://via.placeholder.com/150)

Figure 1. Multilevel assessment of student achievement (after Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002).

Multi-faceted stems from a framework in which achievement is conceptualized based on types of knowledge: (1) Declarative knowledge or knowing that – Factual, conceptual knowledge (e.g., What is an organism?). (2) Procedural knowledge or knowing how – Knowledge focusing on the sequence of operations or actions (e.g., Calculate the density of an
object given this X mass and Y volume). (3) Schematic knowledge or knowing why – Knowledge used to reason about, predict, and explain things in nature (e.g., Why do you think larger volume of water leads to more erosion? How do you know that?).

Purpose of the Study

The DEISA project aims at filling a gap in research: how to develop instructionally sensitive assessments rather than judging assessment already developed, the approach used since the 1980’s (Cooley & Leinhardt, 1980; Crocker, Miller, & Franks, 1989; Haladyna & Roid, 1981; Leinhardt, 1983; Muthén, Kao, & Burnstein, 1988; Schmidt et al., 1983). To this end, DEISA proposes three steps: (1) mapping to collect information about the intended curriculum which will help in developing assessments at different distances; (2) developing items by using the information from the map; and (3) validating the items based on judgmental and empirical evidence. Judgmental evidence comes from asking assessment specialists and practitioners with knowledge and experience of the modules (but whose classrooms will not be administered the assessment) to evaluate the proximity of the items. Empirical evidence comes from multiple sources of information at different times of the process: (a) Administering the assessments to all the students before and after instruction to test if the change of item p values correlates with the item proximity; we hypothesize that the gain of student performance on items will decrease as items become more distal. (b) Using a pretest1-pretest2-posttest design which allows comparing the gain from pretest1 to pretest2 as the un-instructed group with the gain from pretest2 to posttest as the instructed group, we assume that differences between the two sets of gains should be largest for close items and least for distal items. (c) Interviewing students and surveying teachers about perceived item proximity based on their classroom experiences. And (d) analyzing classroom videotapes and artifacts to identify if the implementation and quality of the unit explains the differentiated gains of student performance on items varying in proximity.
This paper focuses on the last piece of validity evidence: judgment of the sensitivity of items developed using teachers and students as source of information. We discuss interview and focus group data collected from teachers and students about the item characteristics and from teachers about their formative value. Information from teachers and students helps to understand the connection between the items and instruction and to identify factors (e.g., topics of the science module) that might influence their perceptions about the items.

**Methods**

The DEISA approach is based on varying levels of assessment sensitivity to the enacted curriculum (see Ruiz-Primo et al., 2002) and different facets of achievement conceptualized as type of knowledge. Distances or proximities to the enacted curriculum may account for the spread of differential effects of instruction on student learning. Three research questions guided this specific study within the larger project: (1) What is the agreement between item developers, teachers, and students about the level of instructional sensitivity of the items? (2) What factors influence the proximity judgments by teachers and students? And (3) Do teachers find the information provided by their students’ results useful for planning adjustments to their teaching?

**Participants**

*Item Development and Item Review Teams.* We worked with science coordinators, science coaches, and lead science teachers from participating districts to generate close and proximal multiple-choice items. Items were reviewed and edited by an external review panel formed by content experts, veteran science educators who are familiar with the target modules, and a linguist.
Teachers. A total of nine teachers, eight fifth-grade teachers and one fourth-grade teacher, participated during the initial year of the study. The teachers were from three different school districts in the Denver metro area, one urban, one suburban, and one rural, representing different distributions of ethnicity (33-74% non-white), socio-economic status (free/reduced 20-86%), and linguistic background (19-35%). A different science module was studied in each district. Three teachers each taught Heat and Change from the BSCS, Landforms from the FOSS series, and the Environment module from the FOSS series.

Students. Two hundred and sixty two students, taught by the nine teachers, were administered pre- and post-assessments (136 for Heat and Change, 65 for Landforms, and 61 for the Environments module). Six students randomly selected from each classroom were interviewed on selected items.¹ Twelve students were interviewed for the Heat and Change module (6 males, 6 females).² Eighteen students were interviewed for both the Environments (10 males, 8 females) and Landforms (10 males, 8 females) modules.

Data Collection and Analysis

Item development

The project team shared and discussed with the item developers the documents which explained both the mapping process and the guidelines for item development based on the frameworks in this project. Item developers worked closely with the unit modules in order to develop items. As each item was generated, developers initially indicated not only the proximity of the item based on his/her consideration of the degree to which the item content matched that of the module directly or was more indirect, but also the expected difficulty (i.e., ‘easy’,

---

¹ Students were selected randomly from each class roster after students designated as the English language learners were excluded since the interviews require fluent English.
² Students in one class were not interviewed.
‘medium’ or ‘difficult’). All the items were then reviewed and edited by the research team as well as by an external review panel formed by content experts, veteran science educators who are familiar with the target modules, and a linguist. During the item development and review process, each item was classified regarding its proximity as close or proximal by following the operational rules for item development (for specifics, see Ruiz-Primo et al., 2010).

Distal items were selected from large-scale assessments (i.e., international, national, and state assessments) based on the Colorado state science standards and the National Science Educational Standards on the content and processes of scientific investigations that are relevant to the modules at hand.

Data Sources

Student pre- and post-assessments. Students were administered a pre-assessment at least one week prior to instruction and the same test as a post-assessment within a week after the module was completed. Both assessments were administered by a graduate student following the test administration protocol developed in this DEISA project. These assessments included close and proximal items developed as discussed above, as well as distal items pulled from national and state assessments. Two booklets for each module were created, with approximately 20 common items across the two booklets and a total of 30 items of each booklet. Students in each of the classrooms received either Booklet 1 or 2 by random assignment (same booklet in pre-test and post-test).

Teacher questionnaire. After the post-test was administered to students, teachers were provided with their students’ results at the aggregate level for the total and sub-scale scores in terms of type of knowledge and learning goals of the module, and a written questionnaire. The questionnaire presented a sample of 30 close and proximal items (no distal items were included
on the questionnaire) and asked teachers questions pertaining to the instructional sensitivity as well as item difficulty and other item characteristics for each item. In this paper, we focus on teaches’ responses to one particular question: (1) Did you teach the topic addressed in this item? In addition, for the Landforms teachers, the questionnaire was revised to also include one question more explicitly about item proximity: (2) Do you think this item is close, proximal, or distal?

Teacher interview. An individual interview was conducted with each teacher. The focus of this study is teachers’ reports on the formative value of information gained from the instructionally sensitive assessments. The questions included, “As a result of teaching this module and seeing your students’ assessment data, what have you learned?” “Before teaching this module again, what might you need to know more about?” “After teaching this module and seeing your students’ assessment results, the next time you teach this module…a. what will you continue to do? b. what will you change?”

Teacher focus groups. Each group of teachers who taught the same module was brought together for a half-day focus group interview to discuss their students’ performance. Teachers were provided with a matrix which had all the close and proximal items in the two booklets. For each item, information about the percent of students who selected each option was provided, including the correct option. Teachers were asked to classify, before the meeting, each item as close or proximal based on their instruction of the module. During the focus group meeting, each item was discussed around questions focusing on the characteristics of the items. Specifically, the teachers were asked to explain how they would classify an item as close or proximal and describe the reasons for their judgment.

---

3 For the Landforms module, data from only one teacher interview was available.
Student interviews. To gather evidence about the response processes of students related to the types of knowledge and the proximity and difficulty of the items, within a week after the posttest was administered, a sample of students were asked to talk aloud as they responded to a triad of selected assessment items. The items included close, proximal and distal assessment items involving different types of knowledge. Once the student completed the think aloud process, the researcher posed a series of questions focused on whether or not he or she was taught what was needed to respond to each item (i.e., Did you learn about this topic in your science class? What did you do in your science class when you studied this topic?). We sought student responses on 32 items for the Environment module, 28 items for Heat and Change, and 6 items for Landforms. Over the course of the year, we modified the student interviews to interview more students on a smaller set of items. Thus, for the Landform module student interviews, we interviewed a total of 18 students across only 6 items whereas in the other two modules, fewer students were interviewed with a larger set of items.

Data Analysis

Data collected from various sources were analyzed quantitatively and qualitatively to approach the three research questions. Using teacher questionnaires (for Landforms) and teacher focus group data (for the other modules), we compared the proximity judgments of teachers to the classification of item proximity by item developers and researchers. Using student interview data, we compared students’ reports of whether or not item content was taught to the classification of item proximity by item developers and researchers. We performed logistic regressions using the agreement between judgments as the response variable. We used module and type of knowledge (as judged by researchers) as predictors of whether judgments for an item matched, and considered whether these had significant effects. In addition, we closely examined
the transcripts of student interviews and teacher focus groups to elaborate reasons for identifying the proximity of items. Data on the formative value of information gained from the assessments was gathered from teacher interviews.

The first level analysis of the transcriptions from students’ responses to the open-ended interview questions was approached from a descriptive coding approach (Miles & Huberman, 1984). The researchers reviewed the data and applied words or phrases to capture the topic of respondents’ answers as opposed to the content. Building on this first level of coding, we then began to look for commonalities across the words or phrases that captured the topics which emerged.

**Results**

*Question 1: What is the agreement between the item developers, teachers, and students about the level of instructional sensitivity of the items?*

Using teacher questionnaire, teacher focus group, and student interview data, we calculated the agreements of the judgments of teachers and students in comparison to the classification of item proximity by item developers and researchers. We chose to use the rating as the unit of analysis, rather than the consensus rating among teachers or students for each item, to evaluate the agreement of judgment since in many cases, only one or two students or teachers rated the item.

Table 1 shows the degree of agreement regarding item proximity between item developers/researchers, the teachers, and the students by module, from students’ interviews, questionnaire responses and focus groups. In the first comparison, it is important to remember
that teachers were asked to indicate whether items were close or proximal. Teachers’ judgment was consistent with the classification of researchers and item developers only approximately half the time. In the next section, we use focus group data to explore teachers’ conceptions of the item proximity in comparison to the researchers and item developers.

Table 1. Agreement of Item Proximity between Teachers, Researchers, and Students

<table>
<thead>
<tr>
<th></th>
<th>Teacher vs. Researcher</th>
<th>Student vs. Researcher</th>
<th>Teacher vs. Student</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Agreement</td>
<td>n of Ratings</td>
<td>Agreement</td>
</tr>
<tr>
<td>Environments</td>
<td>66%</td>
<td>93</td>
<td>69%</td>
</tr>
<tr>
<td>Heat &amp; Change</td>
<td>51%</td>
<td>55</td>
<td>59%</td>
</tr>
<tr>
<td>Landforms</td>
<td>45%</td>
<td>60</td>
<td>54%</td>
</tr>
</tbody>
</table>

In the second set of comparisons, it is important to remember that students were not directly asked about the abstract concept of item proximity, instead, they were asked to indicate if the content tested by an item was taught or not. A follow-up question asked them to describe how the content was taught. For close items, we assume that students should recognize that they learned the content; students should report that they did not learn the specific content of an item if it is a proximal item. Generally the students did not yield a higher agreement with the researchers and item developers when compared to their teachers. Taken together, teachers and students appeared to perceive the item proximity in ways differently from researchers and item developers.

The last comparison shows the agreement between teachers and students about whether the item content was taught. These agreements are generally quite high (except for the small number of Environments ratings). For example, only 12% of student responses overall (close and
proximal items) indicated that the content was not taught. We speculated that the high agreement was due to the fact that students and teachers rarely reported that item content was not taught.

**Question 2. What factors influence the proximity judgments by teachers and students?**

2.1 Proximity judgments by teachers

In order to identify the factors that may influence teacher judgment of item proximity, a mixed-methods approach was employed. First, logistic regressions were performed using the judgment of item proximity classified by researchers and perceived by teachers as the response variable. We used module and type of knowledge (as judged by researchers) as predictors for whether teacher and researcher proximity judgments for an item matched. In addition, we closely examined the transcripts of teacher interview and focus groups for teachers’ reasons for their proximity judgments. Table 2 shows the regression results.

For items judged by researchers to be proximal, agreement was less likely than for close items; the odds of agreement for a proximal item is about half that for a close item. In other words, teachers were more likely to indicate that a proximal item was close, than to say that a close item was proximal. The regression results also indicate that there was a greater likelihood of disagreement for the Landforms items compared to the Environments items. However, no statistical difference is found between Heat & Change and Environments. Lastly, the regression also suggests that type of knowledge did not have a significant effect on agreement between teachers and researchers.
We took a closer look at the interview data to explore what factors teachers considered when they evaluated the item proximity. Below we summarized the findings of interview data for disagreements in two categories: items classified as close by researchers and developers but perceived by teachers as proximal, and items classified as proximal by researchers and developers but perceived by teachers as close.

2.1.1 Close items that teachers judged as proximal

It is important to mention that the instruction to teachers was to judge the items based on their instructional experience – the ‘enacted’ curriculum; yet item developers drew these same judgments on the ‘intended’ curriculum. As reflected in the focus group discussions, teachers
judged an item as proximal if some feature of the item differed even slightly from what they had taught in class. If a word used in the items was not used during instruction, they tended to classify the item as proximal. They did not consider as a primary criterion that the word used in the item was indeed used in the curriculum materials. For instance, when an item referenced the word ‘container,’ teachers declared the item proximal because they used the word ‘cup’ in their teaching. Similarly, teachers classified an item as proximal because a data table used in the item included two additional data points from the table used in their instruction.

A side conversation among the two Heat and Change teachers in the focus group involved logistical limitations that may affect item proximity between the intended and enacted instruction. For example, in this unit module there are investigations that involve heating and cooling liquids at different intervals and with differing levels and observing changes over time. Without a refrigerator/freezer in the classroom, this is impractical to carry out and demonstrate; as one teacher said, “Logistically it did not work for me to bring my 30 plus children down the hall to the staff lounge freezer,” and another teacher concurred, “We have a refrigerator across the hall but it was really impractical.”

Teachers also took account of students’ background knowledge and experiences when judging the items. For example, Heat and Change teachers declared some items proximal because the reading load seemed heavy to them. Another example comes from a Landforms teacher. The term ‘beach’ in Colorado generates a very different visual image than for a person from California. In this module, though the term ‘beach’ was in the curriculum, the teacher did not focus on the term for two reasons: (1) because the teacher was from California she associated ‘beach’ with ‘ocean,’ and did not see Colorado as having ‘beaches’, and (2) the use of stream tables in the labs did not easily allow the concept of ‘beach’ to be created. Though the teacher
was being considerate of students’ context on the one hand, since the term appeared in items (and was ‘close’ to the curriculum), student results indicated this could have been a source of confusion.

2.1.2 Proximal items that teachers judged as close

As noted in the regression results, teachers were more likely to say that a proximal item was close, than to say that a close item was proximal. This part of analysis involved mainly the analysis of the interviews of focus groups with the Environments and Landforms teachers. Teacher judgments reflected features of the enacted curriculum. For example, in the focus groups and interviews, Environments teachers frequently mentioned that experiments with plants and animals provided in the kit did not work as the curriculum materials indicated. In describing why one item was considered close, one teacher noted, “This was one of the ones that actually the plants did what they were supposed to do so I think it really clicked with them.” Also, Environments teachers judged one item, involving a controlled experiment, to be close because they drew a connection with a different FOSS module, Variables, which had been taught prior to the Environments module. They saw the experiments in the Environments module as a continuation of what they did during the Variables module.

Teachers sometimes said that an item was close although its context differed from contexts presented in the unit, believing that their students had had so much exposure to the content that it would be easy for students to transfer their learning to this new context. For example, one Environments teacher, speaking of an item involving organisms not used in the unit, said, “I explicitly teach living and non-living so they should be able, any item I give them, tell me if it's living or non-living.” A Landforms teacher said of an item, “You had to know that
the closer together the contour lines were [on a topographic map], the steeper it is, which is exactly what we did in the FOSS kit.” Therefore the item was judged to be close, even though the teacher recognized that ELL students might not know the term “climbing ropes” used in the item.

### 2.2 Students’ reports of what content was taught

In interviews, students were not asked about the “proximity” of items; they were asked instead whether the content of the item was taught in class. Logistic regressions were used to explore factors influencing whether this judgment agreed with that of researchers (i.e., researchers considered item close and student said content was taught). Predictor variables were module, type of knowledge (as judged by researchers), proximity as perceived by researchers, and proximity as perceived by teachers. We closely examined the transcripts of student interviews for students’ descriptions of how item content was taught in class.

Table 3 shows the effects of these predictors. The effect of researcher proximity is very significant: students were much less likely to agree with researchers about proximal items. (As previously mentioned, students seldom reported that content was not taught.) Also, agreement was less likely for Procedural and Schematic items (borderline significant), compared to Declarative items. Variables that did not have a significant effect included student performance on the post-test relative to other students, and whether the student answered the particular item correctly on the post-test or in the interview.
Table 3. Logistic Regression for the Agreement of Student and Researcher Proximity Judgments

(*n = 72 ratings*).

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard error</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.33</td>
<td>1.62</td>
<td>3.29</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Item Characteristics:*

- Proximal (vs. Close)     | -4.20    | 0.88           | -4.78| .000    |

*Type of Knowledge (researcher):*

- Procedural (vs. Declarative) | -3.24    | 1.28           | -2.54| .011    |
- Schematic (vs. Declarative)  | -2.13    | 1.17           | -1.82| .069    |

*Module:*

- Heat & Change (vs. Environments) | -1.36    | 1.12           | -1.21| .227    |
- Landforms (vs. Environments)    | -0.06    | 1.01           | -0.55| .583    |

Data from student interviews shows students recognizing that even though the exact content of the item was not taught, the items were still connected with what they studied in class. For example, several comments from Environments students illustrate transfer of concepts to animals not used in class, e.g. “We didn’t use snakes or anything. We used kinda brine shrimp and we changed the environment to find out what the optimum environment was with the salinity of the water.” Two other students connected a class experiment with beetles and “roly polys” to another item involving snakes.
**Question 3. Do teachers find the information provided by their students’ results useful for planning adjustments to their teaching?**

Regarding the third dimension of sensitive assessment (i.e., having formative value for teachers’ instruction), we wondered whether knowing about their students’ results on the pre- and post-assessments would be useful to teachers in their subsequent teaching of the same module. To answer this research question we analyzed data from the teacher individual interviews and focus groups.

Teachers suggested a variety of adjustments based on the assessment information, including:

*Lesson content.* Teachers easily identified and reflected on the module activities that they had or had not carried out. For example, one Environments teacher reported that she did not include one activity of asking students to write a letter which summarizes the findings from the lab when analyzing her students’ performance. One Landforms teacher realized items on which students did not perform well assessed content from lesson activities that were not included in her instruction. On the other hand, another Landforms teacher noted that she had used the results of her own prior formative assessment to add an additional lesson to the module, and she felt it had potentially contributed to her students’ results on specific items on our assessment.

*Vocabulary/academic language.* Upon seeing their students’ assessment results, Environments teachers talked about focusing more on academic language and making sure that they covered the appropriate vocabulary with students, such as terrestrial. Also, both Landforms teachers began to consider whether or not they had built sufficient background knowledge for their second language learners for particular vocabulary used in the module.
Teaching procedural knowledge. Items involving graphs led Heat and Change teachers to discuss focusing more on graph interpretation by having more conversations about graphs, what they mean, and how to approach the interpretation of different graph types. A Landforms teacher realized that for many of the labs, she focused more on the mathematics of the content and not the procedural aspects of conducting labs. One Environments teacher attributed her students’ success on the items related to procedural knowledge to her inclusion and enhancement of the design of experiments in each lesson. She mentioned that the mapping experience made her realize the lack of relevance she had put into scientific processes in her teaching.

Making learning goals explicit for students. A Heat and Change teacher indicated that the next time this module is taught, the teacher would be more explicit about sharing the module learning goal and lesson learning targets, focusing on stating them in “kid-friendly” language. Heat and Change teachers also wanted to be more intentional about lesson closure – referring students back to the daily learning target and discussing the links between the class activities and how those connected to the targets.

Transfer. Several teachers wanted to focus more on generalization -- helping students broaden what they learned directly in class to other environments, conditions, materials, etc. They suggested making a greater effort to connect concepts across varying experiments, bringing in additional materials such as videos or a wider range of physical materials (sand, dirt, gravel, pebbles, rocks, volcanic dust) to illustrate different concepts. They also discussed linking the concepts taught to real-life applications and helping students think more like ‘real scientists.’

Finally, one teacher reflected on how the assessment results helped her to uncover assumptions she made when teaching. She said she made assumptions about what students probably already knew, and therefore she didn’t spend time developing that understanding, and
yet when analyzing their performance it became clear to her that they had not known what she assumed they had.

**Discussion and Conclusions**

DEISA item developers and researchers determined proximity based on the ‘intended’ curriculum in the published modules. Teachers, on the other hand, judged the items based on their instructional experience – the ‘enacted’ curriculum. This tension between ‘intended’ and ‘enacted’ curriculum raises a critical issue for instructionally sensitive assessments: whose criteria for sensitivity should be used when creating and validating assessments? One important consideration is the audience for the assessment results. The DEISA project’s perspective is essentially district-level – curriculum implementation and standards alignment across multiple classrooms and schools. Our perspective also overlaps with that of curriculum designers, in that we expect our assessments to reflect how well the curriculum highlights the “big ideas” for teachers and students. Given this perspective, basing proximity on the intended curriculum seems appropriate to us.

The current study illustrates several complications in validating instructional sensitivity with teachers and students. Our student interview data clearly shows students making connections between material taught in class and proximal items – they responded that item content was taught in class even when curriculum and item contexts differed. Also, teachers were more likely to consider proximal items close than vice versa, suggesting that they too had transfer in mind. On the other hand, teachers also frequently considered close items to be proximal, suggesting that teacher validation may be very sensitive to idiosyncrasies of instruction, such as the vocabulary they choose to teach.

We feel that additional research is needed on teachers as sources of information about sensitivity. A promising future research direction involves analyzing the proximity judgements of teachers
whose students are not being assessed in the study and are not currently teaching the module. This might reduce sensitivity to idiosyncracies of instruction. Also, these teachers would not have a ‘stake’ in the assessment results, and so might be less motivated by justification, e.g. calling an item proximal if they were unsatisfied with the way they taught it, or their students performed poorly.

References


