Writing-to-Learn

David Galbraith – University of Southampton
d.w.galbraith@soton.ac.uk

Veerle Baaijen – University of Groningen
v.m.baaijen@rug.nl

Klein (1999)
Processes in writing-to-learn

• Spontaneously generating ideas "at the point of utterance" (Britton, 1980/1982)
• Externalize ideas in text, then reread them to generate new inferences (Young and Sullivan, 1984)
• Use genre structures to organize relationships among units of text, and hence among ideas (Newell, 1984)
• Set rhetorical goals, then solve content problems to achieve these goals (Bereiter and Scardamalia, 1987; Flower and Hayes, 1980a)

Meta-analysis of writing-to-learn interventions

• School-based interventions
  – Effects measured on achievement in final exams / standardised tests
• Significant but small effect sizes
• Need for more immediate measures of developing understanding
• Linking process to academic achievement?

Current research programmes

• Forthcoming edited volume
  – Writing as a learning activity (2014)
    • Klein, Boscolo, Kirkpatrick and Gelati
• Learning protocols and learning journals
  – Nückles, Renkl and colleagues
• Writing-to-learn in science
  – “Natural science, cognitive science and pedagogical influences on science literacy: empowering research and informing instruction”
Today’s focus

• Processes (learning?)

• Knowledge transforming model
  – Bereiter and Scardamalia (1987)

• Dual process model
  – Galbraith (2009)

Writing as problem solving

(Hayes, 1996; Bereiter and Scardamalia, 1987)

• Writing as a deliberate, controlled process
  – The thinking behind the text
  – Text production as passive output process

• Knowledge telling v knowledge transforming
  – Writing directed to communicative goals
  – Managing cognitive load (Kellogg, 1994)
  – Knowledge transforming goes “all the way down”

• Teaching
  – Goals embodied in specific genres
  – Strategies to enable more goal-oriented writing

Self-monitoring

(Snyder, 1987; Gangestad & Snyder, 2000)

• High self-monitors
  – “particularly sensitive to the expression and self-presentation of relevant others in social situations and use these cues as guidelines for monitoring (that is regulating and controlling) their own verbal and non-verbal self-presentation”.
  – Assume that they are more likely to direct their writing towards rhetorical goals.

• Low self-monitors
  – “expressive behaviour is controlled from within by their affective states (they express it as they feel it) rather than moulded and tailored to fit the situation”.
  – They are more likely to express their ideas directly as they unfold.

New ideas as a function of self-monitoring and mode of writing

Galbraith (1992, 1996)
Dual process model
(Galbraith, 2009)

- Explicit planning process (=knowledge transforming)
  - Retrieval of fixed ideas from episodic memory (hippocampus)
  - Manipulation in working memory to create rhetorically appropriate global model (spatial component of WM)
  - Leads to creation of single knowledge object in episodic memory (but not understanding)

- Knowledge-constituting process
  - Synthesis of ideas guided by semantic memory constraints (neo-cortex)
  - Dispositionally guided text production
  - Leads to formulation of ideas corresponding to writer’s implicit understanding of the topic

Relationship between the 2 processes

- Both processes required for effective writing
- Fundamental conflict because processes are optimised under opposing conditions
- Low and high self-monitors prioritise different components
  - Low SM prioritise knowledge-constituting and use bottom-up strategy
  - High SM prioritise knowledge-transforming and use top-down strategy

Knowledge transforming during planning
Galbraith, Hallam, Olive & le Bigot (2009)

- 96 low and high self-monitors writing article about pros and cons of legalising cannabis.
- Three phases
  - listing ideas (phase 1) (5 minutes)
  - constructing outline (phase 2) (10 minutes)
  - writing article (30 minutes)
- Secondary tasks loading on different components of working memory during phase 2
  - Control
  - Visual
  - Spatial
  - General interference

Sketch of knowledge-constituting process
### Effects on planning

- High self-monitors change content (LSA) more than low self-monitors during outlining
- Spatial interference reduces number of new ideas added during outlining
- Supports change in ideas due to rhetorical goals
- Indicates creation of “knowledge object” in spatial WM.

### Relationship between planning and text quality

- Text quality correlated with 3 features of the outlines, accounting for 21% of the variance in text quality
  - Similar content in list and outline (LSA measure)
  - Number of rhetorical headings in outline
  - Number of new ideas in outline
- Text quality correlated with reorganisation of existing content within new global structure that satisfies rhetorical goals

### Baaijen, Galbraith & de Glopper (2014)

**Writing beliefs**

- White and Bruning (2005)
- Transmissonal beliefs (TM)
  - “Writing’s main purpose is to give other people information”
  - “The key to successful writing is reporting accurately what authorities think”
  - Low TM better quality
- Transactional beliefs (TA)
  - “Writing helps me understand better what I’m thinking about”
  - “My thoughts and ideas become more clear to me as I write and rewrite”
  - High TA better quality

### Design

- Writing beliefs
  - TM (transmissional)
  - TA (transactional)
- Type of planning
  - Outline v synthetic
- Measures
  - Change in understanding
  - Text quality
  - Text modification
Writing beliefs and text quality

Writing beliefs and discovery

WBI conclusions

- Writing beliefs and type of planning moderate development of understanding
  - High transactional beliefs and synthetic planning associated with greater development of understanding
  - Top-down (low transactional) v bottom up (high transactional)
Baaijen & Galbraith
Processes related to discovery and text quality

• High and low self-monitors
• Outline v synthetic planning
• Measures
  – Development of understanding
  – Text quality
  – Keystroke measures of 2 types of processes (PCA)

2 process dimensions

• Global linearity
  – Linear transitions between sentences
  – Few insertions earlier in text
• Controlled sentence production
  – Longer pauses between and within sentences
  – Little revision at leading edge

Training school COWR, Writing-to-Learn, Galbraith & Baaijen, August 2014

Writing processes and discovery

- Synthetic planning
- Outline planning

Change in understanding

Writing processes and text quality

<table>
<thead>
<tr>
<th>Low self-monitors</th>
<th>High self-monitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncontrolled sentence production</td>
<td>Controlled sentence production</td>
</tr>
<tr>
<td>Predicted text quality</td>
<td></td>
</tr>
</tbody>
</table>

Training school COWR, Writing-to-Learn, Galbraith & Baaijen, August 2014
Text quality, processes and discovery

(i) Type of planning

- Synthetic planning
  - goals but not global structure of text
  - consistent with text quality but global structure emerges during text production
- Dispositional spelling out
  - Spontaneous, “shaping at point of utterance”
  - Implicit genre knowledge
- Extraction of global ideas from externalised draft
- Organising and rewriting to satisfy rhetorical goals
  - Explicit genre knowledge

(ii) Self-monitoring

Optimum strategy for reconciling discovery with text quality

- All four of Klein’s processes involved
- Problems for knowledge transforming model
  - Discovery and spontaneous sentence production
  - Conflict between quality and discovery
- Dual process predicts optimum strategy for development of understanding
- How is development of understanding related to learning?

Conclusions
Workshop Writing-to-Learn

David Galbraith – University of Southampton
Veerle Baaijen – University of Groningen

Program workshop

• Discussion
• Look at data
  – Data David et al; coding change in ideas in outline and comparing these with LSA measures
  – Keystroke data and text analysis data Veerle

Discussion lecture

Coding exercise

• Distinguishing content and rhetorical units
• Identifying new and old ideas
• Relating to latent semantic analysis (LSA)
• Example of idea change before and after writing
Relationship between LSA and old and new ideas

\[ r = 0.49 \ (p < 0.001) \] for old ideas

\[ r = -0.41 \ (p < 0.001) \] for new ideas

\[ R = 0.68, \ R^2 = 0.47 \]

Conclusion

- Manual coding is difficult
- LSA provides objective comparison
  - Reasonably valid
  - Broad measure of similarity of content
  - Greater similarity – better quality text
- Hand coded New and Rhetorical predict extra variance in quality
  - How elaborated the plan is for both content and rhetorical goals

Keystroke logging data

- Preparation of keystroke data important
- Baaijen et al., 2012, in Written Communication
- Our purpose: Investigate what happens at a text production level

Hand-out Poster Baaijen, Galbraith & de Glopper, 2012

Introduction
- Keystrokes are an unobtrusive indicator of writing processes but by themselves do not provide a transparent measure of distinct cognitive processes.
- This poster describes the procedures that we used and the measures that we developed to align keystroke measures with underlying writing processes.

Isolating text production
- Remove titles and instances of explicit planning during writing.
- Separate initial draft from post draft revision.
- Distinguish linear transitions between text units from events. This allowed us to assess revision at different levels by calculating the percentage of linear transitions at different text locations.
- Calculate pause durations between linguistic units rather than keystrokes (e.g. include space bar presses as part of between word pauses).

Measuring burst length

Dimensions of text production principal component analysis indicated that 3 underlying dimensions could be identified in the data.
- Extent to which sentences are planned or unplanned
- Extent to which sentences are revised as they are produced
- Extent to which successive sentences are linearly produced
Examples of different output

Table 1. Summary of Principal Component Analysis with Varimax Rotation for 2-Factor Solution

<table>
<thead>
<tr>
<th>Variables</th>
<th>Component 1</th>
<th>Component 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Linearity index</td>
<td>2.849</td>
<td>2.200</td>
</tr>
<tr>
<td>2. Percent of bursts on l-bursts</td>
<td>-1.930</td>
<td>-1.307</td>
</tr>
<tr>
<td>3. Trend time event</td>
<td>-0.822</td>
<td>-0.609</td>
</tr>
<tr>
<td>4. Percent linear sentence length</td>
<td>0.373</td>
<td>0.473</td>
</tr>
<tr>
<td>5. Number of production cycles</td>
<td>0.389</td>
<td>-0.278</td>
</tr>
<tr>
<td>6. Percentage linear word sequence</td>
<td>0.481</td>
<td>0.303</td>
</tr>
<tr>
<td>7. Percent of leading edge revision</td>
<td>0.115</td>
<td>-0.961</td>
</tr>
<tr>
<td>8. Percent words produced in P-bursts</td>
<td>-0.442</td>
<td>0.209</td>
</tr>
<tr>
<td>9. Percent of cognitive pauses between words</td>
<td>-0.672</td>
<td>-0.791</td>
</tr>
<tr>
<td>10. Test modification index</td>
<td>-0.390</td>
<td>-0.458</td>
</tr>
<tr>
<td>11. Mean pause time between sentences</td>
<td>0.313</td>
<td>-0.536</td>
</tr>
<tr>
<td>Eigenvalues</td>
<td>4.7</td>
<td>2.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% of variance</th>
<th>42.47%</th>
<th>21.40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.80</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Data inspection

- Look at the data and try to estimate how the participants score on variables below

<table>
<thead>
<tr>
<th>Global linearity</th>
<th>Controlled sentence production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity Index</td>
<td>Percent of bursts on l-bursts</td>
</tr>
<tr>
<td></td>
<td>Percent of leading edge revision</td>
</tr>
<tr>
<td>Percent of words produced in P-bursts</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant 206</th>
<th>Non-linear; controlled sentence production</th>
<th>Linear; controlled sentence production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity index</td>
<td>Percent of bursts on l-bursts</td>
<td>Percent of leading edge revision</td>
</tr>
<tr>
<td>Percent of words produced in P-bursts</td>
<td></td>
<td>Percent of words produced in P-bursts</td>
</tr>
<tr>
<td></td>
<td>.65</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55</td>
</tr>
<tr>
<td>Participant 211</td>
<td>Non-linear; controlled sentence production</td>
<td>Linear; controlled sentence production</td>
</tr>
<tr>
<td>Linearity index</td>
<td>Percent of bursts on l-bursts</td>
<td>Percent of leading edge revision</td>
</tr>
<tr>
<td>Percent of words produced in P-bursts</td>
<td></td>
<td>Percent of words produced in P-bursts</td>
</tr>
<tr>
<td></td>
<td>.96</td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>62</td>
</tr>
<tr>
<td>Participant 417</td>
<td>Non-linear; controlled sentence production</td>
<td>Linear; uncontrolled sentence production</td>
</tr>
<tr>
<td>Linearity index</td>
<td>Percent of bursts on l-bursts</td>
<td>Percent of leading edge revision</td>
</tr>
<tr>
<td>Percent of words produced in P-bursts</td>
<td></td>
<td>Percent of words produced in P-bursts</td>
</tr>
<tr>
<td></td>
<td>.81</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54</td>
</tr>
<tr>
<td>Participant 412</td>
<td>Non-linear; uncontrolled sentence production</td>
<td>Linear; uncontrolled sentence production</td>
</tr>
<tr>
<td>Linearity index</td>
<td>Percent of bursts on l-bursts</td>
<td>Percent of leading edge revision</td>
</tr>
<tr>
<td>Percent of words produced in P-bursts</td>
<td></td>
<td>Percent of words produced in P-bursts</td>
</tr>
<tr>
<td></td>
<td>.54</td>
<td>.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
</tr>
</tbody>
</table>