Inside the Black Box of the Classroom

What we are learning about teacher effectiveness from classroom observations in LAC

Barbara Bruns  
Lead Education Economist  
Latin America and the Caribbean Region  
Human Development Network  
The World Bank  
HD Learning Week, February 7, 2013
Education production functions usually estimate a “reduced version”

- Teacher Content knowledge
- Teaching skill and motivation
- Teacher Classroom practice
- Student characteristics
- School characteristics
- Student Learning Outcomes
Classroom observation allows us to measure perhaps *the* key output of the education system.

Teacher Content knowledge

Teaching skill and motivation

Teacher classroom practice

Student characteristics

School characteristics

Student Learning Outcomes
Theory of action: teacher incentives

Any results *caused by* an incentive must come through changes in teacher effort and/or effectiveness in classroom.

...but we lack data on the **key output**
Classroom observations using the "Stallings method"
Classroom observation – alternative tools, depending on purpose

- **Large-scale research application/system diagnosis**
  - Robust data that is representative at the system level
  - Use trained observers in the classroom – single visit to each classroom but multiple classrooms in each school
  - Or video footage analyzed later
  - **Best methods:** Stallings Classroom snapshot, ETP method, SSOS (special strategies observation system), VTBI (Vergilio), COKER, STARS (Spaulding), OCEPT *(Jukes study is a good resource)*
    
    Why? High inter-rater reliability with limited training; published comparators available

- **Individual teacher performance evaluation**
  - Observe and analyze individual teachers’ best practice/formative feedback
  - Used trained observers in the classroom - but multiple visits
  - Or video footage analyzed later
  - **Best methods:** Cincinnati TES, DC Impact, Chile SED, CLASS *(Gates MET study website is a good resource and Kane et al 2009)*
    
    Why? More comprehensive, capture content mastery and quality of curriculum delivery, pre-class preparation, other school contributions
Stallings “classroom snapshot”

What does it measure?
★ Teacher’s use of instructional time
★ Use of materials, including ICT
★ Core pedagogical practices
★ Ability to keep students engaged

Advantages:
★ low inference, high inter-rater reliability (greater than .7)
★ metrics not affected by curriculum, therefore comparable across countries
★ generates robust quantitative data on classroom dynamics

Research protocol:
★ Randomly selected, representative sample of schools (typically 30% of schools in system)
★ Randomly selected classrooms within schools (among targeted grades and subjects)
★ No warning to schools of date and time of observations
★ Teacher anonymity
Caveats

- But...still need careful interpretation of results
  - Hawthorne effects
  - No data yet on noisiness of multiple observations of same teacher
- Non-random assignment of teachers to classes
  - Ability tracked classes
  - Is it a better teacher or better students producing better dynamics?
- Instrument does not capture content
  - Curriculum content
  - Teacher content mastery

“Things still look good, but I’m recommending caution.”
# Stallings’ “Good Practice” Benchmarks

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Stallings Benchmarks for Effective Time Use (from US research)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instruction</strong></td>
<td></td>
</tr>
<tr>
<td>Active Instruction</td>
<td>85%</td>
</tr>
<tr>
<td>Reading</td>
<td>50% or more</td>
</tr>
<tr>
<td>Instruction/Explanation</td>
<td></td>
</tr>
<tr>
<td>Drill and Practice</td>
<td></td>
</tr>
<tr>
<td>Kinesthetic</td>
<td></td>
</tr>
<tr>
<td>Projects</td>
<td></td>
</tr>
<tr>
<td>Passive Instruction</td>
<td>35% or less</td>
</tr>
<tr>
<td>Monitoring Copying</td>
<td></td>
</tr>
<tr>
<td>Monitoring Seatwork</td>
<td></td>
</tr>
<tr>
<td><strong>Classroom Management</strong></td>
<td>15% or less</td>
</tr>
<tr>
<td>Giving Assignments</td>
<td></td>
</tr>
<tr>
<td>Managing with Students</td>
<td></td>
</tr>
<tr>
<td>Disciplining Students</td>
<td></td>
</tr>
<tr>
<td>Managing Alone</td>
<td></td>
</tr>
<tr>
<td><strong>Off-Task</strong></td>
<td></td>
</tr>
<tr>
<td>Socializing with Students</td>
<td>0%</td>
</tr>
<tr>
<td>Socializing with Others</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Students</th>
<th>Off-Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socializing</td>
<td>6% or less</td>
</tr>
<tr>
<td>Uninvolved</td>
<td></td>
</tr>
<tr>
<td>Acting out</td>
<td></td>
</tr>
</tbody>
</table>

Source: Stallings and Knight (2003)
How does it work?

- 10 observations of each class taken at regular intervals and coded using a standardized grid
# Observation Coding Grid

### Activity

<table>
<thead>
<tr>
<th></th>
<th>No Material</th>
<th>Textbook</th>
<th>Notebook</th>
<th>Black Board</th>
<th>Learning Aides</th>
<th>ICT</th>
<th>Cooperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reading Aloud</td>
<td>T 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>S L</td>
</tr>
<tr>
<td></td>
<td>CHECK IF CHORAL READING</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Demonstration/Lecture</td>
<td>T 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>S L</td>
</tr>
<tr>
<td>3. Discussion/Questions and Answers</td>
<td>T 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>S L</td>
</tr>
<tr>
<td>4. Practice &amp; Drill</td>
<td>T 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>S L</td>
</tr>
<tr>
<td>5. Assignment/Class Work</td>
<td>T 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>S L</td>
</tr>
<tr>
<td>6. Copying</td>
<td>T 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>S L</td>
</tr>
<tr>
<td>7. Verbal Instruction</td>
<td>T 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E 1 S L E</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>I 1 S L</td>
<td>S L</td>
</tr>
<tr>
<td>8. Social Interaction</td>
<td>T 1 S L E</td>
<td>I 1 S L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Student(s) Uninvolved</td>
<td>I 1 S L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Discipline</td>
<td>T 1 S L E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Classroom Management</td>
<td>T 1 S L E</td>
<td>I 1 S L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Classroom Management Alone</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Teacher Social Interaction or Teacher Uninvolved</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Teacher Out of the Room</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### FIGURE 1: EXCERPT FROM CLASSROOM SNAPSHOT (APPENDIX 2)

#### CLASSROOM OBSERVATION SNAPSHOT

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>MATERIAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NO MATERIAL</td>
</tr>
<tr>
<td>1. READING ALOUD</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>1 S L E</td>
</tr>
<tr>
<td>I</td>
<td>1 S L</td>
</tr>
</tbody>
</table>

CHECK IF CHORAL READING

- **T line:** Indicates activities that involve the teacher
- **I line:** Indicates activities that involve the student & not the teacher
- **1, S, L, E:** Indicate one individual, a small, large group and entire class respectively
<table>
<thead>
<tr>
<th>Dia da Observação</th>
<th>Nome da escola:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Código da escola (Censo MEC)</td>
<td>Nome do observador</td>
</tr>
</tbody>
</table>

**Disciplina Observada**
- [ ] Língua Portuguesa
- [ ] Matemática
- [ ] Ciências
- [ ] História
- [ ] Geografia
- [ ] Outra disciplina/Especifique

**Nível de Ensino**
- [ ] Fundamental
- [ ] Médio

**Turno**
- [ ] Manhã
- [ ] Tarde
- [ ] Noite

---

**Marca apenas uma opção**

**Ano ou Série observada:**
- [ ] 1º
- [ ] 2º
- [ ] 3º
- [ ] 4º
- [ ] 5º
- [ ] 6º
- [ ] 7º
- [ ] 8º
- [ ] 9º
- [ ] 1º Série

---

**Marca aqui se o professor observado está substituindo o professor regular da disciplina**

---

**A que horas a aula (deveria começar/terminar)?**
- [ ] :
- [ ] :

**A que horas a aula realmente (começou/terminou)?**
- [ ] :
- [ ] :

---

**Quantos estudantes estão presentes em sala?**

<table>
<thead>
<tr>
<th></th>
<th>Meninas</th>
<th>Meninos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aula</td>
<td>:</td>
<td>:</td>
</tr>
</tbody>
</table>

---

**Esta é uma aula gabinetada?**
- [ ] Sim
- [ ] Não
- [ ] Não Aplicável

**Há trabalhos de estudantes expostos na sala?**
- [ ] Sim
- [ ] Não
- [ ] Não Aplicável

---

**Há pichações nas paredes da sala (Não inclui grafite)?**
- [ ] Sim
- [ ] Não
- [ ] Não Aplicável

---

**Quantos estudantes possuíam o livro texto e/ou caderno de atividades em mãos?**
- [ ] Toda a turma
- [ ] Mais da metade
- [ ] Menos da metade
- [ ] Ninguém
- [ ] Não Aplicável

---

**Quantos estudantes possuíam o caderno, lápis, caneta, ou material para se escrever?**
- [ ] Toda a turma
- [ ] Mais da metade
- [ ] Menos da metade
- [ ] Ninguém
- [ ] Não Aplicável

---

**Quantos estudantes usavam fardamento?**
- [ ] Toda a turma
- [ ] Mais da metade
- [ ] Menos da metade
- [ ] Ninguém
- [ ] Não Aplicável

---

**Durante a aula, ocorreu alguma das seguintes situações?**
- [ ] O professor mencionou uma avaliação externa como SAEPE, Prova Brasil ou SAEB
- [ ] Os alunos fizeram uma atividade em que estavam explicitamente sendo avaliados
- [ ] Os alunos fizeram exercícios de preparação para uma avaliação, como um teste simulado, por exemplo

---

**Durante a aula, houve algum tipo de ação física, verbal, ameaça ou intimidação?**
- [ ] Do professor para o aluno
- [ ] De aluno(s) para aluno(s)
- [ ] De um aluno para professor
- [ ] De uma pessoa externa para alunos e/ou professor
- [ ] De mais de um aluno para o professor
- [ ] Outro, Especifique: ____________________________

---

**Comentários:**

---

**Banco Mundial LAC**

**oportunidades para todos**
Stallings Sample 2009-2012
(new requests from Argentina (BA province and Paraguay)
<table>
<thead>
<tr>
<th>Country</th>
<th>Region</th>
<th>Year</th>
<th>Grades</th>
<th>Schools</th>
<th>Classrooms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>Distrito Federal</td>
<td>2011</td>
<td>1-9</td>
<td>201</td>
<td>2335</td>
</tr>
<tr>
<td>Honduras</td>
<td>National</td>
<td>2011</td>
<td>3,6,9</td>
<td>153</td>
<td>754</td>
</tr>
<tr>
<td>Jamaica</td>
<td>National</td>
<td>2011</td>
<td>4,5,9,10</td>
<td>201</td>
<td>1,000</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>Pilot</td>
<td>2012</td>
<td>1-8</td>
<td>10</td>
<td>51</td>
</tr>
<tr>
<td>Peru</td>
<td>Callao</td>
<td>2011</td>
<td>2</td>
<td>65</td>
<td>313</td>
</tr>
<tr>
<td></td>
<td>San Martin</td>
<td>2011</td>
<td>2</td>
<td>67</td>
<td>322</td>
</tr>
<tr>
<td>Brazil</td>
<td>Minas Gerais</td>
<td>2009</td>
<td>5,9,11</td>
<td>75</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2011</td>
<td>5,9,11</td>
<td>604</td>
<td>3,020</td>
</tr>
<tr>
<td></td>
<td>Pernambuco</td>
<td>2009</td>
<td>2-5,8,9</td>
<td>220</td>
<td>1100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2010</td>
<td>2-5,8,9</td>
<td>301</td>
<td>1,505</td>
</tr>
<tr>
<td></td>
<td>Rio de Janeiro (municipality)</td>
<td>2010</td>
<td>2-5,8,9</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2011</td>
<td>2-5,8,9</td>
<td>100</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>Rio de Janeiro Pilot (state)</td>
<td>2012</td>
<td>5,9,11</td>
<td>60</td>
<td>300</td>
</tr>
<tr>
<td>Colombia</td>
<td>National</td>
<td>2011</td>
<td>5,9,11</td>
<td>200</td>
<td>1091</td>
</tr>
<tr>
<td></td>
<td>Antioquia</td>
<td>2011</td>
<td>5,9,11</td>
<td>84</td>
<td>995</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td>2441</td>
<td>14,011</td>
</tr>
</tbody>
</table>

Source: Authors
Stallings classroom observations

Results to date from Latin America
#1 LAC countries are far from the Stallings Good Practice Indicator (85% of time on instruction)

Average time on instruction in LAC countries

- Mexico: 52% Academic Activities, 39% Classroom Management, 9% Teacher Off-Task
- Peru: 55% Academic Activities, 32% Classroom Management, 13% Teacher Off-Task
- Jamaica: 61% Academic Activities, 28% Classroom Management, 11% Teacher Off-Task
- Rio de Janeiro: 61% Academic Activities, 30% Classroom Management, 9% Teacher Off-Task
- Brazil: 64% Academic Activities, 27% Classroom Management, 10% Teacher Off-Task
- Minas Gerais: 65% Academic Activities, 26% Classroom Management, 9% Teacher Off-Task
- Pernambuco: 65% Academic Activities, 22% Classroom Management, 13% Teacher Off-Task
- Honduras: 64% Academic Activities, 24% Classroom Management, 12% Teacher Off-Task
- Colombia: 65% Academic Activities, 25% Classroom Management, 9% Teacher Off-Task
- Stallings Good Practice Indicator: 85% Academic Activities, 0% Classroom Management, 15% Teacher Off-Task
#2 Instructional time matters for learning

Test scores positively correlated with time on instruction in Minas Gerais and Rio de Janeiro

Minas Gerais, 2011

\[ r^2 = 0.0325 \]

Rio de Janeiro, 2011

\[ r^2 = 0.0239 \]
#2 Instructional time matters for learning

Negative correlation between time off task and student achievement

**Rio de Janeiro, 2011**

**Mexico (DF), 2012**
#3 Teachers have difficulty engaging the entire class

Teacher time on instruction, and with entire class engaged

<table>
<thead>
<tr>
<th>Country</th>
<th>Mexico</th>
<th>Peru</th>
<th>Jamaica</th>
<th>Honduras</th>
<th>Brazil</th>
<th>Minas Gerais</th>
<th>Pernambuco</th>
<th>Rio de Janeiro</th>
<th>Colombia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of Total Class Time</td>
<td>52%</td>
<td>55%</td>
<td>61%</td>
<td>64%</td>
<td>64%</td>
<td>65%</td>
<td>65%</td>
<td>61%</td>
<td>65%</td>
</tr>
<tr>
<td>Time on Instruction</td>
<td>21%</td>
<td>14%</td>
<td>21%</td>
<td>19%</td>
<td>23%</td>
<td>16%</td>
<td>31%</td>
<td>23%</td>
<td>19%</td>
</tr>
<tr>
<td>Time on Instruction with Entire Class Engaged</td>
<td>31%</td>
<td>23%</td>
<td>19%</td>
<td>16%</td>
<td>16%</td>
<td>16%</td>
<td>19%</td>
<td>23%</td>
<td>19%</td>
</tr>
</tbody>
</table>
#4 Students are unengaged

Share of Total Class time at least some students are not engaged

Percent of Total Class Time

- Peru: 52% (17% not engaged)
- Brazil: 56% (26% not engaged)
- Pernambuco: 50% (21% not engaged)
- Rio de Janeiro: 56% (29% not engaged)
- Minas Gerais: 63% (28% not engaged)
- Mexico DF: 61% (23% not engaged)
- Jamaica: 63% (20% not engaged)
- Honduras: 65% (22% not engaged)
- Colombia: 66% (26% not engaged)

Share of class time with some students not engaged
Share of class time with large group (6 or more students) not engaged
#5 Large variance in average practice across schools
#6 Large variance in instructional practice within schools

**Within-school** variance two-thirds as large as variance across entire system

Within median school (61%)
- 70% of time on instruction in best classroom
- 52% in worst classroom

= 36 days/year more instruction in best classroom(s)
#6 Large variance in instructional practice *within* schools

message: good schools are better because their teachers perform more consistently

Colombia, average variation in time on instruction from best to worst classroom *within* schools
The World Bank and State Secretariat of Education are trying to control teacher practices inside the classroom.

"Stopwatches can be used to evaluate the teachers"
Reactions in Mexico, 2012

“With Stallings...we will treat schools like factories”

“...it is an ‘inquisition’ of educators”
Atenção: Orientação do Sepe para os professores da rede estadual sobre o início do trabalho dos avaliadores nesta segunda (dia 12/11)

A direção do Sepe, tendo em vista o início previsto para a próxima segunda-feira (dia 22/11) do trabalho dos supervisores que o projeto SEEDUC/Banco Mundial vai implementar primeiramente em 100 escolas da rede estadual, orienta os professores destas unidades para que procedam da seguinte maneira:

- em primeiro lugar, nós, os professores, somos os "Regentes de Turma". Ou seja, os profissionais são as autoridade máxima em sala de aula.

- também não existe qualquer documento oficial que tenha instituído a obrigatoriedade da presença desses "supervisores" nas salas de aula.

- nesse sentido, nenhum professor poderá ser obrigado a aceitar qualquer interferência desses avaliadores em sua sala de aula.

- o Jurídico do Sepe está dando tratamento e acompanhamento a essa questão.
4. Policymakers’ reactions and next steps
A unique source of quantitative, system-wide feedback for managers

“No one is going to stop me from going into my schools to figure out how to improve their quality”
Governor Eduardo Campos, Pernambuco

“We know much of this intuitively, but this is the first time we have hard data we can take back to our regional offices and schools to discuss and work on.”
Secretary of Education, Minas Gerais

“This is one of the most useful collaborations we have ever had with the World Bank.”
Permanent Secretary of Education, Jamaica
A resource for teacher training

Jamaica

Rio Municipality

“Before the Stallings observations, I didn’t know where my best teachers were! Now we can go back to their classrooms and film their practice to use in our training courses. This will be something completely new.”

--Claudia Costin, Secretary of Education Rio Municipality

Dr. Barbara Bruns
Lead Education Economist,
Latin America and the Caribbean
The World Bank
1818 H Street, NW
Washington, DC 20433
USA

Dear Dr. Bruns:

Re: Requesting Permission to Use Stallings’s Observation Method

Having participated in the Stallings’s Classroom Observation Methodology training in Jamaica earlier this year, we believe that it will add value to our inspection process by providing a more quantitative measure of time management in the classroom. We will therefore train our Lead Inspectors to gather data using this method. Your permission is being sought to:

i) Replicate and use the Stallings’s procedural manual and coding sheet;
ii) Incorporate the expertise of Mr. Javier Luque to address the NEI’s trainees regarding the results of the use of the methodology in Jamaica and the other countries that have benefitted; and
iii) Allow Mr. Giorgio Valentini to observe the training session.

We look forward to your prompt and favourable response as we seek to collaborate in the interests of education in our countries.

Sincerest regards,

[Signature]

Chief Inspector of Education
Thank you!
Gracias!
Obrigada!

bbruns@worldbank.org