Academic writing performance measured for research and instruction

Olga Firsova, Open University of the Netherlands, January 19, 2012
As an introduction

My questions:
What makes academic writing* difficult?
What makes academic writing an interesting research object?
What makes academic writing a difficult research object?

Looking for answers while
.. teaching post-graduate level distance learners (OU)
.. designing instruction in competence-based curricula (OU)
.. doing research of academic writing & writing assessment

* Advanced (graduate & post-graduate) level
Academic writing requires ..

.. a solid knowledge base
  discipline & domain
genre(s)
general & domain-specific conventions

.. advanced skills
  language(s)
  argumentation
  use of sources
  critical & reflective thinking
  clear reasoning

.. attitude
  accuracy
  ethics
Academic writing ..

.. takes years to master [deliberate practice, learning by doing]
.. is an object of learning, instruction, feedback & assessment
.. is product & reader-oriented
.. is a tool for communicating with the academic professional community

=> .. is a tool for mastering the discipline & demonstrating competence growth
Academic writing in research…

- Writing strategies & productivity
- Writing approaches & writing quality
- Writing beliefs & writing quality
- Strategy instruction
- (Peer) feedback & writing
- Discipline-based academic writing
- …

=> How do student writings demonstrate the attained competence level [=integration of general, domain & discipline based knowledge, skills, attitudes]?
Assessment of academic writing..

.. takes place both in professional & instructional environments
.. follows criteria that may be are known yet difficult to master
.. varies in goals, forms and results
... used for learning?!

• Holistic assessment: a single score based on an overall impression
• Analytic (multi-trait): traits scored separately
• Goals, context & task complexity level determine the choice
Quality issues

Issues:
- reliability
- judges biases
- construct validity
- predictive validity
- generalizability ...
- practicality
- authenticity

Quality ensured by:
- scales
- rubrics
- scoring scripts
- benchmarks
Scale & rubrics – 1 (Weigle, 2002, p.120)

4 An essay at this level
- addresses the writing topic adequately but may slight parts of the task
- is adequately organized and developed
- uses some details to support a thesis or illustrate an idea
- demonstrates adequate but possibly inconsistent facility with syntax and usage
- may contain some errors that occasionally obscure meaning

3 An essay at this level may reveal one or more of the following weaknesses:
- inadequate organization or development
- inappropriate or insufficient details to support or illustrate generalizations
- a noticeably inappropriate choice of words or word forms
- an accumulation of errors in sentence structure and/or usage

2 An essay at this level is seriously flawed by one or more of the following weaknesses:
- serious disorganization or underdevelopment
- little or no detail, or irrelevant specifics
- serious and frequent errors in sentence structure or usage
- serious problems with focus
### MICHIGAN WRITING ASSESSMENT SCORING GUIDE

**English Composition Board: Criteria for Reading the Assessment**

<table>
<thead>
<tr>
<th>Ideas and Arguments</th>
<th>Rhetorical Features</th>
<th>Language Control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6</strong> The essay deals with the issues centrally and fully. The position is clear, and strongly and substantially argued. The complexity of the issues is treated seriously and the viewpoints of other people are taken into account very well.</td>
<td>The essay has rhetorical control at the highest level, showing unity and subtle management. Ideas are balanced with support and the whole essay shows strong control of organization appropriate to the content. Textual elements are well connected through logical or linguistic transitions and there is no repetition or redundancy.</td>
<td>The essay has excellent language control with elegance of diction and style. Grammatical structures and vocabulary are well-chosen to express the ideas and to carry out the intentions.</td>
</tr>
<tr>
<td><strong>5</strong> The essay deals with the issues well. The position is clear and substantial arguments are presented. The complexity of the issues or other viewpoints on them have been taken into account.</td>
<td>The essay shows strong rhetorical control and is well managed. Ideas are generally balanced with support and the whole essay shows good control of organization appropriate to the content. Textual elements are generally well</td>
<td>The essay has strong language control and reads smoothly. Grammatical structures and vocabulary are generally well-chosen to express the ideas and to carry out the intentions.</td>
</tr>
</tbody>
</table>
SOLO for rubric construction-1

Structure of Observed Learning Outcomes (SOLO) (Biggs & Collis, 1982, Biggs, 1992)

.. operationalizes learning outcomes in terms of structural organisation of knowledge through development levels
The SOLO Taxonomy with sample verbs indicating levels of understanding
SOLO for academic writing (general)

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unistructural: <em>seeing the trees but not the wood,</em></td>
<td>Bits and pieces of unconnected information, listing of facts, little if any understanding and processing of information demonstrated</td>
</tr>
<tr>
<td>Multistructural: understanding but no integration</td>
<td>Assignment-driven writing with information reproduced including both relevant and superfluous material, abundant listing, data-driven text organisation, vague conclusions, unclear underpinning, incomplete argument</td>
</tr>
<tr>
<td>Relational: a coherent whole</td>
<td>Problem-driven writing, problem as a backbone, linear text organisation, underpinning of conclusions and argument, balanced structure, listing if any used functionally</td>
</tr>
<tr>
<td>Extended abstract: affords transfer</td>
<td>A sound chain of reasoning, valid arguments, clear relations, coherent structure, explicit structural cues for the reader. Making interdisciplinary connections.</td>
</tr>
</tbody>
</table>
SOLO for rubric construction-2

.. is context free and used in education and in research studies in different domains (mathematics, statistics, communication, biology, writing)

.. requires & affords specificity (domain, & discipline, scope)

.. can be used to define and describe complex constructs and constituent traits in the same reference frame
Exploring & analyzing (SOLO-based) complex latent constructs: Rasch model

.. to explore the construct and constituent items
.. to produce an interval measure
.. to validate the construct

Rasch model (Bond & Fox, 2007; Linacre, 2011)
Rasch measurement model

- views items and persons in the same frame of reference (item ‘difficulty’ vs ‘persons’ ability)
- tests unidimensionality (tests & demonstrates if …)
- supports calibration of items (scale construction) and persons, positioning both on the same scale
- affords anchoring of new data
- is invariant (robust): produces test free and sample free measures
- data must fit the model instead of the other way around
- provides visual support of data analyses and a variety of diagnosing tools.
Exploring & analysing writing assessment results: feedback from the Rasch model

Corpus
Course-based writing assignments (reviews, ca 4000 ww x 3)
Distance education Masters’ in Educ. Science
n=258

Materials
An authentic academic writing task (conference review)
10 course-based criteria: integrating discipline & general academic writing aspects
A SOLO-based rubric

Procedure
Authentic assignments scores used
<table>
<thead>
<tr>
<th>PERSON</th>
<th>258 INPUT</th>
<th>258 MEASURED</th>
<th>INFIT</th>
<th>OUTFIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>COUNT</td>
<td>MEASURE</td>
<td>REALSE</td>
<td>IMNSQ</td>
</tr>
<tr>
<td>MEAN</td>
<td>36.1</td>
<td>10.0</td>
<td>1.62</td>
<td>.59</td>
</tr>
<tr>
<td>S.D.</td>
<td>6.0</td>
<td>.1</td>
<td>1.70</td>
<td>.15</td>
</tr>
<tr>
<td>REAL RMSE</td>
<td>.61</td>
<td>TRUE SD</td>
<td>1.59</td>
<td>SEPARATION</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ITEM</th>
<th>10 INPUT</th>
<th>10 MEASURED</th>
<th>INFIT</th>
<th>OUTFIT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>COUNT</td>
<td>MEASURE</td>
<td>REALSE</td>
<td>IMNSQ</td>
</tr>
<tr>
<td>MEAN</td>
<td>931.2</td>
<td>257.9</td>
<td>.00</td>
<td>.11</td>
</tr>
<tr>
<td>S.D.</td>
<td>35.8</td>
<td>.3</td>
<td>.40</td>
<td>.00</td>
</tr>
<tr>
<td>REAL RMSE</td>
<td>.11</td>
<td>TRUE SD</td>
<td>.38</td>
<td>SEPARATION</td>
</tr>
</tbody>
</table>

```
VAR1_1
VAR2_1 VAR3_1
VAR1_2
VAR3_1 VAR4_1
VAR1_3
VAR3_1
VAR1_4
VAR3_1
VAR1_5
VAR3_1
```
• thresholds
Feedback from Rasch model analysis on

.. the validity of the construct
.. possible gaps
.. the scope of ability of the population on this construct
.. traits or items that (might) need improvement
Implications

.. the validity of the construct
.. possible gaps
.. the scope of ability of the population on this construct
.. traits or items that (might) need improvement