The role of different components of working memory in writing

David Galbraith (d.galbraith@staffs.ac.uk)
Jenny Hallam (j.l.hallam@staffs.ac.uk)
Centre for Educational Psychology Research, Staffordshire University
College Road, Stoke-on-Trent, ST4 2DE, UK

Thierry Olive (thierry.olive@univ-poitiers.fr)
Nathalie Le Bigot (nathalie.le.bigot@ univ-poitiers.fr)
CeRCA/MSHS,
99 avenue du Recteur Pineau, 86000 F - Poitiers France

Abstract

This paper describes the results of an experiment designed to assess the effects of self-monitoring and 3 different kinds of secondary task on the development of ideas during planning, and on the quality of subsequently produced text. An LSA-based measure was used to assess the development of ideas during planning and Coh-Metrix was used to assess effects on text. The results suggest that the spatial component of working memory plays an important role in the development of ideas during planning, and that this affects the quality of the final text. Individual differences in self-monitoring also affect the extent to which content develops during planning and are associated with differences in the coherence of the final text.

Keywords: Writing; working memory; text production; latent semantic analysis; Coh-Metrix

Introduction

Writing is commonly characterized as a knowledge-transforming process, in which writers actively transform their thought in order to satisfy rhetorical goals. Typically, this is attributed to high level problem-solving (Bereiter and Scardamalia, 1987). However, there has been relatively little research which directly assesses the effects of writing on knowledge. Research which has (see Galbraith, in press), has found some support for this, but has in addition claimed that these models of writing overemphasize the effects of high level thinking processes and neglect the more implicit processes involved in text production. Galbraith has suggested instead that writing should be characterized in terms of a dual process model.

According to the dual process model, two distinct processes are responsible for the generation of novel content during writing. The first – *knowledge retrieval* – process involves retrieving already-formed "ideas" from an explicit store of knowledge in long term memory, and either translating these directly into text (what Bereiter and Scardamalia would characterize as "knowledge telling") or the goal-directed evaluation and manipulation of ideas prior to translating them into text (what Bereiter and Scardamalia would characterize as "knowledge transforming"). By itself, however, this can only lead to the reorganization of existing knowledge or to the selection of different items of existing knowledge which are more appropriate for the rhetorical context. In order to create new content, the writer has to engage in a different – *knowledge constituting* – process, which involves the synthesis of content guided by the connections between subsymbolic units stored in an implicit semantic memory system. Although this process can be prompted by higher level problem solving, the content produced by it is the product of the implicit organization of content in semantic memory, rather than the explicit manipulation of content in working memory (WM). The result is new content that can be added to the store of existing knowledge in explicit memory.

A key feature of this model is that the two different processes are optimized under different conditions. The knowledge retrieval process involves the creation of a mental model of global structure designed to satisfy the writer's rhetorical goals, and operates best when content is represented economically in WM, and the writer can focus on evaluating and manipulating it rather than on articulating it in full text. By contrast, the knowledge constituting process operates best when thought is articulated in full text, and text production is guided by the implicit organization of semantic memory rather than by explicit rhetorical goals. This leads to a fundamental conflict between explicit organizing processes guiding the evaluation and selection of content in WM and implicit organizing processes guiding the synthesis of content in semantic memory.

Writers vary in the extent to which they prioritize the two processes. In particular, Galbraith has argued that low self-monitors, whose self presentation is guided by dispositional goals, prioritize the knowledge-constituting process, whereas high self-monitors, whose self presentation is guided by rhetorical goals (in the context of writing), prioritize the knowledge retrieval process (see Snyder& Gangestad, 1986, for a review of research on self-monitoring, and Galbraith, in press, for a review of research into the conditions under which writers develop new ideas). For present purposes, the key finding of this research is that low self-monitors tend to develop new ideas as a consequence of spontaneous text production, but not as a consequence of planning in note-form, whereas high self-monitors develop new ideas as a consequence of planning in note-form, but not as a consequence of spontaneous text production.

In order to resolve the conflict between the two processes writers resort to different forms of drafting strategy. In an

outline planning strategy, for example, the writer focuses on knowledge retrieval and generates and organizes their ideas in note-form prior to translating them into text. Kellogg (1988) has found good support for the effectiveness of the outline planning strategy, and that it is effective because it frees WM resources to focus on generating and organizing content during outlining and then to focus on translating this content during text production. However, although this study demonstrated that outlining has a beneficial effect on the quality of the final text, it did not examine the processes involved during outlining. In particular, it did not examine whether outlining was more effective when it involved the transforming of knowledge during outlining, and it did not examine how WM resources were allocated during outlining.

In order to investigate this, Galbraith, Ford, Walker and Ford, 2005) asked writers to make outlines while at the same time carrying out secondary tasks designed to load on different components of WM. They found that a secondary task loading on the spatial component of WM reduced the extent to which writers developed new ideas during outlining and that this was associated with a reduction in the quality of the final text. Galbraith (in press) has cited this as evidence in favour of the claim that the knowledge transforming process involves the construction of a mental model of the global structure of the text, and suggested that a spatial representation of ideas is required in order to be able to evaluate the global organization of simultaneously represented ideas. In particular, he claims that writers need to be able to construct a spatial representation of the text in order to identify where new content may be required in order to satisfy rhetorical and organizational goals. However, various features of the design of the Galbraith et al (2005) experiment meant that alternative explanations could not be ruled out. In particular, because the secondary tasks were imposed throughout the planning process, it could not be established that the effect was a consequence of outlining per se, and secondly because it was not clear that the secondary tasks loaded exclusively on the relevant components of WM, it could not be conclusively established that the effect was specifically a consequence of the spatial component of WM.

In this study, therefore, we had the following goals. First, we wanted to use better controlled secondary tasks, applied at more specific points during planning, to confirm whether a secondary task loading on the spatial component of WM does reduce idea development during outlining, and whether this has an effect on the quality of the text that is then produced. Second, we wanted to examine whether these effects varied as a function of individual differences in self-monitoring. We expected, on the basis of previous research, that high self-monitors would develop their ideas more as a consequence of outlining than the low self-monitors would. Third, we wanted to explore the potential of an alternative measure, based on latent semantic analysis, for capturing the extent to which knowledge is transformed during writing. Fourth, we wanted to assess the effects of outlining under different secondary task conditions on text quality by using a more detailed and objective set of measures of text characteristics.

Method

Participants

96 undergraduate students at Staffordshire University, the majority of whom (80%) were women, volunteered to participate in the experiment in return for credits in the Psychology department's research participation scheme. Their average age was 22.6 years (SD=6.8). Participants were pre-selected using Snyder's revised 18 item self-monitoring scale (Snyder and Gangestad, 1986) to form two groups of high and low self-monitors. They were classified as *low self-monitors* (n = 48) if they scored between 0 and 8 on the scale, and as *high self-monitors* (n=48) if they scored between 10 and 18 on the scale.

Design and procedure

Low and high self-monitors were randomly allocated to one of four conditions, varying in the nature of the secondary task imposed during the primary writing task.

Writing task In all four conditions, participants were asked to write an article for the university newspaper discussing the pros and cons of legalizing the use of cannabis and justifying their own opinion of the matter. They were reminded that they should be careful to consider both sides of the issue and that they should try to produce a reasoned argument justifying their position. They were told that they would have 45 minutes to do this in, and the time would be divided into 3 phases.

In *phase 1*, they were given 5 minutes to write down all the ideas they could think of about the topic, with each idea being in the form of a brief note or phrase, no longer than a sentence in length.

In *phase 2*, they were given 10 minutes to work out an organized outline of the article. They were instructed to think about its overall structure, considering what order to put their ideas in and how they could be grouped together. They were allowed to use their initial list of ideas for reference, but were free to change their ideas if they wanted to. As with the initial list, each idea was to be expressed as a brief note or phrase no longer than a sentence in length.

In *phase 3*, they were given 30 minutes to write the article itself, and were reminded that they should consider both sides of the issue, with the aim of producing a reasoned argument justifying their position.

Secondary tasks In the *control* condition, participants carried out the three phases of the writing task as normal without any secondary task. In the remaining conditions, participants were asked to carry out one of 3 secondary tasks during *phase 2* of the writing task. In each case, they were given a brief time to practice the secondary task before starting the writing task. During phase 2, secondary task

probes were presented and participants were asked to make a judgment about the stimuli, indicating their response by clicking a button held in their non-writing hand. Probes were presented randomly at intervals with a mean of 30 s and a range of 15 s to 45 s. After participants had completed the writing task, baseline data were collected using the same procedure and schedule.

In the *visual* condition, following a warning tone, participants were presented with three shapes in sequence but at different locations on a computer monitor in front of them, and asked to judge whether the shapes were the same or different to those presented on the previous trial, regardless of the locations they had appeared in. Same responses were indicated by a button press; different responses were indicated by the absence of a button press. Stimuli were presented randomly at intervals with a mean of 30 seconds and a range of 15 to 45 seconds. In the *spatial* condition, the same set of stimuli were randomly presented on the same schedule, but participants were required to judge whether the stimuli appeared in the same locations on the screen as on the previous trial regardless of whether they were the same shapes as before.

In the *interference* condition, participants were presented with a warning tone occurring on the same interval schedule as in the visual and spatial conditions, and asked to indicate whether or not the screen "whited out" following the tone.

Measures

Development of ideas "Ideas" were defined as separate chunks of text within the initial list produced during phase 1 or the outline produced during phase 2. Content ideas were differentiated from rhetorical headings (e.g. "Introduction", "Pros", "Cons") and phrases consisting of instructions to the writer without specific content (e.g. "Deal with potential objections here"). The following scores were calculated: number of ideas in initial list; number of rhetorical units in initial list; number of "old" ideas retained from initial list in outline; number of "new" ideas introduced into the outline; number of rhetorical units in outline. Two judges scored all the lists and outlines. Inter-rater reliabilities for these measures ranged between R=.81 and R =.95 ($p < .001$ in all cases). The means of the two judges' scores for each measure were used for analysis.

Semantic similarity of lists and outlines Latent semantic analysis (LSA) was used to compare the semantic similarity of the content in the lists produced in phase 1 and the outlines produced in phase 2. LSA represents the meaning of a text as a vector in a high-dimensional semantic space constructed by singular value decomposition from word co-occurrences in a large sample of documents. Text representations are constructed by summing the vector representations of a text's constituent words. Similarities in the meaning of texts are represented by the cosines of the angles between the vectors representing the two texts, which vary between -1 and 1 and are ordered in the same manner as correlations. (See Landauer, McNamara, Dennis and Kintsch, 2007, for a collection of papers about the underlying theory and its applications).

Comparisons of the lists and outlines were made using the LSA website at Colorado University (http:lsa.colorado.edu). The college reading space, with 300 factors defining the dimensions of the semantic space, was used to make document to document comparisons of each participant's initial list and outline. The resulting cosines were used to represent the extent to which an initial list of ideas had been transformed during the construction of the outline (with high scores corresponding to a strong similarity in content and low scores representing a greater change in content).

Coh-Metrix analysis of texts Coh-Metrix is a tool for the automated analysis of text developed at the University of Memphis. It calculates a wide range of indices of the linguistic and discourse features of a text, including indices of: general text and word properties; syntactic complexity; pronoun use; positive and negative connectives of different types; anaphoric references between sentences; coherence; and a range of situation model dimensions. (See Graesser, McNamara, Louwerse and Cai, 2004, for details of the indices and the architecture of the text analysis tools).

All texts were analyzed using Coh-Metrix 2.0 at http://cohmetrix.memphis.edu/cohmetrixpr/index.html. Principal components analysis of the raw set of 53 initial indices followed by varimax rotation suggested that the original set could be reduced to 8 orthogonal components, with the first component explaining 18% of the variance, and the eighth component explaining 5% of the variance; cumulative variance accounted for was 78%. Table 1 shows the factor labels and gives an indication of their interpretation. (Details are necessarily extremely brief because of space limitations). Factor scores for each of these components were used to assess effects on text properties.

Table 1: Text components and their interpretations.

	Component	Interpretation
1	Local coherence	High argument overlap and semantic similarity between sentences
2	Anaphoric reference	High use of anaphoric reference and pronouns in complex sentences
3	Causal argument	High use of causal connectives & verbs; high ratio of causal particles to verbs
4	Syntactic repetition	Syntactically similar sentences
5	Elaborated text	Longer texts including negation and logical operators
6	Two-sided argument	High use of negative additive and logical connectives
7	Logical argument	High use of logical operators and positive additive connectives
8	Global coherence	High semantic similarity between paragraphs with conditional content

Subjective rating of text quality Two judges rated the quality of the texts on a 9 point scale. They were asked to imagine that they were the editor of the student newspaper and consider the extent to which they would publish the

article. Their judgments should be based on the coherence of the overall argument, its originality, and the appropriateness of the tone and the relation to the readership of an article in the student newspaper. The correlation between the two judge's ratings was R=0.74, p< .001).

Results

Secondary task performance

There was a highly significant reduction in accuracy ($F(1, 66)=20.9$, $p<.001$, $\eta^2=.24$) and increase in response time ($F(1,66)= 20.2$, $p<.001$, $\eta^2 =.24$) when the secondary tasks were carried out at the same time as outlining, as opposed to during the baseline task, indicating that outlining interfered with participants' ability to perform the secondary task. However, this did not interact with either self-monitoring or condition ($p>.52$, $\eta^2<.02$ for all tests), indicating that same amount of attention and effort were paid to outlining in all the secondary task conditions. Participants performed more accurately ($F(2,66)=14.41$, $p<.001$, $\eta^2=.30$), and responded faster ($F(2,65)=7.5$, $p=.001$, $\eta^2=.19$), to the secondary task in the interference condition – where all they had to do was indicate whether the screen had "whited out" – than they did in either the spatial or the visual conditions ($p<.005$ in both cases). However, there were no significant differences in accuracy or response time between the visual and spatial conditions or between low and high self-monitors. Taken together, these results imply that the visual and spatial tasks were of equivalent difficulty, and that any difference in their effects on writing can be attributed to differences in the components of WM they require rather than to a difference in task difficulty.

Development of ideas during planning

There were no significant differences between the conditions or between low and high self-monitors in the number of ideas or in the total number of words produced in the initial list ($p>.21$, $\eta^2<.02$, for all effects). There were also no significant differences in the number of old ideas retained in the outline, or in the number of rhetorical headings included in the outline, during phase 2 ($p>.22$, $\eta^2<.02$, for all effects). This suggests that self-monitoring and the different types of secondary task had no effect on the ability to generate ideas or on the selection of ideas from the initial list for incorporation in the outlines.

A two-way between subjects ANCOVA with self-monitoring and secondary task condition as independent variables, and the number of ideas in the initial list as a covariate, showed a significant main effect of secondary task condition on the number of new ideas introduced during the construction of the outline ($F(3, 85)=3.30$, $p=.024$, $\eta^2=.10$). (Note that the distribution of the new ideas was positively skewed so logs were taken to normalize the distributions). Figure 1 shows the mean number (logs) of new ideas produced in each condition. Post–hoc comparisons using Tukey's HSD showed that the spatial condition produced significantly fewer new ideas than the control condition ($p=.04$) and the interference condition ($p=.006$) but that the difference with the visual condition was not significant ($p=.16$)

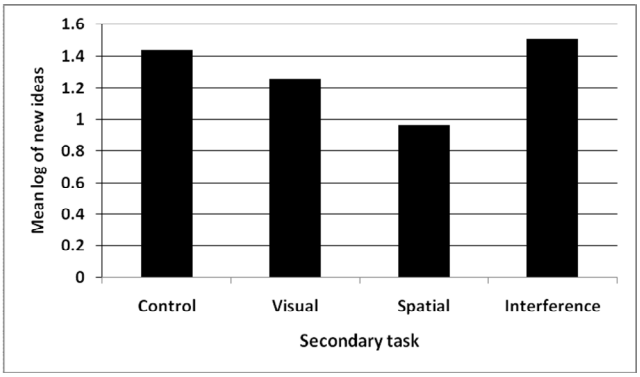

Figure 1: Mean number of new ideas (log) as a function of secondary task condition.

Semantic similarity of lists and outlines

The distribution of the cosines representing the semantic similarity of the lists and the outlines was negatively skewed so they were reflected and square rooted in order to normalize the distribution. In order to assess the relationship between this measure of semantic similarity and the analysis of changes in ideas during outlining, the transformed cosines were correlated with the number of old ideas retained in the outlines and the number of new ideas introduced in the outlines, controlling for the number of ideas initially produced. These showed a significant positive correlation between semantic similarity and the number of old ideas retained ($r=.51$, $p <.001$) and a significant negative correlation with the number of new ideas introduced ($r=-.56$, $p<.001$). In combination, these two variables showed a multiple correlation of $R=.68$ ($p< 001$) with the LSA scores. This suggests that the LSA-based measure is a valid measure of the extent to which content has been transformed during outlining.

In order to assess the effects of self-monitoring and secondary tasks on the extent to which semantic content of was transformed during outlining, a 2-way between subjects ANOVA was carried out on the LSA scores. This showed a significant main effect of self-monitoring ($F(1,78)=4.00$, $p=.049$, $\eta^2=.05$) but no effect of secondary task ($p=.63$, $\eta^2=.02$) and no interaction between self-monitoring and secondary task condition ($p=.57$, $\eta^2=.02$). High self-monitors' outlines were less similar to their initial lists (median=.72, interquartile range = .22) than the low self-monitors' (median=.79, interquartile range = .19) were. In other words, high self-monitors changed the semantic content of their outlines more than low self-monitors did.

The fact that these results are so different to those for analysis of the number of new ideas suggests that the LSA measure captures different aspects of change in content. In particular, we suspect that it captures changes to implicit semantic content, whereas the coding of ideas captures how

content is explicitly represented as separate ideas. In order to test this hypothesis we repeated the two-way between subjects ANCOVA assessing the effects of self-monitoring and secondary task condition on the development of new ideas with the LSA measure added as a covariate. This showed a similar main effect of secondary task ($F(3, 76)=4.04$, $p=.01$, $\eta^2=.14$) but, in addition, post hoc comparisons showed that the spatial condition produced significantly fewer new ideas than each of the other conditions. For the comparison with the control condition, $p=.007$; with the visual condition, $p=.04$; and with the interference condition, $p=.002$. This implies that a more refined measure of the number of new ideas relative to the amount of change in semantic content provides a more sensitive indicator of the effect of the spatial secondary task.

Text quality

Subjective ratings A 2-way between subjects ANOVA was used to assess the effects of self-monitoring and secondary task on the subjective ratings of the quality of the text produced in phase 3. This showed a significant main effect of secondary task ($F(3, 88)=3.58$, $p=.02$, $\eta^2=.11$) but no effect of self-monitoring or interaction with self-monitoring. Texts produced in the control condition ($M=5.94$, $sd=2.16$) were rated of higher quality than in all the secondary task conditions, with the difference being significant for the comparisons with the interference condition ($M=4.15$, $sd=2.04$, $p=.02$), marginally significant for the comparison with the spatial condition ($M=4.46$, $sd=2.15$, $p=.07$), and non-significant for the comparison with the visual condition ($M=4.62$, $sd=1.75$, $p=.12$). There were no significant differences between the secondary task conditions.

Coh-Metrix Multiple regression was used to assess the relationship between the factors identified using Coh-Metrix and subjectively rated text quality. This showed a significant multiple correlation between text quality and four of the factors ($R=.51$, $p < .001$), which accounted for 23% of the variance in the quality ratings. The four factors which made significant independent contributions to this relationship were: *anaphoric reference* (a negative correlation) (partial $r= -.30$, $p<.004$); *elaborated text* (partial $r=.36$, $p<.001$); *logical argument* (partial $r=.21$, $p<.05$); and *global coherence* (partial $r=.23$, $p<.05$). This suggests that longer more elaborated texts, which included opposed but logically coherent arguments, integrated by co- rather than anaphoric reference, were given higher quality ratings.

We then examined what effect self-monitoring and secondary task condition had on these objective features of the text. There were 3 significant effects here. First, there was a significant effect of self-monitoring on the local coherence of the text (($F(1, 88)=10.88$, $p=.001$, $\eta^2=.11$), with low self-monitors ($M=.32$, $sd=1.10$) producing more locally coherent text than high self-monitors ($M=-.32$, $sd =.80$).

Second, there was a marginally significant effect of secondary task condition on anaphoric reference (($F(3, 88)=2.55$, $p=.06$, $\eta^2=.08$), with all secondary task conditions showing higher levels of anaphoric reference than the control condition ($M=-.038$, $sd=.78$) but with only the difference in the visual condition ($M=.37$, $sd=1.15$) showing a significant effect on post-hoc comparisons ($p=.04$). The spatial ($M=-.05$, $SD=1.1$, $p=.62$) and interference conditions ($M=.06$, $SD=.88$, $p=.38$), though higher, were not significantly different to the control or visual conditions. These results are important insofar as this factor showed a significant relationship with subjective ratings of quality. They suggest that this may be an important feature of how planning affects the quality of the final text. However, it is less clear what aspect of planning is responsible for the effect. It may be specific to the visual condition, but could also be a more general effect of disrupting planning.

Third, there was a significant effect of secondary task on two-sided argument (as indicated by a high incidence of negative additive and logical connectives) ($F(3, 87)=3.55$, $p .02$, $\eta^2=.11$). As can be seen in figure 2, the spatial condition showed less use of these kinds of connectives than the control ($p=.03$) and interference ($p=.03$) conditions, and less than in the visual condition, but not significantly so ($p=.16$).

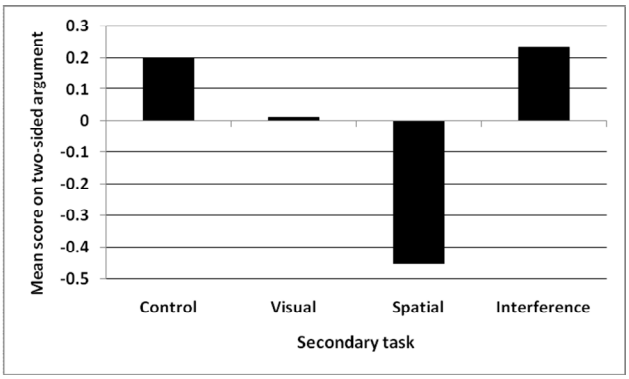

Figure 2: Mean score on two-sided argument factor as a function of secondary task condition.

Conclusions

Our first conclusion is that a secondary task loading on the spatial component of WM does reduce the extent to which writers develop new ideas during outlining. This strengthens the findings of Galbraith et al (2005) because the effect persisted even with these better controlled secondary tasks. It is worth noting, however, that the initial analysis failed to show a significant difference between the spatial and the visual conditions. Although this is probably simply a matter of statistical power, which could be tested by a replication with a larger sample, it may also be a question of how the effect is defined. When the LSA measure of change in semantic content was added as a covariate, the effect increased, and the difference with the visual condition became significant. We think that this is because adding the LSA measure effectively converts the score to a ratio measure: it reflects the extent to which a given amount of change in semantic content has been

differentiated into separate ideas. If this is correct, it implies that the spatial task reduces the extent to which writers are capable of differentiating new content into distinct ideas, and that the function of this component of WM during outlining is to enable the writer to articulate a global change in semantic content into discrete ideas.

Our second important finding was that, although the spatial secondary task was associated with a reduction in text quality, this was also true of the other secondary task conditions. In particular, the interference task was also associated with a reduction in quality. This implies that the effect could be a consequence of a general disruption of planning rather than a specific effect of the spatial component of WM. This possibility is partly ruled out by the fact that the spatial task did have a specific effect on the incidence of negative additive and logical connectives. Since negative connectives of this type are associated with the presentation of two-sided arguments, this implies that writers in the spatial condition were less able to generate counterarguments to their position, perhaps because they were less able to simultaneously represent different views on a common topic. However, since this effect was not significantly different to the visual condition, we cannot be absolutely sure that this is a distinctive effect of the spatial component of WM. Further research with a larger sample is required to establish whether this effect is unique to the spatial condition.

The addition of self-monitoring as an independent variable in this experiment revealed several new findings. First, the high self-monitors changed the semantic content of their outlines more than the low self-monitors, which is consistent with the findings of previous research showing that high self-monitors develop their ideas more during planning than low self-monitors do (Galbraith, in press; Galbraith, Torrance and Hallam, 2006). This is assumed to be a consequence of high self-monitors adapting the content of their texts more to rhetorical goals. However, the fact that this greater change in semantic content was not associated with a concomitant difference in the number of new ideas produced by low and high self-monitors suggest that there may also be differences in the extent to which low and high self-monitors differentiate semantic content into separate ideas during planning. Second, there was a marked difference in the local coherence of the texts produced by low and high self-monitors. The natural explanation for this is the greater amount of change in semantic content for high self-monitors during planning. This would also fit with Galbraith et al.'s (2006) finding that new ideas developed by high self-monitors after writing planned texts were less coherently related to one another than new ideas generated by low self-monitors (note though that this was a different kind of measure and was not text-based). However, since there was no direct correlation between the extent to which semantic content changed during planning and the local coherence of the resulting text, it may be that this difference is a consequence of a difference in how the two groups construct adjacent sentences during text production, rather than of the difference in the effects of outlining. Whatever the precise explanation of these differences, they do provide further strong evidence that self-monitoring has important effects on both planning and text production, and are compatible with the assumptions that the dual process model makes about the way that low and high self-monitors prioritize planning and text production processes.

Finally, this study demonstrates the utility of both LSA and Coh-Metrix as tools for investigating writing. In this experiment, LSA not only correlated well with human coding of development in ideas but also provided direct evidence of a difference - the effect of self-monitoring on changes in content during planning - that was not detected by other measures. From a theoretical point of view, it also promises to provide a means of capturing effects of implicit semantic memory processes on writing. Coh-Metrix also proved to be a valuable tool for identifying the effect of the independent variables on specific text features. Global quality ratings, though useful in providing a general indication of whether a particular strategy is effective or not, do not provide information about how specifically the strategy has its effect on the text. With larger samples than used in this experiment, we should be able to use path analysis to identify the relationships between experimental manipulations, their effects on planning processes, and specific features of the resulting texts.

References

Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition.* Hillsdale, NJ: LEA.

Galbraith, D. (in press). Writing as discovery. *British Journal of Educational Psychology Monograph Series II, Number 6 - Learning and Teaching Writing.*

Galbraith, D., Ford, S., Walker, G. & Ford, J. (2005). The contribution of different components of WM to knowledge transformation during writing. *Educational Studies in Language and Literature, 15,* 113-145.

Galbraith, D., Torrance, M. & Hallam, J. (2006). Effects of writing on conceptual coherence. *Proceedings of Annual Conference of the Cognitive Science Society,* 1340-1345.

Graesser, A., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments and Computers, 36,* 193-202.

Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 355-365.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2007). *Handbook of Latent Semantic Analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. *Journal of Personality and Social Psychology, 51,* 125-139.