Background, Methods, and Results of a 7-year Longitudinal Assessment of Undergraduate Business Writing

Drexel University’s Fourth Annual Conference on Assessment September 14, 2017

Scott Warnock
Associate Professor of English
Director of the University Writing Program
Drexel University

Plan for today

• Provide institution background
• Describe assessment methods and process
• Discuss data from 3,500+ assessment records from 2007-13
• Describe two curricular interventions
  – Some good stories
• Analyze disparities between assessors with different backgrounds

Institutional context

• Drexel University
  • Private, urban, co-op, 25,000 students
• LeBow College of Business and AACSB
  • Driven by accreditation, in 2005 LeBow identified “writing skills” as one of four “learning goals”
  • Problem solving, knowledge of economics, career learning
  • LeBow developed assessment methods for each
  • Faculty recognized “the critical importance of written communication in business today” (Linnehan)

Project design/methods

• Pilot in 2006 (WRAB I & II, Warnock JBTC 2009)
• Each summer from 2007 to 2013 (WRAB III; Warnock et al. JBTC 2017)
  • 30 assessors (paid $100* honorarium by LeBow)
  • 15 industry background, 15 English background
  • Recruitment
• 300 documents from LeBow courses
  • Gathered by LeBow
  • Scrubbed of identifiers and coded
  • Distributed to assessors using array generated by statistician
• Training
  • No norming
  • One-hour training session
  • Initially f2f option, then just video

Waypoint Outcomes

• Web-based rubric-creation software
• Assessors used LeBow rubric
• Assessors’ responses recorded
  • Qualitative: Written comments managed as well
• Data managed through Waypoint Outcomes/Access
  – Statistician Nicholas Rouse was excellent

Project design/methods

• All communications/interactions electronic
  • Student—research fellow or research co-op—distributed documents to assessors
• Waypoint Outcomes used for assessments and to aggregate date
• Assessors have six-week deadline
LeBow Writing Rubric

2 Binary Trait-Decision
(Yes/No)
1. Gross mistakes
2. Ethics

8 Scaled Trait-Decision
(Excellent-Good-Fair-Poor-Unacceptable, defined in-context)
3. Purpose/main point
4. Audience
5. Organization
6. Evidence
7. Sentence style: Flow of writing
8. Correctness: Grammar and writing measures
9. Document design/appearance
10. Visuals (if applicable)

Data challenges: Data scrubbing and crunching
• Change in operations in 2010
  • Assessors “category” changed in Waypoint
  • Lack of consistency between pre-2010 and post-2010 data
    • Difficult to combine into data base
• Solutions
  • Clear “game plan” from the beginning!
  • Know a statistician!

Theoretical approach
• Distributive evaluation
  • If you have a large enough group of readers who can, in the aggregate, “make a useful, representative statement about the way writing will be received by audience members for whom it may be important” (Warnock)
  • “…reliability and validity are troublesome criteria, inadequate and too limited for the distinctive task of evaluating writing” (Whithaus)
  • Readers’ instinctive, visceral responses can be mined as an asset, not necessarily a problem to “solved” through calibration
  • Orthopedic pain scale

Overall results from 2007-2013
• 1,988 student documents
• 3,706 total assessments
• 2009 and 2010: Curricular intervention for first-year business course
  • As reflected in 2009 & 2010 summer assessments
• 2011 and 2012: Curricular intervention about evidence
  • As reflected in 2012 & 2013 summer assessments
• English vs. business assessors
Mean Element Score by Year

2007-2013 mean Element scores

• For the assessment, 0.75 is “good”
• Closely grouped in most cases from year to year
• Evidence always low
• Ethical Issues and Gross Mistakes close to the 0.85

Big question:
If you are the LeBow faculty, what do you do with this data about your student writers?
Experiment/intervention 2009-10

- Plan for “treatment” group
  - Treatment sections capped at 20 (vs. 30)
  - Teaching circles
  - Similar writing assignments as controls, but embedded in a recursive sequence
  - Time reallocated from content to peer review
  - Faculty were available to review drafts
  - Students not aware they were in an “experiment” until very late in the process, if then

Curricular intervention experiment in BUSN 101

BUSNC versus BUSNT: By year

What happened?: There’s a story...

Curricular intervention: Evidence

- Evidence annually one of the lowest Elements
- Focused WAC-type work with ORGB and BUSN in fall 2011 and winter 2012
  - Workshops with faculty
  - Focused in-class citation and evidence instruction
  - Heightened citation and evidence instruction requirements
Mean overall Evidence by year

Mean Evidence score by document type: Grouped by year

Mean Evidence score by document type: Grouped by document

Mean Element Score by Year: Business vs. English

Two “types” of assessors

- Each year we recruit 15 business assessors and 15 English assessors
  - Background is loosely defined
- We have tracked their responses

Business vs. English: By year
Data and (campus) conversations

- Data can help build relationships
- English department became a key part of writing assessment efforts
- During reviewer campus visits during Drexel’s 10-year Middles States accreditation review, reviewers were enthusiastic about these efforts
- Other programs?

Further questions

- What can you do with data?
- Why did the “experiment” disappear?
- What’s happening with evidence?
- Why are business faculty evaluations higher?
- Is the writing *good* enough?
  - A decision for LeBow faculty and other stakeholders
- How well do the students write in your program?
  - A conversation driven largely by anecdote