# The development of understanding through writing

Baaijen, Veerle Marije

# THE DEVELOPMENT OF UNDERSTANDING THROUGH WRITING

**Veerle Baaijen**



The work in this thesis has been carried out under the auspices of the Research School of Behavioural and Cognitive Neurosciences (BCN) and the Center for Language and Cognition (CLCG). Both are affiliated with the University of Groningen.

Groningen Dissertations in Linguistics 107

Veerle M. Baaijen
The development of understanding through writing

ISBN: 978-90-367-5842-0
ISBN: 978-90-367-5843-7 (electronic version)



Cover design: Onno Kramer
Printed by Off Page Amsterdam, the Netherlands

**RIJKSUNIVERSITEIT GRONINGEN**

# The development of understanding through writing

## Proefschrift

ter verkrijging van het doctoraat in de
Letteren
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
donderdag 22 november 2012
om 16.15 uur

door

**Veerle Marije Baaijen**

geboren op 7 april 1982
te 's-Gravenhage

| | |
|---|---|
| Promotor: | Prof. dr. C.M. de Glopper |
| Beoordelingscommissie: | Prof. dr. J.R. Hayes |
| | Prof. dr. C. Jansen |
| | Prof. dr. G.C.W. Rijlaarsdam |

External supervisor: David Galbraith
University of Southampton (UK)

# Contents

# Chapter 1

*Outline of the Thesis*

## 1.1 INTRODUCTION

The topic of this thesis – the development of understanding through writing – straddles two different aspects of research on writing. On the one hand, there is the applied field of writing-to-learn research. This field of research is primarily concerned with the use of writing as a tool for learning in educational contexts, and treats learning as a relatively general effect of different kinds of writing tasks. In this more general context, writing is seen as having a number of potential benefits for learning. For example, it may help learners to test their understanding of concepts and to consolidate material that they have learned in memory. In addition, it may encourage students to think more critically about a topic and to construct new knowledge (Ackerman, 1993; Applebee, 1984; Klein, 1999). However, despite a widespread assumption that writing can have these beneficial effects, empirical studies of the value of writing as a tool for learning have produced inconclusive and contradictory results (Bangert- Drowns, Hurley & Wilkinson, 2004; Klein, 1999; Kieft, Rijlaarsdam & van den Bergh, 2006). Klein (1999) has suggested that part of the reason for these conflicting results is that research on writing-to-learn has taken a relatively broad view of "learning" and has paid insufficient attention to the specific cognitive processes involved in writing, and the way these vary across different contexts.

On the other hand, there is more basic research into the cognitive processes involved in writing. A focus of this research has been on the way in which writers actively develop their ideas in the course of writing, rather than simply translating pre-existing ideas into language. This has led to writing being characterised as a process of discovery or invention (Flower & Hayes, 1980). Although this clearly involves learning, it is a much more specific kind of learning than is assumed in writing-to-learn research, focussing on the way that the writer develops their subjective understanding of a topic as they write, rather than on whether this leads to, for example, better retention of material, or better objective performance on tests of learning outcomes in educational contexts.

The research described in this thesis focuses on the second of these two aspects. Thus, although its findings may have implications for research on writing-to-learn, the focus of the research will be on the cognitive processes involved in the development of understanding during writing, and on how this may, or may not, relate to the quality of the texts that writers produce, rather than on testing the value of a school-based writing-to-learn intervention. This will involve comparing, and testing, the claims made by two different cognitive accounts of how writing facilitates the development of understanding.

The most widely accepted account of this process is the problem-solving, or meta-cognitive (Klein, 1999), approach exemplified in the theories of Hayes and Flower (1980; Flower & Hayes, 1980b) and Bereiter and Scardamalia (1987; Scardamalia & Bereiter, 1987). Both these theories claim that the construction of new ideas depends on the extent to which writers adapt their thought to rhetorical goals. Hence, these theories attribute learning to a problem solving activity in which students have to set explicit rhetorical goals and pursue goal directed strategies to transform knowledge (Hayes & Flower, 1980; Bereiter & Scardamalia, 1987; Klein, 1999; 2004). The notion that writers can change their knowledge through writing is derived from the knowledge-transforming model of Bereiter and Scardamalia (1987). They describe expert writing as a knowledge-transforming process and contrast it with the knowledge-telling process employed by novices. Even though this model suggests that expert writers will develop their understanding when they are involved in knowledge-transforming processes, it should be noted that knowledge change is rarely empirically tested. Typically, studies inspired by meta-cognitively oriented theories (Hayes & Flower, 1980; Bereiter & Scardamalia, 1987) focus on the underlying cognitive processes involved in writing, rather than on the effects of writing on knowledge change.

Recently, Galbraith has investigated more directly how writers develop understanding through writing (Galbraith, 1992; Galbraith, Torrance & Hallam, 2006). This is shown most clearly in an early experiment performed in 1992. Within this research Galbraith administered Snyder's self-monitoring scale (Snyder & Gangestad, 1986) to distinguish between writers whose writing is presumed to be directed towards rhetorical goals (high self-monitors) or dispositional goals (low self-monitors). Galbraith then measured development of subjective understanding under different writing conditions (notes versus full prose). This setup allowed Galbraith to measure the extent to which writers developed new ideas as a function of writing in these different writing conditions. Galbraith assumed that, if Bereiter and Scardamalia were right about the knowledge-transforming process, then high self-monitors, who are supposed to take rhetorical goals into account, would develop more new ideas compared to low self-monitors, who would only be engaged in a knowledge-telling process. In fact, although he found that high self-monitors did indeed produce more new ideas than low self-monitors after writing notes, these ideas were not associated with increases in subjective understanding. Furthermore, after writing full prose, the low self-monitors produced more new ideas than the high self-monitors, and these new ideas were associated with increases in subjective understanding. The results of Galbraith's research, then, appear to contradict important features of Bereiter and Scardamalia's knowledge-transforming model, since not only the assumption that

increases in understanding depend on adaptation of content to rhetorical goals, but also the assumption that low self-monitors' writing is a simple matter of 'knowledge-telling', were challenged.

Although, Galbraith did not investigate the underlying cognitive processes involved, he nevertheless proposed that a dual process model of writing was needed to explain these findings and the relationship between writing processes and knowledge change. In his dual process model, Galbraith (Galbraith, 1992; Galbraith et al., 2006) claims that, while writing, two processes occur simultaneously. One is a rhetorical process like the knowledge-transforming process described by Bereiter and Scardamalia, the other is a spontaneous text production process. Galbraith describes this as a knowledge constituting process in which knowledge is constituted in the process of language production, and in which idea generation involves the synthesis rather than the retrieval of content.

In summary, while Galbraith (1992) has measured development of understanding and gives some suggestions for processes which could be involved in discovery, he has not tested this directly. On the other hand, while Bereiter and Scardamalia have directly observed the processes involved in their knowledge-transforming model, and assume that these lead to development of understanding, they have not directly measured knowledge change. In order to investigate how writing contributes to the development of understanding through writing it is, therefore, of key importance to address the question of which processes are related to discovery. The overarching aim for this thesis is to bring those two approaches together and to test how processes of writing relate to the development of understanding through writing.

### *1.1.1 Outline of the thesis*

In what follows, this thesis will continue with a more detailed discussion of the main claims of two widely different cognitive accounts to the question of how writing facilitates learning. The claims of the meta–cognitively oriented theories (Hayes & Flower, 1980; Bereiter & Scardamalia, 1987) will be compared and contrasted with the predictions of the dual process model from Galbraith (2009). This will then be followed by four empirical studies and a concluding chapter discussing the main findings and implications of the present research. Finally, at the end of the thesis, there will be an extended summary of the research in Dutch.

There is some overlap in the literature discussed in the different chapters of this thesis. This overlap is due to the decision to write the empirical chapters of this thesis as manuscripts for separate journal articles. Chapter 4 is based on a publication

in *Written Communication*, chapter 3 has been submitted for publication and chapter 5 and 6 will be submitted for publication in the near future. The advantage for the reader is that each chapter can be read on its own.

**Chapter 2** gives an overview of the different cognitive models involved. The Hayes and Flower (1980) cognitive model of writing and the knowledge-transforming model of Bereiter and Scardamalia (1987) and the predictions that they make about the development of understanding through writing will be compared and contrasted with the dual process model from Galbraith (2009). This chapter will consider the previous attempts to test discovery through writing empirically and will conclude with an overview of issues that will be addressed in the following empirical studies.

**Chapter 3** evaluates the Writing Beliefs Inventory designed by White and Bruning (2005). This individual difference measure was selected in addition to the self-monitoring scale that Galbraith has used in his research. This chapter will report a study that tests the effect of implicit writing beliefs on writing processes, text quality and development of understanding. The results of this chapter show that the interpretation of White and Bruning's scale is not very clear. The chapter will provide suggestions for a different interpretation of the scales and stresses that further research is required on the development of the questionnaire if it is to be used as a valid and reliable writing construct. For this reason it was decided not to include this individual difference measure in the other analyses reported in the thesis.

**Chapter 4** discusses the methodological issues involved in trying to turn raw keystroke logs into usable measures that can be related to models of writing processes. It identifies problems that arise when trying to capture writing processes on the basis of keystroke logging data and it suggests ways in which these problems can be solved. First, it considers ways in which the text production process itself can be isolated from other components of the writing process. Having done so, it will provide an overview of analyses investigating how pauses, bursts and revisions that occur during text production can be individually classified, and how these measures can be combined together to create global measures that can be used to characterise texts and writing processes. Slightly modified versions of these measures will be used in the following chapters as indicators of writing processes.

**Chapter 5** presents the results of the main study of this research. It reports about development of understanding through writing, shows how this is related to process measures and idea change measures and finally relates this to text quality. Overall,

the results show that writers are more likely to experience increases in understanding after synthetically planned writing than after outline planned writing, and that these increases in understanding are associated with unplanned sentence production combined with less linear production of text. It also shows clear differences between low and high self-monitors in the way these processes relate to text quality. High self-monitors produce better texts when they produce linear text, whereas low self-monitors produce their best texts when planned sentence production is combined with non-linear text production. In general, text quality is not related to the development of understanding. These results therefore support the broad claim of the dual process model that the development of understanding involves two kinds of processes.

**Chapter 6** reports a study which assesses whether ideas produced before and after writing can be located in the process of writing. It then uses generalized additive mixed effect regression modelling to assess the distribution of old and new ideas within the texts produced in the different writing conditions, and evaluates the relationships between these patterns and the development of understanding through writing. The results of this study show that development of understanding depends on the use of ideas in the course of text production. Furthermore, the results show that the process by which ideas are produced in the course of writing show different patterns for the different planning conditions.

**Chapter 7** summarizes the main findings from the work presented in the foregoing chapters. It then presents the main interpretations of these results and discusses methodological issues. Finally, the chapter reflects on the implications of these findings and suggests a new approach to the study of writing.

**The Dutch summary** provides an overview of the presented studies and a summary of the overall conclusions of the research.

# Chapter 2

*Introduction*

## 2.1 WRITING RESEARCH

In the writing-to-learn literature, it is assumed that writing can help students to acquire content knowledge, to infer new relationships between ideas, to think critically about new content and hence, to develop their understanding of a topic (Klein, 1999). Writing-to-learn assignments often involve tasks that require students to explore the relationships between ideas (Langer & Applebee, 1987). The primary function of such tasks is to foster critical thinking and to enable students to order and represent experiences in terms of their own understanding. However, there is a discrepancy between writing-to-learn research and the claims and predictions of the cognitive models in writing. Writing-to-learn research is typically concerned with the effect of different writing tasks on academic learning outcomes rather than on the detailed cognitive processes involved in writing. Cognitive models of writing are concerned, by contrast, specifically with the processes involved in writing, and with how these processes develop and vary with expertise. Although these models do claim that writing is associated with the development of understanding, these claims are typically presented as an incidental consequence of the general character of the expert writing process, rather than being the explicit focus of the research.

The research presented in this thesis draws in part on both the writing-to-learn approach and the cognitive process approach, but does not fully coincide with either. Thus, although it does explicitly examine the conditions under which developments of understanding occur, it does so by examining the way in which ideas change as a consequence of writing, and the extent to which these changes are associated with the subjective experience of development of understanding. This is primarily assumed to reflect the nature of the writing processes involved rather than the objective effects of writing on learning. Although the development of subjective understanding may often be related to improved academic performance, it need not be.

The development of understanding in writing is described in a variety of different ways in the literature. For example, Galbraith (2009) refers to 'discovery through writing'; Flower and Hayes (1980b) refer to 'invention'; whereas Bereiter and Scardamalia (1987) refer to the 'transformation of knowledge'. Regardless of the different terms that are used, we assume that these notions refer to the same underlying construct of 'discovery'. According to Klein (1999), this is fundamentally different to the forms of learning that are used in writing-to-learn research. Typically, writing-to-learn research is concerned with the rehearsal of subject matter, conceptual change and the retention of material. Since retention is integral to learning, writing-to-learn involves relative permanent changes in behaviour (Klein, 1999; Friedenberg &

Silverman; 2006). By contrast, the notions of discovery used by cognitive models of writing refer to the subjective experience of understanding. As Flower and Hayes put it, this is related to moments of insight, to the feeling of "Eureka, now I see it" (Flower & Hayes, 1980b, p.21).

In the next section, we will outline how the different cognitive models account for this form of discovery through writing.

### *2.1.1 Meta-cognitive theories*

Within the meta-cognitively oriented theories writing is seen as more than translating preconceived ideas into written text. Instead it is seen as a problem solving activity. Flower and Hayes (1980a; Hayes & Flower, 1980; Hayes & Flower, 1986) have characterized writing as 'juggling constraints', since writing involves thinking about and acting on many different things at the same time. According to their theory, during writing, writers have to think about their knowledge of a certain topic and at the same time they have to keep in mind how much their audience knows about this topic. In addition, writers have to think about their goals for writing a specific text, about text and genre demands and about how they can translate their ideas into written language. Thinking about all of these constraints at the same time takes up a lot of the available working memory capacity and, therefore, writers face the difficult task of coordinating the different processes effectively while writing. In order to prevent cognitive overload writers use different strategies to organise the different processes effectively (Flower & Hayes, 1980b; Hayes & Flower, 1980).

These basic notions about writing characterise the way in which the meta-cognitively oriented theories claim that writing can be seen as a process of discovery. First of all, they assume that writing is a strategic problem solving activity in which knowledge is transformed in order to satisfy rhetorical goals. In addition, these theories claim that this process of transforming knowledge is most effective when the writer is able to concentrate on strategic thinking, separated from text production (Flower & Hayes, 1980; Bereiter & Scardamalia, 1987; Kellogg, 1990).

### *2.1.2 Hayes, Flower and others*

The most influential models of writing research are the cognitive writing models that emerged in the 1980's. Hayes and Flower (1980) are the most cited researchers within this field. Their view of writing as a problem solving activity led to the development of a general model of the processes involved in writing and to a theory of writing expertise (Hayes & Flower, 1980; Hayes and Flower, 1986; Galbraith & Torrance,

1999a). The model distinguishes between three basic processes: planning, translating and revising. Planning has generating ideas, organisation and goal setting as its components. Translating means translating plans into text. It refers to constructing sentences and to actual language production. Revising includes reading and editing as its components. It involves evaluating the text or plans for the text, so it refers to both mental as well as written evaluations. These processes operate upon a representation of the task environment and on knowledge stored in the long term memory. The representation of the task environment consists of the writing assignment and the text produced so far. The long term memory consists of things such as topic knowledge, knowledge of the audience, a writing plan, and rules for text production and knowledge of text standards (Galbraith & Torrance, 1999a). The model of Hayes and Flower (1980) is shown in figure 2-1.

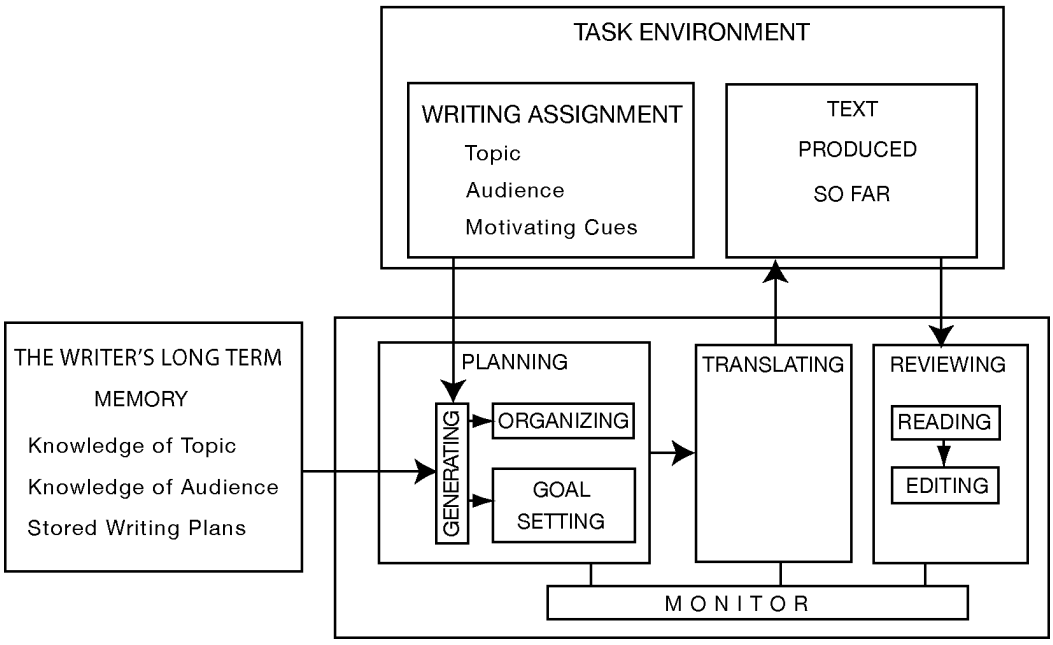


**Figure 2-1** Flower and Hayes' cognitive model of writing. Taken from Hayes & Flower (1980, p.11).

A key feature of this model is that it does not view writing as a linear process of plan-write-edit, but rather as a recursive process in which planning, translating and revising can occur at any moment during the writing task. Basically, it is assumed that the writer develops an initial definition of the rhetorical problem, which may be more or less fully developed, for the text, and then sets sub goals to satisfy these global plans for the text. When working out how to satisfy these sub goals for the text, the writer might need to adjust the more global plans for the text as a whole. Therefore,

both the goals for the text as well as the text itself evolve during writing. In this way, Hayes and Flower (1980) stress the recursive nature of writing. Cognitive processes such as planning, translating and reviewing are repeatedly called upon during the production of text. Hayes and Flower argue that these cognitive processes compete for cognitive resources and that it is the responsibility of the monitor to coordinate the different processes (Bochardt, 1984; Hayes & Flower, 1980).

This research has been mainly conducted by comparing the writing processes of expert and novice writers. Hayes and Flower (Flower and Hayes, 1980b; Hayes & Flower, 1986) found that expert writers construct a more elaborated representation of their goals than novice writers do. In particular, expert writers explicitly set rhetorical goals for the text as a whole, whereas novice writers rely on more concrete content goals. In consequence, experts develop more elaborate plans. The more elaborate conceptual representation of goals of the text enables these writers to revise more extensively and to evaluate their text in terms of its underlying function with respect to their goals, rather than simply considering if the text is appropriately expressed as novice writers do. In contrast, novice writers generate content in response to the topic alone (Hayes & Flower, 1980; Galbraith & Torrance, 1999a).

In summary, expert writers are assumed to elaborate their rhetorical goals to adapt to the knowledge of their audience and the requirements of the task. To fulfil these top level rhetorical goals, writers generate sub goals. To satisfy these sub goals, writers generate ideas and organize their knowledge. This process can deliver ideas that did not previously exist or which were not part of the original plan for the text. These new ideas are then incorporated within the initial plan for the text and hence, the global plan for the text as a whole is adjusted. It is this restructuring of the global plan for the text that is assumed to account for discovery through writing (Hayes & Flower, 1980; Flower & Hayes, 1984; Klein, 1999). By contrast, for novices, the writer's goal is simply to say what he or she knows about the topic. Novices will generate information that is directly relevant to the topic, rather than selecting and organizing knowledge tailored to the reader's needs. Therefore, Flower and Hayes assume that this mode of writing does not contribute to the development of the writer's understanding of the topic (Hayes & Flower, 1980; 1986; Flower & Hayes, 1980b).

One limitation of Hayes and Flower's initial model (1980) is the absence of a short-term memory component. However, in Hayes' revision of the model (Hayes, 1996, see figure 2-2), working memory is explicitly included in the model. The description of working memory in Hayes' revised model draws heavily on Baddeley's (1986) general model of working memory and it likewise includes a phonological loop and a visual-spatial sketchpad.

In the revised model, the processes of text production and reflection replace the components of planning and translating in the original Hayes and Flower model. Reflection includes the cognitive process of problem solving, decision making and inferencing and can essentially be seen as equivalent to the original planning component, but should be interpreted as a more general cognitive process. Furthermore, the new text production component implies that expression in language is not just a simple translation process.

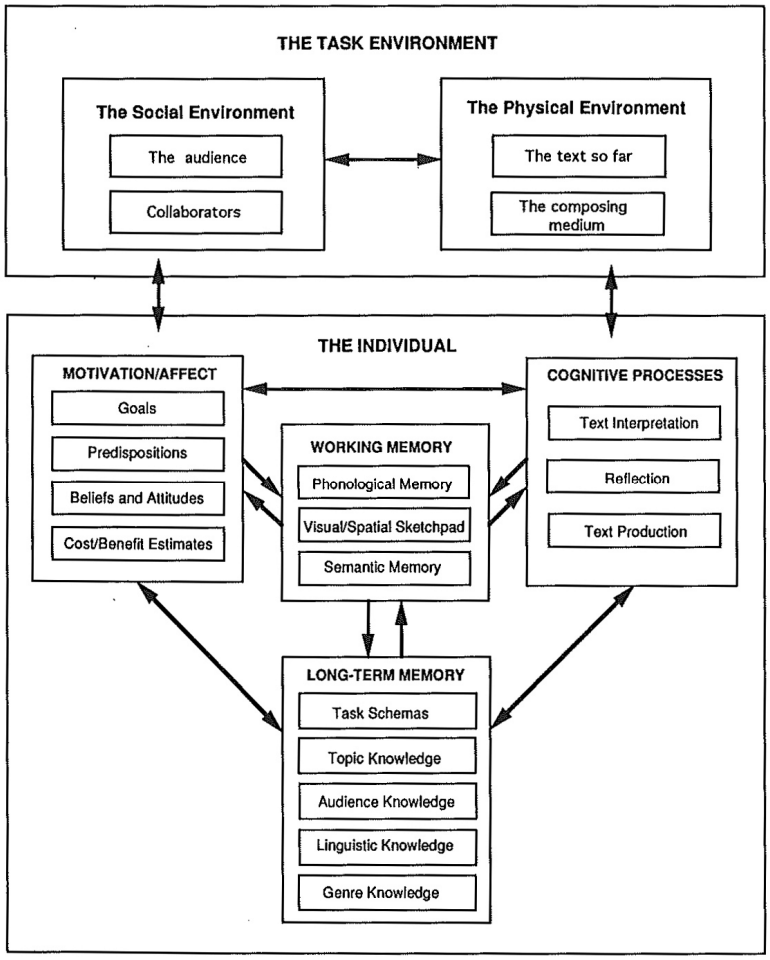


**Figure 2-2** Hayes' (1996, p.4) framework for understanding cognition and affect in writing.

In addition, Hayes has suggested that, based on insights from research from Hayes, Flower, Stratman and Carey (1987), revision should not be treated as a separate, basic process, but instead as a combination of text interpretation, reflection and text production. The key feature of the new text interpretation component is that revision does not involve a process of reading to comprehend, but that it is, instead, concerned

with the detection of problems in the text. Revision, in the revised model, is now seen as a more elaborate process of evaluation and modification of text and plans for the text. Finally, Hayes has stressed the importance of social and motivational factors in his new model.

#### *2.1.2.1 Model of text production*

On the basis of research from Kaufer, Hayes and Flower (1986), Hayes proposed a provisional model of text production (see Hayes 2009 for a review). A sketch of the model, taken from Chenoweth and Hayes (2003) is shown in figure 2-3.

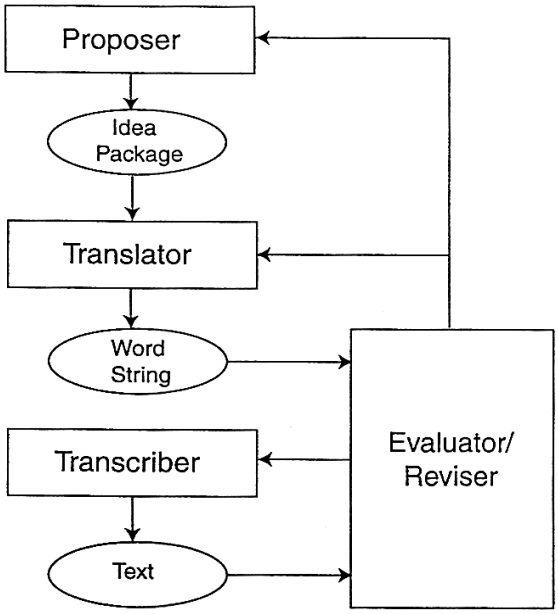


**Figure 2-3** Model of the text production process from Chenoweth and Hayes (2003, p. 113).

The model distinguishes between four different processes. The proposer proposes ideas for expression. This component is assumed to include the higher level processes involved in planning and reflection as characterised in global models of the writing process, and is responsible for creating an idea package to be formulated in language. Essentially it is responsible for deciding - in the context of the writer's goals, the text he has written so far, and the specified writing task - what to say next. The translator is responsible for converting this message into linguistic strings. The transcriber is responsible for converting linguistic strings into written text. A strong initial assumption of the model is that these processes occur in sequence, with each process completing before the next one begins (although see Hayes, 2009, and some comments in chapter 4 of this thesis for some discussion of this issue). Finally, the evaluator /

reviser is responsible for monitoring and evaluating word strings and text as they are produced, and for revising them when this is required.

It is important to note that these processes are embedded within the overall model of the writing process, and that the way they are combined will therefore depend on the writer's task schema, as well as on the availability of working memory resources. In this context, task schemas refer to packages stored in long term memory that included information about how a particular task can be carried out. Thus, for example, one can imagine writers who prefer to work out their ideas in detail before starting to write full text, running through lengthy cycles of translation, which are only followed by transcription when important points are identified. Or perhaps this process would be better characterised as lengthy cycles of reflection, with only salient parts being formulated in explicit word strings. Similarly, the extent to which the reviser is involved will depend on whether the writer is trying to produce well-formed text or a rough draft that will be evaluated and revised later. Chenoweth and Hayes (2003), for example, point out that the task schema that children have of writing may not include revision, and that they therefore may proceed directly from the expression of ideas to proposing the next idea.

To summarise, the original model of Hayes and Flower focused on the goal-directed nature of thinking behind the text, and treated the translation of ideas into text as a passive process (Hayes & Flower, 1980, Flower & Hayes, 1980b). More recent research has paid more attention to the processes involved in text production (Kaufer, Hayes & Flower, 1986; Chenoweth & Hayes, 2003). This has led to a revision of the original Hayes and Flower model in which processes are presented as more interrelated with the processes in other cognitive tasks, such as reading for comprehension (Hayes, 1996; Hayes, 2009). Thus, planning has become one component of a more general 'reflection' component. Translation is subsumed under the more general text production process and reflects a less passive role of text production. Finally, revision is not included as a separated process in its own right, but is incorporated within the basic processes of text interpretation, text production and reflection. Finally, working memory has been given a central role in the new model.

We will now turn to the description of Bereiter and Scardamalia's (1987) writing process theory, which also emphasis problem-solving processes as the key feature for development of understanding through writing.

### *2.1.3 Bereiter and Scardamalia*

Bereiter and Scardamalia (1987) contrast a knowledge-telling mode of writing, which is used by children and novice writers, and a knowledge-transforming mode of writing typically employed by more expert adults in which writers engage in the sort of problem definition and goal setting described by Flower and Hayes (Keys, 2000; Galbraith & Torrance, 1999a). According to Bereiter and Scadamalia's model, the development of ideas during writing depends on the extent to which the retrieval of content is strategically controlled in order to satisfy rhetorical constraints.

Novice writers are assumed to employ a knowledge-telling strategy (see figure 2-4). Basically, knowledge-telling is a method of composing in which writers retrieve content and ideas associatively and directly afterwards translate it into the textual form. The resulting text reflects the structure of the writer's knowledge and this knowledge is only modified as much as is required in order to conform to text demands. Since these writers do not adopt higher level goals they only pay a lot of attention to surface features of the text.

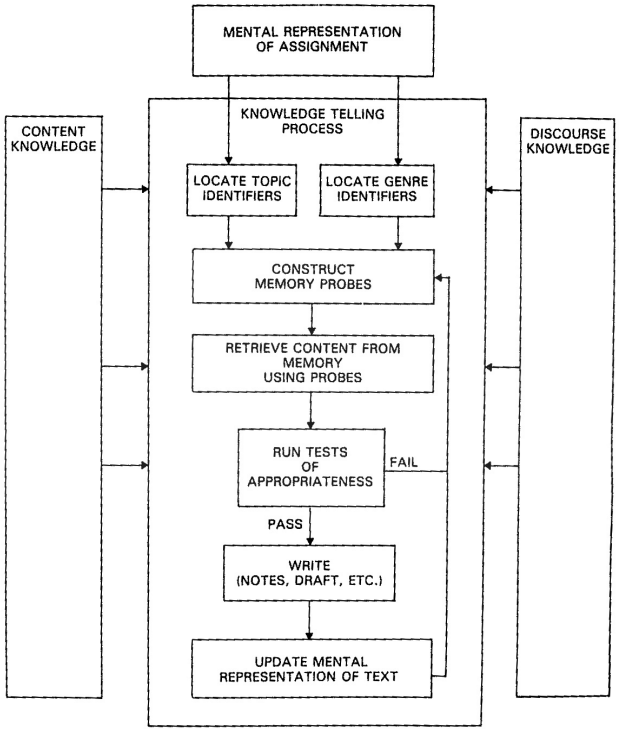


**Figure 2-4** Bereiter and Scardamalia's knowledge-telling model. Taken from Bereiter & Scardamalia (1987, p.8).

In contrast, knowledge-transforming strategies involve actively designing a text to satisfy communicative goals with respect to the reader. Planning becomes much more elaborate. And revising is more extensive because they are directed towards the writer's underlying goals (Galbraith & Rijlaarsdam, 1999).

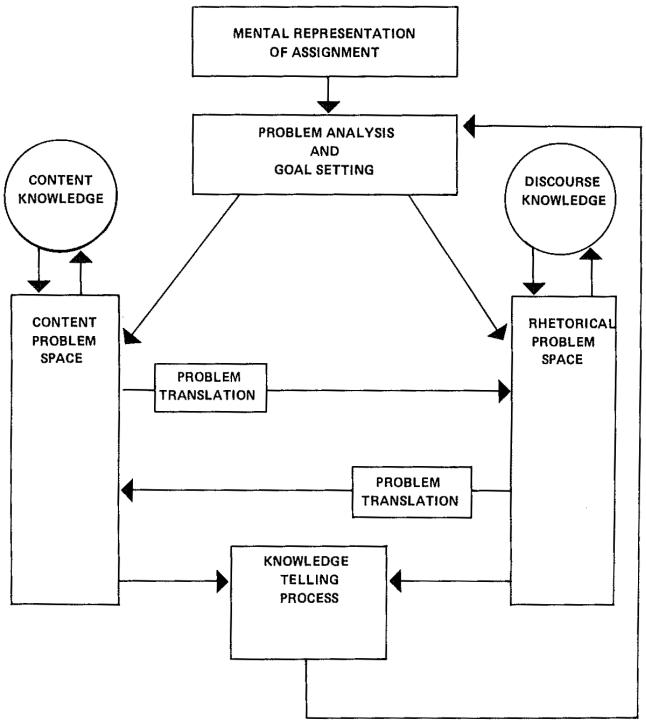


**Figure 2-5** Bereiter and Scardamalia's knowledge-transforming model. Taken from Bereiter & Scardamalia (1987, p. 12).

The most distinctive feature of the knowledge-transforming model is the dialectic between two problem spaces (see figure 2-5). On the one hand, Bereiter and Scardamalia (1987) distinguish the content space – in which knowledge states can be characterized as beliefs and in which writers pursue content goals related to the question '*what do I mean*'– and on the other hand they distinguish a rhetorical space – where knowledge states are represented as mental representations of actual or intended text. In the rhetorical space writers address compositional goals which are related to the question '*what do I say*'. This rhetorical space, therefore, contains knowledge states, which can be seen as plans for achieving various purposes in composition (Flower & Hayes, 1980; Bereiter & Scardamalia, 1987). The dialectic

between these two spaces is intended to reflect two important features of the writing process: (i) ideas are represented in relation to their rhetorical function in the text (rhetorical space) and not just as a direct translation of the writer's knowledge (content space); (ii) writing is not just a process of adapting content to rhetorical goals, but is, instead, presented as an emergent process in which both the text as well as goals for the text are developed as the text is produced.

According to Bereiter and Scardamalia (Bereiter & Scardamalia, 1987) knowledge-transforming through writing happens because writers generate sub goals in each of these problem spaces that provide input to the other problem space. In other words, the dialectic nature of their model - ideas are translated from the rhetorical space into the content space and vice versa - leads to the development of the writer's understanding through writing (Bereiter & Scardamalia, 1987; Klein, 2004). To give an example of this rather abstract notion: According to Bereiter and Scardamalia expert writers set rhetorical goals, such as the idea to convince the reader about a particular claim. In order to satisfy these goals they then set sub goals in the content problem space to serve these rhetorical goals and eventually develop sub goals in the content problem space to find evidence for their claim. Next, the writer has to work out this content problem by finding information that supports his or her idea and thereby the writer develops his or her understanding by linking new information to it. This new understanding is then transformed back to the rhetorical problem space which makes the model dialectic (Klein, 1999).

In their empirical research, Bereiter and Scardamalia (1987) found that expert writers plan more extensively before they start composing in comparison with novice writers. Additionally, the plans that both expert and novices make differ in scope. Expert writers plan at such a level of abstraction that they will be able to use the ideas in the text whereas novice writers normally make more general plans that are less useful to guide the writing effectively or their plans are so specific that they largely duplicate the final text. Similarly, during revision, novice writers focus more on low level revisions whereas expert writers focus, in addition, on more global features of the text such as organisation and coherence. Since expert writers revise more extensively, higher order demands are placed on these writers. These higher order demands have different consequences for the strategies that these writers apply. This relates to another aspect of meta-cognitive theories of writing research, namely the effective orchestration of the different processes involved, to which we will now turn.

### *2.1.4 Outlining*

Another recurring ingredient in successful writing, within the context of the meta-cognitive oriented theories, is being able to manage the wide range of processes involved within the constraints of a limited capacity of working memory (Hayes, 1996; Kellogg, 1994). Therefore, a dominant theme of research on writing has been on identifying effective strategies that help writers to reduce the cognitive load during text production (Kellogg, 1988). According to Kellogg (1988) both the written product as well as the processes suffer when the writer is overloaded. He claims, therefore, that planning and translating processes are carried out more effectively when they are carried out separately from each other (Kellogg, 1994). Studies into drafting strategies have focused on the effects of separating higher level thinking processes from the linguistic processes of writing in order to ease up the cognitive load for writers (Kellogg, 1988; Galbraith, 2009). Kellogg (1988, 1990, and 1994) conducted a series of studies from which he concluded that the benefits of outlining are twofold: it enables writers to organise ideas before writing and it makes more resources available for text production. Kellogg (1988, 1990) also found evidence that making an outline led to higher quality texts, because this allows writers to focus on the generation and organisation of ideas before shifting to translating these ideas into written text. In one of these studies, Kellogg (1988) used directed retrospection to identify the distribution of different cognitive processes while writing and he divided time spent on writing into thirds. The results of this study showed that students who made an outline prior to writing devoted less time to planning during the writing task itself and that those writers, subsequently, devoted more time to translating.

To conclude, the meta-cognitive processes described attribute the development of understanding through writing to the deliberate adaptation of thought towards rhetorical goals. Thus, both the knowledge-transforming model of Bereiter and Scardamalia (1987) and the Hayes and Flower model (1980) assume that writing can lead to increases in knowledge as a consequence of engagement in problem solving processes (Klein, 1999). The Flower and Hayes model assumes learning through dialectical processes within a content space (Hayes & Nash, 1996). In contrast, Bereiter and Scardamalia assume that this problem solving process involves a two-way dynamic interaction between a content problem space and a rhetorical problem space. Nevertheless, both models assume that writers construct knowledge by setting rhetorical goals, then from these writers derive content sub goals and finally, writers revise their top level rhetorical goals to accommodate these content goals (Klein, 1999). In addition, discovery is seen as a consequence of an effective way of

orchestration the different processes involved in writing. The most effective way to do this is when the higher order processes are separated from the lower order text production processes (Kellogg, 1988; Galbraith, 2009).

As noted before, research concerning these two models compares the writing strategies from expert and novice writers. Both models explain the creation of rhetorically well written texts and both models compare and contrast the writing strategies of expert and novice writers, but the assumption that the writing strategies which expert writers use actually contribute to knowledge change has not been tested empirically (Klein, 1999).

### *2.1.5 Text production oriented theories*

In contrast to the meta-cognitively oriented theories, there are theories that focus on the spontaneous text processes of writing. These theories claim that the development of understanding is inherent in the processes of written expression (Klein, 1999; 2004). Ideas emerge in the text as it is produced instead of being something that directs text production. For these theories, the characterization of writing as a process of discovery is, therefore, derived from the idea that a writer can use composition in order to find out what to say (Klein, 1999).

Britton (1982) proposed the 'shaping at the point of utterance' hypothesis. This hypothesis assumes that writer's tacit knowledge becomes explicit to them at the point of utterance. Hence, they only find out what they think when they write it down. Shaping at the point of utterance is an aspect of expressive or free writing during which the cognitive load is minimal and which is accessible for all kinds of writers. In general, Britton preferred expressive writing over drafting and revision as a mode for learning, because expressive writing does not require a specific form of language and can therefore easily be used to generate new meanings (Klein 1999). Britton (1982) believed that processes that intervene with the spontaneity of language, such as structure or rhetorical goals, would affect the power of writing as a tool for learning. Within this context, Wason (1980) also claimed that discovery through writing is most likely to occur when no sequential organisation other than the one that occurs spontaneously is imposed on the writing of an initial draft.

It is important to note here that most of the evidence for theories that focus on spontaneous text production processes is anecdotal. In various research papers expressions of famous writers are quoted to support the claim that free writing contributes to the generation of new content and subsequently to new knowledge (Murray, 1978; Galbraith, 1999; Klein, 1999). A famous example is the quote from Wright Morris "*the language leads, and we continue to follow where it leads*". (In: Murray,

1987, p. 101-102). Another famous quote exemplifying the claim that ideas do not exist prior to their verbal expression comes from E.M. Forster who said: "*How do I know what I think until I see what I say*" (In: Murray, 1987, p 101-102).

Advocates of spontaneous text production theories stress the importance of free writing for the development of understanding through writing. Free writing is a writing technique in which a person writes continuously for a set period of time without regard to spelling, grammar, or topic (Elbow, 1981). In general, Elbow suggests that writing involves two different kinds of process: creating and criticizing. According to Elbow, these two processes conflict with each other. Consequently, Elbow claims that the writing process should be separated into two stages. First, writers should write everything down as it comes to mind and without any constraints. Secondly, writers have to critically reconsider what they want to say. Therefore, Elbow stresses the importance of intuitive processes in the first half of the writing process and explicit, reflective processes in the second half of the writing process.

Overall, then, Elbow (1981) suggests that writers should first externalise their ideas spontaneously in text and only turn to critically assessing the text with respect to rhetorical constraints once this has been done Although this approach has been highly influential (Klein, 1999), there are only a few empirical studies that have tested whether spontaneous writing does in fact lead to conceptual change (Galbraith, 1992). The evidence that spontaneous text production theories contribute to learning, therefore, remains limited (Klein, 1999).

#### *2.1.5.1 Testing discovery through writing*

Galbraith (1992, 1996), set out to test this account of discovery through writing empirically. According to Galbraith, discovery occurs when writers attempt to express their disposition in language spontaneously, free from external constraints. This relates to the theories which assume that discovery is a consequence of the dialectical interaction between the writer's implicit disposition towards the topic and the emerging text as articulated by the spontaneous expression of ideas in language.

Galbraith (1992) compared two groups of writers - low and high self-monitors– who were asked either to write an essay or to make notes in preparation for an essay. Galbraith administered Snyder's self-monitoring scale (Snyder & Gangestad, 1986) to distinguish between high and low self-monitors. Snyder (1974) showed that individuals differ in the extent to which they monitor their expressive behaviour and in which they control their self-presentation. High self-monitors are people who are particularly sensitive to the expressive behaviour and/or presentation of others in social situations and, therefore, high self-monitors will use cues from the social

context to control their own expressive behaviour. Low self-monitors, on the other hand, are people who are less concerned about the appropriateness of their self-presentation and, therefore, pay less attention to the presentation of others. Their self-presentation or expressive behaviour will more likely be controlled from within their affective state and hence, they will express their behaviour as they feel it. Galbraith applied the self-monitoring scale to writing, to distinguish between writers whose writing is supposed to be more or less directed to rhetorical goals (high self-monitors) or dispositional goals (low self-monitors), where dispositional goals refer to direct reflections of the internal state of the writer.

The self-monitoring scale consists of 18 statements which participants have to rate as either true or false. Examples of statements are '*I have trouble changing my behaviour to different people and situations*' which the high self-monitor is expected to deny, or '*I would make a good actor*', which the high self-monitor is expected to confirm. Low self-monitors are supposed to answer these statements the other way round, so they would confirm the first statement and deny the second statement. This setup allowed Galbraith to measure the extent to which both high and low self-monitors developed new ideas as a function of writing in these two different writing conditions (notes versus full text production).

Before writing, the participants were asked to list all the ideas they could think of, to rate the importance of these ideas and to indicate how much they felt they knew about the topic. Immediately after writing, Galbraith (1992) asked them again to indicate how much they felt they knew, to list their ideas and to rate the importance of these ideas. If discovery through writing is - as the meta-cognitive theories claim - a consequence of active rhetorical problem solving, one would expect the high self-monitors to produce more new ideas than the low self-monitors, since these writers are assumed to redirect their goals towards rhetorical goals. This difference should be most clear in the notes condition, because in this condition writers can focus on planning without having to focus on text production. However, if discovery through writing is – as the text production theories claim – a consequence of finding out what you think in the course of writing, one would expect the low self-monitors to produce more new ideas than the high self-monitors, since these writers are assumed to prioritise the direct expression of their ideas in writing.

The results of this study showed that high self-monitors produced more new ideas after writing notes than low self-monitors did, which is in agreement with the knowledge-transforming model. However, these ideas were not associated with increase in understanding, contrary to what the knowledge-transforming model claims. In addition, Galbraith found that low self-monitors produced twice as many new ideas in the text production condition as the high self-monitors did. These ideas

were - contrary to the new ideas from the high self-monitors in the notes condition - associated with increases in subjective understanding about the topic. These results supported Galbraith's view that discovery depends on the extent to which text production is dispositionally driven (low self-monitors) rather than directed towards rhetorical goals (high self-monitors). Hence, it depends on whether writers express their implicit knowledge about the topic by means of spontaneous text production processes. By contrast, although rhetorical planning does lead to the retrieval of different content (high self-monitors' notes), this is not associated with increases in understanding. Galbraith explains this by suggesting that this involves the retrieval of different existing ideas from episodic memory once rhetorical goals have been imposed.

The results of these initial studies of Galbraith, then, appear to contradict important features of Bereiter and Scardamalia's knowledge-transforming model, since not only the assumption that knowledge-transforming depends on adaptation of content to rhetorical goals, but also the assumption that low self-monitors' writing is a simple matter of 'knowledge-telling' were challenged.

Galbraith's 1992 study did not manipulate the extent to which participants planned their writing in advance. In order to assess this, he set out to test the effects of pre-planning on the development of understanding (Galbraith, 1999). In this study he distinguished between three different planning conditions; (i) non-outline or synthetic planning, which involved writing down a single sentence summing up one's overall opinion before writing, (ii) outline planning, which involved making an organized outline before writing, and (iii) an unplanned condition, which involved starting writing immediately, without pre-planning, and writing out ideas as they came to mind in connected sentences but without worrying how well the ideas were expressed.

The results of this study showed that, similar to the 1992 findings, low self-monitors produced more new ideas after writing than high self-monitors and this was most pronounced in the synthetic planning condition. Furthermore, this was the only condition in which new ideas were associated with the development of understanding. Nevertheless, the results also showed that writers in the other conditions also produced new content. Galbraith, therefore, concluded that new content can be produced by both explicit planning processes as well as dispositionally driven text production, but that the effects of dispositional text production are reduced by outlining, whereas the effects of explicit rhetorical planning are enhanced by outlining.

Galbraith, Torrance & Hallam (2006) repeated the above experiment with some minor changes to the methodology. The results of this experiment supported the

results of the previous experiments and also showed that low self-monitors produced more new ideas within the synthetic planning condition than the high self-monitors did, supporting the assumption that discovery depends on dispositionally guided text production, and is reduced when external rhetorical constraints are imposed. Furthermore, the high self-monitors produced more new ideas in the outline planning condition than in the synthetic planning condition, which supports the assumption outlining facilitates explicit planning designed to satisfy rhetorical goals.

An important feature to note about this experiment is that the synthetic and outline planning conditions differed in a number of respects. In the synthetic planning condition, participants were instructed to write a rough draft. Writing a rough draft involved writing down thoughts as they came to mind, in continuous prose, but without worrying about how well organised or well expressed their texts were. In contrast, in the outline planning condition, participants were instructed to write a well organised essay, and communicating ideas to readers as clearly as possible. This means that differences between the synthetic and outline planned conditions could be a consequence of either the difference in the pre-planning activities or the differences in the nature of the draft that was then produced. Furthermore, in Galbraith's studies in general, the rhetorical context is typically left relatively unspecified. Participants are asked to write an essay, or a rough draft, but the particular rhetorical context is left unspecified. These issues will be taken up in the empirical element of this thesis.

Thus, with these studies Galbraith (Galbraith 1992; Galbraith, 1999; Galbraith, et al., 2006) investigated more directly how writers develop their understanding through writing. These appear to show consistent differences between low and high self-monitors, and between outline and synthetically planned texts in the extent to which new ideas are generated after writing. Furthermore, they suggest that new ideas are not necessarily associated with the development of writer's understanding. In the light of this Galbraith has developed an alternative dual process model of writing.

#### *2.1.5.2 The dual process model*

The dual process model of writing (Galbraith, 2009) suggests that, while writing, two processes occur simultaneously. One is a rhetorical process similar to the knowledge-transforming process described by Bereiter and Scardamalia (1987). In addition, the model includes a more spontaneous text production process which allows sentences to be produced by synthesizing information within semantic memory. Galbraith (2009) describes this as a knowledge constituting process in which content is synthesized in the process of language production.

The dual process model of writing suggests that both rhetorical processes as well as dispositional guided text production processes are required for effective communication. The problem solving part of the dual process model works on the principles of a knowledge retrieval process. Content is assumed to be retrieved from episodic memory and can be manipulated in working memory in order to create rhetorically appropriate content. This aspect of the dual process model, therefore, overlaps with Bereiter and Scardamalia's knowledge-transforming model of writing (1987). It also acknowledges that explicit rhetorical planning is an important aspect of writing, since it requires writers to collect together a disparate set of ideas and organise them in relation to the rhetorical constraints of the writing task. This aspect of writing is tied to working memory limitations and Galbraith, consequently, argues that it operates best when it is not combined with full text production.

However, a more distinctive feature of the dual process model is its focus on the interaction between the writer's implicit disposition towards a topic and the explicit text. Galbraith (2009) describes this as a knowledge constituting process. He claims that dispositionally guided text production involves parallel constraint satisfaction within the writer's semantic memory.

This assumes that knowledge is implicitly stored in a semantic network which works according to connectionist principles. The connections between the units within the semantic network are assumed to correspond to the writer's disposition towards the topic. This disposition is fixed, but it can produce varying output depending on the input to the network. The crucial thing about this disposition it that knowledge in this network is stored implicitly and hence, writers are unable to actively retrieve this knowledge from memory. Depending on the input to the network, writers compose successive sentences in which parts of their implicit knowledge are constituted. Since the feedback to this semantic network consists of an inhibition process each further utterance can be composed of other aspects of the disposition that were previously suppressed. This process will stop when the writer has sufficiently covered the topic of the writing task and hence, when the output that is produced constitutes the writer's understanding of the topic.

To make this process a bit more comprehensible, a sketch of the knowledge constituting process in shown in figure 2-6 (Galbraith, 2009). It should be noted that Galbraith stresses that this representation of the knowledge constituting process is only intended as a rough description of processing principles. It should not be taken as a detailed description of a set of explicit processing stages. In figure 2-6, the writer's disposition is represented by the network of interconnected units in the middle of the diagram. Knowledge in this kind of network is stored implicitly as the connections between the units. The inputs to this network are the external constraints. In this case

the topic of the writing task and the task specifications. These external constraints activate the units within the network, which then pass activation to one another. Over a series of cycles within the network, in which activation is passed back and forth between the units, the network gradually settles into a stable state. It is this process of parallel constraint satisfaction which Galbraith describes as a synthetic process, and which is contrasted with the retrieval process involved in accessing content in episodic memory. The final state of activation of the units corresponds to the message that the writer wants to convey. This output is then formulated into language as presented by the arrows labelled with A. It should be noted that the content of this message only becomes apparent to the writer at the point of utterance. Here, tacit knowledge has been turned into explicit knowledge. The way in which this message is transformed into language is not formally explained in the sketch of the model. It could be produced by any text production process model available. The crucial thing being that the process involves the synthesis of ideas guided by the implicit organization of semantic memory (as represented by the strength of the connections between units within the network) rather than the retrieval of already formed ideas from episodic memory. In the diagram, B represents the utterance that is produced by this process. This utterance it then, via C, fed back into the network. Together with the topic and task specifications it activates the network of connections in order to produce further content. This new message is represented in D of the diagram. This cycle is repeated in a sequential manner. What is produced in B, D, or F is considered new content when the message that is produces does not relate to ideas stored in episodic memory.

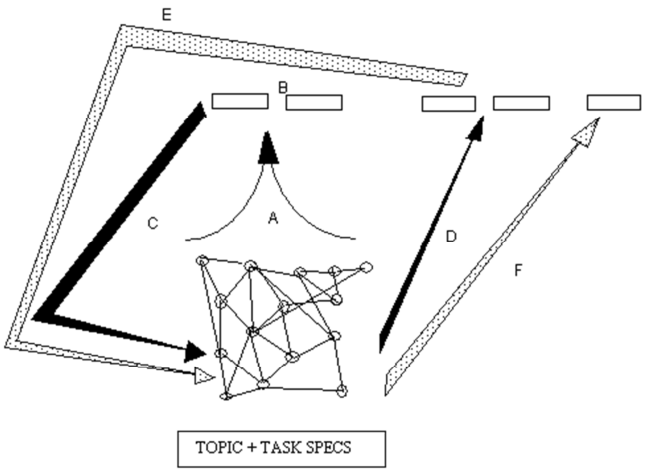


**Figure 2-6** Visual representation of writing as a knowledge constituting process (Galbraith, 1999).

An important thing to note is that the language production component of this knowledge constituting process has limited capacity and, therefore, any specific utterance only partially represents the implicit knowledge stored in the network. When this content is fed back into the semantic network, it is assumed to involve inhibition of the content corresponding to the message that has just been formulated, and hence the writer will produce a message corresponding to the remainder of the content implicit in semantic memory. Therefore, this additional run through the network will produce a message that is a refinement or modification of what has been said before. This inhibitory process, therefore, allows one to produce further content without the need for further explicit planning. Depending on the activation of the units in the network this content can contain a more specific message of the previous message or instead composes content which was hidden before. The crucial feature of this process is that the writer's implicit knowledge is constituted across a series of synthetically produced utterances.

Overall, then, the dual process model incorporates an explicit organising process and an implicit text production process. The crucial feature being that the explicit organising process is assumed not to lead to changes in the understanding of the writer. This contradicts important features of the knowledge-transforming model. The 'knowledge retrieval' process draws on an explicit store of pre-existing ideas in episodic memory, and hence, can, at most, lead to the reorganisation of existing knowledge. Nevertheless, Galbraith assumes that this represents a vital part of the writing process, because this process will enable writers to take a set of ideas and organise them in a coherent manner in respect to the rhetorical goals of the writing task. However, because this system operates on the basis of existing knowledge, it is, unable to lead to discovery through writing.

Galbraith contrasts this knowledge retrieval process with the knowledge constituting process shown in figure 2-6. This implicit text production process is guided by the implicit organization of knowledge stored in semantic memory and hence, involves the synthesis, rather than retrieval, of content. Furthermore, the process that guides text production is assumed to produce further content through inhibitory feedback from preceding utterances and hence, enables writers to produce sequential output without explicit goals to drive the process. This process is seen as intrinsically a process of discovery, since it makes the tacit knowledge of the writer explicit and accessible to the writer. The extent to which this process leads to developments of understanding depends on the extent to which content that is produced relates to content stored in episodic memory. When it produces content that is already stored in memory, the writer will not experience developments of

understanding; when the content is unrelated to ideas already stored in episodic memory, then the writer will experience developments of understanding.

Finally, Galbraith argues that the explicit organising process and the implicit text production process are both required for effective communication. The explicit planning process is required to impose structure on the text, to set explicit goals, and to tailor the text to the needs of the reader; the implicit text production process is required to constitute the writer's personal understanding of the topic. However, because the two processes operate according to different principles, they lead to a fundamental conflict in writing between top-down, controlled processes, which operate best on a fixed, economical representation of ideas, and bottom-up, spontaneous processes, which operate best when ideas are constituted, discursively, in the course of text production. The model assumes that this conflict can be reduced by employing different kinds of drafting strategies (see Galbraith and Torrance, 2004): outlining is assumed to optimise explicit organising processes, whereas rough drafting is assumed to optimise knowledge-constituting processes, and writers are assumed to switch between these two strategies depending on which process is currently prioritised. Furthermore, Galbraith suggests that different writers prioritise the two processes differently. Low self-monitors are assumed to prioritise dispositional goals when writing and are, therefore, more likely to express their ideas directly as they unfold, and to then to re-organise this content to satisfy rhetorical goals. By contrast, high self-monitors are assumed to prioritise rhetorical goals and are, therefore, more likely explicitly organising their ideas before formulating them in full text.

To conclude, in order to account for the patterns of change in ideas observed in Galbraith's research, and the way that these varied under different writing conditions Galbraith proposed the dual process model. To date, however, the predictions of the model about the processes involved have not been directly tested.

### *2.1.6 Conclusion*

In the literature review, it became clear that Bereiter and Scardamalia (1987) and Hayes and Flower (1980) provided the field with detailed descriptions of the cognitive processes of expert and novice writers, but never tested directly whether the processes of expert writers, even though they are expected to, contribute to the development of understanding through writing. Galbraith (2009) has measured development of understanding through writing subjectively and found consistent differences as a function of self-monitoring and type of pre-planning, differences which he claims are incompatible with the assumptions of the knowledge-transforming model about the

relationship between processes and knowledge change. Galbraith therefore proposed a different model of the processes involved. These approaches attribute the epistemic effects of writing to different processes. Galbraith (2009) claims that development of understanding depends on the extent to which text production is dispositionally driven. On the other hand, the meta-cognitive theories claim that development of understanding depends on the deliberate adaptation of thought to rhetorical goals.

The aim of this thesis is to use keystroke logs as an indicator of processes to explicitly test the contrasting predictions of the two models about the relationships between processes and the development of understanding.

In order to test the claims of the different models, this thesis will describe the results of an experiment based on the design of the experiment by Galbraith et al. (2006). This experiment manipulated two variables - self-monitoring and planning - and assessed their effects on writing processes, as well as measures of idea change and the development of subjective understanding. Writing processes were assessed using the keystroke logging program, Inputlog (Leijten & van Waes, 2006).

Over the last 10 years, the focus on the underlying writing processes has become an essential part of writing research. Different keystroke logging programs have been developed that register and reconstruct the writing processes of writers who compose a text on the computer. These programs log the writing process in great detail without being intrusive. Keystroke logging tools register all keystrokes and mouse clicks and movements and enable researchers to analyse the pausing, translating and revision behaviour of writers. Inputlog (Leijten, & van Waes, 2006) – one of the keystroke logging tools available nowadays – enabled us to compare and contrast the writing processes of high and low self-monitors producing text in different writing conditions. This method will therefore allow us to test the claims of different writing accounts and to investigate in more detail which processes are responsible for the epistemic effect of writing.

# Chapter 3

## *The Effects of Writing Beliefs and Planning on Text Quality and Understanding through Writing*[1]

[1] Submitted as: Baaijen, V.M., Galbraith, D., & de Glopper, K., *Effects of beliefs and planning on text quality and discovery through writing*.

## *Abstract*

White and Bruning (2005) identified two independent sets of beliefs about writing – transactional beliefs and transmissional beliefs – and suggested that they are associated with important effects on writing performance. This paper describes an experiment designed to investigate the effects of these beliefs on text quality, the development of writer's understanding, and the extent to which text is modified during its production. In addition, we examined whether these effects are moderated by outline planning, and compared an outline planned condition with a non-outline planned (or synthetically planned) condition. The results show that the effects of writing beliefs vary depending on type of planning. Within the synthetic planning condition, high transactional and low transmissional writers produced better quality text, and high transactional writers experienced greater increases in understanding than low transactional writers. By contrast, in the outline condition, low transactional writers produced text similar in quality to the high transactional writers, and high transactional writers were less likely to experience increases in understanding. Further analysis suggested that, for writers with low transactional beliefs, text modification was associated with the production of lower quality text, and with only moderate increases in understanding; for writers with high transactional beliefs, text modification was strongly related to increased understanding but unrelated to text quality. Overall, these results confirm that writing beliefs have strong effects on text quality, and in addition suggest that they are related to the development of understanding during writing. However, we suggest that the nature of these beliefs needs to be reinterpreted and that the scales used to measure these beliefs need to be refined.


*Keywords:* writing beliefs, writing strategies, text quality, writing processes, development of understanding

## 3.1 INTRODUCTION

Performance of most higher level cognitive tasks is likely to be strongly influenced by people's conceptions of what the task involves. This is particularly true of a complex task like writing. At its most basic, writing involves knowing how to represent spoken language in a visual form: the writer has to know what symbols can be used to represent the sounds making up spoken language (for alphabetic scripts at least) and how to use punctuation to mark the boundaries between different conceptual and linguistic units. Someone who knows these conventions would, in a strict sense, know how to write and be able to produce "text". However, if writing is to be a full-fledged meaningful activity, the writer will have to manage a wide range of further cognitive processes. Hayes and Flower (1980) identified three different kinds of basic writing processes: (i) planning (setting goals, and generating and organising content); (ii) text production – formulating content in words and sentences); and (iii) reading and revising the text that has been produced. The ability to manage the wide range of processes involved within the constraints of a limited capacity working memory (Hayes, 1996; Kellogg, 1996) is a key ingredient of successful writing. But, how these processes are managed depends ultimately on how the writer defines writing as an activity in itself and on the goals towards which he believes writing can be directed.

Although this has long been recognised (e.g. Bereiter and Scardamalia, 1987; Hayes, 1996), there has been relatively little research into individual differences in conceptions of writing, or into how any such differences in conception relate to either the process of writing or to the effectiveness of instructional interventions. Recently, however, White and Bruning (2005) have developed a writing beliefs inventory designed to capture individual differences in writers' implicit beliefs about writing, and found that these are strongly related to differences in the quality of the writing that individuals produce. Our aim in this paper is to explore this relationship further, focussing in particular on how differences in these beliefs interact with different writing strategies. In the remainder of the introduction, we will first describe the nature of the writing beliefs identified by White and Bruning and point out some problematic features of the scales that they developed. We will then consider how these beliefs map onto models of the cognitive processes involved in writing, focussing particularly on the processes involved in the development of understanding through writing. We will then describe the results of an experiment designed to assess the relationship between implicit writing beliefs, different forms of pre-planning, and three different aspects of writing performance: (i) text quality; (ii) the development of understanding during writing; and (iii) the extent to which text is modified during its production.

### *3.1.1 Implicit writing beliefs*

White and Bruning (2005) proposed a distinction between transactional and transmissional beliefs about writing and suggested that these have important effects on writers' performance. This distinction had its roots in earlier research by Schraw and Bruning (1996, 1999) investigating the effect of implicit beliefs on reading behaviour. In this context, transmissional beliefs were assumed to involve a perception of text as an objective reflection of the writer's intentions, and the role of the reader as being able to extract the information contained in the text. Readers with these beliefs showed less cognitive and affective engagement during reading and a less developed understanding of the text. By contrast, transactional beliefs were assumed to involve a conception of reading as an act of interpretation in which meaning is constructed through relating the text to the reader's own beliefs and understanding. Readers with these beliefs show higher levels of cognitive and affective engagement with the text and developed a more complex understanding of the meaning of the text.

White and Bruning (2005) investigated whether writers have corresponding transmissional and transactional beliefs about writing and created a writing beliefs inventory designed to identify them. In their words, "writers with predominantly transmissional writing beliefs . . . envision writing as a way to transfer information from authoritative sources to the reader in a manner that limits how the writer's ideas are reflected in the text." (White and Bruning, 2005, p. 168), and are assumed therefore to show lower levels of cognitive and effective engagement during the writing process. By contrast, "writers with predominantly transactional writing beliefs . . . view the purpose of writing as a way to personally and critically construct the text by actively engaging their own thinking into the process." (ibid, p. 168), and are assumed therefore to show higher levels of cognitive and affective engagement during the writing process.

The writing beliefs inventory consists of two scales - a transmissional scale and a transactional scale – which are uncorrelated with one another. This feature, which is shared with the corresponding reading beliefs scales, suggests that it is possible for writers to hold both sets of beliefs at the same time. While this is not in itself problematic, it does make the interpretation of the scales complex. The definitions provided above refer to particular combinations of scores on the two component scales. "Predominantly transmissional beliefs" are equated with a combination of high scores on the transmissional sub-scale and low scores on the transactional sub-scale. "Predominantly transactional beliefs" are equated with a combination of high scores on the transactional sub-scale and low scores on the transmissional scale.

Conceptually, this is rather problematic as the two overall configurations are treated as if they are the opposite ends of a single dimension, and at first sight this seems to imply that the two sub-scales should be negatively correlated with one another, particularly as the other possible configurations of beliefs (high transactional-high transmissional and low transactional-low transmissional) are not given any interpretation and are not associated with clear cut predictions about relationships with writing performance.

Having developed and refined these measures of transmissional and transactional beliefs, White and Bruning (2005) then assessed how they related to the quality of the text produced in a task involving the discussion and analysis of a short narrative story. They also collected data on a number of other scales, and responses to a range of questions about the participants' writing experience. Their main finding was that writers with high transactional beliefs produced better quality text than writers with low transactional beliefs, and that writers with low transmissional beliefs produced better texts than writers with high transmissional beliefs. There was no interaction between these variables. In other words, the two sets of beliefs had independent and additive effects on writing quality, as predicted by their assumption that each type of belief makes an independent contribution to the writer's cognitive and affective engagement with the task. In addition, they found that writers with high transactional beliefs had greater confidence in their writing ability (efficacy), were more likely to view writing as a means of self-expression, and more likely to write for pleasure than writers with low transactional beliefs. Writers with high transmissional beliefs were less likely to write for pleasure than writers with low transmissional beliefs.

### *3.1.2 Relationship with writing processes*

White and Bruning's main aim was to establish whether different writing beliefs could be identified and then whether these had consequences for writing performance. They therefore had relatively little to say about how transactional and transmissional beliefs lead to differences in text quality, beyond the suggestion that high transactional and low transmissional beliefs should be associated with higher levels of cognitive and affective engagement. In this section, we briefly consider current cognitive models of writing, focussing in particular on how they account for the development of personal understanding during writing. Since this is an explicit feature of the transactional beliefs scale but not of the transmissional scale, we will focus primarily on relationships between cognitive processes and the transactional

scale, and restrict ourselves to incidental comments about relationships with the transmissional scale.

Within research on the cognitive processes involved in writing, there is a general consensus that effective writing is a matter of actively developing thought in the course of writing rather than of simply transcribing preconceived thought into words. To that extent, there is a clear correspondence between the processes emphasized by cognitive models of writing and the transactional beliefs described by White and Bruning. However, within this general consensus, there are two contrasting accounts of how writing leads to the development of understanding.

Problem solving models of writing, as exemplified by Bereiter and Scardamalia (1987), focus on explicit problem solving processes designed to satisfy rhetorical goals. Thus, Bereiter and Scardamalia (1987) contrast two conceptions of writing. In the knowledge-telling model of writing, typically associated with younger, less expert writers, the writer is assumed to focus on translating material available in long term memory directly into words, guided by the structure of the material stored in memory, and unmediated by communicative goals. In the knowledge-transforming model of writing, the generation of content is assumed to be mediated by the writer's communicative or rhetorical goals. Existing content is evaluated, and new content is developed, in order to satisfy the writer's communicative goals. This is assumed to have wide-ranging consequences: (i) the writing process becomes more elaborate and recursive as the writer proposes and evaluates content before writing (pre-planning), during text production, and during revision; (ii) the text is tailored to the needs of the reader rather than simply reflecting the structure of existing content in the writer's long term memory; and (iii) the writer, in trying to satisfy their rhetorical goals, adapts and develops their understanding of the topic they are writing about.

Given that the transactional beliefs scale contains items that refer to the development of understanding and extensive revision of text, there appears to be a straightforward correspondence with a knowledge-transforming model of writing. Indeed, Schraw and Bruning (1999, p.288) explicitly acknowledge this when they characterise the type of engagement associated with transactional beliefs as involving "Knowledge transformation (our emphasis); constructed text meaning; subjective analysis of text content and structure; personalized response; strong emphasis on holistic interpretation of text and reader affect; aesthetic appreciation". This would imply, therefore, that writers with high transactional beliefs are more likely to engage in deliberate problem solving to satisfy rhetorical goals whereas writers with low transactional beliefs are more likely to engage in knowledge-telling. The one major difficulty with this being that the transactional scale does not include items referring to rhetorical goals or to the writer's interaction with the reader. Indeed, the emphasis

is, if anything, on the writer's personal view of the topic, rather than on their communicative goals.

Turning to transmissional beliefs, Schraw and Bruning (1999, p. 288) also explicitly relate this to the knowledge-telling / knowledge-transforming contrast when they say that: "Transmission beliefs emphasize the role of the author and text in understanding. Readers engage in an objective, decontextualized understanding of the text and author that corresponds to knowledge transmission (Bereiter & Scardamalia, 1987)." In the context of writing, then, White and Bruning appear to be claiming that high transmissional writers, with their emphasis on accurately reflecting external sources, are more likely to engage in the less effortful form of writing associated with knowledge-telling than low transmissional writers.

Overall, then, it seems plausible to equate transactional beliefs with a knowledge-transforming approach to writing, and transmissional beliefs with a knowledge-telling form of writing. The one major dissimilarity being the absence of items on the transactional scale that relate to the writer's goals with respect to the reader, and a corresponding greater emphasis on the writer developing a personal view of the topic, as exemplified in Schraw and Bruning's (1999) characterisation of the type of engagement associated with transactional beliefs in reading.

Recently, Galbraith (1999, 2009) has proposed an alternative, dual process account of discovery through writing, which places a similar emphasis on the writer's personal view of the topic, and which therefore suggests an alternative account of the relationship between writing beliefs and writing processes. This claims that two different kinds of process are involved in effective writing. The first is an explicit problem solving process in which pre-existing ideas are retrieved from episodic memory and evaluated and organized in working memory in order to satisfy the writer's rhetorical goals. This is treated as equivalent to the knowledge-transforming model and plays an important role in organizing content and tailoring it to the needs of the reader. The key difference is that the model assumes that, by itself, this process does not lead to developments in understanding, but only involves the reorganization of existing content. The second process is an implicit, knowledge-constituting process in which content is synthesized according to the constraints within semantic memory and then transcribed as text. Over a series of cycles, this allows the writer to constitute their implicit knowledge in the text. The key feature of this process is that, because of its implicit nature, the writer is unable to predict the content that will be produced, and hence, in order to constitute their knowledge in the text, the writer has to allow text production to unfold free from external constraints. It is this process which is responsible for the development of the writer's personal understanding during writing.

According to the model, both processes are required for effective writing. The explicit organising process is required to impose structure on the text, to set explicit goals, and to tailor the text to the needs of the reader; the implicit text production process, which is guided by the implicit structure of semantic memory, is required to constitute the writer's personal understanding of the topic. However, because the two processes operate according to different principles, they lead to a fundamental conflict in writing between top-down, controlled processes, which operate best on a fixed representation of ideas, and bottom-up, spontaneous processes, which operate best when ideas are constituted, discursively, in the course of text production. The model assumes that this conflict can be reduced by employing different strategies at different points in writing (see Galbraith and Torrance, 2004). Thus, outlining is assumed to optimise explicit organising processes, whereas spontaneous text production is assumed to optimise knowledge-constituting processes, and writers are assumed to switch between these two strategies depending on which process is currently prioritised. Furthermore, it assumes that differences in the extent to which writers prioritise these processes are a key source of individual differences in approaches to writing.

The dual process model, therefore, suggests an alternative account of the relationship between writing beliefs and writing processes. Writers with high transactional beliefs, who emphasize the development of understanding during writing, extensive revision and a personal approach to the topic, are assumed to prioritise the implicit knowledge-constituting process. Writers with low transactional beliefs, by contrast, are assumed to prioritise explicit organising processes. This interpretation suggests that the key features of White and Burning's transactional scale are its emphasis on the personal nature of writing and it corresponding lack of emphasis on the need to take the reader into account. In the dual process model, it is precisely the implicit structure of semantic memory which is the source of the writer's personal understanding of the topic. The transmissional scale, by contrast, has a less clear interpretation in terms of the model. However, to the extent that high transmissional beliefs are associated with a less personal approach to writing, the dual process model would predict that high transmissional writers would be more likely to prioritise explicit organising processes.

Overall, then, the problem solving model and the dual process model provide two different accounts of the nature of engagement in writing, and hence provide different interpretations of the nature of transactional and transmissional beliefs about writing. The problem-solving model implies that the essential feature of engagement is the extent to which the writer engages in knowledge-transforming processes. By

contrast, the dual process model implies that the essential feature is the extent to which writing is driven by the writer's implicit personal understanding of the topic.

The key feature here is the dual process model's claim that outlining has a differential effect on the explicit organising process and the implicit knowledge-constituting process. Outlining is assumed to enhance the explicit organising process and hence, other things being equal, to lead to improvements in text quality (see also Kellogg, 1988, 1994). But it is also assumed to inhibit the implicit knowledge-constituting process and hence to reduce the development of understanding. Since, this is also assumed to contribute to text quality, this means that the overall effect of outlining on text quality will be different for writers who prioritise the two processes differently. Low transactional writers, who are assumed to prioritise the explicit organising process, and hence not to experience developments in understanding to the same extent as high transactional writers, should write better when they are able to make an outline before writing than when they have to try to organise at the same time as producing the text itself. By contrast, high transactional writers, who are assumed to prioritise the implicit knowledge-constituting process, and hence to experience developments in understanding during writing, should experience more ambiguous effects of outlining. It would be expected straightforwardly to reduce the extent to which they develop their understanding during writing, since, by pre-determining the order in which ideas are produced, it would reduce the extent to which writers would be able to follow the path of their understanding as it is constituted in the text. But the effects on text quality should be mixed, with increases arising from an enhancement of the explicit organising process, but decreases arising from the inhibition of the knowledge-constituting process. In general, therefore, the dual process model would predict that effects on text quality and the development of understanding should be dissociated from one another.

By contrast, the problem-solving model assumes that the development of understanding and text quality should be related to one another, since the process of knowledge-transforming is assumed to be responsible both for the development of understanding and the production of higher quality text. In general, on the assumption that high transactional beliefs are associated with a knowledge-transforming writing process, it would predict that high transactional writers should produce both better quality text and experience greater developments in understanding than low transactional writers.

### *3.1.3 Aims of the experiment*

Our first aim in the following experiment was to assess whether the relationship between writing beliefs and text quality that White and Bruning have found for narrative writing generalizes to an argumentative writing task. Our second aim was to investigate how writers' beliefs are related to the development of understanding during writing. We therefore asked writers to provide subjective ratings of their understanding before and after writing. Third, we wanted to assess whether writing beliefs are associated with different degrees of text modification during writing. For that reason, we used a simple measure of the extent to which text has been modified during the course of text production (collected using keystroke logging).

Our fourth aim was to assess how the performance of writers with different beliefs varies depending on the kind of planning they carry out before writing. Writers were asked either to make an outline before writing (outline planning) or to spend the same amount of time trying to sum up their overall opinion of the topic (synthetic planning). The synthetic planning condition was designed to occupy the same amount of time before writing as outline planning but not to involve specifying the order in which ideas would be produced in the text.

On the assumption that synthetic planning is essentially a non-outline planning condition, and hence is similar to the conditions under which White and Bruning's participants wrote, we expected, on the basis of their previous research, that in this condition text quality should vary as a function of both transactional and transmissional beliefs, and that these beliefs would have independent and additive effects. In addition, we expected that in the non-outline planning condition writers with high transactional beliefs would experience greater increases in understanding after writing than writers with low transactional beliefs. We do not expect there to be differences between high and low transmissional writers for this measure. Finally, to the extent that high transmissional beliefs are associated with knowledge-telling, we expected them to modify their text less during writing than low transmissional writers when writing in the synthetic planning condition. Since high transactional writers are assumed to revise their text more, we expected high transactional writers to modify their text in this condition more than low transactional writers.

Effects of outlining were expected to vary as we have described above. We expected that, if the problem solving account of the processes associated with writing beliefs is correct, the effects of outlining should be the same for the different measures. Given Kellogg's (1988, 1994) research on outlining, outline planned texts should be of higher quality than non-outline or synthetically planned texts, should show more evidence of increases in understanding, and should show less evidence of modifying

text during writing. By contrast, if the dual process model's account of the processes involved is correct, outlining should have different effects for low and high transactional writers, and should have differential effects on text quality and the development of understanding. Low transactional writers, who are assumed to prioritise the explicit planning process, would be expected to benefit from outlining because it would enable them to generate and organize their ideas separate from text production. This should not be associated with changes in understanding, and there should be less evidence of text modification. By contrast, high transactional writers, who are assumed to prioritise the implicit knowledge-constituting process, we do not have expectations with respect to text quality, but they would be expected to experience less development of subjective understanding, and engage in less text modification in the outlining condition than in the synthetic planning condition.

## 3.2 METHOD

The data that we deal with in this study were collected as part of a larger-scale study which included another individual difference variable – self-monitoring (Snyder, 1974) – and which will ultimately involve a more detailed analysis of keystroke logs, which is still currently under development, as well as a range of other variables. Initial analysis of the data showed that there were no direct relationships between self-monitoring and writing beliefs (for transactional beliefs and self-monitoring $r$=.02; for transmissional beliefs and self-monitoring, $r$=.07). Since the sample size ($N$=84) is too small to meaningfully explore higher level interactions between self-monitoring, writing beliefs and different types of planning – several cell sizes approach zero - we have focused in this paper on a simple measure of text modification during writing, and have excluded self-monitoring from the analysis.

### *3.2.1 Participants*

Overall, 84 participants were recruited to take part in the experiment. They were all bachelor or master students at the University of Groningen and were all native speakers of Dutch. They received € 10 for their participation and carried out the experiment individually. One participant was removed from the sample because they had a very low score (over 3 SDs below the mean) on the prior understanding rating.

### *3.2.2 Writing Beliefs Inventory*

We administered the Writing Beliefs Inventory (White & Bruning, 2005) to measure two writing beliefs variables: high versus low transmissional beliefs and high versus low transactional beliefs. White and Bruning reported a Cronbach's $\alpha$ of .72 and .76 for the transmissional and transactional scales of the WBI, respectively. These Cronbach alphas are reported on the basis of their 20-item inventory. However, they advise to exclude item 7 for the final Writing Beliefs Inventory due to the ambiguous loadings of this item on the two belief factors across the different experiments that they conducted. We therefore administered the version of the Inventory consisting of 19-items and in which item 7 is excluded. White and Bruning don't report Cronbach alpha's for this selection of items. Examples of items loading on the two scales can be seen in table 3-1. This version of the inventory was then translated into Dutch, with the adequacy of translation being checked by both a Dutch and English native speaker. The transmissional scale of the Inventory that we administered consisted of 7 items and had a Cronbach's $\alpha$ of .57. The transactional scale consisted of 12 items and had a Cronbach's $\alpha$ of .66. Our analysis showed that these scales were, as found by White and Bruning, uncorrelated ($r$=.016, NS).

**Table 3-1** Examples of transmissional beliefs items (1-3) and transactional beliefs items (4-6) from the writing beliefs inventory (White and Bruning, 2005).

| | **Transmissional beliefs** |
|---|---|
| **1.** | Good writers include a lot of quotes from authorities in their writing |
| **2.** | Writing's main purpose is to give other people information |
| **3.** | Writing should focus around the information in books and articles |
| | **Transactional beliefs** |
| **4.** | Writing requires going back over it to improve what has been written |
| **5.** | Writing helps me to understand better what I'm thinking about |
| **6.** | Writers' views should show through in their writing |

Participants were asked to rate how much they agreed with the different statements about writing on a Likert-Scale of 1 to 5, with 1 representing "strongly disagree" and 5 representing "strongly agree". We administered the Writing Beliefs Inventory at the

beginning of the writing experiment, even before the participants were taken through the procedure of the experiment. This was done to minimise the extent to which conceptions about the task that they were going to perform in the experiment would influence the ratings of their implicit beliefs about writing.
In order to analyse the effects of transmissional and transactional beliefs we took the median as cut point for the scores on both factors and classified participants as low or high on either of the beliefs. On the basis of these divisions participants either have a high transmissional (HTM) - low transactional (LTA) configuration, a low transmissional (LTM) –high transactional (HTA) configuration or the more complicated configurations of high transmissional - high transactional, and low transmissional – low transactional beliefs. Table 3-2 shows the distribution of the participants over the different beliefs model configurations.

**Table 3-2** Distribution of participants over the different beliefs configurations.

| | **Transactional beliefs** | | |
|---|---|---|---|
| **Transmissional beliefs** | **LTA** | **HTA** | **Total** |
| **LTM** | 28 | 16 | 44 |
| **HTM** | 21 | 18 | 39 |
| **Total** | 49 | 34 | 83 |

### *3.2.3 Procedure*

Participants were randomly allocated to one of two different planning conditions. We distinguished between *outline* planning, which involved making a structured outline before writing, and *synthetic* planning, which involved writing down a single sentence summing up one's overall opinion. The purpose of the synthetic planning condition was to prevent participants from making a mental outline of the text that they were going to write. The crucial difference being that the synthetic planning condition does not specify the order in which content should be produced during text production. In both planning conditions, participants had 5 minutes to either make a structured outline or to write down a single sentence. After the planning task, participants were given 30 minutes to write an article for the university newspaper. They wrote their papers on the computer. This enables us to log the writing process with the use of

Inputlog (Leijten & van Waes, 2006). During writing, participants were allowed to consult their written outlines.

### *3.2.4 Text modification index*

In order to assess the process by which writing was carried out we used a global measure of text modification (Baaijen, Galbraith & de Glopper, 2010). For the text modification index, the total number of process words recorded by Inputlog is divided by the total words appearing in the final text. When writers transcribe their thoughts directly into text without modifying the content the index should be 1: all the words that are written down during text production will be included in the final text. To the extent that the writer changes the way that they express their ideas during text production, the index should increase: writers will produce more words during the process of text production than appear in the final text.

### *3.2.5 Development of understanding*

In order to assess the development of understanding we asked participants to rate their understanding about the topic both immediately before and after writing. This involved a simple subjective measure in which participants were asked to indicate how much they felt they knew about the topic on a 7-point scale ranging from 1=very little to 7= a great deal.

### *3.2.6 Text quality*

The quality of the texts was rated by two independent judges on a 9-point scale. The first rater was the principal researcher and the second rater was a trained research assistant. The written articles were rated on the coherence of the overall argument, the originality of the ideas, and the appropriateness of the tone and the relation to the audience for an article in the university newspaper. Independently from each other, both raters first read all articles to get an overall impression of the quality of written outputs and then divided them up in three categories; poor, medium and good quality texts. Then, again independent from each other, all texts were read for a second time and the three categories of poor, medium and good quality texts were sub-divided into three further sub-groups of poor, medium and good quality texts. The end result being a nine point scale, with a rating of 1 corresponding to the poorest texts, and a rating of 9 corresponding to the best texts. Then, all texts were read for a third time and raters were allowed to reallocate essays to a different category if

necessary. Inter-rater reliability for this measure proved to be satisfactory ($r$=.84, $p$<.001). The means of the two judges' scores were used for analysis.

## 3.3 RESULTS

### *3.3.1 Text quality*

A three-way (2*2*2) between subject ANOVA with transactional beliefs, transmissional beliefs and type of planning as independent variables and text quality as dependent variable revealed a significant main effect of transmissional beliefs on text quality ($F$(1,74) = 7.07, $p$=.01, partial $\eta^2$=.09). This reflects the fact that participants with low transmissional beliefs produced text of higher quality ($M$=5.60, $SD$=1.91) than participants with high transmissional beliefs ($M$=4.67, $SD$=1.99), and confirms White and Bruning's (2005) findings for this writing belief. There was also a significant interaction between type of planning and transactional beliefs ($F$ (1, 74) = 4.49, $p$=.04, partial $\eta^2$=.06). The means and standard errors for text quality in each condition are plotted in figure 3-1.

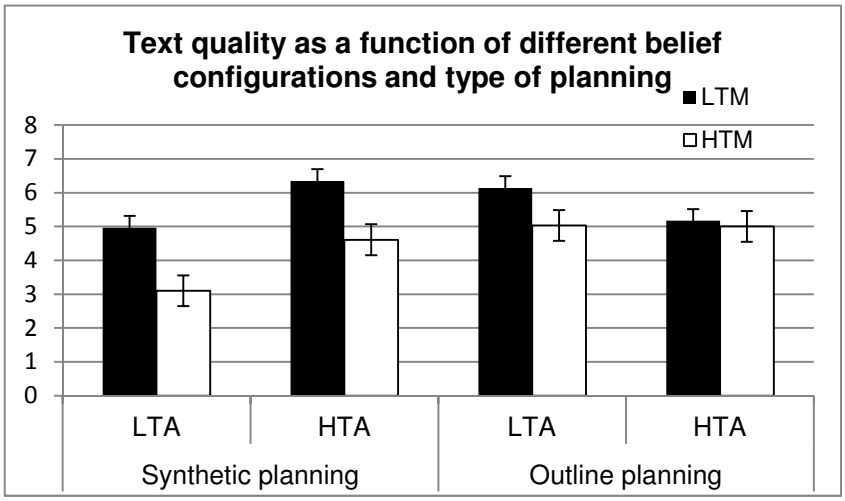


**Figure 3-1** Mean text quality for different belief configurations within synthetic and outline planning.

In order to determine the source of this interaction, we carried out separate two-way between subjects ANOVAs within the two planning conditions and for the low and high transactional participants. Within the synthetic planning condition, these showed a significant main effect of transactional beliefs ($F$ (1, 37) = 10.02, $p$ =.003, partial $\eta^2$ =

.21) and a significant main effect of transmissional beliefs ($F$ (1, 37) = 6.42, $p$ =.02, partial $\eta^2$ = .15), with no evidence of an interaction between the two sets of beliefs ($F$ (1, 37) = .01, $p$ = .90, partial $\eta^2$ < .001). As can be inferred from figure 3-1, within the synthetic planning condition, participants with high transactional beliefs (HTA: $M$=5.53, $SD$=2.03), produced text of higher quality than participants with low transactional beliefs (LTA: $M$ = 4.55, $SD$ = 1.63), and participants with low transmissional beliefs (LTM: $M$ = 5.48, $SD$ = 1.77) produced text of higher quality than participants with high transmissional beliefs (HTM: $M$ = 4.07, $SD$ = 1.76). These results correspond directly with those found by White and Bruning, and support our interpretation that the synthetic planning condition in this experiment is comparable with the writing condition used by White and Bruning, in which pre-planning was not explicitly manipulated. An important feature to note here is that, as in White and Bruning's results, the two sets of beliefs have additive effects within synthetic planning and do not interact with each other.

By contrast, there were no significant effects within the outline planning condition, the closest to significance being the main effect of transmissional beliefs ($F$(1, 37) =.79, $p$ = .38, partial $\eta^2$ = .02). In effect, then, instructing participants to make an outline before writing appears to remove the differences in quality arising from participants' transmissional and transactional beliefs within the synthetic planning condition.

In order to determine the source of this effect, we carried out two-way between subjects ANOVAs for the low and high transactional participants separately. These suggest a more nuanced interpretation of the effect of outlining. For participants with low transactional beliefs, there was a significant main effect of transmissional beliefs ($F$ (1, 45) = 5.85, $p$ = .02, partial $\eta^2$ = .11) and a significant main effect of condition ($F$ (1, 45) = 6.34, $p$ = .015, partial $\eta^2$ = .12). As can be seen in figure 3-1, these effects reflect the fact that outlining increases the quality of the text produced by writers with low transactional beliefs, bringing it up to a similar level to the texts produced by writers with high transactional beliefs in the synthetic planning condition. This suggests that outlining can compensate for the effects of low transactional beliefs on text quality. The fact that there is still a difference between writers with high and low transmissional beliefs within the outline condition, suggests, however, that outlining does not compensate for the negative effects of high transmissional beliefs, and reinforces the conclusion that the two sets of beliefs are independent of one another.

By contrast, outline planning had no significant effect for participants with high transactional beliefs ($F$ (1, 29) = 0.34, $p$ = .56, partial $\eta^2$ = .012), with the mean for the outline planning condition ($M$ = 5.07, $SD$ = 1.91) being lower than the mean for the synthetic planning condition ($M$ = 5.53, $SD$ = 2.03). As can be seen in figure 3-1, there

is a suggestion in the data that this decline in quality for outline planned texts, is restricted to the sub-group of high transactional writers who also have low transmissional beliefs. However, although the interaction between transmissional beliefs and type of planning has a small to medium effect size (Cohen, 1988), it is not statistically significant ($F$ (1,29) = 1.33, $p$ = .26, partial $\eta^2$ = .04). We conclude therefore, that outline planning has no significant effect on text quality for participants with high transactional beliefs, but that future research with a larger sample should investigate whether it may have negative effects for participants who also have low transmissional beliefs. Retrospective power analysis (Cohen, 1988) suggests that a large sample with 70 participants within each condition would be required to achieve a power of .80 for this analysis.

Taken together, these results replicate White and Bruning's (2005) findings that high transmissional and low transactional beliefs have independent and negative effects on text quality when writing is not explicitly planned. The results suggest, in addition, that outline planning may compensate for these negative effects for writers with low transactional beliefs but has no effect (or possibly negative effects) on the quality of texts produced by writers with high transactional beliefs. Outline planning has no effect on the difference between high and low transmissional beliefs within each condition, suggesting that it does not affect the way in which transmissional beliefs contribute to differences in text quality.

### *3.3.2 Development of subjective understanding*

The next step was to assess the effect of transmissional and transactional beliefs on the development of subjective understanding through writing. There is a range of ways in which these data could be analysed, including: (i) analysis of covariance, using ratings after writing as the dependent variable, with prior ratings of understanding as a covariate; (ii) repeated measures analysis of variance on rating of understanding before and after writing; and (iii) analysis of variance on change in understanding scores. For ease of presentation, we will describe the results in terms of change scores. The alternative analyses do, however, produce the same effects.

In order to assess the effects of writing beliefs and type of planning on changes in understanding we carried out a three way (2*2*2) between subjects ANOVA with transmissional beliefs, transactional beliefs and type of planning as factors, and change in understanding as dependent variable. This revealed a significant main effect of transactional beliefs ($F$(1,75)=6.93, $p$=.01, partial $\eta^2$=.08), and a marginally significant main effect of type of planning ($F$(1,75)=3.38, $p$=.07, partial $\eta^2$=.04). However, there was also a significant two-way interaction between transmissional

beliefs and type of planning ($F(1,75)=6.33$, $p=.014$, partial $\eta^2=.08$), a marginally significant interaction between transactional beliefs and type of planning ($F(1,75)=3.79$, $p=.055$, partial $\eta^2=.05$), and a significant three-way interaction between transactional beliefs, transmissional beliefs and type of planning ($F(1,75)=5.79$, $p=.02$, partial $\eta^2=.07$).

Figure 3-2 shows the mean change in understanding (with standard error bars) as a function of transmissional beliefs (LTM v HTM) and transactional beliefs (LTA v. HTA) plotted separately for the synthetically planned and outline planned conditions.

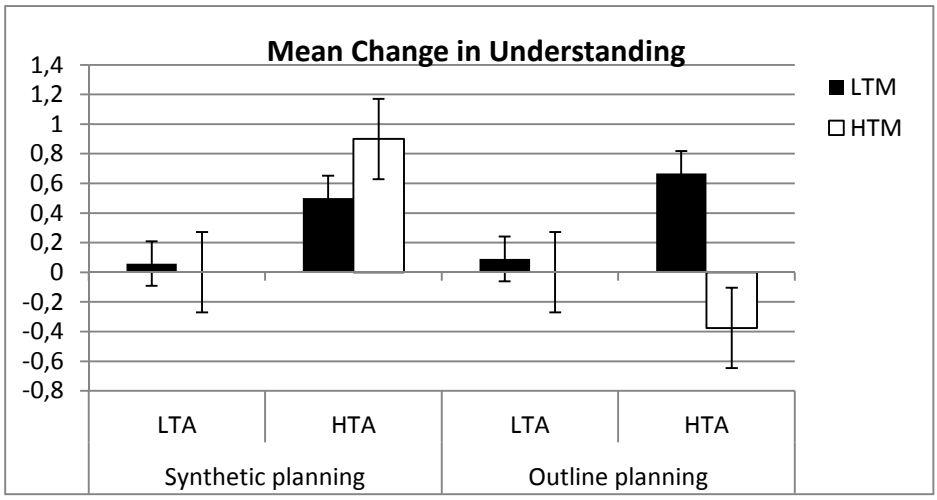


**Figure 3-2** Development of understanding as a function of different writing belief configurations and type of planning.

As can be seen in figure 3-2, there is little evidence that writers with low transactional beliefs experience changes in understanding in either synthetic or outline planned conditions. However, there are effects for writers with high transactional beliefs which vary depending both on the writers' transmissional beliefs and the type of planning that they engaged in before writing. In order to determine the source of these effects, we first carried out two-way between subjects ANOVAs within the synthetic and outline planned conditions.

Within the synthetic planning condition, there was a highly significant main effect of transactional beliefs ($F(1,38)=12.07$, $p=.001$, partial $\eta^2=.24$) but no significant effect of transmissional beliefs ($F(1,38)=0.78$, $p=.38$, partial $\eta^2=.02$), and no significant interaction between transactional and transmissional beliefs ($F(1,38)=1.41$, $p=.24$, partial $\eta^2=.04$). A repeated measures comparison within the high and low transactional writers showed that writers with high transactional beliefs experienced

significant increases in understanding ($F(1, 19)=22.70$, $p<.0005$, partial $\eta^2=.54$) whereas writers with low transactional beliefs did not experience any change in understanding ($F(1, 21)=0.19$, $p=.66$, partial $\eta^2=.01$). These results provide strong support for the hypothesis that, for synthetically planned writing, high transactional beliefs are related to the extent that writers experience increases in understanding during writing, but suggest that, contrary to White and Bruning's assumption, transmissional beliefs are unrelated to changes in understanding.

By contrast, in the outline planning condition, there was a significant main effect of transmission beliefs ($F(1,37)=6.62$, $p=.014$, partial $\eta^2=.15$), but no main effect of transactional beliefs ($F(1,37)=0.21$, $p=.65$, partial $\eta^2=.01$) and a significant interaction between transmissional and transactional beliefs ($F(1,37)=4.66$, $p=.04$, partial $\eta^2=.11$). Repeated measures analysis within each condition revealed that writers with high transactional beliefs and low transmissional beliefs experienced significant increases in understanding ($F(1,5)=10.00$, $p=.02$, partial $\eta^2=.67$), whereas writers with low transactional beliefs experienced no change in understanding ($p>.50$ for both comparisons), and writers with high transactional beliefs and high transmissional beliefs experienced a non-significant decrease in understanding ($F(1,7) = 2.03$, $p = .20$, partial $\eta^2=.22$). Although the decrease in understanding for writers with high transactional and high transmissional beliefs was not significant, the effect size was very large, so the lack of significance may be a consequence of low statistical power. Retrospective power analysis suggests that 17 participants within each condition would be required to achieve a power of .80 for this comparison.

In order to assess the direct effects of outline planning and how these varied depending on writers' beliefs we then carried out two-way between subjects ANOVAs for the low and high transactional writers separately. For the writers with low transactional beliefs, there were no effects of either outline planning or transmissional beliefs ($p > .69$ in both cases). These writers, then, do not experience significant changes in understanding (see above) and this is the same regardless of whether writing is synthetically or outline planned and whether they hold high or low transmissional beliefs. By contrast, for writers with high transactional beliefs, there was a significant main effect of type of planning ($F(1,30) = 5.95$, $p = .02$, partial $\eta^2=.16$), and a significant interaction between type of planning and transmissional beliefs ($F(1,30)=10.06$, $p = .003$, partial $\eta^2=.25$). Individual comparisons of synthetic and outline planning for writers with high and low transmissional beliefs showed that, for writers who also held high transmissional beliefs, outline planning led to significantly less increased understanding than synthetic planning ($F(1,16) = 13.17$, $p = .002$, partial $\eta^2=.45$). By contrast, for writers who also had low transmissional beliefs, type of

planning made no difference to the extent to which they increased their understanding after writing ($F(1,14) = 0.38$, $p = .55$, partial $\eta^2$=.03).

Overall, then, these results suggest that increases in understanding as a function of writing depend on the writer's transactional beliefs rather than their transmissional beliefs, with high transactional beliefs being necessary for significant increases in understanding to occur. This effect is absent for the sub-group of writers with high transactional beliefs who also hold high transmissional beliefs when they make an outline before writing. Future experiments with larger samples are needed to establish whether, for this sub-group, outlining simply prevents writers from developing their understanding, or whether it actively leads to decreases in understanding after writing.

### *3.3.3 Amount of text modification*

This analysis had two aims: (i) to assess the effects of writing beliefs on the amount of text modification carried out during writing (see Baaijen et al., 2010); and (ii) to assess the relationships between text modification and text quality and between text modification and changes in understanding, and whether these varied as a function of writers' beliefs. High scores on the text modification index reflect the extent to which the text is modified during text production: writers will produce more words during the process of text production than appear in the final text.

Preliminary analysis of the text modification index revealed 1 outlier with a score more than 3 SD's above the mean which was removed from all process analyses. Inspection of the keystroke logs showed that this participant had deleted substantial sections of text which they then re-wrote in very similar words. The deletion of substantial sections of text inevitably means that the text modification index will be very high, but that this will not directly reflect the extent to which text has been modified during production. Furthermore, the process data from 1 participant were lost due to technical problems with the computer. Therefore, the overall sample of participants for this measure consists of 81 participants.

### *3.3.4 Effects of writing beliefs on amount of text modification*

A three–way (2*2*2) between subject ANOVA with transactional beliefs, transmissional beliefs, type of planning as independent variables and TMI as dependent variable revealed a significant main effect of transmissional beliefs ($F(1,73)=4.79$, $p$=.03, partial $\eta^2$=.06) and a significant main effect of type of planning ($F(1,73)=6.75$, $p$=.01, partial $\eta^2$=.08). Participants modified their text more following

synthetic planning (*M*=1.30, *SD*=.18) than they did following outline planning (*M*=1.24, *SD*=.12), and participants with high transmissional beliefs modified their text more (*M*=1.31, *SD*=.16) than participants with low transmissional beliefs (*M*=1.25, *SD*=.16).

There was, however, also a marginally significant interaction between transactional and transmissional beliefs (*F* (1,73)=3.34, *p*=.07, partial $\eta^2$=.044), which suggests that the main effect of transmissional beliefs should be treated with caution. Individual comparisons showed that, for participants with low transactional beliefs, those with high transmissional beliefs (*M*=1.34, *SD*=.17) had significantly higher scores (*F* (1,44) = 8.49, *p*=.006, partial $\eta^2$=.16) on the text modification index than participants with low transmissional beliefs (*M*=1.23, *SD*=.14). However, this difference was not significant for participants with high transactional beliefs (*F*(1, 29) = 0.64, *p*=.80, partial $\eta^2$=.002), with the mean amount of text modification for participants with high transmissional beliefs (*M*= 1.26, *SD*=0.12) being minimally different to that for those with low transmissional beliefs (*M*=1.27, *SD*= 0.18). As can be seen in figure 3-3, which shows the mean amount of text modification as a function of transactional and transmissional beliefs. the increase in the level of text modification appears to depend on a combination of transmissional and transactional beliefs rather than being a main effect of transmissional beliefs.

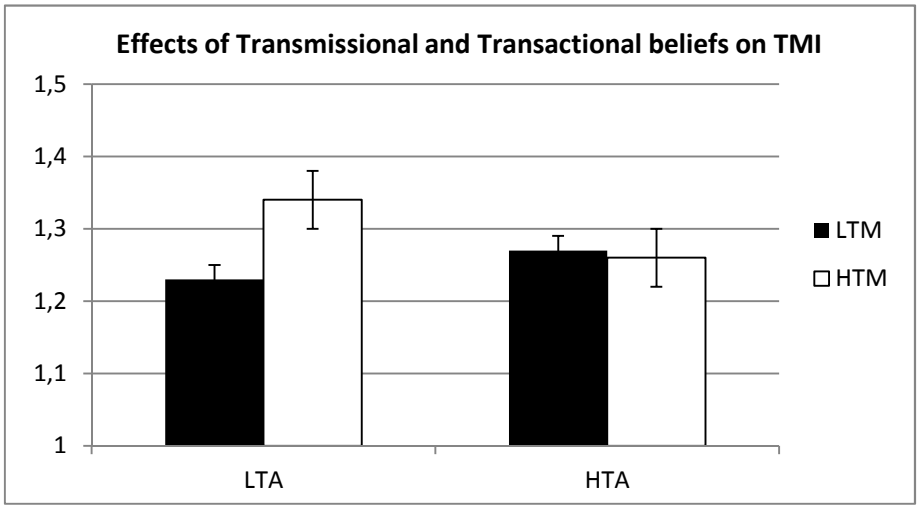


**Figure 3-3** Effects of transmissional (LTM v HTM) and transactional beliefs (LTA v HTA) on amount of text modification (TMI).

Overall, we draw two main conclusions from this analysis. First, outline planning tends to reduce the extent to which the text is modified while it is being produced,

and this effect applies across all writers regardless of the beliefs. Second, participants with a combination of low transactional and high transmissional beliefs tend to modify the text more than other writers. Note that the fact that it is high transmissional rather than low transmissional writers who modify their text more. This contradicts the assumption that high transmissional beliefs are associated with a simple knowledge-telling process of writing. Instead, our data suggest that the process associated with high transmissional beliefs involves an active process of modifying the text in order to ensure that it complies with the writer's intentions. Furthermore, it this group of writers for whom outline planning had the most beneficial effects on text quality.

### *3.3.5 Relationships between amount of text modification, text quality and changes in understanding.*

In order to assess these relationships we carried out two multiple regressions, one with text quality as the dependent variable, and one with change in understanding as the dependent variable. In order to prevent problems of multi-collinearity and to ease interpretation of interactions, the transactional and transmissional beliefs scores were centred around the mean, and the text modification variable was standardized; type of planning was dummy coded, with 0 corresponding to synthetic planning, and 1 corresponding to outline planning. For each analysis, we used a combination of forward and backward regression. In order to observe the effects of interactions, we entered the variables into the regression model in the following sets: (i) main effects of transactional beliefs, transmissional beliefs, type of planning, and amount of text modification were entered at step 1; (ii) terms corresponding to the two-way interactions between the variables were added at step 2; (iii) three-way interactions were entered at step 3; and (iv) the four-way interaction was entered at step 4. We then removed sets which did not add significant $R^2$ change when they were entered. The models were then simplified by progressively removing non-significant coefficients, unless they contributed to higher order interactions (Cohen, 1988; Cohen, Cohen, Aiken & West, 2003). We describe the final, minimal models below (Crawley, 2005).

### *3.3.6 Relationships with text quality*

The final model for this analysis accounted for 20% of the variance in text quality ($F$ (5, 75) = 3.64, $p < .0005$), and is shown in table 3-3. As can be seen in the table, the final model has two significant coefficients: the interaction between type of planning and

transactional beliefs that we have already noted in the analysis of the effects of writing beliefs on text quality, and an interaction between amount of text modification and transactional beliefs.

**Table 3-3** Predicting text quality from transactional beliefs, type of planning, and amount of text modification.

| | **B** | **SE** | **β** |
|---|---|---|---|
| Constant | 5.20 | 0.31 | |
| Transactional beliefs | -1.02 | 0.74 | .20 |
| Type of planning | -0.27 | 0.43 | -.07 |
| Text modification | -0.40 | 0.23 | -.20 |
| Type of planning * Transactional beliefs | 2.45 | 1.10 | .32* |
| Transactional beliefs * Text modification | 1.23 | 0.58 | .25* |

*Note: $R^2$ = .20, Adjusted $R^2$ = .14, ($F(5, 75) = 3.64$, $p = .005$); * $p < .05$.*

The relationship between amount of text modification (standardized scores) and text quality (adjusted predicted values) is plotted in figure 3-4, with lines of best fit shown separately for participants with low transactional beliefs (dashed line and crosses) and high transactional beliefs (solid line and circles).

The first feature to note in the figure is that there is one extremely low quality score. In order to check whether this extreme score was responsible for the effects, we removed it and re-ran the multiple regression analysis. The model remained significant ($R^2$ = .17, ($F(5, 74) = 3.13$, $p$ = .013), as did the two coefficients for the interaction terms (for the interaction between type of planning and transactional beliefs, $\beta$ =.34, $p$ = .02; for the interaction between transactional beliefs and text modification, $\beta$ =.25, $p$ = .025). The second feature is that the relationship is negative for low transactional writers, with greater amounts of text modification being associated with lower quality text ($r$ = -.83, $N$ = 48, $p < .0005$) but there is no relationship for high transactional writers ($r$ = .08, $N$ = 33, $p$ = .66).

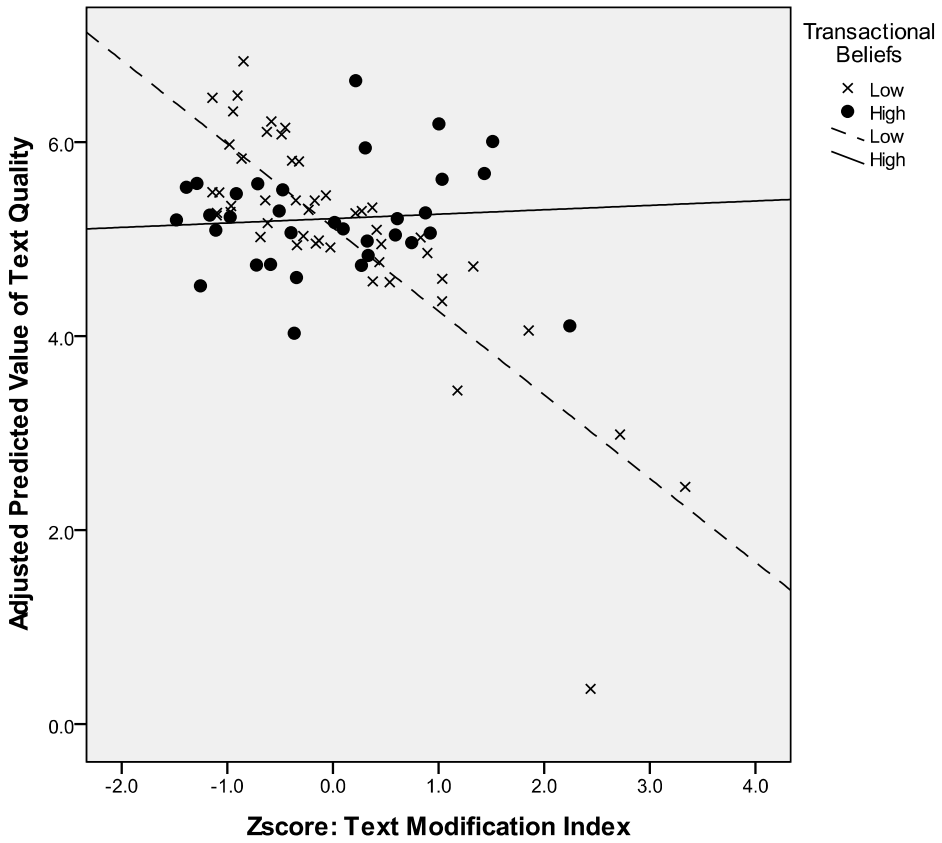


**Figure 3-4** Relationship between amount of text modification and predicted values of text quality as a function of low and high transactional beliefs.

**Table 3-4** Predicting increases in understanding from transmissional and transactional beliefs, type of planning, and amount of text modification.

| | **B** | **SE** | **β** |
|---|---|---|---|
| Constant | 0.35 | 0.08 | |
| Transactional beliefs (TA) | 0.30 | 0.14 | .22* |
| Transmissional beliefs (TM) | -0.50 | 0.16 | -.47*** |
| Text modification | 0.18 | 0.06 | .35*** |
| Type of planning | 0.03 | 0.11 | .03 |
| Type of planning * TM | 0.80 | 0.23 | .53*** |
| Text modification* TA | 0.32 | 0.14 | .23* |

*Note: $R^2$ = .60; Adjusted $R^2$ = .36, ($F(6, 67) = 6.24$, $p < .0005$); * $p < .05$. ** $p < .01$, *** $p < .005$.*

### *3.3.7 Relationships with changes in understanding*

For the purposes of this analysis, we excluded participants who reported decreases in understanding because they appeared to show a different relationship with text modification compared to the remainder of the sample, but were too few in number to enable reliable estimates to be made of any non-linear relationship with text modification. The analysis is therefore restricted to the relationship between amount of text modification and increased or unchanged understanding.

The final model for this model accounted for 36% of the variance in increased understanding ($F$ (6, 67) = 6.24, $p < .0005$), and is shown in table 3-4. As can be seen in the table, there are significant main effects of both transmissional and transactional beliefs, and a significant interaction between transmissional beliefs and type of planning. These effects correspond to the effects we have already analysed above. In addition, there is a significant interaction between amount of text modification and transactional beliefs, indicating that the relationship between amount of text modification and change in understanding varies depending on whether the writer has high or low transactional beliefs. This interaction is shown in figure 3-5, where amount of text modification is plotted against the predicted values for change in understanding, with separate lines of best fit for low (dashed line and crosses) and high transactional writers (solid line and circles).

As can be seen in the figure, both high and low transactional writers show positive relationships between amount of text modification and increased understanding. However, the relationship is significantly stronger for the writers with high transactional beliefs, ($r$ = .81, $p < .0005$) than it is for the writers with low transactional beliefs ($r$ = .34, $p$ = .03).

Taken together, these analyses suggest that text modification is associated with different effects for writers with high and low transactional beliefs. For writers with low transactional beliefs, a greater amount of text modification is associated with the production of lower quality text, and with only moderate increases in understanding. By contrast, for writers with high transactional beliefs, a greater amount of text modification is unrelated to text quality, but strongly related to increased understanding.

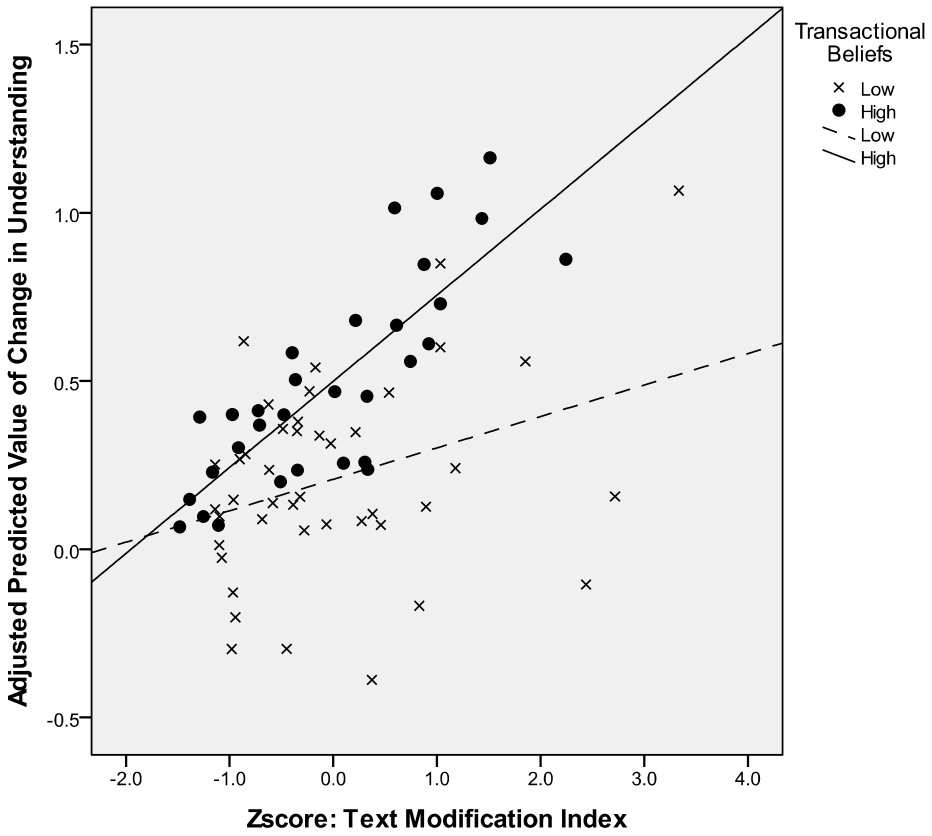


**Figure 3-5** Relationship between amount of text modification and changes in understanding as a function of low and high transactional beliefs.

## 3.4 DISCUSSION

These results provide strong support for White and Bruning's (2005) empirical findings. We found, like they did, that the transmissional and transactional components of writers' beliefs are independent of one another. Furthermore, within the synthetic planning condition, which is the condition most similar to their original research, we replicated their finding that low transmissional and high transactional beliefs are associated with higher quality writing, and that these are additive, independent effects.

In addition, these findings extend White and Bruning's research by showing that writers' beliefs are also related to differences in the extent to which writers experience developments of understanding through writing. Writers with high transactional beliefs do indeed experience greater developments of understanding after writing than writers' with low transactional beliefs, as would be expected from White and Bruning's definition of the construct and the items on the transactional scale. Furthermore, the fact that transmissional beliefs were not associated with

differences in the development of understanding supports the assumption that this is an independent belief about a different aspect of writing. It suggests that high transmissional writers can nevertheless (at least under synthetic planning conditions) develop their understanding of this information just as readily as writers with low transmissional beliefs so long as they also hold high transactional beliefs. An important question for future research is whether the differences we have found for subjective developments in understanding are also apparent for more objective measures of learning through writing.

### *3.4.1 Text modification*

The results for the text modification measure were not as we expected, and we will argue that in order to account for this the interpretation White and Bruning give to the transactional beliefs scale needs to be refined. Based on their definition of the construct, and on the fact that several items on the scale refer to the extent that writing is revised, we expected that writers with high transactional beliefs would show more evidence of modifying text during writing than low transactional writers. Similarly, on the assumption that high transmissional beliefs are related to knowledge-telling, and the fact the scale includes items such as "A primary goal of writing should be to have to make as few changes as possible", we expected these beliefs to be associated with less text modification. In fact, we found that it was precisely the writers with a combination of low transactional and high transmissional beliefs who modified their text most in the course of production. This contradicts Bruning and White's assumption that this combination of beliefs is associated with lower levels of engagement in writing. On the contrary, these writers appear to be the most actively engaged of all the writers in the sample in modifying the text as it is produced.

We want to suggest therefore that the distinction between low and high transactional writers is not in how much text modification is carried out, but rather in the goals towards which it is directed. The key findings here are that, for high transactional writers, greater amounts of text modification were associated with increased understanding but unrelated to text quality, whereas, for low transactional writers, they were only moderately related to changes in understanding and negatively correlated with text quality. This suggests that, for high transactional writers, text modification is primarily related to the goal of developing understanding, rather than with maintaining text quality. By contrast, for low transactional writers, text modification is less related to the goal of developing understanding, and is instead a response to a perceived mismatch between text as it is produced and the writer's original intentions – it represents an attempt to make the

text conform to the writer's original intentions. This provides an explanation for the interaction with transmissional beliefs for amount of text modification. High transmissional writers focus on external content as can be seen from four of the five items that define the transmissional scale in experiment 2 in White and Bruning (2005). Because their information is external they may understand it less well. When problems in understanding or expressing external information occur, text modification will be required. By contrast, writers with low transmissional beliefs - who are expressing personal content rather than information from external sources – may experience less mismatches between intended content and text, with the result that there is less requirement to engage in text modification.

This interpretation of the results implies that the definition White and Bruning of the transactional beliefs scale needs to be refined. At present this has an essentially unipolar definition – writers with high transactional beliefs are assumed to believe that writing involves the development of understanding during writing (and hence involves modifying the text as it is produced), and writers with low transactional beliefs are assumed not to have the same motivation to develop their understanding (and hence not to modify the text to the same extent). We want to suggest, instead, that writers with low transactional beliefs differ in their beliefs about *when* developments in understanding should occur. They believe, essentially, that writing involves "getting one's thoughts straight" *before* writing the text, and hence that understanding is not something that should develop during text production, and that writing need not involve extensive revision of thought. By contrast, high transactional writers believe that understanding is constituted in the course of producing text, and hence that it will be updated and developed in the course of text production. This redefinition, it seems to us, remains consistent with the items on the scale relating to emotional aspects of writing. One would expect writers who see writing as involving the development of their understanding to view it as pleasurable, and writers who see it as a matter of trying to capture pre-determined thought to experience it as frustrating.

The key change in interpretation that we are suggesting, then, is that the low end of the transactional scale should be viewed as a belief that text production should be actively controlled in order to ensure that it corresponds to the writer's pre-determined intentions. This could be tested by adding items to the scale which actively expressed this alternative view, rather than leaving it as an implication of a negative response to statements asserting that writing involves the development of understanding. If we are correct, then these items should correlate well with items on the existing scale, and the resulting scale should have better predictive validity than the present version. We would also predict that, under natural conditions, writers

with low transactional beliefs will – in order to be able to control text production - be more likely to engage in planning before writing than high transactional writers.

### *3.4.2 Effects of outlining*

Our second set of findings was about the differential effects of outlining for writers with different writing beliefs. The most straightforward finding, here, was that the low-transactional writers wrote better in the outline planned condition than they did in the synthetic planning condition. This fits well with previous research on outlining (Kellogg (1994) and is compatible with the assumption that in separating the generation and organisation of ideas from text production writes are able to carry out both more effectively. Furthermore, the fact that text modification was reduced in the outline conditions adds further support to this assumption by confirming that writers were able to produce "cleaner" text following outlining than after synthetic planning. Since text modification was also negatively correlated with text quality for these writers, the results suggest that the reason outlining improves quality for these writers is because it enables them to produce text that corresponds more directly with their intentions. Overall, then, the results for the low transactional writers fit extremely well with the typical explanations for the benefits of outlining.

We would like to make two further observations about these results. First, the beneficial effect of outlining on text quality for low transactional writers is precisely what would be expected if these writers have the goal of trying to control text production so that it conforms to their initial intentions, as we suggested in the previous section. Second, outline planning did not remove the difference in text quality between low and high transmissional writers. This is further support for the assumption that the transmissional and transactional scales capture different components of writer's beliefs. If, as the content of the items that define the transmissional scale in experiment 2 of White and Bruning (2005) suggest, transmissional beliefs are primarily about the source of the content to be written about, then this suggests that the negative effects of these beliefs are not alleviated by outlining: outlining appears to have its effects after the source of content has been selected.

The most important finding here was that high transactional writers did not produce better text after outlining than after synthetic planning. The dual process model would argue that this is because, although outlining may facilitate the explicit planning process, it also inhibits the knowledge constituting process, and since both processes are assumed to contribute to text quality, the end result is no overall improvement in quality. The problem with this interpretation is that one can conceive

of a range of other explanations for this effect. For example, it could be that high transactional writers are better at managing cognitive load during writing – perhaps they have larger working memory capacities than low transactional writers – and hence have less need to outline before writing. If this were the case, then one could argue that high and low transactional writers employ the same writing processes but high transactional writers are able to carry them out more efficiently; outlining would, in effect, compensate for this and bring the low transactional writers up to the same level as the high transactional writers. In order to provide stronger evidence that the processes are different, therefore, one would need to be able to show that outlining had a negative effect on high transactional writers. And, although the direction of the difference was in fact negative for the high transactional writers, this was not a significant difference. Given that the sample size was relatively small for these comparisons it would be worth examining this further in future research with larger samples.

The strongest support for dual process model's claim, therefore, comes from the fact that high transactional writers did show significant reductions in the development of understanding in the outline planning condition compared to the synthetic planning condition. However, this is only partial support because it only applied to those high transactional writers who also held high transmissional beliefs. This suggests that the process by which understanding is developed interacts with the source of the content that the writer is writing about. This in itself is actually quite compatible with the dual process model, since it claims that the knowledge-constituting process consists of constraint satisfaction within semantic memory, and this is assumed to correspond to the writer's implicit personal understanding of the topic. By contrast, planning is assumed to operate on an explicit store of ideas in episodic memory. If writers with low transmissional beliefs typically draw on their personal understanding when writing rather than on information from external sources, then it may be that this enables them to overcome the negative effects of outlining on the development of understanding. By contrast, if high transmissional writers typically draw on their explicit store of ideas when writing, then outlining may reinforce this sufficiently for it to reduce the extent to which the knowledge-constituting process is active during text production. The point being that, according to the dual process model, both processes interact in guiding the path of text production, and the extent to which either predominates is influenced by a range of factors. For present purposes, however, in the absence of more detailed data about processes, we can only conclude that for a sub-group of transactional writers, outlining does indeed have a negative effect on the development of understanding.

Overall, then, our conclusion about this aspect of the data is that, although they are problematic for the problem-solving model, in that they do not show a uniform effect of outlining across writing beliefs, they do not provide conclusive evidence in support of the dual-process model either. Further research is needed with larger samples, and with more transparent measures of writing processes, to establish the nature of the processes involved. For the present, we will simply conclude that measuring individual differences in writing beliefs enables one to identify differential effects of outlining, and that for some sub-groups of writers it may reduce the extent to which they develop their understanding through writing.

### *3.4.3 Evaluation of writing beliefs inventory*

Our findings suggest that writing beliefs have an important relationship with writing performance and with the effects of writing on the development of writers' understanding. The scales that make up the writing beliefs inventory do, therefore, appear to have good predictive validity. As found by White and Bruning, responses to statements about writing are reliably associated with the quality of the text that writers subsequently produce. In addition they are also associated with variations in the extent to which writers experience an increase in understanding through writing. However, they also suggest a number of changes that need to be made to the way beliefs are measures and to how they are interpreted.

First, both scales – the transmissional scale in particular - have relatively low reliabilities (Cronbach's alpha = .57 for the transmissional scale, .66 for the transactional scale). This may, in part, be a consequence of cultural differences between Dutch and US students and of contextual differences between our study and White and Bruning's. The transmissional scale in particular would benefit from an increased number of items. Improving the reliability of the scales should, by increasing their discrimination, increase the statistical power of research using them.

Second, as we indicated in the introduction, the underlying constructs represented by the scales are not perfectly clear. In our view, a better way of characterising the scales would be to make a clear distinction between the underlying dimension of *personal* engagement and two factors that contribute to it. To us, the underlying dimension appears to be the level of personal cognitive and affective engagement the writer has with writing – the extent to which they view writing as an active process of making personal meaning rather than conveying external information. The contributing factors, that are represented by the different scales, can be seen as being about two independent aspects of writing.

Inspection of the items on the two scales, however, suggests the following interpretations. The "transmissional" scale is primarily concerned with the *source of the content* that is expressed in writing. Items on the scale include: "Good writers include a lot of quotes from authorities in their writing"; "Writing should focus around the information in books and articles"; and "The most important reason to write is to report what authorities think about a subject". We therefore assume that writers with transmissional beliefs believe that the source of content is external to the writer and that the writer's responsibility is to transmit this content as accurately as possible, as is expressed in the following item: "Writing's main purpose is to give other people information". The "transactional" scale is primarily concerned with the *process of writing* and with the extent to which this involves the development of the writer's understanding. Items on this scale include: "Writing helps me to understand better what I'm thinking about"; "Writing is a process involving a lot of emotion"; and "Writing requires going back over it to improve what has been written". We therefore assume that writers with transactional beliefs believe that writing is an emotional experience involving the development and revision of thought in the course of writing.

We believe that this way of defining the two scales helps to clarify how the two sets of beliefs can, at the same time, be independent of one another but contribute to the same underlying dimension of cognitive and affective engagement. One can readily imagine a writer, in a strongly academic context, for example, believing that it is important to faithfully represent the views of external authorities while at the same time believing that their understanding of those views is something that will develop in the course of writing. Furthermore, one can see how a writer might find expressing an external authority's views less motivating than expressing their own personal opinion, while at the same time finding the development of their own understanding during writing a more motivating experience, and hence, how attitudes to the two different aspects of writing could contribute to a single underlying engagement dimension.

Thus, a high transactional score would be seen as increasing personal engagement, which would be further increased by a low transmissional score, or reduced by a high transmissional score. This would have the virtue that other combinations would have a natural interpretation as involving intermediate levels of personal engagement. A key feature of this interpretation is that implies that the main way in which these beliefs affect writing is through their effect on how the writer goes about writing. These are specific effects tied distinctively to the different types of writing belief. Beliefs about the source of content affect which content available in

memory is paid attention to, and beliefs about the process of writing affect how writers combine the different component processes of writing.

Third, an important way in which the scales could be further clarified would be to make the negative poles of each scale more explicit. At present, all the items on each scale are defined in terms of the positive end of the scale, with the result that when writers disagree with an item it is not clear what beliefs they are actively endorsing. We have argued, in particular, that low transactional writers have active beliefs about the way in which text production should be controlled and that these can account for the differences we have observed in writing performance. These could be made more explicit by including items like, for example "It is important in writing to get one's ideas straight before one starts to write". Similarly, when writers disagree with statements about the importance of external authorities, it is not clear what they believe instead. Including items like, for example, "Good writers explain ideas in their own terms" would help define the beliefs better, and should therefore produce more reliable scales with better predictive validity.

Finally, neither of the scales currently includes statements about the writer's relationship to the reader. Since this plays a central role in models of writing, it would be valuable to develop a further scale assessing writers' beliefs about this aspect of writing. We would anticipate this being concerned with the extent to which writing should be a matter of personal expression as oppose to being tailored to the needs of the audience.

### *3.4.4 Educational implications*

These results have a number of important educational implications. First, in showing that writers' beliefs about writing have an important influence on their performance and on their ability to develop their understanding, they suggest that an important focus for writing instruction should be the often hidden beliefs that students have about writing. Writing instruction tends to focus either on the product – how best to express one's thoughts in language – or, more recently, on the process by which text is produced – planning and revising drafts of text. Although this may indirectly relate to writers' beliefs about writing, the present findings suggest that it may also be useful to elicit and discuss students' beliefs about writing.

Second, an important question for further research is how students develop the beliefs that they have about writing and the appropriateness of these beliefs in different contexts. For example, the belief that writing is about citing authoritative sources seems to us to be a natural inference from the emphasis given to this in academic contexts. The fact that this belief was associated with the production of

poorer quality text in both this study and in White and Bruning's study suggests that this issue needs to be carefully judged and elucidated for students. Feedback consisting simply of brief comments like "reference?" or "evidence" may lead to oversimplified beliefs about the role of external authorities in writing and have the opposite to the desired effect on the quality of students' writing. Clearly, though, this effect needs to be explored in other more academic forms of writing. Would there be the same negative relationship between high scores on the "transmissional" scale and the quality of more explicitly academic writing? Or are these beliefs – transferred perhaps from more academic contexts - simply less appropriate in the kinds of writing examined in White and Bruning and this paper?

Finally, our results suggest that standard advice to make an outline before one writes may not be useful to all writers. While some writers – those with low transactional beliefs – may benefit from outlining, others may either not benefit or perhaps even write less well when encouraged to make an outline. In particular, our results suggest that, for some writers, outlining can prevent the development of understanding during writing. An important question here is whether this is only associated with the writer's subjective experience of understanding, as measured here, or whether this also has effects on more objective measures of a student's learning.

# Chapter 4

## *Keystroke Analysis: Reflections on Procedures and Measures*[2]

[2] A slightly adapted version of this chapter appeared as: Baaijen, V.M., Galbraith, D., & de Glopper, K. (2012). Keystroke analysis: Reflections on procedures and measures. *Written Communication, Special Issue on Writing and Cognition: In Honor of John R. Hayes*, *29* (3), 246-277.

## *Abstract*

Although keystroke logging promises to provide a valuable tool for writing research, it can often be difficult to relate logs to underlying processes. This paper describes the procedures and measures we developed to analyse a sample of 80 keystroke logs with a view to achieving a better alignment between keystroke logging measures and underlying cognitive processes. We used these measures to analyse pauses, bursts and revisions, and found that: (i) burst lengths vary depending on their initiation type as well as their termination type, suggesting that the classification system used in previous research should be elaborated; (ii) mixture models fit pause duration data better than unimodal central tendency statistics; and, (iii) individuals who pause for longer at sentence boundaries produce shorter, but more well-formed, bursts. A principal components analysis identified three underlying dimensions in these data: (i) planned text production; (ii) within sentence revision; and, (iii) revision of global text structure.


*Keywords:* writing models, text production, revision, language bursts, pause analysis

## 4.1 INTRODUCTION

Keystroke logging promises to be a productive tool for research into writing, and has grown in popularity over recent years as a research tool (for recent reviews, see Sullivan and Lindgren, 2006; Waes, van, Leijten, Wengelin and Lindgren, 2012). A major attraction of the method is that it provides a non-intrusive record of the moment by moment creation of the text. However, in its raw form this record only provides information about empty time (when no keys are being pressed) and filled time (when, and which, keys are being pressed). In order to make sense of this, the researcher needs to relate this to a model of the different types of processes involved in writing. Furthermore, the complex and recursive nature of the writing process means that undifferentiated measures of global properties of keystroke logs are likely to be extremely insensitive measures of underlying writing processes. Indeed, some authors have claimed that this may be a fruitless enterprise. Schilperoord (2001, p.67) has suggested that

> *"if the writing process that one wants to study is characterized by intensive problem-solving and massive editing on the part of the writer, with actual language production being but one aspect of these processes, it [may] prove impossible to relate pause data to ongoing processes".*

In this paper, we describe how we took up this challenge, and tried to derive measures from the raw output of keystroke logs that could be related to cognitive models of writing. These logs were collected as part of a larger project investigating the effects of self-monitoring and planning on the development of understanding in writing. We will focus exclusively here on the procedures we used to construct measures from the keystroke logs, and on examining the properties of these measures across the sample of texts as a whole. In future papers, we plan to investigate whether these characteristics vary as a function of self-monitoring and type of planning, as well as how they relate to the development of understanding and text quality. Our reason for focussing on the methodological and conceptual issues here is because, although there is a growing body of research using keystroke logs to examine specific aspects of writing (see Sullivan and Lindgren, 2006), there has been relatively little research investigating the writing process as a whole across larger numbers of writers. Furthermore, the research that has been published in this area (e.g. van Waes & Schellens, 2003) tends to focus on the results of the research rather than the methods

used to achieve them. Our aim therefore is to describe the procedures and measures that we have developed. First, we will discuss some of the issues that arise in interpreting the basic units of analysis available from keystroke logs – pause, burst and revisions – in terms of cognitive models of writing. We will then outline our broad approach to addressing these issues. In the analytic section, we will describe the procedures that we have developed for preparing keystroke logs for analysis, and examine the properties of the measures that we constructed. Finally, we will assess whether these measures can be combined together to create global measures that can aid our understanding of writing processes.

### *4.1.1 Keystrokes and models of writing*

Hayes and Flower's (1980) model of the writing process identified three different kinds of basic process and suggested different ways in which they could be coordinated during writing. These included: (i) planning, which included setting goals, and generating and organising ideas to satisfy those goals, (ii) translating, and (iii) revision. The coordination of these processes was controlled by a monitor and different configurations of the monitor were assumed to correspond to individual differences in writing strategy (see table 4-1). The four different configurations differ in whether planning is carried out before, or at the same time as, text production, and in whether revision is carried out at the same time or after text production.

More recently, Hayes (1996) has suggested that revision should not be a considered a basic process in its own right, but should instead be seen as involving the recursive application of cycles of reading, reflection and text production. Furthermore, Hayes and Nash (1996) have refined the characterisation of planning, distinguishing between process planning (focused on the management of the process itself), abstract planning (concerned with goal setting and content generation), and language planning (concerning the formulation of content in language).

Even this brief sketch of some basic components of a cognitive model of writing makes it clear that there are major problems in tying keystroke units of analysis to specific components of the writing process. Pauses may reflect not just various different levels of planning and reflection, but also re-reading and text production during revision. Bursts of language may reflect an initial formulation of thought or an attempt to improve previously formulated text. Revisions may reflect semi-automatic correction of errors or a systematic attempt to modify content.

**Table 4-1** Overview of writing processes mapped onto stages of writing

| | Pre-planning | Drafting text | Revision |
|---|---|---|---|
| **Depth first (single draft)** | | Generating / organising /translating / reviewing | |
| **Get it down as you think of it (rough draft)** | | Generating / organising / translating | Review |
| **Perfect first draft (outline / polished draft)** | Generating / organising | Translating / reviewing | |
| **Breadth first (outline /rough draft)** | Generating / organising | Translating | Review |

### *4.1.2 Hayes' Model of Text Production*

In this section, we will assess how the basic units of keystroke analysis can be interpreted in terms of the most fully developed current model of text production (see Hayes, 2009, for a review). A sketch of the model, taken from Chenoweth and Hayes (2003) is shown in figure 4-1.

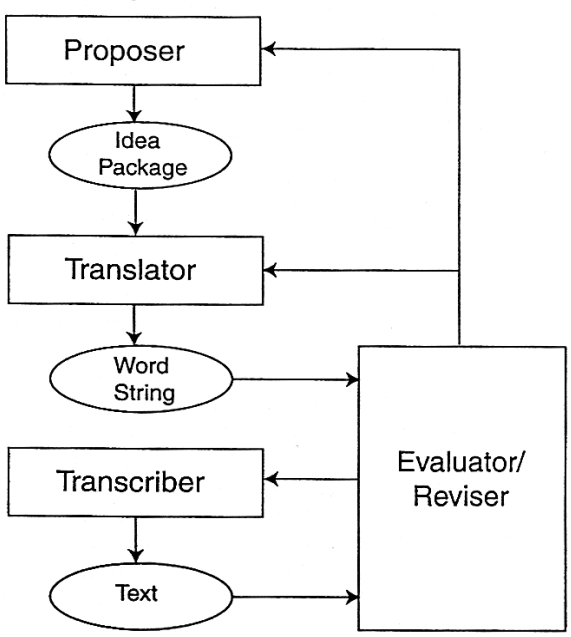


**Figure 4-1** Model of the text production process from Chenoweth and Hayes (2003, p. 113).

The model distinguishes between four different processes. The proposer proposes ideas for expression. This component is assumed to include the higher level processes involved in planning and reflection as characterised in global models of the writing process, and is responsible for creating an idea package to be formulated in language. Essentially it is responsible for deciding what to say next. The translator is responsible for converting this message into linguistic strings. The transcriber then converts this linguistic string into written text. Finally, the evaluator/ reviser is responsible for monitoring and evaluating word strings and text as they are produced, and for revising them when this is required.

In its general form, this model corresponds directly with what has become the standard model of spoken language production (Levelt,1989, 1999; Vigliocco & Hartsuiker, 2002). We assume, since the output from the translator process in Hayes' model is a phonological representation, that up to this stage of the process it is identical to a model of spoken language production. There are then a range of differences between speech and writing. Of these, two seem to us to be particularly important. First, the external representation of language in writing means that the writer has access to the preceding text in planning coherent continuations of what they have already said, and is not restricted to the fading trace of speech in short term memory. This means that pauses between segments of text could reflect re-reading of previous text as well as planning of the unit of text itself. Second, there may be important differences in the extent to which internal and external monitoring can be carried out. Vigliocco and Hartsuiker (2002) have suggested that, although inner speech can in principle be monitored, this cannot be done at the same time as speaking, arguing that externally articulated speech may overwrite the representation of inner speech in short term verbal memory. In our view, this may be an important difference between speech and writing. Writing may allow the monitoring of inner speech at the same time as transcription, and hence may allow for greater revision prior to transcription. This means that pauses during writing may reflect revision of planned language as well as the mental planning of the next unit of language.

This receives some support from the study carried out by Kaufer, Hayes and Flower (1986), which compared thinking-aloud protocols produced by expert and novice writers. We will describe this in detail to illustrate how bursts in text production are typically analysed. Kaufer et al. segmented the protocols into language bursts. Bursts that ended in a pause of two or more seconds were classified as P-bursts; bursts that were terminated by an evaluation, revision or some other grammatical discontinuity were classified as R-bursts (Chenoweth & Hayes, 2001). Figure 4-2 shows an illustrative example adapted from Kaufer et al. (1986) of the process of producing a sentence, broken down into segments and bursts. We have

typed in bold text that we assume was transcribed, though the moment at which it was transcribed is not indicated in the original paper.

| Segment | **Protocol content** (Pauses >2 seconds are represented by dashes) | Process | Model component | Burst type |
|---|---|---|---|---|
| **1** | **The best thing about it is that –** | Proposed part | Translator | P-burst |
| **2** | what? | Goal setting | Proposer | |
| **3** | Something about using my mind - | Goal setting | Proposer | |
| **4** | **it allows me the opportunity to** – uh - | Proposed part | Translator | P-burst |
| **5** | I want to write something about my ideas - | Goal setting | Proposer | |
| **6** | to put ideas into action or to - | Proposed part | Translator / Evaluator | P-burst |
| **7** | develop my ideas into - | Proposed part | Translator / Evaluator | P-burst |
| **8** | what? – | Goal setting | Proposer | |
| **9** | into a meaningful form? | Proposed part | Translator | R-burst |
| **10** | Oh, bleh! - | Evaluation | Evaluator / Reviser | |
| **11** | say it allows me - | Re-reading | Evaluator / Reviser | |
| **12** | to use - | Proposed part | Translator | P-burst |
| **13** | na! - | Evaluation | Reviser | |
| **14** | allows me - | Re-reading | Reviser | |
| **15** | scratch that –(Deletion of "the opportunity to" in 4) | Evaluation | Reviser | |
| **16** | The best thing about it is that it allows me to use - | Re-reading | Reviser | |
| **17** | **my mind and my ideas in a productive way**. | Proposed part | Translator | P-burst |

**Figure 4-2** Process of the production of a sentence broken down in bursts and segments. Taken from Kaufer, Hayes and Flower (1986, pp.125-126).

The writer starts by proposing language which is then transcribed. Because the part appears in the think-aloud protocol we take this as representing the output of the

translator rather than the transcriber. Since the part ended in a pause, Kaufer et al. classified this as a P-burst, and assume that this represents the full output of the content entered into the translator. The next two segments represent operations involved in creating content and reflect the output of the proposer. In segment 4, a new sentence part is proposed and transcribed. In segments 5, 6 and 7, the writer first searches for content and then proposes two alternative sentence parts. Neither of these is written down, so we assume that this is output from the translator which is rejected by the evaluator. Note, though, that both sentence parts were classified as P-bursts despite that fact that they were not transcribed. By contrast, the proposed sentence part in segment 9 is classified as an R-burst because it is followed by an explicit evaluation (segment 10), and this occurs within 2 seconds, and so is taken as a possible interruption of the output of the translator. The distinction appears to be that segments 6 and 7 are rejected after a delay, whereas segment 9 may have been interrupted before the output of the translator has been completed. The aim of the classification is to distinguish cases where the burst reflects the capacity of the translator from cases where the output of the translator has been interrupted before it is completed.

It is important to note that the observations are based on a spoken protocol. This means that a great many of the bursts that are taken to reflect the output of the translator will not be apparent in writing since they are rejected before they are written down. Thus, the protocol above would, in a keystroke log, reduce to: ***"The best thing about it is that*** PAUSE ***it allows me the opportunity to*** PAUSE deletion of "the opportunity to", replacement with ***to use*** PAUSE ***my mind and my ideas in a productive way."*** This raises an important problem for the interpretation of burst lengths in writing. Hayes (2009) has argued that the length of P-bursts may reflect in part the capacity of the translator, and, more generally, individual differences in language skills. While this may be true of bursts in spoken protocols it is less clear that this is the case for written output. Burst lengths may be strongly affected by writing strategy. Thus, a writer who tries to produce well-formed sentences before they transcribe them, may try out and revise several bursts of language mentally before then producing a fluent P-burst. By contrast, a writer who transcribes the output of the translator directly, and then revises as they go along may produce fewer P-bursts shorter in length. The point being that variations in P-burst length may reflect the writer's overall strategy for combining translation and revision, rather than the capacity of the translator.

Finally, there is a general conceptual issue with the fact that bursts are only defined by how they are terminated. This means that bursts which follow a revision are treated as equivalent to bursts which follow a pause. It seems possible, however,

that bursts which follow a revision involve modifying the previous burst instead of producing new content, and hence may differ systematically in length from "pure" P-bursts which begin and end with a pause. This would be particularly important if writers differ strategically in the extent to which they revise bursts as they are produced. In what follows, therefore, we will analyse bursts in terms of both how they are initiated and how they are terminated, and test whether this is associated with systematic differences in length.

### *4.1.3 Aims of the analysis*

In our view, the fundamental issue that arises in trying to use keystrokes as indicators of underlying cognitive processes is one of alignment. As we have seen, if one were to take the raw keystroke log of a writing session and then analyse it as an undifferentiated whole, the characteristics of the pauses, burst and revisions that one would measure would potentially reflect a very heterogeneous range of underlying processes. (This is, in fact, not infrequently what is actually done (see Wengelin, 2006, for a review)). Our general approach to the analysis, therefore, had the overriding aim of trying to improve the mapping of measures onto hypothetical underlying processes as much as possible. This can be broken down into 4 complementary aims.

First, we used the final product to sort the keystroke log into different kinds of activities. Our main aim here was to isolate text production in its "purest", most linear form from other activities. In brief summary, this involved the following distinctions. (i) Distinguishing the initial draft of text from post-draft revision, on the grounds that text production during post-draft revision may be different in form from text production during the initial draft. (ii) Within the initial draft, distinguishing between linear and non-linear transitions between units of text in order to separate text produced as part of the forward progression of the text from revision activities. (iii) We then distinguished within the non-linear events between revision at the leading edge and revision elsewhere in the text. In the analysis section we will describe the procedures that we used to do this, and elaborate on some of the more fine-grained distinctions that emerged in the course of our engagement with the data.

Our next aim was to develop more differentiated measures of pauses and bursts than have previously been used. Pauses have typically been measured in one of two ways (Wengelin, 2006). First, by imposing a threshold – typically of 2 seconds – and then defining only the intervals that last longer than this threshold as pauses. Analysis then typically involves comparing the frequency with which these pauses occur at different text locations or the frequency with which they are produced by different writers or under different conditions. The problem is that this restricts

analysis to longer pauses which presumably reflect higher level processes, and ignores the shorter pauses involved in more linguistic processing. The alternative approach is to use the raw intervals between events and calculate measures of central tendency for different kinds of pause durations. The problem with this is that the resulting distribution of pause durations is extremely heterogeneous– it mixes together the exceptional, longer pauses and the more routine, shorter pauses in the same distribution. This can be reduced by grouping the pauses according to the location at which they occur (e.g. within words, between words, sentences and paragraphs), but the resulting distributions are still extremely skewed and represent a heterogeneous mixture of different kinds of pauses. In what follows we will use mixture models (McLachlan and Peel, 2000) to identify sub-components of these distributions and to estimate separate measures of their central tendency. Our aim, was to try to achieve a better match between the pause measures and the underlying cognitive processes.

As we have seen, bursts have typically been classified as either P-bursts or as R-bursts. The aim being to isolate bursts assumed to reflect the operation of the translator from bursts which may also involve the reviewer. On the basis of the issues we raised in discussing his model, we will apply a more differentiated classification with a view to improving the mapping between bursts and different types of process. We will assess the value of this by testing whether the resulting burst types differ in length.

Finally, we will use principal component analysis to assess the relationships between different measures and the correspondence between the separate components and components of cognitive models of writing.

## 4.2 METHOD

The data we will be discussing here were collected as part of a larger study investigating the effects of self-monitoring and planning on writing. As a part of this study we collected keystroke logs using Inputlog (Leijten & van Waes, 2006). The texts were written by 80 participants recruited at the University of Groningen. Participants were asked to plan and write an article for the University newspaper discussing whether 'our dependence on the computer and the Internet is a good development or not'. The experiment was divided into three phases. During the first and third phase we administered a collection of measures that investigated the development of understanding through writing (see Baaijen, Galbraith & de Glopper, 2010, for more information and a full description of the experiment). In the writing phase of the experiment, participants were given 5 minutes to plan the writing assignment using

pen and paper. Half of the participants had to write down a single sentence summing up their overall opinion (synthetic planning) and the other half of the participants had to construct a structured outline (outline planning). They then had 30 minutes to write the article on a computer. It was stressed that they had to produce a well-structured and complete article in the time available. Participants were allowed to consult their written outlines.

Our aim is to describe the procedures we used to construct measures from the keystroke logs, and to examine the properties of these measures across the sample of texts as a whole. In future papers, we plan to investigate whether these characteristics vary as a function of self-monitoring and type of planning, as well as how they relate to the development of understanding and text quality.

### *4.2.1 Data preparation and coding*

Our first aim in data preparation was to isolate "pure" text production - linearly produced text – from revision and a number of other types of output.

First, we excluded titles and treated these as a separate category. Although these are clearly text production, they are very different to other examples. They often involve considerably longer pauses, may be produced at the end or in the middle of the writing session, and following lengthy re-reading episodes.

Second, we categorised text produced as part of explicit planning separately from text production. Some writers in our sample (20%) broke off from producing full text to make a plan on screen. Quite apart from the point that explicit planning is a distinctively different process and, therefore, should be analysed separately from text production, these episodes also provide good examples of why the raw output of Inputlog should always be checked closely. When making explicit plans, participants often press ENTER to mark the beginning of a new point. Inputlog automatically codes these as between paragraph boundaries and doesn't distinguish these from paragraph boundaries that are intended for inclusion in the text.

Third, we distinguished text produced during an initial draft from text produced during a revision draft. Many writers (65%) show evidence of writing an initial draft and then of going back systematically through the text, editing and revising the initial draft. Although this may sometimes involve producing extended chunks of text, we did not include this in the analysis of text production on the grounds that it may be different in character to text produced as part of the initial draft. We defined end revision as occurring when an individual made revisions outside the final paragraph while writing the final paragraph. In the majority of cases, this amounted to revisions made after the final sentence. But there were some cases

where individuals broke off to make revisions and then returned to write a final summary sentence or two. In these cases these sentences were also excluded from text production analysis.

We then distinguished between linear text production and revision during the initial draft by making a distinction between linear transitions and events. Linear transitions were defined as 'empty' thinking episodes before the continuation of text production; events were defined as episodes that include other material or operations before the continuation of text production. These included scrolling and other movements which might indicate re-reading or evaluating previously written text, or the insertion of text away from the leading edge. This was done, on the basis of the automatic output from Inputlog, by flagging objective locations where other material or movements were executed before the continuation of text production. As well as allowing us to assess pause durations from of the linear transitions alone, this also enabled us to quantify the amount of revision at different text levels by calculating the percentage of linear and non-linear transitions at different pause locations.

Finally, we defined pauses in a way that corresponds to a model of spoken language production rather than in terms of the keystrokes. Typically, a between word pause in the keystroke log consists of two pauses, one pause before the between word marker (SPACE) and one pause before the first letter of the next word. These are often classified separately (e.g. Wengelin, 2006). In Inputlog these separate locations are coded with the same code, so that mean between word pause durations provided in the automatic summary output represent a hybrid of before and after <SPACE> pause durations. We decided to take the sum of these two pauses together on the assumption that space presses between words and sentences are relatively automatic motor activities, and that the interval between the end of one word and the beginning of another is a better reflection of the cognitive processes between these units. We applied the same principle to the calculation of pauses at other locations.

Overall, then, we distinguished the following pause locations: within words (WW), between words (BW), between sub-sentences (BSS) (indicated by commas, semicolons and colons), between sentences (BS) and between paragraphs (BP). These correspond to the boundaries used in other research (e.g. Wengelin, 2006) but differ in the way they are calculated.

## 4.3 RESULTS

### *4.3.1 Coding language bursts*

This analysis was carried out on the sections of the keystroke logs corresponding to the initial draft. First, we identified automatically all interruptions longer or equal to two seconds, and then classified these as either (i) P-boundaries, when they were associated with a linear transition, or as (ii) R-boundaries, when they were associated with an event. Second, we scanned all other automatically flagged events to identify whether they were associated with revisions of text. A number of important features should be noted.

First, typos are extremely common in keystroke logs. Following conventional practice (see Wengelin, 2006), we excluded these as indicators of revision and included the relevant sections of text as part of a burst. We classified typos as corrections of errors within a word, which leave the word otherwise unchanged, and which do not result in a break of more than two seconds before the next keystroke.

Second, another characteristic of keyboarded texts is that writers are much more able to modify previously produced text. In our coding, instances where writers leave the leading edge to insert text elsewhere are automatically flagged as non-linear transitions. We classified bursts terminated by insertions as a separate category of revision burst ((RI) – **R**evision involving **I**nsertion), in contrast to bursts terminated by revision at the leading edge of the text ((RL) – **R**evision at **L**eading edge). Since some writers occasionally did both at the same time, we added a combined code for these cases (RLI). A key question is whether RL-bursts and RI-bursts differ in length. If RI-bursts involve the completion of a full burst followed by evaluation leading to insertion, one might expect them to be similar in length to P-bursts and longer than RL-bursts, where evaluation appears to interrupt the production of a burst.

Third, we classified the insertion-bursts themselves as a separate kind of burst taking place during revision. Since these are produced in the circumscribed context of existing text, we would expect them to be shorter than P- or R-bursts. Overall, we distinguish between three different kinds of insertion-bursts: (i) IG-bursts reflect within-sentence revision; (ii) IR-bursts reflect end-of-sentence revision; and (iii) IB-bursts reflect revision over sentence boundaries (see table 4-2 for a detailed definition of bursts). For all bursts, length was calculated by counting all words within the burst, including any partial completions of words.

The final distinction was designed to enable us to test whether there is a difference between bursts initiated after a pause and bursts initiated after a revision. We classified pure P-bursts (PP-bursts) as bursts which begin and end with a 2 second

pause, and distinguished these from RP-bursts, which begin after a revision and terminate with a 2 second pause. We distinguished three different types of RP-bursts. (i) RP1-bursts are bursts which consist entirely of new language terminated by a pause. These typically occur following an insertion elsewhere in the text, and we would expect these to be similar in length to PP-bursts. (ii) RP2-bursts involve the replacement of preceding text, followed by further text production terminated by a pause. (iii) RP3-bursts are bursts which only involve the replacement of the preceding text and then pausing before continuing with text production. We would expect these to be shorter than PP-bursts. In principle, we think that PP-bursts should be separated from RP-bursts if they are to be used as estimates of the language capacities of writers.

Finally, we applied the same principle to R-bursts, and classified these separately as PR-bursts and RR-bursts. Table 4-2 shows a list of the different categories of bursts, along with a brief definition.

**Table 4-2** Definition for different language bursts distinguished on the basis of keystroke logging data

| **Burst type** | **Definition** |
|---|---|
| PP | Burst initiated after and terminated in a pause of at least two seconds |
| RP1 | Burst initiated after a revision and terminated in a pause of at least two seconds which involves the production of new language only |
| RP2 | Burst initiated after a revision and terminated in a pause of at least two seconds which involves the replacement of previously written text as well as the production of new language |
| RP3 | Burst initiated after a revision and terminated in a pause of at least two seconds which only involves the replacement of previously written text |
| PRL | Burst that is initiated after a pause of at least two seconds and which is terminated by a revision at the leading edge |
| PRI | Burst that is initiated after a pause of at least two seconds and which is terminated by a revision insertion, i.e. a revision carried out elsewhere in the text either within the same sentence or elsewhere in the text |
| PRLI | Burst which is initiated after a pause of at least two seconds and which is terminated by a revision at the leading edge as well as a revision carried out elsewhere in the text |
| RRL | Burst that is initiated after a revision and terminated by a revision at the leading edge |
| RRI | Burst that is initiated after a revision and terminated by a revision insertion, i.e. a revision carried out elsewhere in the text either within the same sentence or elsewhere in the text |

| Burst type | Definition (continued) |
|---|---|
| RRLI | Burst that is initiated after a revision and terminated by a revision at the leading edge as well as a revision carried out elsewhere in the text |
| IG | Insertion burst which is initiated in the middle of sentence production and which involves a revision within the current sentence |
| IR | Insertion burst which is initiated at the end of sentence completion and which involves a revision within the current sentence that has just been produced |
| IB | Insertion burst which is initiated either in the middle of sentence production or after sentence completion but which involves a revision carried out over sentence boundaries, i.e. elsewhere in the text and is no longer than a sentence in length |

### *4.3.2 Length of different burst types*

We tested first whether there was an overall difference in length between PP, RP, PR, RR and I-bursts using a one-way within subjects ANOVA with burst type as the independent variable and mean burst length as the dependent variable. Because the data lacked sphericity degrees of freedom were adjusted using the Greenhouse-Geisser correction. The results showed a highly significant difference in burst length depending on burst type ($F(3.49, 272.44)=144.00$, $p<.0005$). Planned pair-wise comparisons, with a Bonferroni adjustment for multiple comparisons, revealed no significant difference in length between PP ($M$=5.87 words, $SD$=1.87) and RP ($M$=5.58, $SD$=1.26) bursts ($t(78)=1.71$,$p$=.90) and no significant difference between PR ($M$=4.43, $SD$=1.25) and RR ($M$=4.65, $SD$=1.25) bursts ($t(78)$ =1.32, $p$=.1). These initial results suggest that whether a burst follows a pause or a revision has no effect on the length of the burst. However, there were highly significant differences between P, R and I-Bursts. R-bursts were significantly shorter in length than both P-bursts ($p<.0005$ for all comparisons); I-bursts ($M$=2.22, $SD$=1.45) were significantly shorter than all other burst types ($p<.0005$ for all comparisons). These results support the assumption that R-bursts are bursts that have been interrupted before they have been fully executed. They also suggest that insertions in earlier text tend to consist of short, partial bursts modifying earlier bursts rather than the production of full bursts of new language.

The next step in the analysis was to investigate sub-categories within the RP, PR and RR-categories. A breakdown of the sub-categories is shown in figure 4-3.

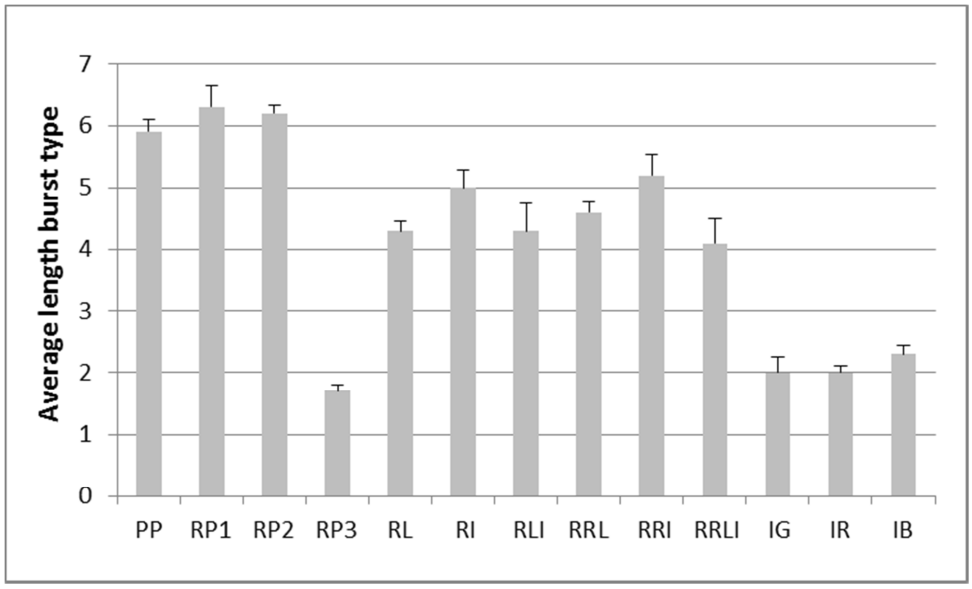


**Figure 4-3** Average lengths for different burst types with standard error bars

First, we used a one-way within subjects ANOVA to compare the burst lengths of the different RP-types with pure P-bursts. This showed a significant overall effect of burst type ($F$(1.90, 136.75)=143.65, $p$<.0005) with the RP3-burst type ($M$=1.70, $SD$=.97) being significantly shorter than all the other bursts ($p$<.0005 for all comparisons). These results suggests that RP3-bursts, which consist simply of the replacement of previous text, are much shorter than bursts where new language is being added, and should not be included in estimates of P-burst length.

Second, in order to test whether revision bursts terminating with revision at the leading edge (RL) differ from revision bursts terminating with an insertion elsewhere (RI), and whether this is affected by how the burst is initiated (following a pause (PR), or following a revision (RR)), we carried out a 2-way within subjects ANOVA with termination type and initiation type as independent variables and mean burst length as the dependent variable. We excluded the mixed category (PRLI and RRLI) from this analysis because we wanted to compare pure examples of each type. This showed a significant effect of termination type ($F$(1,46)=6.17, $p$=.017), but no significant effect of initiation type ($p$=.36). The results confirm the earlier finding that R-bursts do not vary in length depending on whether they are produced following a pause or a revision. However, they also suggest that RL-bursts ($M$=4.31, $SD$=.15) are shorter than RI-bursts ($M$=4.95, $SD$=.24). One possible explanation for this is that RIs occur when evaluation is applied after the burst has been produced, and hence a burst is more fully completed than when evaluation is applied during burst production.

Note that RI-bursts are not as long as pure P-bursts, suggesting that these bursts are not as fully planned as a normal P-burst.

### *4.3.3 Pauses*

Pausing durations were measured from all the linear continuation pauses at each different location. As can be seen in the histogram of the between word pause durations for one of the participants (P307) in our data (figure 4-4), their distribution is heavily skewed.

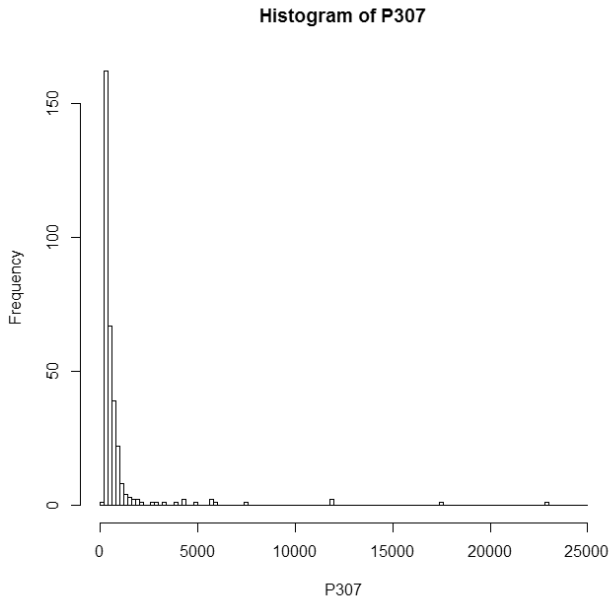


**Figure 4-4** Histogram of participant 307 showing the particularly skewed distribution of between word pauses.

A standard transformation to reduce positive skew is to take logs. The histogram of the resulting distribution is shown below (figure 4-5).

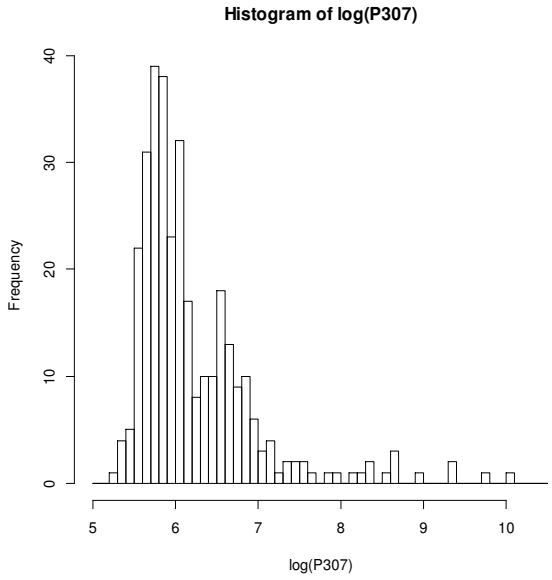


**Figure 4-5** Histogram of log transformation of between word pauses for participant 307.

Although this does pull the extreme pause durations in towards the main body it is clear that the distribution is still extremely skewed. Furthermore, the left hand of the distribution is noticeably bimodal, which implies that the median is a poor estimate of central tendency. This has also been observed by Kirsner, Hird and Dunn (2005) in between word pauses for spoken language production. They suggest that, rather than treating this as a single skewed distribution, a better approach is to assume that the distribution is a mixture of several components and to fit mixture models in order to estimate the parameters of the distributions. We used the R-package EMMIX developed by McLachlan and Peel (2000) to fit a series of models to the data, and estimated the goodness of fit of these models using the Bayesian Information Criterion (BIC). The best fitting distribution for this participant (P307) for the between word pauses is shown in figure 4-6.

This model had a BIC value of 476.33, which was overwhelmingly a better fit than a single normal distribution (BIC=5871.08), a single lognormal distribution (725.16), a double lognormal distribution (550.25) and the model with four normal distributions (627.65). (Lower values are superior, and Kass and Raftery (1995, p.777) suggest that a difference greater than 3.7 can be considered a positive difference, and a difference greater than 20 can be considered a strong difference).

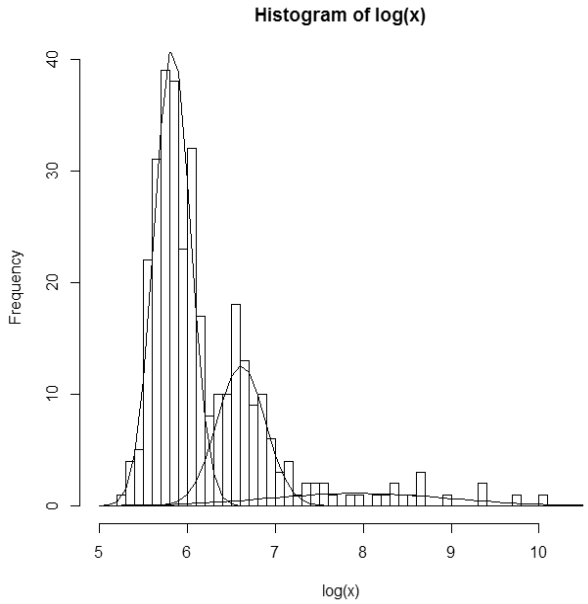


**Figure 4-6** Histogram showing three fitted models for participant 307.

For this particular participant, there appears to be a good case for suggesting that there are three different distributions within the data for the between word pauses. One might speculate that the left hand distribution (65% of the pauses) represents word retrieval processes, the middle distribution (26% of the pauses) represents phrase boundary processes, and that the right hand distribution (9% of the pauses) represents higher level message planning or reflection. Note, though, that some of the more extreme pauses are extremely long (the longest pause for this participant is almost 23 seconds), which suggests that more is involved than a higher level of structural planning. Indeed, it is doubtful whether the right hand distribution should be treated as a single normal distribution since, as shown in figure 4-6, the symmetry assumption implies that it also includes a proportion of extremely short pauses. Our conclusions for this participant are: (i) that there are two normally distributed groups of pauses and that these would be better represented by two lognormal means than by a single median; (ii) that the more extreme pauses should be treated as a miscellaneous group identified by a threshold determined by the tail of the second lognormal distribution. In this case, 3 standard deviations above the mean the threshold would be log 7.43, which is equivalent to a pause duration of 1686 milliseconds.

To assess the generality of these findings, we fitted the same series of models to the pause durations for all participants, and assessed the goodness of fit of the models using the BIC values. This analysis showed that for 58% of the participants the three

distribution model was the best fitting, with 19% of the participants being best fit by a two distribution model, and the remaining 23% being indeterminate between two and three distributions.

We repeated this analysis for the other pause locations. For the pauses between sentences and sub-sentences, single lognormal distributions generally had the best fit. It was harder to determine which distribution was the best overall fit for the within word pauses. Generally, three distributions were a better fit than four distributions (54% of the participants), and therefore we selected these lognormal scores for the analysis of the within-word pauses.

To assess the differences in pause duration for different text locations a one-way within subjects ANOVA was conducted comparing the means for the left hand distributions at each text location. There was a significant main effect of pause duration (Greenhouse Geiser $F(2.56, 77) = 1378.68$, $p<.001$). Planned comparisons, with a Bonferroni adjustment for multiple comparisons, showed that pause durations at all locations were significantly different from one another ($p<.0005$ in all cases). These results (see table 4-3 for an overview of pause times at different locations) confirm previous findings suggesting that pause length increases for higher level locations within the text (Wengelin, 2006).

Further analysis of the subdivisions within specific pause locations showed that the means from these distributions were still significantly lower ($p<.0005$) than the pause times for the next level up. This suggests that these sub-divisions correspond to sub-components of the distributions at particular locations rather than to overlap with higher level text locations.

**Table 4-3** Means and log scores for different pause locations.

| Pause locations | Mean log (SD) | Equivalent pause time |
|---|---|---|
| Within words | 4.5 (.44) | 90.02 |
| Between words | 5.6 (.25) | 270.43 |
| Between sub-sentence | 7.1 (.43) | 1211.96 |
| Between sentences | 8.0 (.54) | 2980.96 |

We then assessed the relationship between pause duration and burst length. A natural assumption is that the longer one pauses, the longer the burst length. First, we correlated the mean (log) pause durations for each individual at between word,

between sub-sentence and between sentence locations with their mean PP-burst and mean R- burst lengths. As can be seen in table 4-4, these were negative in all cases, and very significantly so for between sub-sentence and between sentence locations. This suggests that the average length of time an individual pauses at grammatical boundaries is, in fact, negatively associated with the average length of the language bursts they produce. In order to assess this within individuals, we then calculated the correlations at between word and between sentence levels for all an individual's pauses and the associated burst lengths in their text production samples. Although a few individuals showed significant correlations, these were not consistently positive or negative, and the vast majority of individuals (85% for the sub-sentence level and 96% for the sentence level) showed no significant relationship between pause duration and pause length.

**Table 4-4** Pearson product-moment correlations between mean pause length and burst characteristics.

| Pause location | P-burst length | R-burst length | Percentage of P-bursts |
|---|---|---|---|
| Between words | -.15 | -.21 | .14 |
| Between sub-sentences | -.48* | -.50* | .21 |
| Between sentences | -.41* | -.56* | .52* |

**p<.001 (2-tailed)*

An alternative possibility is that longer pause durations are associated with producing "better bursts", perhaps through mental evaluation and revision prior to production. To assess this we calculated the correlation between mean (log) pause duration and the percentage of P- bursts for across participants. As can be seen in the table, there was a highly significant positive correlation for the between sentence level pauses, with longer pauses being associated with a higher percentage of P-bursts.

### *4.3.4 Revision*

We assessed revision in four ways. First, we calculated a number of global indicators of the extent to which the text was modified in the course of creating the final product. These included: (i) the Text Modification Index, which consists of the ratio of the total number of words in the process log to the total number of words in the final product,

with higher numbers indicating greater amounts of revision. (ii) Measures of the total amount of text deleted as a percentage of the total number of process words.

Second, in order to assess changes made during post-draft revision, we calculated the number of words that were modified during post-draft-revision (as a percentage of the total number of words) and the number of words that were added during post-draft revision (again as a percentage of the total number of words).

Third, in order to distinguish revision at the leading edge from revision elsewhere, we assessed these separately. To assess revision at the leading edge we counted the total number of words modified at the leading edge as a percentage of the total number of process words produced during the initial draft, and the mean amount of deletion carried out in each of these instances. To assess non-local revision, we counted the total number of words added during insertions (see earlier burst analysis section) as a percentage of the total number of process words produced during the initial draft, and the mean size of these insertions.

Fourth, we used the distinction between linear transitions and event-filled transitions to assess the extent to which the linear progression of text production was disrupted at different levels of text structure. To do this, we calculated the percentage of linear transitions compared to non-linear transitions (events) at each different text level. These measures provide a fine grained measure of the forward movement of the text, and, by measuring this separately for word, sentence and paragraph transitions, enabled us to assess different levels independently. Note that these measures capture linearity in processes as well as product. Thus, a non-linear transition between words can, in principle, involve anything that is not a straightforward continuation to the next word: it could be a matter of scrolling elsewhere in the text to re-read, and then returning to the leading edge to produce text, or it could involve revising at the leading edge, or at another location within the text. In order to assess linearity at a more global level and in a way that captured product changes alone we also calculated a sentence linearity index. This involved comparing the order of sentences in the final product with the order that they were produced in. Each sentence in the final product was numbered by when it was produced during text production, and the percentage of non-linear sentences was calculated.

### *4.3.5 Interrelationships between pauses, bursts and revision during text production*

In order to assess the relationships between these different measures, we carried out a principal component analysis (PCA) on 16 variables selected to cover the range of different activities and because they satisfied the assumptions required for PCA. The

set of measures had a Kaiser-Meyer-Oklin sampling adequacy of .71 and Barlett's test of sphericity was highly significant ($p$<.001), indicating that principal component analysis is appropriate for these data (Field, 2009). Varimax rotation was used to extract orthogonal components from the data.

We selected the five components with eigenvalues over Kaiser's criterion of 1 and which, taken together, explained 73.6% of the variance. Table 4-5 shows the loadings of each variable on these components after rotation, as well as the amount of variance that each component accounts for, and Cronbach's alpha for each component.

The components can be described in the following ways:

*Component 1 - Planned sentence production:* This component consists of a combination of longer pauses at grammatical boundaries (sentences and sub-sentences), as well as the exceptionally long pauses between words, along with short bursts.

*Component 2 - Within sentence revision:* This component consists of high levels of revision at the leading edge of the text combined with non-linear transitions at the word and burst levels. This suggests that it reflects the extent to which different writers revise sentences as they are produced.

*Component 3 - Revision of global text structure:* This component represents the extent to which the global plan of the text is revised as captured by the between sentence linearity. Note that it captures a revision at a higher level than the preceding component.

The remaining 2 components are much lower in reliability, with few item loadings, making them harder to interpret. Their main interest is in suggesting further independent features of the process for which additional measures could be developed. We have tentatively labelled these as:

*Component 4 - Post draft revision:* reflecting the extent to which revision is postponed until after the initial draft) and

*Component 5 - Careful word choice:* reflecting perhaps the extent to which words are chosen carefully.

Component 5 is mainly of interest in demonstrating that different aspects of between-word pauses load on different components, supporting the suggestion that these should be measured separately.

**Table 4-5** Summary of PCA with varimax rotation for 5 factor solution

| Variables | Rotated Factor Loadings | | | | |
|---|---|---|---|---|---|
| | **Component 1** Planned sentence production | **Component 2** Within sentence revision | **Component 3** Revision of global text structure | **Component 4** Post draft revision | **Component 5** Careful word choice |
| Overall length R-bursts | **-.890** | -.110 | -.001 | .183 | -.071 |
| Overall length P-bursts | **-.839** | -.036 | -.152 | .205 | -.018 |
| Percentage of between-word pauses above threshold of 2 seconds | **.823** | -.248 | .044 | -.083 | .204 |
| Average pause time between sub-sentences (log) | **.715** | -.185 | .069 | .252 | .067 |
| Average pause time between sentences (log) | **.680** | -.179 | -.388 | -.052 | .154 |
| Percentage of leading edge revision | -.099 | **.867** | .154 | .016 | .098 |
| Text modification index | -.230 | **.836** | .101 | .259 | .101 |
| Percentage of burst as P-burst | .204 | **-.738** | -.386 | .026 | .176 |
| Percentage linear word transitions | -.278 | **-.627** | **-.464** | .075 | .094 |

(continued)

| Variables | Rotated Factor Loadings | | | | |
|---|---|---|---|---|---|
| | **Component 1** Planned sentence production | **Component 2** Within sentence revision | **Component 3** Revision of global text structure | **Component 4** Post draft revision | **Component 5** Careful word choice |
| Percentage of insertion revision | -.009 | .105 | **.820** | .309 | .077 |
| Percentage linear sentence transitions | -.076 | -.260 | **-.743** | -.020 | -.215 |
| Percentage of end revision | -.014 | -.059 | .206 | **.845** | -.104 |
| Percentage of words produced during text production | .183 | -.240 | .131 | **-.672** | -.102 |
| Average pause time between words 2 (log) | .194 | -.109 | .009 | .110 | **.827** |
| Percentage linear paragraph transitions | -.152 | -.266 | -.358 | .234 | **-.640** |
| Eigenvalues | 4.34 | 3.57 | 1.51 | 1.25 | 1.07 |
| % of variance | 27.17 | 22.48 | 9.43 | 7.82 | 6.70 |
| $\alpha$ | .87 | .84 | .78 | .58 | .47 |

*Note that factor loadings over .40 appear in bold.*

## 4.4 DISCUSSION

We set out on this analysis with four main aims, which we will now discuss in turn.

Our first aim was to establish a set of procedures for separating out different components of the writing process in order to be better able to relate keystroke measures more directly to specific cognitive processes. Overall, we think that the procedures that we employed did enable us to separate "pure" text production from other processes, and to distinguish between different forms of revision. This is an important initial step in aligning keystroke units of analysis more directly with the

cognitive components of writing process models. We would not wish to claim, however, that the procedures we have used are necessarily the best way of doing this, nor that they are the only procedures that need to be applied. We have focussed almost entirely on procedures for isolating "pure" text production from other activities and on procedures that can be carried out automatically or quasi-automatically. Our main point is to emphasize the need to carry out these kinds of procedures, and the general principle that the procedures should be designed to isolate components of the keystroke log that can be mapped on to components of process models of writing. Our impression from reading reviews of research in this area (e.g. Sullivan and Lindgren, 2006) is that it is not uncommon for researchers to use the automated output from a software analysis program to analyse undifferentiated logs of the whole text.

Perhaps the most important of these procedures was the separation of linear transitions from event transitions, which had a number of important benefits. First, it enabled us to separate "pure" text production within the initial draft from other processes, and hence to base our analysis of pauses solely on those associated with the forward progression of the text. Although these may in part reflect re-reading, the fact that they are associated directly with the following unit of text, makes it much more plausible to interpret them in terms of the planning of text production. In addition, this enabled us to distinguish between revision at the leading edge of the text and revision elsewhere in the text, and – in the form of percentages of linear continuations at different text levels - provided us with global measures of the extent of revision at different levels of text production.

Finally, our experience of coding the logs reinforced the need (stressed by the developers themselves) to be wary of the automated output provided by the software. Logs, and associated automatic statistical analyses, really do need to be closely screened before analysis begins.

### *4.4.1 Pauses*

Our second aim was to use mixture modelling to estimate the parameters of pauses at different locations. Our results suggest that this has a range of advantages, and we would recommend that this become standard practice. First, it extends the general principle of differentiating different components of the writing process down to the pause level of analysis. Modelling the pauses between words, for example, enabled us to distinguish between routine transitions between words, a further distribution centred round a longer pause duration, and exceptional cases where some non-routine and presumably higher level reflective activity is taking place. We have

suggested that these might reflect lexical retrieval, phrase structure processing, and higher level message planning, but do not want to make strong claims about this. Identifying the source of these differences should clearly be an important topic for future research. The key point is that, without modelling the distribution of pauses, these distinctions would remain completely unobservable. A second benefit of doing this is that it enables one to assess effects on the different components of the distributions separately. We found, for example, that the group of exceptionally long pauses loaded on the component in the PCA corresponding to planned sentence production along with other measures of higher level sentence processing, whereas the more routine group of pauses did not. This approach also enables one to take a more principled and contextualised approach to establishing thresholds for identifying non-routine pauses at different locations. Furthermore, thresholds can be established relative to the particular characteristics of individual writers, rather than against a general criterion. Finally, we believe that this approach can be extended to pauses at a higher level than just within-word and between-word pauses. Although we found that a single lognormal distribution fitted the data best for between sentence pauses, this is probably a reflection of the relatively low number of such pauses present in the relatively brief texts that we analysed. We would expect analysis of much longer texts to reveal the same kinds of multimodal distributions for between sentence pauses as we have found for within- and between-word pauses.

### *4.4.2 Bursts*

The principal aim of this analysis was to establish how best to divide up the bursts of language occurring during text production. However, in the course of doing this we also observed that writers who pause for longer on average before writing sentences tend to produce shorter rather than longer PP-bursts, and that instead they tend to produce a higher percentage of clean PP-bursts in their text. It seems to us that if, as Hayes assumes, PP-bursts are a direct reflection of the capacity of the translator component of text production, then one would expect writers who typically pause for longer at grammatical junctures to also produce longer PP-bursts. The fact they don't, and instead produce a higher percentage of clean bursts, suggests that, the bursts that appear in written output reflect not only planning of content, but also the extent to which the writer is concerned with producing well-formed text, and hence the operation of the reviewer component of the model. This could be tested by assessing whether writers demonstrating this feature in keystroke logs also show more evidence of revision prior to transcription in think-aloud protocols. More generally, these results imply that research into the relationship between bursts and the capacity of the

translator may be more valid if they are carried out on spoken output or on relatively spontaneous kinds of writing where the use of strategies is minimised.

This conceptual issue aside, our aim with this analysis was to establish whether a more refined classification system is required. We evaluated this by examining how the burst types within this classification system varied in length. Our first finding was that insertion bursts were much shorter than normal P- or R-bursts. This provides strong support for our general assumption that text produced as part of the forward progression of a draft should be analysed separately from text production occurring during revision. We should stress, however, that we do not regard this necessarily as an intrinsic feature of insertion bursts. Our texts were collected from writers producing a completed text in half an hour. Under other circumstances, writers might engage in much more extensive revision and produce larger I-bursts. Rather, our point is that for analytic purposes these bursts should be classified separately; future research is needed to establish how these bursts vary for different writers and under different conditions.

Our second main finding was that RP3-bursts, which only involve the modification of previous output, were much shorter than other bursts. These bursts which terminate with a pause, but which are initiated following a revision, would be classified as P-bursts using the standard P- and R-burst classification (see Chenoweth and Hayes, 2003, for an example) and hence, would seriously contaminate the collection of P-bursts. For analytic purposes we, therefore, recommend that bursts should be classified both in terms of how they are initiated and in how they are terminated.

Finally, we found that revision bursts terminated by revision at the leading edge were significantly shorter than revision bursts terminated by an insertion. It is possible that the relative frequency of these two types of revision bursts may vary between writers employing different drafting strategies, with some writers deliberately postponing revision to the end of sentences and others evaluating and revising bursts as they are produced. Future research should assess these differences.

### *4.4.3 Revision*

The approach to revision that we have taken has focussed less on different types of revisions (e.g. Witte & Cherry, 1986; Lindgren & Sullivan, 2006; Stevenson, Schoonen & de Glopper, 2006) and more on when they occur and on simple quantitative measures of how much they occur. In part this was a deliberate decision to focus on measures that might in principle be automated. But it was also a side effect of the fact that, in focussing on isolating "pure" text production from text production produced

during revision, and then on developing more fine-grained measure of pauses and bursts during "pure" text production, we paid less attention to the revision component of writing. We have, for example, not actively examined the role that pauses play in revision. Clearly, this is an area that should be developed further in future research.

That said, there are two features of the measures that we used that should be noted. First, in line with the general principle underlying our approach, we have distinguished between the contexts in which revision occurs – after a draft has been completed, at the leading edge of text production, and as non-linear departures from the leading edge - rather than treating them all as examples of an undifferentiated revision process. The fact that these three forms of revision load on different components in the PCA suggests that they are independent activities. Amount of revision at the leading edge is not, for example, inversely related to amount of post-draft revision. This suggests that there is not a simple trade-off between revision at the same time as producing text and revision after a draft has been completed as is implied by the different configurations of the monitors described in the original Hayes and Flower model. Rather these may involve different kinds of revision and need to be analysed separately.

Second, the revision measures that we have used capture complementary features of the process. The measures we have just discussed capture the amount of revision at the leading edge or elsewhere; the linearity measures (percentage of linear transitions at word, sentence and paragraph levels) capture the extent to which this happens at different levels of the text. Furthermore, since non-linear events included simple scrolling, these measures also capture the minimal component of revision involved in re-reading previous text. In combination, these measures provide information about both the amount of revision at different places in the text, and the point at which this is initiated. The fact that linearity of sentence transitions loads on the same component in the PCA as amount of revision insertion, whereas linearity of word transitions loads on the same component as revision at the leading edge, suggests that revision elsewhere in the text is more likely to take place at sentence boundaries than within sentences.

### *4.4.4 Interrelationships between pauses, bursts and revision*

Our aim with the PCA was to identify groupings of interrelated measures, considering both how different measures were related to one another, and the range of independent characteristics of text production that they captured. As we have seen already in the discussion, this played a valuable role in helping to interpret measures

in terms of potential underlying processes. We should stress, however, that this is not a full factor analysis. Given the small sample of texts, the results of the PCA should only be taken as an economical summary of the relationships between the measures within our sample. An important goal for future research is to collect data using a much larger sample of texts with a view to identifying a factor structure and scales that can be generalised across studies. One consideration we have had in mind in the course of developing the different measures described here has been to consider how readily they can be automatically assessed, and hence would be practical to use in collecting large samples of data.

The three main components that we have identified correspond fairly directly to different components of the writing process: planned sentence production and within sentence revision appear to capture the extent to which different writers plan and revise at the sentence level; revision of global structure appears to capture the extent to which different writers move back and forth in the text altering the linear ordering of sentences. An important feature of these components is that they are orthogonal to one another. In other words, writers who score highly on planned sentence production do not necessarily score higher or lower than other writers on the other components. These measures should therefore enable us to identify writers who combine planned/unplanned text production with immediate revision or not, and assess how this relates to other variables. At the same time, the combination of complementary measures in higher order dimensions helps consolidate our interpretations in a way that would not be possible with single measures alone.

### *4.4.5 Conclusion*

We hope that we have demonstrated that the procedures and analytic techniques that we have described here help to improve the alignment between measures derived from keystrokes and potential underlying processes. We do not claim to have exhausted all the possible ways in which this could be done. In particular, we think that a range of more fine-grained analyses of revision could be carried out. However, we do think that the general strategy of sorting the keystroke log according to actions in relation to the final product and then classifying measures of pauses, bursts and revisions either statistically or conceptually is a productive guiding principle. A principle which goes some way to addressing Schilperoord's doubts about the possibility of analysing text production in a context replete with intensive problem solving and massive editing.

# Chapter 5

*The Process of Discovery through Writing*

## *Abstract*

Even though writing is often characterised as a process of discovery there are contrasting conceptions about what is responsible for this epistemic effect of writing. Problem solving models of writing (e.g. Flower & Hayes, 1980; Bereiter and Scardamalia, 1987) attribute development of understanding through writing to deliberate planning designed to satisfy rhetorical goals. By contrast, Galbraith (2009) claims that discovery depends on the extent to which text production is dispositionally driven. This paper focuses on the fundamentally different role that these two theories attribute to text production processes and assesses both the development of understanding through writing as well as the processes responsible for it. The results show that discovery through writing has two accounts: strategic rhetorical planning processes and implicit text production processes. We will argue that this result is supporting Galbraith's dual process model of writing which incorporates both of these processes within its framework. More specifically, the results show that development of understanding is unrelated to text quality and that the processes responsible for these outcomes can be explained by the dynamic interaction between planned sentence production and global linearity which operate on different principles across different writing conditions.


*Keywords:* writing models, planning, development of understanding, writing processes, keystroke logging, text production

## 5.1 INTRODUCTION

Over the last thirty years, the dominant problem solving models in research on writing have treated expert writing as involving knowledge-transforming and hence, as a process of discovery. Typically, this research has focused its attention on the reflective thinking processes associated with writing, i.e. the thinking behind the text, rather than on the processes involved in translating ideas into written output. The essential claim has been that changes in understanding occur as a consequence of strategic rhetorical planning before, during or after text production (Flower & Hayes, 1980; Bereiter & Scardamalia, 1987). These problem solving models of writing, therefore, attribute the epistemic effect of writing to explicit planning processes in which writers evaluate their existing knowledge and modify this knowledge in order to satisfy rhetorical constraints. In other words, writing requires the active organisation of one's personal understanding. Once writers have externalized their own understanding of the topic in written output, this consequently enables them to reflect on their current understanding in relation to the audience for their text and to evaluate the expression of their ideas in respect to the requirements of the intended text. This explicit reflection process is assumed to lead to discovery, especially when it leads to the expression of new content. Bereiter and Scardamalia (1987) have described this process in their knowledge-transforming model. This is the most fully developed cognitive problem solving account of writing as a process of discovery.

In their research, Bereiter and Scardamalia focussed on the differences between expert and novices writers and distinguished between a knowledge-transforming model and a knowledge-telling model of writing. They claim that the knowledge-transforming model accounts for 'the peculiar value that many have claimed for writing as a way of developing one's understanding' (Bereiter and Scardamalia, 1987, p.302). In order for writing to function in this way, writers have to define the rhetorical context by establishing what their goals are for the text and how they want to present their ideas in the emerging text. These goals will then function as constraints within which the writer searches memory for relevant content and develops ideas for inclusion in the text. This process is described as a dialectical interaction between a content problem space that is concerned with the question "*What do I mean*", and a rhetorical problem space that is concerned with the question "*What do I say?*" (Bereiter and Scardamalia, 1987, p.303) and it is this interaction that is assumed to be responsible for discovery through writing. Bereiter and Scardamalia (1987) contrast this with a knowledge-telling process employed by novice writers, in which a fixed store of ideas in long term memory is translated directly into text. Essentially, this model only reflects the processes carried out within the content space

of the previous model and, therefore, does not include the interaction between the two problem spaces that is held responsible for discovery through writing.

More recently, the dual process model of writing (Galbraith, 2009) has shifted the view of writing from a problem solving approach towards a view of writing as text production. Galbraith (2009) claims that discovery through writing occurs during text production and hence, treats text production not as a passive process of translating preconceived ideas into written output, but as an active knowledge constituting process which itself involves the generation of ideas. The dual process model acknowledges the existence of an explicit rhetorical component within writing, but claims that this component only involves the retrieval of existing knowledge from long term memory and, therefore, does not lead to changes in understanding. The implicit text production process that is assumed to lead to discovery through writing, and which Galbraith has defined as dispositionally driven text production, operates on different principles than the reflective processes which are involved during knowledge-transforming. It is assumed that it involves a synthesis process rather than a retrieval process and this synthesis process leads to discovery when the ideas it produces are different from the explicit ideas retrieved from long term memory. Galbraith suggests that dispositionally guided text production is optimised under different conditions than the more reflective processes and that it will be at a maximum when text production is free from external planning. In so far, Galbraith argues that writing can function as a process of discovery when text production is a dialectical interaction between the text written so far and the writer's disposition towards to topic. We will turn to more specific assumptions of this model and the previously described problem solving accounts later.

For present purposes, we want to stress that both approaches present writing as a process of discovery, but that they attribute the epistemic effect of writing to different processes. The problem solving accounts (Flower & Hayes, 1980; Bereiter & Scardamalia, 1987) claim that development of understanding depends on the deliberate adaptation of ideas to rhetorical goals, whereas the dual process model of writing (Galbraith, 2009) claims that development of understanding depends on the extent to which text production is dispositionally driven. In addition to these different conceptions about writing as a process of discovery, these two approaches use different methodologies to investigate the development of understanding through writing. The problem solving accounts of writing suggest discovery by comparing the processes of novice and expert writers, but they do not provide empirical evidence that these distinct processes actually lead to differential development of understanding. Galbraith (2009), on the other hand, explicitly assesses how much writers feel they know about a topic. He assumes that different processes are

responsible for these different writing outcomes, but Galbraith does not investigate the underlying writing processes. The aim of the current paper is, therefore, to assess the relationships between writing processes and discovery and to investigate how writing processes are related to text quality. We will report the results of an experiment that assessed outline planned and non-outline planned text production from two different groups of writers, high and low self-monitors. Below both the predictions from problem solving accounts of writing and the dual process model of writing will be reviewed and assessed.

### *5.1.1 Problem solving account*

The original cognitive model of Flower and Hayes (1980) describes the different cognitive processes in writing and distinguishes between three basic processes: (i) planning; which involves generating content, organising and goal setting; (ii) translating, which involves translating ideas in text, and revising (iii); which includes reading and editing as components. An important feature of this account is the recursive nature of these processes. Planning, translating and revision can, in principle, occur at any moment during writing. The way in which these processes are combined is controlled by a monitor and different configurations of these processes are assumed to reflect different writing strategies. Hayes and Flower (1986) used this descriptive model of processes involved in writing to compare the writing processes of novice and expert writers thinking aloud while writing. They found that expert writers set more explicit rhetorical goals for their text and that experts created more connections between their goals than did the novices. Flower and Hayes, therefore, argue that expert writers construct a more elaborate representation of their communicative goals and that they use these goals to guide the retrieval of information.

As such, an important feature of writing research influenced by the problem solving models of writing is that is has primarily been concerned with the assessment of higher level thinking processes. This strand of research assumes that variations in the way ideas are expressed in writing are a consequence of variations in the thinking behind the text, i.e. the reflective processes that are held responsible for the generation of these ideas. The main body of results suggests that experts are involved in more overt-problem solving which leads to the production of better quality texts (Rijlaarsdam & van den Bergh, 1996; Sanders & van Wijk, 1996; Sanders & Schilperoord, 2006). These results are related to the assumption that these expert writers will change their ideas and consequently their understanding about the topic. This assumption, however, has rarely been empirically tested. More specifically, it is

often taken for granted that when writers show evidence of reflective thinking processes, this will then, as a natural and logical consequence of this involvement, lead to changes in ideas and changes in understanding. However, instead of assessing changes in understanding or changes in ideas, research from the problem solving perspective has tended to test the effects of problem solving processes on the texts that writers produce. The text quality of the resulting text is then taken as an indicator of the effectiveness of the orchestration of processes (Rijlaarsdam & van den Bergh, 1996). Furthermore, previous research has equated effective communication with experience and has suggested that experienced writers are good writers, whereas novice writers were categorised as weak writers (Torrance, 1996). This has led to the further assumption that experienced writers will produce better quality texts than novice writers and hence that high quality texts are evidence of higher order reflective processes.

Although this is not necessary always explicitly stated, the overall implication of these models is therefore that text quality and development of understanding are linked. In a recent study, Klein and Kirkpatrick (2010), hypothesized that text quality and development of understanding through writing are linked and examined the relationships between strategic rhetorical planning processes, text quality and discovery through writing. Their results showed that rhetorical planning processes improve text quality, which on its turn, leads to the development of understanding through writing. These results seem to be in line with the general conceptions about these problem solving models.

Even though the main focus of these problem solving accounts is on the higher order thinking processes behind the text, it should be noted that strategic rhetorical planning processes will also affect the way in which knowledge is presented in language (Flower & Hayes, 1984). According to Flower and Hayes (1984) expert writers evaluate and reflect not only their initial rhetorical problem for the text, but also the "*framing, focus and word choice*" (Flower & Hayes, 1984, p.154). This suggests that rhetorical planning goes 'all the way down' from the thinking behind the ideas to the wording of the ideas in the text. Furthermore, Flower and Hayes (1986) showed that expert writers also revise their text more extensively than novice writers do, because expert writers will evaluate whether the text produced so far is capturing the goals that they have set for the text, even on word choice level. We, therefore, assume that these problem solving models would predict more planned sentence production which might reflect careful word choice and more involvement in revision for writers who are involved in knowledge-transforming processes.

One important last point to make in this context is that the extent to which writers are able to engage in this sort reflective evaluation of content depends on how

they manage to control the interaction between higher order reflective processes and text production processes. Translating processes and higher order thinking processes are assumed to compete for limited cognitive resources, and hence it is assumed that writers will be less able to engage in rhetorical problem solving the more they try to carry out translating at the same time as generating content. It is this potential conflict which is assumed to be addressed by outlining prior to writing. The beneficial effect of outlining prior to writing is that writers are able to generate ideas separately from involvement in demanding translating processes. In a series of experiments, Kellogg (1988) has provided convincing evidence that outlining is associated with the production of better quality text, and that this is because it enables writers to clearly separate the reflective processes involved in generating, organizing and evaluating ideas from the processes involved in formulating these ideas in well-formed text.

To sum up, the problem solving accounts suggest that discovery through writing is a consequence of strategic modification of content in order to satisfy rhetorical goals, and this will be enhanced when the writer is able to focus on higher level thinking about text structure and content free from the demands of simultaneously formulating full text.

### *5.1.2 Dual process model of writing*

The dual process model of Galbraith (2009) attributes the effect of writing on knowledge change to both explicit planning processes as well as to more implicit text production processes. Galbraith (2009) argues that the problem solving models of writing overemphasize the role of explicit thinking at the expense of more implicit text production processes. In contrast with the problem solving models of writing, Galbraith (2009) treats translating as a more active process and, therefore, he treats discovery as intrinsic to the process of written expression itself. Galbraith set out to test directly how writing contributes to development of understanding. In a series of experiments examining the conditions under which writers develop their understanding, Galbraith and his colleagues have found different patterns of knowledge change through writing than would be expected on the basis of the problem solving models and, following from that, suggested that implicit text production processes can lead to the creation of novel content which leads to changes in the writer's understanding.

The first study that we want to consider, Galbraith (1992), set out to assess the main claims of the knowledge-transforming model; (i) whether discovery through writing is a consequence of rhetorical guided writing and (ii), whether discovery through writing is more likely to occur when writers are able to separate higher level

thinking processes from low level translating processes. In order to test these claims, Galbraith selected high and low self-monitors (see Snyder & Gangestad, 1986, for a detailed review of differences between high and low self-monitors). Self-monitoring is a personality scale that measures the extent to which people are concerned with the impression they make on others. According to Snyder, high self- monitors will use cues from the rhetorical context to guide their behaviour, whereas low self-monitors are less inclined to tailor their expression so that it fits the situation. Galbraith distinguished between high and low self-monitors, since he expected that high self-monitors would generate content to satisfy rhetorical goals whereas low self-monitors were assumed to prioritise the direct expression of their belief. By making a distinction between high and low self-monitors, Galbraith manipulated the extent to which writing was presumably dispositionally or rhetorically guided. Furthermore, Galbraith varied the extent to which writers had to produce full text or write notes in preparation for an article. According to Galbraith, this manipulation enabled him to test whether dispositionally guided or rhetorically guided text production processes are more likely to occur when the writer is separating idea generation from translating processes, which maps onto the predictions of the knowledge-transforming model, or when the writer spells out their ideas in full text, which maps onto the predictions of the dual process model.

Galbraith (1992) found that high self-monitors produced more new ideas after writing notes than low self-monitors which is in agreement with the knowledge-transforming model. However, Galbraith also found that low self-monitors produced a high number of new ideas after writing continuous prose, which was unexpected and contradicts a central claim of the knowledge-transforming model of Bereiter and Scardamalia (1987), namely that the thinking behind the text rather than the text production itself will lead to changes in content. More importantly, he found that only the new ideas from the low self-monitors were associated with changes in knowledge. From this experiment Galbraith concluded that development of understanding is more likely to happen when text production is dispositionally driven, i.e. when writing is not directed towards rhetorical goals and when the process of text production is not pre-determined by an explicit plan. Later experiments (Galbraith, 1999; Galbraith et al, 2006) explored these relationships in more detail and also investigate the effects of pre-planning on the development of ideas and understanding.

In 1999, Galbraith set out to test the effects of pre-planning on dispositionally guided text production by distinguishing between three different planning conditions; (i) non-outline or synthetic planning, which involved writing down a single sentence summing up one's overall opinion, (ii) outline planning, which

involved making an organized outline, and (iii) an unplanned condition, which involved writing out ideas as they come to mind without worrying how well they are expressed. All writers were asked to produce full text instead of writing notes as in the 1992 study. However, it was still stressed that writers should not worry about how well expressed their text was. The results of this study showed that, similar to the 1992 findings, low self-monitors produced more new ideas after writing than high self-monitors and this was most pronounced in the synthetic planning condition. Furthermore, this was the only condition in which new ideas were associated with the development of understanding. Nevertheless, the results also showed that writers in the other conditions also produced new content. Galbraith, therefore, concluded that new content can be produced by both explicit planning processes as well as dispositionally driven text production, but that new content needs to be guided by one's implicit disposition towards the topic rather than by a linear sequence of ideas stored in memory in order to lead to the development of the writer's understanding.

In 2006, Galbraith and colleagues repeated the above experiment with some minor changes to the methodology. The results of this experiment supported the results of the previous experiments and also showed that low self-monitors produced more new ideas within the synthetic planning condition than the high self-monitors did, supporting the assumption that discovery depends on dispositionally guided text production, and is reduced when external rhetorical constraints are imposed. Furthermore, the high self-monitors produced more new ideas in the outline planning condition than in the synthetic planning condition, which supports the assumption outlining facilitates explicit planning designed to satisfy rhetorical goals. So again, these results confirm that the generation of new ideas can be explained by two different processes; one which involves implicit guided text production processes in the case of low self-monitors synthetically planned texts and explicit planning process combined with text production in the case of high and low self-monitors outline planned texts. It should be noted though that Galbraith, although he explicitly assesses the way in which writers develop their ideas through writing, did not collect process data and therefore can only make indirect claims to the processes responsible for the generation of these new ideas.

Nevertheless, when this evidence is taken together, the results suggest that for low self-monitors outlining can reduce, rather than enhance, the development of ideas and that effects of writing on the development of understanding vary as a function of individual differences in self-monitoring. On the basis of these experiments, Galbraith has suggested an alternative dual process account of discovery through writing which identifies two conflicting processes in writing. The first of these is an explicit planning process. This involves the retrieval of content from an explicit store of ideas and the

goal-directed manipulation of these ideas in working memory designed to create a coherent knowledge object that satisfies rhetorical goals. The nature of this process is essentially equivalent to the knowledge-transforming process of Bereiter and Scardamalia (1987). However, according to Galbraith, this process only involves the reorganization of existing knowledge and therefore, the crucial difference is that, by itself, this process is not associated with development of understanding. The second is an implicit text production process. This implicit text production process operates on an implicit store of conceptual knowledge in semantic memory, which Galbraith defines as the writer's disposition towards the topic, and involves synthesizing content during text production. The key feature of this process is that the content that it produces is unpredictable, since the process is guided by the implicit organization of material in semantic memory. Ideas are synthesized in the course of formulation rather than being directly retrieved from memory and translated into text. According to Galbraith, this process leads to changes in knowledge when the content it produces is different from the explicit content stored in memory.

Thus, Galbraith claims that both processes are required for effective writing: the explicit planning process is required to set goals and to ensure that the global structure of the text is coherent and satisfies the writers' rhetorical goals; the implicit text production process is required for the writer to develop their implicit understanding in explicit text. However, because they operate best under different conditions this leads to a fundamental conflict in writing. The explicit planning process operates best when writing is goal directed and when content is manipulated in working memory; the text production process operates best when thought is formulated in full text, and when the sequence of sentences is allowed to unfold in correspondence with the implicit organization of semantic memory. The crucial consequence of this is that the two processes would be expected to conflict with one another. Conditions which favour explicit planning should enable the production of more controlled and more explicitly organized text but at the expense of learning; conditions which favour implicit text production should enable changes in knowledge but at the expense of producing coherent text, with the result that text has to be modified in order to produce coherent text.

### *5.1.3 Aim of the experiment*

The aim for the present study is to assess empirically which text production processes are involved during the development of subjective understanding through writing and what their effect is on text quality. We have manipulated self-monitoring so that we can test whether discovery is a consequence of dispositionally or rhetorically

guided text production. Moreover, we have manipulated planning to assess whether these different processes are more likely to occur when writing is free of external planning or not. In previous experiments, Galbraith (1992) has distinguished between writing full prose and notes, because those writing conditions mapped onto the predictions of the different theoretical conceptions most directly. In later studies, Galbraith (2006) varied both planning as well as the constraints placed on the writer during writing. Therefore, it is still not clear whether the typical effects that Galbraith found are a consequence of the planning manipulation or of the constraints placed on the writing task. In contrast with Galbraith, it is our aim to investigate the underlying text production processes and to assess writing quality, and therefore, an important criterion of the setup of this experiment is to keep the writing tasks comparable over the writing conditions. For this reason, we will manipulate planning as a variable. We will distinguish between an outline planning and a non- outline planning or synthetic planning condition, because we assume that these different planning tasks will provoke the use of different text production processes during writing, which, in addition, can also be related to the different claims of the cognitive models that we want to test. For outline planning, it is assumed that writers will be able to rely on ideas that they have organised during pre-planning and hence, that their outline plan will guide the generating of content during writing, whereas for synthetic planning it is assumed that ideas, since pre-planning did not enable them to organise the order of ideas, will emerge during text production. In order to investigate text production processes during writing we will collect keystroke logging data. The keystroke logging data will provide us with non-intrusive logs of the writing processes involved.

Text production processes are generally investigated in terms of pausing behaviour, burst behaviour and revision behaviour (see Sullivan & Lindgren for a review). In a previous study (Baaijen, Galbraith & de Glopper, 2012), we have shown that analysing these features on the basis of undifferentiated keystroke logs potentially reflect a very heterogeneous range of underlying writing processes. When using keystrokes as indicators of underlying cognitive processes one has to carefully consider the alignment between the writing processes and the keystroke measures. In our previous study we suggested ways in which keystroke logging data can be prepared in order to relate keystroke logging data more precisely to the underlying text production processes. On the basis of a principal component analysis it was shown that a variety of measures derived from keystroke logs can capture local organisation of sentence planning and more global organisation of text structure. In this current paper, we will analyse the text production processes on the basis of the measures described in Baaijen, et al. (2012). We hypothesise that high self- monitors

show more evidence of reflective processes both in terms of sentence production, i.e. more planned sentence production, and in terms of the overall text production, i.e. more revision. Furthermore, we hypothesise low self-monitors, who are expected to prioritise the direct expression of their disposition, to show more evidence of less planned sentence production and more revision.

To sum up, we anticipate that high and low self-monitors prioritise the explicit planning process and the implicit text production process differently and that these types of writers use different strategies to reconcile the conflict between the explicit planning processes and the implicit text production processes. It is hypothesised that high self-monitors will employ a top down process of text production which will be facilitated by the outline planning condition, whereas we hypothesize that the low self-monitors will employ a more bottom up approach to writing which will flourish within the synthetic planning condition. We assume that the writers employing the most pure top down process (.i.e. high self-monitors in outline planning) first develop an explicit mental model for the text designed to satisfy rhetorical goals and then use this model to guide the realisation of their ideas during text production. Furthermore, we assume that writers who employ the most pure bottom-up process (low self-monitors in synthetic planning) will articulate their emerging understanding of the topic during spontaneous text production and that the resulting content will then be organized into a coherent text.

We will examine the effect of planning and self-monitoring on the way in which writers develop ideas through writing and we will assess how the development of ideas is related to text quality as well as to development of understanding. We hypothesize that strategic rhetorical planning is beneficial for the production of good quality text, but that it is limiting the extent to which writers can use writing as a process of discovery. This implies that, if the dual process model is right, outline planning will increase text quality and the retention of ideas, whereas synthetic planning will lead to changes in understanding and the production of more new ideas, especially for low self-monitors in the synthetic planning condition.

## 5.2 METHOD

### *5.2.1 Participants*

84 students from the Faculty of Arts of the University of Groningen were recruited to participate in the experiment. They were all native Dutch speakers and received € 10 for their participation. Their average age was 22.2 years (*SD*=3.8). Participants were pre-selected using Snyder's revised 18 item self-monitoring scale (Snyder &

Gangestad, 1986), and were classified as high self-monitors (HSM, *N*=42) if they scored between 11 and 18 on the scale and as low self-monitors (LSM, *N*=42) if they obtained a score between 0 and 8 on the scale.

Four participants were dropped from the initial sample. Three were excluded because they did not complete one or other of the idea change measurement tasks correctly; one because he/she had very low knowledge about the topic (3 SD's below the mean knowledge rating). Taken together, we therefore have a sample of 80 participants.

### *5.2.2 Design and procedure*

The two groups of low and high self-monitors were randomly assigned to either an outline planning or synthetic planning condition. All participants were tested individually.

In all four conditions, participants were asked to plan and write an article for the university newspaper discussing whether "*our growing dependence on computers and the Internet is a good development or not*". The session was divided into three different components.

- **Before writing**

Participants were given 10 minutes to list all the ideas they could think of relevant to the topic. The only restriction was that each idea should be no longer than a sentence in length. They were then asked to rate the importance of these ideas on a 5-point scale, where 5 = more important idea, 1= less important idea. Finally, the participants were asked to rate how much they felt they knew about the topic on a 7-point scale, where is 1 = very little and 7 = a great deal.

- **Writing task**

In the synthetic planning condition, participants were given 5 minutes to think about the topic and to write down a single sentence summing up their overall opinion of the topic. Participants in the outline planning conditions were given 5 minutes to construct a structured outline. Following this initial planning phase, participants had 30 minutes to write a well-structured article for the university newspaper. It was stressed that they had to produce a reasoned argument reflecting their own opinion about the matter. During this task, all participants were allowed to consult their written outlines. It was emphasized that they had to budget their time so that they could produce a completed article in the time available. During writing, keystrokes were logged with the use of Inputlog (Leijten & van Waes, 2006). Inputlog is a logging tool that logs all keystrokes and movements of the mouse and runs in a Word

environment. As a consequence, participants were able to produce their articles in a familiar environment and had access to all Word processing functions. The program starts recording as soon as the start button is clicked within the Word document. The researcher would always start the program after the instructions for the writing task were given and questions about the assignment were answered.

- **After writing**

Immediately after writing, participants were asked again to rate how much they felt they knew about the topic. Then, they were again given 10 minutes to list of the ideas they could think of relevant to the topic and to rate the importance of these ideas. Finally, participants were presented with the two lists that they had produced and were asked to rate the extent to which ideas on the second list corresponded with the ideas on the first list on a 6-point scale ranging from 1= identical point to 6 = no correspondence. On a different sheet they were presented with three columns. In the first column they found the numbered list of ideas produced on the second list. In the second column they were asked to list corresponding ideas from list 1. This could involve a single idea or a group of ideas. Then, in the last column they were asked to rate the correspondence between these ideas or this group of ideas. Ideas on the second list were classified as new ideas if participants did not identify a corresponding idea on the first list and hence rated the correspondence between the ideas with a score of 4 to 6 whereas ideas from the first list that persisted onto the second list were classified as preserved ideas when they received a score of 1 to 3 on the correspondence rating.

### *5.2.3 Idea change measures*

#### *5.2.3.1 Development of understanding*

On the basis of the above described procedure we measured discovery through writing with a subjective rating of knowledge of the topic. We have established increases in understanding when participants evaluated their knowledge after writing with a higher ranking than their rating of knowledge before writing.

#### *5.2.3.2 Development of ideas*

In order to assess the development of ideas through writing different calculations were carried out on the lists produced before and after writing. The following scores were calculated: number of ideas produced before and after writing (list 1 and 2); mean length of ideas (mean number of words per idea) produced before and after writing; mean importance of ideas produced on list 1 and 2.

In addition, we calculated the number of preserved and new ideas on list 2; the mean length of preserved and new ideas on list 2; and the mean importance of preserved and new ideas produced on list 2. In order to control for variations in the number of ideas generated in list 1, the length of ideas on list 1 and the mean importance of ideas on list 1, the number of ideas on list 2, the length of ideas on list 2 and the importance of ideas produced on list 2 will be analysed with the equivalent score from list 1 as a covariate. For the analyses in this paper we used the ratings of importance also to divide up new and preserved ideas into major ideas (those which received a rating of 4 or 5 on the importance scale) and minor ideas (those which received a rating of 1, 2 or 3) on the importance scale.

### *5.2.4 Writing process measures*

In order to assess the underlying writing derived a variety of measures from our keystroke logging data. For a discussion of the development of these process measures, see Baaijen et al. (2012). Below, only a short overview is given of the data preparation and the variables that are included within the current process analysis. To reduce the number of variables we have performed a principal component analysis which will be discussed in the result section. In order to link the process measures described below to the process measures included in the principal component analysis we will number the variables as a reference (1-14). The basic aim for the principal component analysis was to identify components that capture different conceptual levels of the underlying writing processes. We set out to operationalise the two processes by which ideas can be formulated: (i) dispositionally guided text production which is at its highest when text production is spontaneous and unplanned and which can represent a bottom-up writing approach and; (ii) a more explicit planning process which is at its highest when it is directed towards rhetorical goals and which represents a more top-down control of text production processes.

#### *5.2.4.1 Pause behaviour*

In order to assess pause behaviour we have, first of all, distinguished between pauses and events at different pause locations (see Baaijen et al., (2012) for more details). Pauses are defined as 'empty' thinking pauses before the continuation of text production whereas events are defined as episodes that include pauses as well as other operations or material before the continuation of text production. Pause behaviour is assessed on the basis of the linear transitions only. In addition, this way of annotating the data allows us to investigate the percentage of linear and non-linear transitions at different pause locations.

For the principal component analysis we have included a score capturing the total time spent pausing, which we have defined as total time spent on reflection (4). We have also included an overall score of time spent on events, which we have defined as total time spent on revision (10). In addition, we have included scores that capture the percentage of linear transitions at word (14) and sentence boundaries (11). For the assessment of pause behaviour we have included a score capturing the percentage of cognitive pauses (>.2 seconds) between words (5) and sentences (13) and, finally, we have included an indication of mean pause length between sentences (7). These pause times are lognormal pause times based on mixture models (see Baaijen et al., (2012) for a detailed description).

#### *5.2.4.2 Burst behaviour*

In order to assess text production we have classified text production in terms of bursts (Chenoweth & Hayes, 2003). We distinguished between P-Bursts (Pause), R-bursts (Revision) and I-bursts (Insertion). P-bursts can be defined as bursts initiated after and terminated by a pause of at least 2 seconds. R-bursts can be defined as bursts that are terminated in a revision either at the leading edge or by the writing moving away from the leading edge for an insertion elsewhere in the text. R-bursts can either be initiated after a revision or a pause of at least two seconds. I-bursts can be characterised as text production that is executed away from the leading edge. (See Baaijen et al., (2012), for a detailed overview of burst definitions).

For the principal component analysis we have included the percentage of bursts as P-bursts (3) and R-bursts (1). We have also included a measure that captures the number of R-bursts that are terminated at the leading edge specifically (2) and we have included the percentage of bursts that was executed as I-bursts (8) as an indicator of how often participant leave the leading to insert text elsewhere. In Baaijen et al. we also looked at length of different types of bursts, but since the length of burst measures were highly correlated with the different kind of bursts we did not included them in the analysis of this paper.

#### *5.2.4.3 Revision behaviour*

In order to assess revision behaviour we have computed a number of variables that measure the extent to which participants produce the text in one piece or produce the text in a more recursive nature. Furthermore, we have tried to capture what sorts of revisions are carried out. Already in the time spent on events (10) we are capturing some aspects of revision. The same counts for the number of bursts that are revised at the leading edge (2).

For the principal component analysis we have also included some variables that solely provide information about aspects of linearity and revision. We have included a score that captures the number of cycles in which the text is produced (12), where cycles refer to sequences of activities that are executed in a linear order at the leading edge. Each break away from the leading edge is defined as the start of a new cycle. Furthermore, we have included a score that measures the sentence linearity (9). This involved comparing the order of sentences in the final product with the order that they were produced in. Finally, we have included a score that solely captures the percentage of process words that are produced during writing but that do not end up in the final product and are deleted in the course of writing, which we have defined as percentage of revision deletions (6).

### *5.2.5 Text quality*

The quality of the texts was rated by two independent judges on a 9-point scale. The written articles were rated on the coherence of the overall argument, the originality and the appropriateness of the tone and the relation to the audience for an article in the university newspaper. Inter-rater reliability for this measure proved to be satisfactory ($r$=.84, $p$<.001). The means of the two judges' scores were used for analysis.

### *5.2.6 Analysis*

For the multiple regression analyses presented in this paper we have used the strategy recommended by Crawley (2005) for all the analyses. We used a combination of forward and backward regression. In order to observe the effects of interactions, in the forward steps, we entered the main effects of variables at step 1, the two-way interactions at step 2 and the three way interactions at step 3. Since Crawley (2005, p. 204) recommends that one should not "try to estimate more than n/3 parameters during a multiple regression", three-way interactions are the highest level that were modelled with the number of parameters in our analyses and our sample size. For some analyses, the analysis will be restricted to two-way interactions for the same reason. We then progressively removed non-significant terms, starting with the highest level interactions. In the results section we will present the final, minimal models resulting from this process. These consist of the significant terms in the model plus terms contributing to higher level interactions (Cohen, Cohen, West & Aiken, 2003).

For analyses of the development of understanding, we will regress knowledge after writing on prior understanding at step 1, then enter the main effects at step 2,

followed by higher level interactions at subsequent steps. By controlling for prior understanding, this means that we will be, in effect, assessing the relationships between the variables and increases in understanding. For these models we will present the final model in two steps, with prior understanding at step 1, and the significant predictors at step 2. $R^2$ change between step 1 and 2 will be used to assess whether the predictor variables are significantly related to increased understanding.

To describe the results of these regression models we will plot the predictor variables against the adjusted predicted values of the relevant dependent variable and we will plot the regression lines.

## 5.3 RESULTS

### *5.3.1 Development of subjective understanding*

A two-way (2*2) between subjects ANCOVA with self-monitoring and type of planning as factors and with prior understanding as a covariate revealed a significant main effect of type of planning on subjective understanding after writing ($F$ (1,75) = 4.67, $p$ = .03, $\eta^2$ =.03), with the mean rating of understanding after writing in the synthetic condition ($M$ = 5.00, $SD$ = 0.82) being higher than the mean rating of understanding after writing in the outline condition ($M$ = 4.62, $SD$ = 0.84). Planned comparisons comparing mean ratings of understanding before and after writing in the synthetic and outline planned conditions showed that there was a significant increase in understanding in the synthetic planned condition ($M$ before = 4.66, $SD$ = 0.83; $M$ after = 5.00, $SD$ = .82; $t$ (39) = 3.34, 2 tailed test, $p$ = .002) but no significant difference in the outline planned condition ($M$ before = 4.55, $SD$ = 0.81; $M$ after = 4.62, $SD$ = 0.84; $t$ (39) = 0.68, 2 tailed test, $p$ = .50). Overall, then, while the synthetic planning condition was associated with a significant increase in understanding after writing, the outline condition was not.

Table 5-1 shows the number of participants in each condition who experienced decreases in understanding, the number whose understanding remained unchanged, and the number whose understanding increased.

As can be seen in the table, only 7.5% of the participants in the whole sample experienced decreases in understanding, and within individual conditions, the number is low, varying from 0 to 3. Since these numbers are too low to make reliable estimates of the characteristics of participants whose understanding decreased, in the following analyses, when the relationships between the different measures and changes in understanding are being analysed, the analysis will be restricted to assessing relationships for subjective same and increased understanding only.

**Table 5-1** Number of participants who experienced different types of changes in understanding as a function of self-monitoring and type of planning.

| Change in understanding | **Outline** | | **Synthetic** | | |
|---|---|---|---|---|---|
| | High SM | Low SM | High SM | Low SM | Total |
| Decreased | 3 | 1 | 2 | 0 | 6 |
| Unchanged | 12 | 15 | 11 | 13 | 51 |
| Increased | 5 | 4 | 6 | 8 | 23 |

### *5.3.2 Effects of self-monitoring and type of planning on writing processes*

The next step in the analysis was to test the effects of self-monitoring and type of planning on writing processes, as reflected in measures derived from the keystroke logs.

Preliminary screening of the data using the text modification index (a global measure of the extent to which content is modified during text production, see Baaijen, Galbraith & de Glopper, 2010) revealed 1 extreme outlier (defined as a score more than 3 SD's above the mean). Inspection of the keystroke logs showed that this participant – a high self-monitor in the outline condition - had deleted substantial sections of text which he/she then re-wrote in very similar words. The deletion of the text inevitably means that the text modification score will be very high but that this will not directly reflect the extent to which text has been modified during production. This participant was therefore excluded from the process analyses. In addition, the process data from another participant were lost due to technical problems with the keystroke logging software. For the process data we therefore have a total sample of 78 participants.

In order to assess writing processes we conducted a principal component analysis (PCA) on 14 variables selected to cover a range of different process activities, including measures related specifically to sentence production and measures related to more global planning operations. The set of measures had a Kaiser-Meyer-Oklin sampling adequacy of .79 and Bartlett's test of sphericity was highly significant ($p$<.001), indicating that principal component analysis is appropriate for these data. Varimax rotation was used to identify orthogonal components of the data. An initial analysis was run to obtain eigenvalues for each component in the data. Three components had an eigenvalue over Kaiser's criterion of 1 and taken together these

components explained 75% of the variance. However, the scree plot suggested that only retaining two components was also an option. In a second round of the analysis we therefore extracted two components which, taken together, explained 65 % of the variance. We decided to retain only these two components for the final analysis. Table 5-2 shows the loadings of each variable on each component after rotation, as well as the amount of variance that each component accounts for, and the Cronbach alpha's for the two components.

**Table 5-2** Summary of principal component analysis with varimax rotation for 2 factor solution.

| Variables | **Rotated Factor Loadings** | |
|---|---|---|
| | **Component 1** Planned sentence production | **Component 2** Global linearity |
| 1 Percentage of R-bursts | **-.945** | -.025 |
| 2 Percentage of leading edge R-bursts | **-.931** | .133 |
| 3 Percentage of P-bursts | **.834** | .393 |
| 4 Total time spent on reflection | **.710** | .468 |
| 5 Number of pauses >2 seconds between words controlled for text length | **.701** | -.038 |
| 6 Percentage of revision deletions of total process words | **-.620** | -.382 |
| 7 Mean pause time between sentences | **.599** | .258 |
| 8 Percentage of Insertion-Bursts | .103 | **-.851** |
| 9 Sentence Linearity Index | .174 | **.846** |
| 10 Total time spent on revision between sentences, paragraphs and word transitions | -.022 | **-.797** |
| 11 Percentage of linear sentence transitions | .092 | **.775** |
| 12 Number of cycles | -.236 | **-.717** |
| 13 Number of pauses >2seconds between sentences controlled for text length | .455 | **.611** |
| 14 Percentage of linear word transitions | .373 | **.561** |
| Eigenvalues | 6.3 | 2.8 |
| % of variance | 44.98 | 20.24 |
| $\alpha$ | .89 | .87 |

*Note that factor loadings over .5 appear in bold*

Inspection of the loadings on the first component suggests that it reflects longer pauses between and within sentences combined with cleaner bursts (more P-bursts and fewer R-bursts) and less leading edge revision. Since all these variables relate to the sentence level, we have labelled this component *planned sentence production.* High scores on this component reflect longer pauses between sentences combined with less within-sentence revision; low scores reflect brief pauses between sentences combined with more within-sentence revision.

The second component reflects a combination of linear transitions between sentences with fewer insertions earlier in the text and fewer revisions between sentences and paragraphs. Since these variables relate to how linearly writers move between sentences and higher level units of text we have labelled this component *global linearity*. High scores on this component represent linearly produced texts; low scores represent less linear text production and more revision of the global structure of the text. Taken together, these components represent two independent measures of sentence level, text production processes and higher level planning processes.

Table 5-3 shows the mean scores for each of these variables as a function of self-monitoring and type of planning. As can be seen in the table, sentence production is generally more planned for the high self-monitors and in the outline conditions, with the low self-monitors' synthetically planned texts showing the least planned sentence production. For global linearity, the pattern is slightly different, with the low self-monitors' outline planned texts and the high self-monitors' synthetically planned texts showing higher levels of global linearity than the other conditions, but with the low self-monitors' synthetically planned texts again showing the lowest levels of global linearity. Note, however, that the scores are very variable within each condition for both measures.

**Table 5-3** Mean scores for planned sentence production and global linearity as a function of self-monitoring and type of planning.

| | **Outline planning condition** | | **Synthetic planning condition** | |
|---|---|---|---|---|
| | High SM | Low SM | High SM | Low SM |
| Planned sentence production | 0.07 (0.96) | 0.02 (1.1) | 0.08 (1.00) | -0.14 (1.02) |
| Global linearity | -0.06 (1.09) | 0.16 (1.02) | 0.16 (0.63) | -0.24 (1.15) |

In order to assess the effects of self-monitoring and type of planning on planned sentence production and global linearity a two-way (2*2) between subjects MANOVA was carried out. This failed to show any significant effects. Despite the apparent differences in the means, the effect sizes for the interactions between self-monitoring and type of planning were very small (for planned sentence production, $\eta^2$ = .002; for global linearity, $\eta^2$ =.02) and none of the effects for either the multivariate analyses ($p$ > .38 for all effects) or univariate analyses ($p$> .18 for all effects) approached significance.

### *5.3.3 Effects of self-monitoring and type of planning on idea change*

The analysis of idea change will focus on two different aspects of the data. The first aspect is the relative length of the ideas produced before and after writing. We will argue that, although this is derived from the ideas listed before and after writing, it can be used as an indicator of processing, and therefore should be treated separately from the other idea change measures.

The second aspect is the number of new and preserved ideas and their ratings of importance. Our initial analysis of these measures used the overall number of new and preserved ideas as in the previous research from Galbraith and the mean ratings of importance of these ideas. However, in exploring these, we developed an alternative grouping of the data, in which the ratings of importance were used to divide up new and preserved ideas into *major ideas* (those which received a rating of 4 or 5 on the importance scale) and *minor ideas* (those which received a rating of 1, 2 or 3) on the importance scale. This way of grouping the data proved to be a more sensitive way of analysing the data – it produced the same effects of conditions as for the mean importance ratings of old and new ideas, but also revealed a number of other effects. We, therefore, will only describe the analysis of major and minor ideas here.

For each of these aspects of the data, we will present the effects of self-monitoring and type of planning first, and then consider how the measures are related to increases in understanding. Preliminary screening of the data identified 2 outliers (3 SD's above the mean). A high self-monitor in the synthetic planning condition produced an extremely high average number of words per point, indicating that he/she had not complied with the instructions to restrict points to single ideas. Since this affects not just the measure itself, but also distorts counts of new and old ideas, this participant was excluded from the idea change analysis. A low self-monitor in the outline planned condition produced a final list consisting almost entirely of new ideas; since we suspected that this was because he/she had misunderstood the

instruction for listing ideas after writing, we also excluded him/her from the idea change analysis

### *5.3.3.1 Length of ideas produced before and after writing*

The first analysis assessed the mean length of the preserved and new ideas in list 2 compared to the mean length of ideas in list 1. A three-way (2*2*3) mixed ANOVA with 2 between factors (self-monitoring and type of planning) and 1 within factor (type of idea) revealed a significant main effect of type of idea ($F(2,144) = 8.00$, $p = .001$, partial $\eta^2 = .10$) and a significant interaction between this effect and type of planning ($F(2, 144) = 3.45$, $p=.03$, partial $\eta^2 = .05$). The mean length of the different types of ideas in the outline and synthetic planning conditions is shown in figure 5-1.

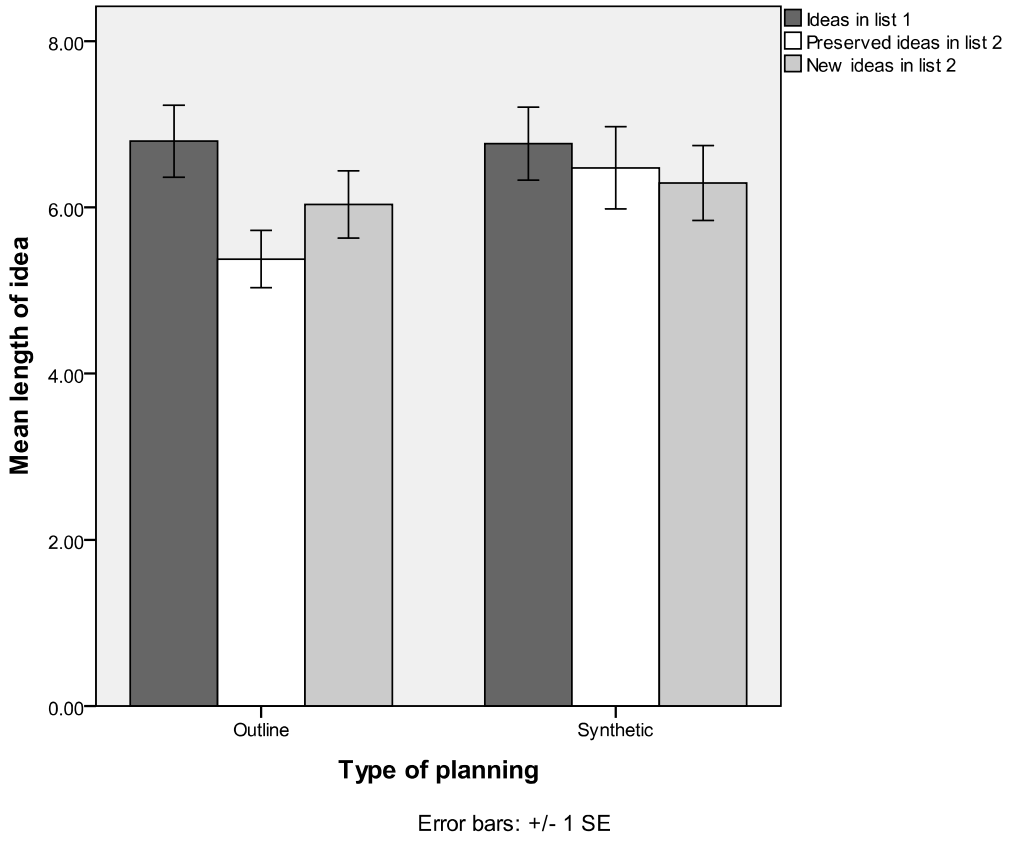


**Figure 5-1** Mean length of ideas in list 1, and preserved and new ideas in list 2, as a function of type of planning

In order to determine the source of the interaction, one-way repeated measures ANOVAs were carried out separately within the outline and synthetically planned conditions. These revealed a significant main effect of type of idea within the outline

condition ($F$ (2, 74) = 9.80, $p$ < .001) but not in the synthetic condition ($F$ (2, 74) = 1.25, $p$ =.29). Post hoc comparisons within the outline condition showed that the preserved ideas in list 2 were significantly shorter than both the ideas in list 1 ($t$(38) = 3.73, $p$= .001, 2 tailed test) and the new ideas in list 2 ($t$(37 = 2.88, $p$ = .007); the new ideas were also significantly shorter than the ideas in the initial list ($t$(37)= 2.39, $p$= .02) but the reduction was smaller than it was for the preserved ideas.

These results suggest that the reduction in length of the preserved ideas compared to list 1 may be a marker of outlining before writing. One possible explanation for this is that when an outline is constructed, writers maintain it in memory and use it to guide text production. When these ideas are reproduced in list 2 they are retrieved directly from memory rather than being formulated afresh and are expressed in an abbreviated form. New ideas may be affected to a lesser extent, depending on the extent to which they are incorporated into a modified outline during writing. If this is the case then this measure may be an indicator of the extent to which writers store their initial ideas during writing, and variations in the extent to which writers do this – in either the outline or synthetic condition – may indicate the extent of *memorisation* during writing. To take account of this possibility, we will therefore include this measure in later analyses of the relationship between writing processes and other measures.

#### *5.3.3.2 Major and minor new and preserved ideas*

This analysis focussed on the number of major and minor ideas from the initial list that were preserved in the list produced after writing, and the number of major and minor new ideas that appeared in the list produced after writing. In order to remove positive skew for all four of these measures, we expressed them as a ratio of the number of ideas in the initial list, and then took the square root. The resulting measures were normally distributed.

We then carried out a two-way (2*2) between subjects MANOVA to assess the effects of self-monitoring and type of planning on the four measures. This revealed a highly significant interaction between self-monitoring and type of planning (Wilk's lambda=.82; $F$(4, 70) = 3.86, $p$ =.007). Follow-up univariate analyses showed significant interactions between self-monitoring and type of planning for the number of major new ideas ($F$ (1, 73)=8.43, $p$=.005, $\eta^2$=.10) and for number of minor preserved ideas ($F$(1, 73) =8.96, $p$=.004, $\eta^2$=.11). The means for each of these measures (with standard errors) are plotted in figure 5-2(i) and figure 5-2(ii).

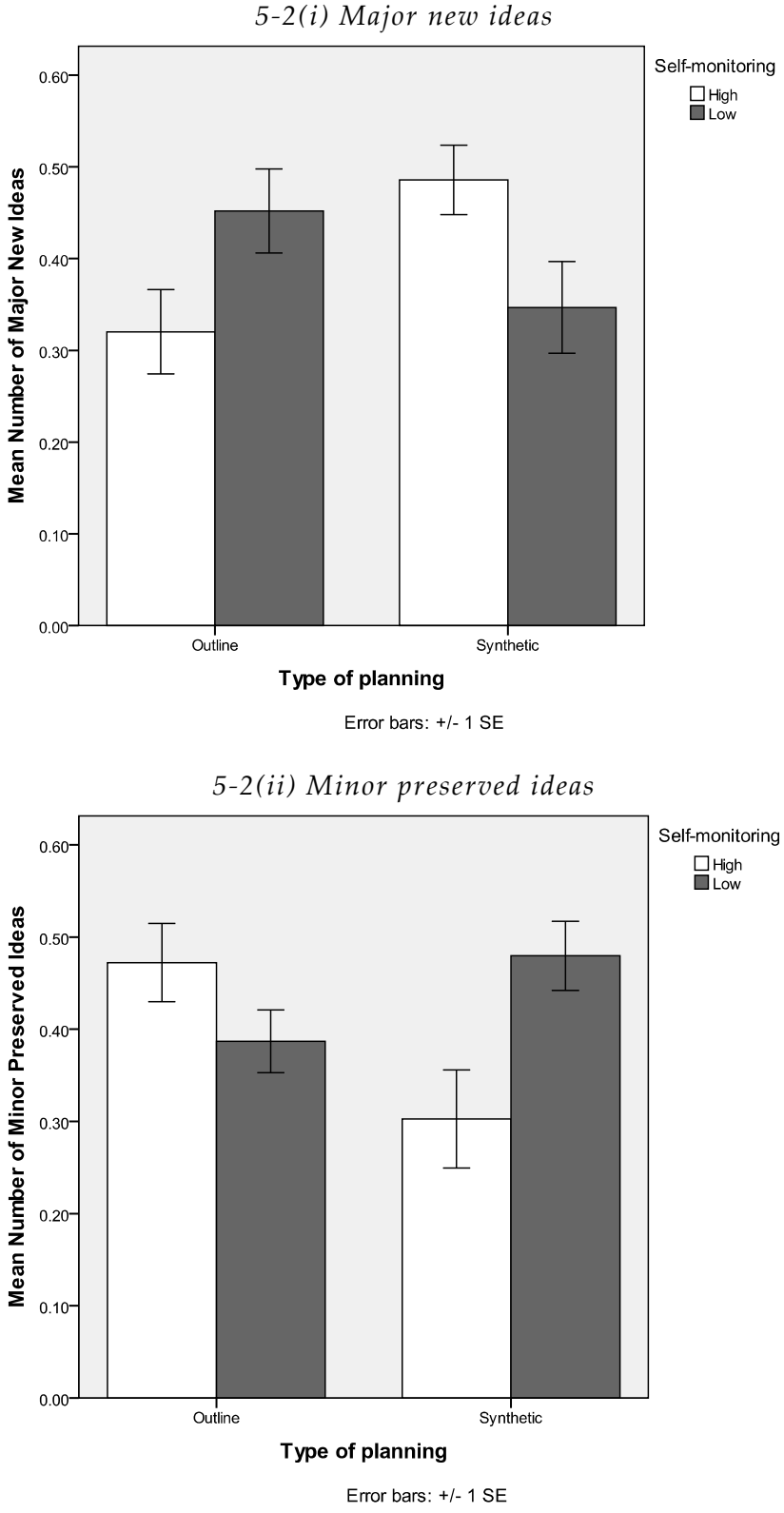


**Figure 5-2** Number of major new ideas and minor preserved ideas (both expresses as the square root of the ratio of ideas in the initial list) as a function of self-monitoring and type of planning.

Individual comparisons showed, as can be seen in figure 5-2(i), that high self-monitors produced significantly more new ideas of major importance in the synthetic planning condition than in the outline planning condition ($t(35) = 2.65$, $p = .012$, 2 tailed test), while the low self-monitors produced more new ideas of major importance in the outline planning condition than in the synthetic planning condition (though this difference was not significant) ($t$ (38) = 1.54, $p$ =.13). These results suggest that outline planning and synthetic planning have opposite effects on the number of new ideas of major importance generated by the low and high self-monitors after writing.

Post hoc analysis of the interaction for preserved ideas of minor importance (see figure 5-2(ii)) revealed a similar effect, with high self-monitors retaining fewer preserved ideas of minor importance in the synthetic planning condition than in the outline planning condition ($t$ (35) = 2.17, $p$=.04, 2 tailed test), while low self-monitors retained fewer such ideas in the outline planning condition than in the synthetic planning condition ($t$ (38) = 2.14, $p$ = .04, 2 tailed test).

### *5.3.4 Relationships with increased understanding*

In this section we describe the results of sequential regressions assessing the relationships between writing processes and increases in understanding, and between idea change and increases in understanding.

#### *5.3.4.1 Relationship between writing processes and increased understanding*

This analysis included memorisation as a process variable along with planned sentence production and global linearity. Memorisation was added to the analysis as a process variable in the light of our interpretation that the reduction in length of preserved ideas represents the extent to which writers maintain an explicit representation of their ideas in memory during writing. This measure was calculated as the mean length of preserved ideas in list 2 divided by the mean length of the ideas contained in list 1. Lower scores are therefore assumed to represent higher degrees of memorisation. (Note, in addition, that square roots were taken to reduce positive skew, and that the scores were standardised before being entered into the regressions). As a preliminary step we examined the correlations between memorisation and extent of planned sentence production and global linearity. This showed that memorisation was uncorrelated with extent of planned sentence production ($r$=-.08) and global linearity of text production ($r$= -.02). Regression analyses examining whether these relationships interacted with self-monitoring and type of planning also failed to show significant effects.

Table 5-4 shows the final minimal model for this analysis. After controlling for prior understanding at step 1, there is a significant increase in variance accounted for at step 2 ($R^2$change =.12, $F$ (9, 60) =2.93, $p$=.006). As can be seen in the table, two terms made a significant independent contribution to the relationship with increased understanding: (i) a three-way interaction between memorisation, planned sentence production and global linearity and (ii) a two way interaction between self-monitoring and planned sentence production. Other significant terms correspond to lower level terms contributing to these interactions.

**Table 5-4** Increased understanding predicted by self-monitoring, memorisation, planned sentence production and global linearity.

| | **B** | **SE** | **β** |
|---|---|---|---|
| **Step 1** | | | |
| Constant | 1.19 | 0.36 | |
| Prior understanding | 0.81 | 0.08 | .78*** |
| **Step 2** | | | |
| Constant | 1.38 | 0.33 | |
| Prior understanding | 0.77 | 0.07 | .75*** |
| Self-monitoring (SM) | -0.03 | 0.12 | -.02 |
| Memorisation | 0.07 | 0.06 | -.09 |
| Planned sentence production | 0.06 | 0.07 | .07 |
| Global linearity (GL) | -0.14 | 0.06 | -.17* |
| SM* Planned sentence production | -0.29 | 0.12 | -.23* |
| GL* Planned sentence production | 0.09 | 0.07 | .09 |
| Memorisation * Planned sentence production | -0.18 | 0.06 | -.21** |
| Memorisation *GL | 0.01 | .07 | .01 |
| Memorisation * Planned sentence production * GL | 0.27 | 0.12 | .18* |

*$R^2$=.61 for Step 1, $R^2$ change=.12 for step 2 (F (9, 60) =2.93, p=.006), * p < .05, ** p < .01, ***p < .001*

The relationship between memorisation, planned sentence production and global linearity is shown in figure 5-3. In order to plot these data, the memorisation variable was divided into two groups of "high memorisation" (more reduced words per point) and "low memorisation" (less reduced words per point) using a median split, and the global linearity variable was divided into two groups of "less linear" and "more linear" global linearity, again on the basis of a median split. The two panels correspond to the different levels of memorisation, and the solid line (dots)

corresponds to more linear global planning, while the dashed line (crosses) corresponds to less linear global planning.

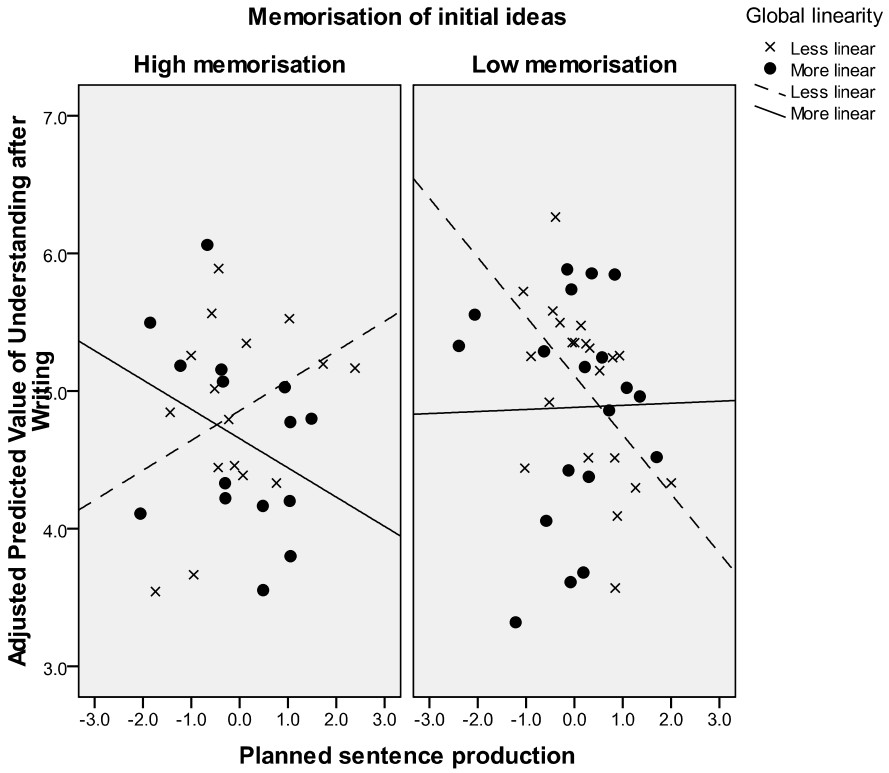


**Figure 5-3** Relationship between planned sentence production and increased understanding as a function of memorisation and global linearity.

As can be seen in the figure, there is a significant negative relationship between planned sentence production and increased understanding for writers who memorised their original ideas less, and who produced text non-linearly (right hand panel, crosses, with dashed lines) ($r$=-.53, $N$=21, 2 tailed test, $p$ = .013). This suggests that, when writers impose less top-down control on text production (low memorisation, right hand panel), increases in understanding are associated with less planned sentence production and less linear global planning. By contrast, when these writers produce more linear text, the relationship is non-significant ($r$ = .02, $N$=19, 2 tailed test, $p$ = .94). Similarly, when more top-down control is imposed (high memorisation, left hand panel), although the relationships vary in direction depending on the linearity of global planning, neither is significant (for less linear global planning, $r$ = .36, $N$=16, 2 tailed test, $p$=.17; for more linear global planning, $r$= -.34, $N$=- 15, 2 tailed test, $p$ =.21).

Figure 5-4 shows the relationships between planned sentence production and increased understanding as a function of self-monitoring.

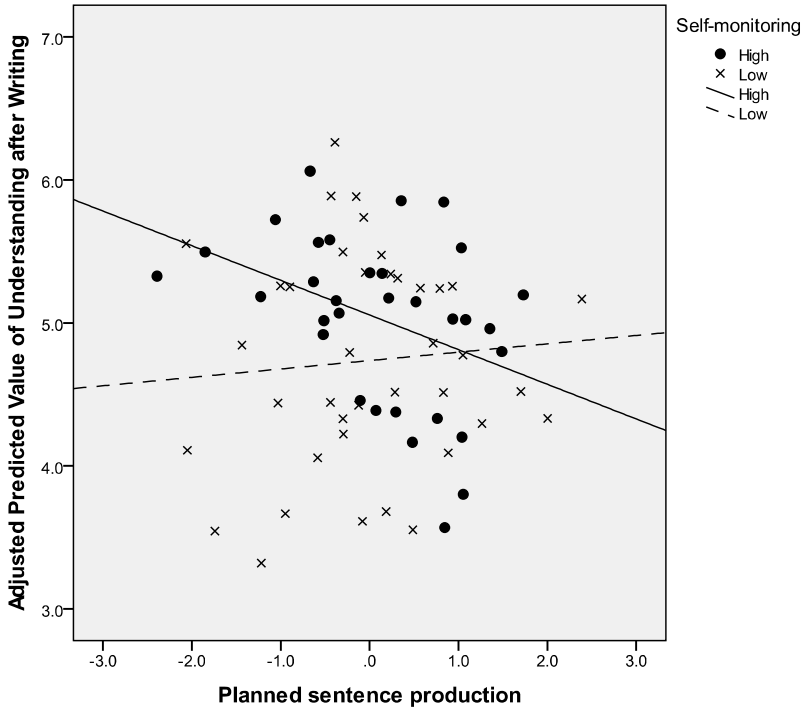


**Figure 5-4** Relationship between planned sentence production and increased understanding as a function of self-monitoring.

As can be seen in the figure, there is a significant negative relationship between planned sentence production and increased understanding for the high self-monitors ($r$=-.39, $N$ =32, 2 tailed test, $p$ =.03), but no relationship for the low self-monitors ($r$ =.08, $N$=39, 2 tailed test, $p$= .63). Taken together with the previous analysis, this suggests that, for high self-monitors, there is a general relationship between unplanned sentence production and increases in understanding that is enhanced when writing is, in addition, less top-down controlled and the global structure of the text is less linearly produced. By contrast, for the low self-monitors, the relationship appears to depend on writing also being less top-down controlled and the global structure of the text being non-linearly produced.

Overall, then, this analysis suggests that increases in understanding are associated with an interaction between memorisation, unplanned sentence production and less linearly produced global text structure. In general, the extent to which increases in understanding are related to less planned sentence production appears to depend on the extent to which writing is less top-down controlled and on the extent

to which the writer revises the global structure of their text. Note, also, that the extent of memorisation varies depending on type of planning, with the outline planned condition producing a greater reduction on the length of preserved ideas than synthetic planning. The fact that memorisation also relates to increases in understanding may also account for the fact that synthetically planned writing was associated with greater increases in understanding than the outline planned condition.

#### *5.3.4.2 Relationships between idea change and changes in understanding*

To assess the relationships between the major and minor preserved and new ideas and increases in understanding we carried out a sequential regression with ratings of knowledge after writing as the dependent variable. Because of the number of variables involved, we restricted the analysis of interactions to two-way interactions between variables. The analysis therefore involved entering prior understanding ratings at step 1 (to control for prior understanding) and then assessing whether entering the idea change variables entered at subsequent steps were associated with a significant increase in variance accounted for.

The final model for this analysis is shown in table 5-5. After controlling for prior understanding, there is a significant increase in $R^2$ at step 2 ($R^2$change =.07, $F$ (6, 64) = 2.25, $p < .05$). Three variables made significant independent contributions to increased understanding: (i) Type of planning; (ii) Number of minor new ideas; and (iii) Number of minor preserved ideas.

The relationship between type of planning and increased understanding simply reflects the finding that we have already described whereby synthetic planning was associated with greater increases in understanding than outline planning. These results suggest, in addition, that increased understanding is associated across conditions with the production of fewer minor new ideas and the retention of fewer minor preserved ideas than when understanding remains unchanged. The fact that the relationship is negative for minor preserved ideas suggests that increases in understanding depend on the extent to which writers change or drop their initial minor ideas. However, the fact that increases in understanding are also associated with the production of fewer minor new ideas suggests that this is not simply a matter of replacing minor preserved ideas with minor new ideas.

**Table 5-5** Predicting understanding after writing from prior understanding, self-monitoring, type of planning, and the number of major and minor new and preserved ideas.

| | B | SE | β |
|---|---|---|---|
| **Step 1** | | | |
| Constant | 1.16 | 0.36 | |
| Prior understanding | 0.82 | 0.08 | .78*** |
| **Step 2** | | | |
| Constant | 1.75 | 0.39 | |
| Prior understanding | 0.71 | 0.08 | .68*** |
| Self-monitoring | 0.11 | 0.12 | .07 |
| Type of planning | -0.27 | 0.12 | -.17* |
| Major new ideas | -0.03 | 0.06 | -.04 |
| Minor new ideas | -0.16 | 0.07 | -.18* |
| Major preserved ideas | -0.00 | 0.06 | -.00 |
| Minor preserved ideas | -0.14 | 0.06 | -.17* |

*$R^2$=.61 for Step 1, $R^2$ change=.07 for step 2 (F(6, 64) = 2.25, p < .05), * p < .05, ** p < .01, ***p < .001*

One possibility is that when new ideas are introduced during writing they are of minor importance to the writer (minor new ideas) but that, in order to lead to increased understanding, these minor new ideas need to be converted into more major new ideas during writing, and that the negative relationship between number of minor new ideas and increased understanding reflects the extent to which this transformation of importance of ideas takes place. This may not be reflected in the relationship between major new ideas and increased understanding, because major new ideas may also be produced by other processes that are unrelated to changes in understanding. Whatever the explanation for this, the fact that the relationships are negative for both measures suggests that the interrelationships between idea change at different levels of importance are complex, and that the relationships between increased understanding and the number of minor preserved and new ideas should be treated as a symptom of more complex underlying changes rather than as reflecting a direct causal relationship with increased understanding.

### *5.3.5 Relationship between process measures and idea change measures*

In this section, we examine the relationships between the process measures and change in ideas. As with the earlier analyses of process measures, we have included the memorisation measure as a process indicator along with the measures of planned

sentence production and global linearity. Because of the number of variables involved, we restricted the analysis of interactions to three-way interactions.

#### *5.3.5.1 Relationships with number of major preserved ideas*

The final model for this analysis accounted for 20% of the variance in the number of major preserved ideas ($R^2$ = .20, $F$ (6, 69) = 2.89, $p$= .014). As can be seen in table 5-6 two terms made significant independent contributions to the model: (i) the interaction between self-monitoring and memorisation; and (ii) the interaction between type of planning and global linearity. Other significant effects reflected lower level terms contributing to these interactions. The interaction between self-monitoring and degree of memorisation is plotted in figure 5-5. Please note that memorisation is reversed scored.

**Table 5-6** Number of major preserved ideas predicted by self-monitoring, type of planning, global linearity and degree of memorisation.

| | **B** | **SE** | **β** |
|---|---|---|---|
| Constant | 0.53 | 0.03 | |
| Self-monitoring | 0.01 | 0.03 | .02 |
| Type of planning | 0.05 | 0.03 | .17 |
| Global linearity | -0.06 | 0.02 | -.43** |
| Memorisation | 0.05 | 0.02 | -.30* |
| Type of planning * Global linearity | 0.08 | 0.03 | .40* |
| Self-monitoring * Memorisation | -0.11 | 0.03 | -.43*** |

*$R^2$ = .20, F (6, 69) = 2.89, p= .014. * p < .05, ** p < .01*** p < .005*

As can be seen in figure 5-5, there is a negative relationship between the memorisation score and number of major preserved ideas for the high self-monitors ($r$ = -.78, $N$= 35, 2 tailed test $p$ < .0005), indicating that a higher level of memorisation is associated with the retention of a higher number of major preserved ideas for the high self-monitors. By contrast, the relationship is positive for the low self-monitors ($r$=.44, $N$=38, 2 tailed test, $p$<.006), indicating that higher levels of memorisation are associated with the retention of a lower number major preserved ideas for the low self-monitors.

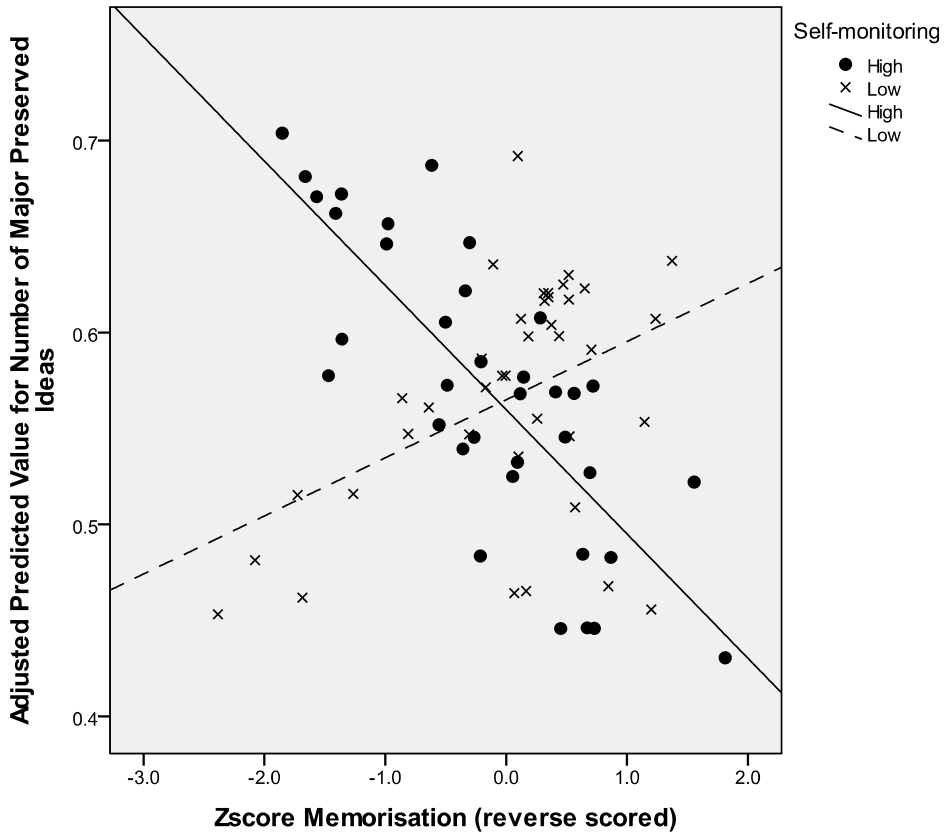


**Figure 5-5** Relationship between memorisation and number of major preserved ideas as a function of self-monitoring.

These data suggest that the mental outline maintained during writing is different for the low and high self-monitors. For the high self-monitors, it appears to include a high number of their original ideas, whereas for the low self-monitors it appears to involve actively reducing the number major preserved ideas maintained during writing. Note that, since there was no significant effect of type of planning on the number of major preserved ideas in the earlier analysis of the effects of self-monitoring and type of planning and idea change, this effect appears to be a consequence of a difference between low and high self-monitors in how they maintain an outline during writing. This is consistent with the suggestion that high self-monitors engage in a more top-down writing process than the low self-monitors.

The interaction between type of planning and global linearity is shown in figure 5-6.

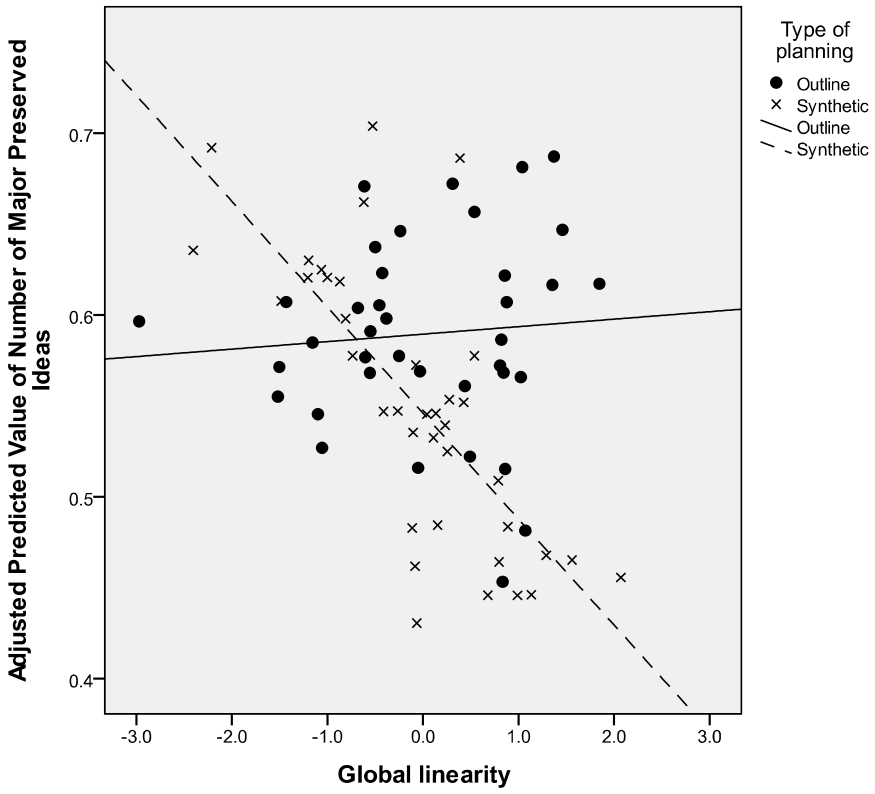


**Figure 5-6** Relationship between global linearity and number of major preserved ideas as a function of type of planning.

As can be seen in the figure, the interaction is a consequence of the fact that there is a negative relationship between the number of major preserved ideas and global linearity within the synthetic planning condition ($r$ = -.73, $N$ = 38, $p$<.0005) but no relationship within the outline planning condition ($r$= .08, $N$ = 37, $p$ =.63). This result implies that when writers produce linear text following synthetic planning they tend to retain less of their major initial ideas than when they outline plan, but that this can be overcome if the writer revises the global structure of the text during text production. In other words, synthetic planning reduces the extent to which writers are able to linearly reproduce their main initial ideas: their writing tends to follow the path of text production unless they explicitly re-evaluate the global structure of the text.

### *5.3.5.2 Relationships with major new ideas*

The final model for this analysis was of marginal significance ($R^2$ =.18, $F$(7,68)= 2.11, $p$ =.054). As can be seen in table 5-7, two terms made significant contributions to the model: (i) a significant interaction between self-monitoring and type of planning; and (ii) a significant three-way interaction between self-monitoring, type of planning and

global linearity. The two-way interaction is simply the result we have already described above, in which high self-monitors' synthetically planned, and low self-monitors' outline planned texts were associated with the production of a higher number of major new ideas than in other conditions.

**Table 5-7** Number of major new ideas predicted by self-monitoring, type of planning, global linearity and degree of memorisation.

| | **B** | **SE** | **β** |
|---|---|---|---|
| Constant | 0.34 | 0.05 | |
| Self-monitoring | 0.13 | 0.07 | .31 |
| Type of planning | 0.11 | 0.06 | .25 |
| Global linearity(GL) | -0.03 | 0.04 | -.13 |
| Self-monitoring * Type of planning | -0.26 | 0.09 | -.53** |
| Type of planning * GL | 0.09 | 0.06 | .31 |
| Self-monitoring * GL | 0.13 | 0.09 | -.37 |
| Self-monitoring * Type of planning * GL | -0.25 | 0.11 | -.64* |

*$R^2$ = .22, F (7, 68) = 2.811, p= .054. * p < .05, ** p < .01*

The three-way interaction is represented in figure 5-7, where the relationship between global linearity of text production and the number of major new ideas is plotted separately for the two planning conditions and for low and high self-monitors.

The figure clearly shows that the interaction between self-monitoring and type of planning for the number of major new ideas that we described in the section analysing idea change is a consequence of a different relationship between the linearity of text production and major new ideas for low and high self-monitors in the two planning conditions. In the synthetically planned condition, there is a positive relationship between linear text production and the number of major new ideas produced after writing for the high self-monitors ($r$ =.95, $N$= 17, $p$<.001), but a negative relationship for the low self-monitors ($r$ = -.81, $N$ =21, $p$<.001). In the outline planned condition, the relationships are in opposite directions (for high self-monitors, $r$ = -.97, $N$= 19, $p$< .001; for low self-monitors, $r$=.92, $N$= 19, $p$< .001). Thus, the fact that low self-monitors' outline planned texts, and the high self-monitors' synthetically planned texts produced a higher number of major new ideas than the other conditions is a consequence of the fact that linear text production is associated with the production of more major new ideas in these conditions, whereas it is associated with the production of fewer major new ideas in the other conditions.

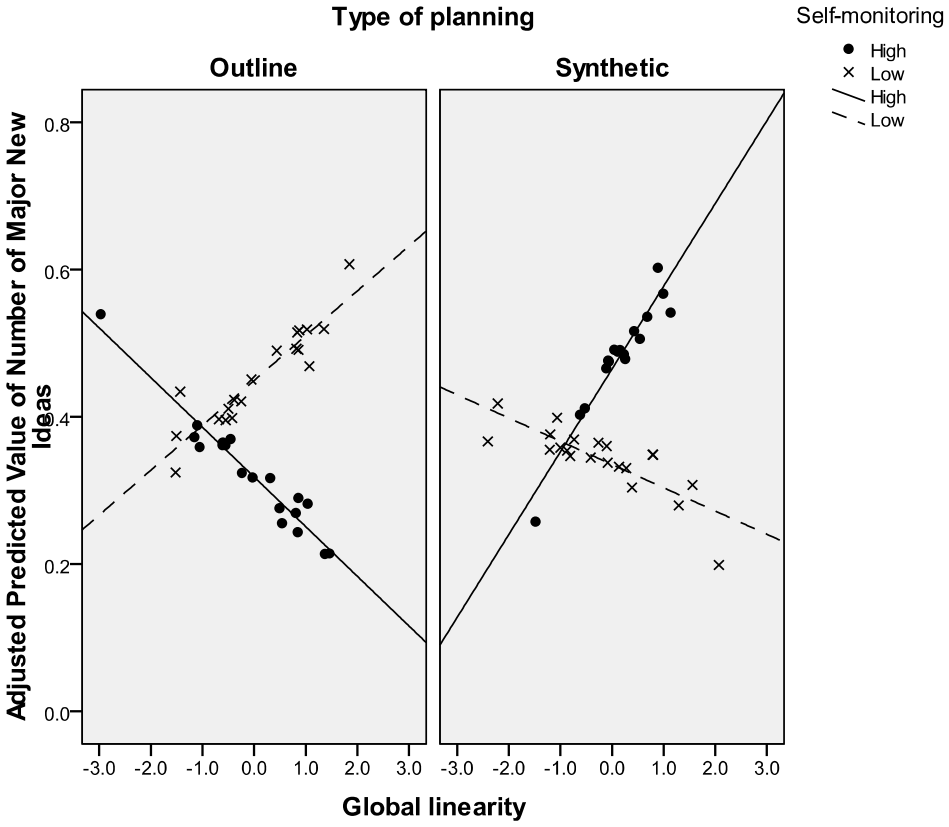


**Figure 5-7** Relationship between major new ideas, global linearity and self-monitoring plotted separately for outline planning and synthetic planning.

Overall, these data provide evidence that the differences in the number of major new ideas in the different conditions are related to differences in the linear planning of text. Our suggestion is that this involves an interaction between global linearity and local sentence production processes, with strong top-down control preventing the production of new ideas, and weak top-down control preventing the conversion of locally produced ideas into global level ideas. On this account, the higher number of major new ideas in the low self-monitors' outline planned texts and in the high self-monitors' synthetically planned texts, is a consequence of the fact that both processes are in operation.

*5.3.5.3 Relationships with minor preserved ideas.*

The final simplified model for this analysis is shown in table 5-8. As can be seen in the table, there is a significant interaction between self-monitoring and type of planning, reflecting the finding we described earlier whereby fewer minor preserved ideas were retained in the low self-monitors' outline condition and the high self-monitors' synthetically planned condition. In addition, there is a significant negative

relationship between degree of memorisation and number of minor preserved ideas, indicating that higher levels of memorisation are associated with the retention of a greater number of minor preserved ideas.

**Table 5-8** Number of minor preserved ideas predicted by self-monitoring, type of planning, and degree of memorisation.

| | **B** | **SE** | β |
|---|---|---|---|
| Constant | 0.50 | 0.04 | |
| Self-monitoring | -0.19 | 0.06 | -.49*** |
| Type of planning | -0.13 | 0.06 | -.33* |
| Memorisation | -0.05 | 0.02 | -.24* |
| Self-monitoring *Type of planning | 0.27 | 0.02 | -.23* |

*$R^2$ = .19, F(4, 67) = 3.85, p = .007. * p < .05, *** p < .005*

#### *5.3.5.4 Relationships with minor new ideas*

There were no significant relationships between any of the variables and the number of minor new ideas. We concluded that the process measures do not predict the extent to which minor new ideas are listed after writing.

### *5.3.6 Text quality*

In this section we consider the effects of self-monitoring and type of planning on text quality, and assess the relationships of writing processes, idea change and change in understanding with text quality.

#### *5.3.6.1 Effects of self-monitoring and type of planning*

Table 5-9 shows the mean ratings of text quality in each condition. Although participants in the outline planning condition produced higher quality text than those in the synthetic planning condition, and high self-monitors produced higher quality text than low self-monitors, neither of these effects was significant ($p$ > .20 in both cases), and the effect sizes were relatively small in both cases (for type of planning, partial $\eta^2$ = .02; for self-monitoring, partial $\eta^2$ = .01)

**Table 5-9** Mean ratings of text quality as a function of self-monitoring and type of planning.

| | **Outline planning** | | **Synthetic planning** | |
|---|---|---|---|---|
| Self-monitoring | High SM | Low SM | High SM | Low SM |
| Text quality | 5.72 (1.79) | 5.17 (2.25) | 5.06 (1.28) | 4.76 (2.18) |

Given that synthetic planning produced significantly more increases in understanding than outline planning, these results suggest that increases in understanding are not necessarily associated with high quality texts.

*5.3.6.2 Relationships with writing processes*

The final model for this analysis accounted for 24% of the variance in text quality scores ($R^2$ = .24, $F$ (5, 70) =4.40, $p$ =.002). The coefficients for the terms in the final model are shown in table 5-10.

**Table 5-10** Text quality predicted by self-monitoring, type of planning, planned sentence production and global linearity.

| | **B** | **SE** | **β** |
|---|---|---|---|
| Constant | 4.91 | 0.28 | |
| Self-monitoring | 0.51 | 0.40 | .13 |
| Planned sentence production | 0.47 | 0.20 | .24* |
| Global linearity | -0.03 | 0.26 | -.01 |
| Self-monitoring * Global linearity | 0.95 | 0.42 | .30* |
| Global linearity * Planned sentence production | -0.80 | 0.26 | -.33*** |

*$R^2$ =.24, F (5, 70) = 4.40, p =.002. * p < .05, *** p < .005*

As can be seen in the table, there are significant interactions between self-monitoring and global linearity, and between planned sentence production and global linearity. These data are plotted in figure 5-8. This figure shows the relationship between planned sentence production and the model's predicted values of text quality, plotted separately for high and low self-monitors, with separate lines for different levels of global linearity (based on a median split of the global linearity measure).

As can be seen in the figure, there is a clear difference for the high self-monitors (the left hand panel) between linearly produced texts (the black dots and

solid line) and non-linearly produced texts (the crosses with a dashed line), with the linearly produced texts being of higher quality than the non-linearly produced texts ($r$ = .84, $N$ =36, 2 tailed test, $p$ < .001). By contrast, for the low self-monitors (the right hand panel), there is no significant relationship between global linearity and text quality ($r$ = .02, $N$ = 40, 2 tailed test, $p$ =.89). In addition, for the linearly produced texts, the high self-monitors produce better quality text than the low self-monitors ($r$ =.80, $N$=37, 2 tailed test, $p$< .001). Overall, then, the high self-monitors produce the best quality text when they write linearly, and this is of higher quality than when the low self-monitors write linearly.

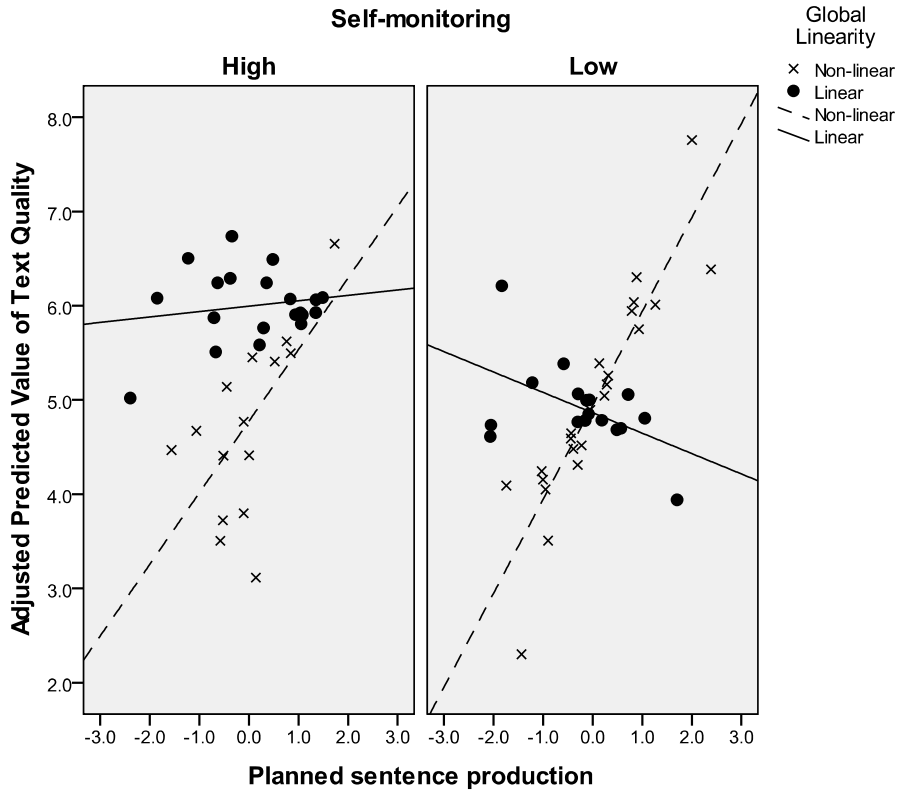


**Figure 5-8** Relationship between text quality, planned sentence production and global linearity plotted separately for high and low self-monitors.

Turning to the relationship between planned sentence production and global linearity, the figure shows clearly that there is a strong positive relationship, for both low and high self-monitors, between planned sentence production and higher text quality when texts are non-linearly produced (crosses and dashed lines) ($r$ = .83, $N$= 39, 2 tailed test, $p$<.001) but no relationship when texts are linearly produced ($r$ = .04, $N$ = 37, 2 tailed test, $p$ = .83). As can be seen in the figure, this reflects the fact that for both

groups, poorer text is produced when sentence production is unplanned and non-linear organisation is imposed during writing. However, if sentence production is planned then non-linearly produced texts are equal in quality to linearly produced texts. Indeed, for the low self-monitors, whose linearly produced texts are of lower quality than the high self-monitors', a combination of planned sentence production and non–linear organisation produces the highest quality text.

#### *5.3.6.3 Relationships with idea change*

The final model for this analysis is shown in table 5-11. This model predicts 25% of the variance in text quality ($F$ (7,68) =3.32, $p$=.004), with two terms making significant independent contributions: (i) a two way interaction between self-monitoring and the number of minor preserved ideas; and (ii) a two-way interaction between the number of major new and old ideas. In addition, there is a marginally significant negative relationship between the number of minor new ideas and text quality ($p$=.063)

**Table 5-11** Text quality predicted by self-monitoring and number of major and minor preserved and new ideas.

| | **B** | **SE** | **β** |
|---|---|---|---|
| Constant | 5.01 | 0.28 | |
| Self-monitoring | 0.49 | 0.42 | .13 |
| Major new ideas | -0.16 | 0.22 | -.08 |
| Minor new ideas | -0.42 | 0.22 | -.21 |
| Major preserved ideas | -0.17 | 0.21 | -.09 |
| Minor preserved ideas | -0.67 | 0.34 | -.34 |
| Self-monitoring * Minor preserved ideas | 1.42 | 0.44 | .55*** |
| Major new ideas * Major preserved ideas | -0.37 | 0.17 | -.24* |

*$R^2$ = .25, Adjusted $R^2$ =.18, F (7, 68) = 3.32, p =.004; * p <.05, *** p <.005.*

The interaction between self-monitoring and the number of minor preserved ideas reflects the fact that there is a positive relationship between the number of minor preserved ideas and text quality for the high self-monitors ($r$= .99, $N$ =37, 2 tailed test, $p$ < .0001), but a negative relationship for the low self-monitors ($r$ =-.98, $N$ = 40, 2 tailed test, $p$ < .0001). High self-monitors wrote better texts the more they retained minor preserved ideas; low self-monitors wrote better texts the less they retained minor preserved ideas. Given that a lower number of minor preserved ideas is also

associated with increased understanding – for both groups – this adds further support to the suggestion that the production of better text is not associated with increased understanding for high self-monitors but may be associated with increased understanding for low self-monitors.

The interaction between the number of major new ideas and major preserved ideas is plotted in figure 5-9. For this figure, the number of major preserved ideas was divided into two groups using a median split, and the relationship between the number of major new ideas and the adjusted predicted values for text quality was plotted separately for more and less major preserved ideas.

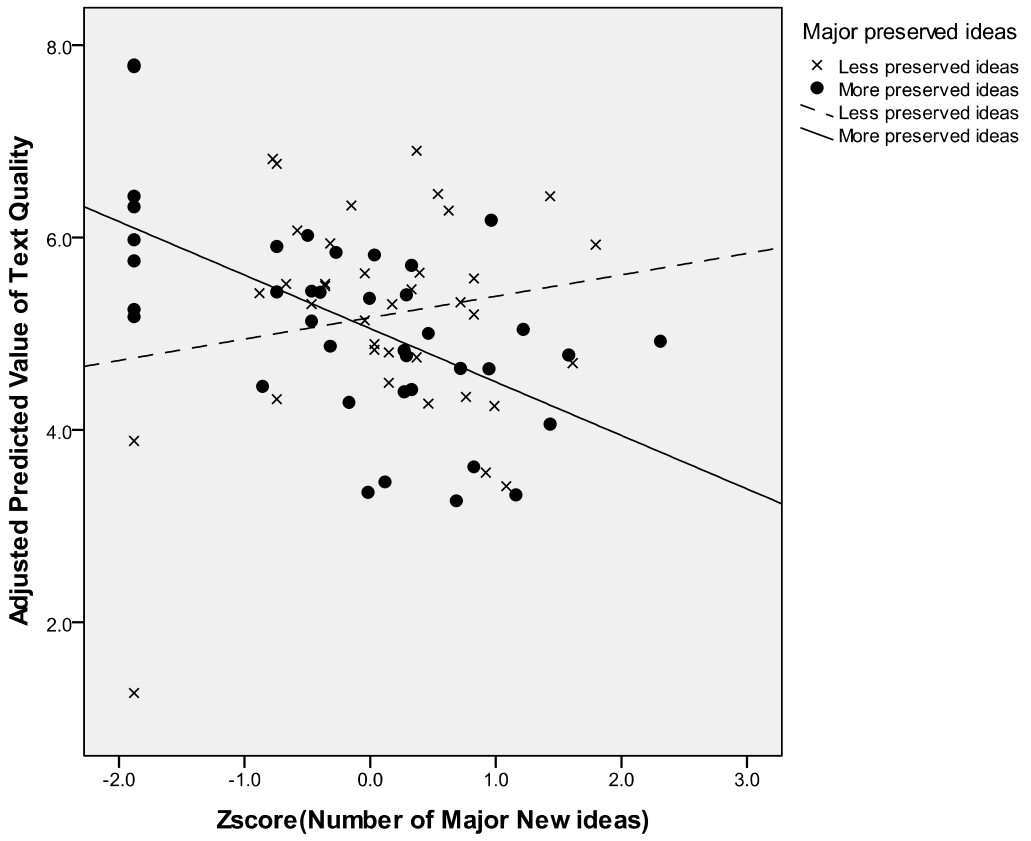


**Figure 5-9** Relationship between number of major new ideas and adjusted predicted values of text quality as a function of the number of major preserved ideas.

As can be seen in the figure, there is a negative relationship between the number of major new ideas and text quality when there are also a higher number of major preserved ideas ($r$ = -.59, $N$ = 39, 2 tailed test, $p$ < .0005). By contrast, when there are fewer major preserved ideas (the dotted line), the relationship is positive, but not significant ($r$=.17, $N$=37, 2 tailed test, $p$ = .32). This suggests that lower quality text is

produced when a high number of both major preserved and major new ideas are produced, perhaps because of a conflict between these ideas. By contrast, higher quality text is produced when the major ideas are either preserved ideas or new ideas.

#### *5.3.6.4 Relationships with increased understanding*

In order to examine the relationship between increased understanding and text quality we first calculated the simple correlation between increased understanding and text quality. Although this was a negative relationship, it was not significant ($r$=-.17, $N$=73, 2 tailed test, $p$=.16). We then regressed text quality on changed understanding (with decreases in understanding excluded), entering main effects of self-monitoring and type of planning and interactions at successive steps. However, the overall model was not significant ($R^2$ = .07, $F$ (7, 65) =0.68, $p$ =.69), and none of the individual coefficients approached significance ($p$ >.35 in all cases).

We conclude, therefore, that there is no evidence of a relationship between increased understanding and text quality.

## 5.4 DISCUSSION

In order to manage the discussion of this complex set of results, we will break it down into a series of progressively more detailed steps. We will begin by discussing the features of the results that are mostly directly related to the broader writing research literature – the effects on text quality and on overall ratings of subjective understanding after writing. We will then consider the idea change data in relation to Galbraith's previous research in this area. Having contextualised the present experiment in terms of these two lines of research, we will then consider the extent to which the inclusion of keystroke measures in this experiment has enabled us to develop our understanding of the processes involved. Finally, we will consider the interrelationships between processes and idea change, and speculate about the way in which these lead to the effects on the development of understanding and on text quality that we have observed. Throughout, we will consider the implications of the results for the problem solving and dual-process models of writing.

### *5.4.1 Effects on development of understanding and text quality*

One of the most well-established findings in the literature is Kellogg's finding that outlining has beneficial effects on text quality (see Kellogg, 1994, for a review). The first important result in the present research, therefore, is the failure to find a beneficial effect of outlining on text quality. Although there was a tendency for better

text to be produced in the outline condition, and for high self-monitors to produce better text than low self-monitors, neither of these effects approached significance, and the effect sizes were small in both cases. Previous research by Kellogg (1988) also found that mental outlining was just as effective as making an external outline. This implies that a possible explanation for our failure to find beneficial effects of outlining could be that writers in the synthetic planning condition made a mental outline during pre-planning. In this experiment we have actively tried to prevent participants in the synthetic planning condition to make a mental outline by instructing them to sum up their ideas in a single sentence before writing rather than to organise their ideas. This does not guarantee that these writers did not make mental outlines, but it will, at least, reduce the extent to which this was done. Furthermore, it does not seem likely that writers in the synthetic planning conditions made a mental outline during pre-planning, since outlining and synthetic planning had different effects on other measures. In particular, outlining was associated with a significant reduction in the length of preserved ideas after writing, and with significantly less increased understanding than synthetic planning. This suggests that synthetic planning involved a genuinely different form of planning rather than simply being a mental equivalent of outline planning. These results, therefore, suggest that synthetic planning can produce texts similar in quality to outline planned texts, and that the use of this form of planning as part of an explicit drafting strategy should be further investigated.

Given this broad equivalence in the effects of synthetic and outline planning on text quality, these results also have important implications for the problem solving and dual process models of writing. First, the fact that increased understanding was not generally associated with higher text quality (if anything, the relationship was negative), calls into question the standard reception of Bereiter and Scardamalia's (1987) knowledge-transforming model. Although they themselves do not claim a direct relationship between text quality and knowledge-transforming processes (see Bereiter, Scardamalia and Steinbach, 1984), suggesting instead that knowledge-transforming processes should be associated with the production of more reflective text, this is a general assumption within the field. It is very common for researchers to attribute differences in quality of texts to a difference in the extent of knowledge-transforming processes, and in some cases – e.g. Klein et al. (2011) – it is explicitly claimed that the development of understanding should be correlated with higher text quality. Our results suggest, then, that the term "knowledge-transforming" should not be used to characterise high quality texts simply on the grounds that they are of higher quality. Our results show, very straightforwardly, that "knowledge transformation" is not necessarily associated with better writing. By contrast, the dual

process model explicitly predicts a dissociation between the development of understanding and the production of high quality text. In particular, it claims that, although outlining may facilitate explicit organising processes and henceforth that these do make a contribution to the quality of the text, it inhibits the implicit organising processes involved in knowledge constituting, and hence reduces the development of understanding. The present results strongly confirm this prediction.

The dual process model goes further than this in that it provides an explicit account of the different kinds of processes involved, and specifies that individual differences in self-monitoring will be related to different ways of combining these processes. It claims that both processes are required in order to produce high quality text, and that low and high self-monitors' high quality texts will therefore show evidence of both kinds of processing, but that these will be combined in different ways by the two groups. High self-monitors are assumed to engage in top-down processing; focussing first on explicit organisation and then on translating content specified by this process into text. Low self-monitors are assumed to engage in bottom-up processing; focussing first on formulating their implicit disposition in connected sentences, and then on adapting the global organisation of the text to the ideas that are constituted in the text.

The principal component analysis that we carried out on measures of pauses, bursts and revisions at different text levels enabled us to distinguish between processes occurring at sentence level and processes involving processing of global text structure. Both of these measures are essentially measures of linearity of text production but differ in whether this occurs at the sentence level or at higher levels of text structure. Planned sentence production, as we have called it, distinguished between writers who typically pause longer before producing relatively clean bursts which are not revised, and writers who pause for briefer periods before producing bursts that are more frequently revised. The global linearity measure distinguished between writers who produced sentences in sequence one after another from writers who scrolled back through the text and inserted text elsewhere. The first group of writers are presumed to produce text in an essentially top-down manner, generating and organising ideas before writing them down whereas the latter group of writers are presumably revising earlier parts of the text more frequently.

The relationships between these measures and text quality provide strong support for the dual process model. First, the interaction between planned sentence production and global linearity in predicting text quality confirms the general claim that higher level explicit planning processes and more local sentence production processes have an interactive effect on text quality. Specifically, this suggests that planned sentence production is only positively related to text quality when the text

that results from planned sentence production is re-organised at a more global level, i.e. the text as a whole is non-linearly produced. Second, the interaction between self-monitoring and global linearity, in which global linearity is positively related to text quality for the high self-monitors but not the low self-monitors, strongly supports the claim that high self-monitors write best when they employ a top-down strategy, planning what they want to write before writing it down. Finally, taken together, the overall pattern of these results show that high self-monitors write best when they produce text linearly, from the top-down, and that they produce better text than the low self-monitors who write linearly. By contrast, low self-monitors write best when they use planned sentence production combined with non-linear development of the global structure of the text. These findings provide strong support for the dual process model, and represent a new finding as this is the first research to directly examine effects of self-monitoring and types of processing on text quality.

Two caveats are necessary here. First, the above interpretation of the data depends crucially on assuming that planned sentence production leads to non-linear global text production. Given that the data are correlational in form, however, it should be kept in mind that the causal relationship between the two variables cannot show directly in the data, but is an inference on our part. This inference needs to be tested further in future research. Secondly, although these data are compatible with the dual process model's claim that bottom-up sentence production is unpredictable and hence conflicts with higher level organisation, they do not provide any evidence about the nature of planned sentence production itself. In fact, a key question for the dual process model's specific claims about sentence production is what distinguishes the kind of planned sentence production associated with non-linear text production from other forms of sentence production, and whether this can be characterised as dispositionally guided sentence production. It therefore elicitates the question what the precise nature of dispositionally guided text production is. Henceforth, an important avenue for future research is to establish what the precise nature of planned text production is and how this relates to dispositionally guided text production as described by Galbraith.

Overall, these results provide strong support for the broader claims of the dual process model. They call into question whether producing high quality text is a knowledge-transforming process. They show that text quality depends on the joint effects of high level planning processes and more local sentence production processes, and suggest that these have conflicting effects. They confirm its prediction that high self-monitors' better quality writing is associated with a top-down process, whereas low self-monitors' is associated with a bottom-up relationship between sentence production and global planning.

### *5.4.2 Effects on idea change*

In Galbraith's previous research (Galbraith, 1999, 2009; Galbraith, Torrance and Hallam, 2006) low self-monitors have typically been found to produce more new ideas after synthetically planned writing than high self-monitors, and these new ideas have typically been associated with increases in understanding for low self-monitors' synthetically planned texts but not for high self-monitors' synthetically planned texts or for outline planned writing by either group. Galbraith (2009) explains these effects by claiming that only ideas produced by dispositionally guided text production (low self-monitors' synthetically planned texts) lead to increases in understanding. Ideas produced by other processes may be "new" in the sense that they appear for the first time in the list produced after writing, but are assumed in fact to consist of different pre-existing ideas retrieved from long term memory in order to satisfy the writer's rhetorical goals. Galbraith assumes that it is only in the low self-monitors' synthetically planned writing that new ideas are predominantly produced by dispositional text production and hence that only in these conditions there is a direct relationship between the number of new ideas and increased knowledge. In other conditions, where new ideas are produced by a mixture of processes, there is no direct relationship between number of new ideas and increased understanding.

The first important feature, therefore, of the present experiment was that it produced a different pattern of results to previous research. In particular, there was no difference between the number of new ideas produced by low and high self-monitors under synthetically planned conditions. In fact, the results were in the opposite direction, with high self-monitors producing significantly more major new ideas than low self-monitors in the synthetic planning condition. Furthermore, although significantly more increases in understanding occurred after synthetic planning than after outline planning, this was true for both the low and high self-monitors, and in neither case was this associated with a higher number of new or major new ideas.

There are two differences in the way the present experiment was carried out that could account for the difference in results. First, participants in this experiment were Dutch rather than of UK origin. It is possible that differences in self-presentation, writing strategies or educational experiences between Dutch and UK students may have caused that self-monitoring does not relate to underlying writing processes in the same way for the two national groups. Second, the rhetorical context was much more fully specified in this experiment than in the earlier experiments. Thus, in this experiment, in order to disentangle the confounding effects of type of planning and of rhetorical constraints that had been present in previous research, all

writers were asked to produce a well-structured article for the university newspaper. In Galbraith's experiments, the context was either much less fully specified – " e.g. write an essay" (Galbraith, 1992, 1999) – or writers were actively instructed to ignore the rhetorical context and to write down their thoughts, in connected sentences, but without worrying about how well organised the text was. According to the dual process model, increasing rhetorical constraints should increase the extent to which high self-monitors produce new ideas (because they are assumed to be more likely to adapt their ideas to rhetorical goals) and reduce the extent to which low self-monitors produce new ideas (because imposing rhetorical constraints is assumed to inhibit the extent to which dispositional text production occurs). If these assumptions are correct, then one would expect to find, as we have, that low self-monitors (because of the imposed rhetorical constraints) should produce fewer new ideas than high self-monitors, and that there should be no direct relationships between the number of new ideas and increases in understanding. Clearly, both these explanations need to be tested in future research (i) by comparing the two national groups writing under identical conditions; and (ii) by comparing the effects of different types of planning, either with rhetorical constraints (as in this experiment) or without rhetorical constraints (as in Galbraith et al., 2006).

Given this difference in the overall pattern of results, however, the question remains of how the pattern of results that we observed in the present experiment can be explained. To do this, we will first summarise the main features of the idea change results, and then consider how these relate to the writing process measures and changes in knowledge.

The first important finding here was that outline planning was associated with a reduction in the length of preserved ideas compared to the length of ideas in the initial list whereas synthetic planning was not. We have interpreted this as an indication of the extent to which writers maintain their initial ideas in memory during writing. The argument being that, when ideas are stored in memory, the writer does not need to formulate them in full after writing, as they would if they were generating them for the first time, but can instead label them in an abbreviated form. This is clearly speculative and needs to be explicitly tested in future research. One way to do this would be to compare the effects of different forms of outlining, with and without a verbal secondary task, on the length of ideas when they are subsequently recalled. For example, writers could be asked either (i) to make an outline as usual without further instruction; (ii) to actively try to remember the outline during writing; or (iii) to write without an outline. They could then be asked either to carry out a verbal secondary task or not. The prediction would be that ideas in the outline would be more reduced in length after active memorising instructions compared to the other

conditions, but that these differences would disappear when a secondary task was imposed.

On the assumption that the reduction in length of preserved ideas is indeed a measure that indicates the extent to which writing is controlled by the writer's initial plan our analyses suggested a number of conclusions about this measure. First, it appeared to be relatively independent of the other process measures in that there were no significant relationships with the other measures. Second, it was correlated with the number of major preserved ideas for the high self-monitors but not the low self-monitors. This suggests that high self-monitors' mental outlines are based more directly on their main initial ideas, and is compatible with the assumption made earlier that high self-monitors' writing is more top-down than low self-monitors. Third, it moderated the relationships between planned sentence production, global linearity and increases in understanding. We will postpone discussion of these effects until we consider the effects of the process measures on idea change. For the moment, we shall simply conclude that this measure may be a useful indicator of top-down processes during writing, and that it needs to be investigated further in future research.

Turning to the effects on the other idea change measures, our second important finding was that low self-monitors' outline planned writing and high self-monitors' synthetically planned writing produced a higher number of major new ideas and lower number of minor preserved ideas than other conditions. Although these differences were clear cut and highly significant, they are hard to explain, particularly for a single process model. For example, if the high self-monitors are assumed to direct their writing towards rhetorical goals then, according to the knowledge-transforming model, the reason that high self-monitors lose more of their original minor ideas and produce a higher number of major new ideas than the low self-monitors in the synthetically planned condition, would be because they had adapted their ideas more to the rhetorical context than low self-monitors. It could then be argued that outlining, by determining the order of ideas, would reduce the extent to which they are able to do the same in the outline planning condition. But if this were the case, how can the low self-monitors' similar pattern of results in the outline condition be explained? By contrast, for the dual process model, the varying patterns are assumed to be a consequence of a complex interaction between explicit planning processes and implicit text production processes, with the particular pattern of change depending on the precise balance between these processes in different conditions. The key question for the dual- process model is to specify how these patterns of change are related to changes in understanding and to the underlying writing processes.

The relationships between change in ideas and increased understanding are also extremely hard to interpret. The main finding was that increases in understanding were associated with the production of fewer minor new ideas and the retention of fewer minor preserved ideas across all conditions. This suggests that increases in understanding may be associated with a general shift towards a more global level of representation after writing where ideas have gained in importance. However, the fact that increases in understanding were not associated with differences in the number of major new or preserved ideas suggests that this is not simply a matter of how important the ideas produced after writing are. Rather, the present results suggest that simply counting the number of new and old ideas at different levels of importance is not sufficient to capture the changes in ideas associated with changes in understanding. Instead, one may need to measure the interrelationships between these ideas, and how these relate to the overall coherence of the writers' ideas. Our main conclusion, then, is a methodological one. Counting number of new and old ideas may be sufficient in conditions where a relatively pure single process is responsible for their production, as may have occurred for low self-monitors synthetically planned texts in previous research where rhetorical constraints have been relaxed. In general, though, when multiple processes are involved, more sophisticated analyses of the structural relationships between ideas may be required.

Overall, then, our main conclusions for the idea change analysis are that the addition of external rhetorical constraints, which was motivated for external validity reasons and to be able to test text quality, alters the effect of writing on idea change compared to that found in previous research, and that idea change by itself is not directly related to increased understanding. Although these findings are both compatible with the dual-process model's claim that idea change is multiply determined, and that increases in understanding depend on the processes by which ideas are produced during writing rather than on the amount of change in ideas, they do suggest also that more sophisticated measures of idea change may be required in order to capture what changes in understanding involve in terms of idea change.

### *5.4.3 Relationships between processes and increases in understanding*

We started out on this analysis with two measures derived from the key stroke logs, one of which – global linearity – is assumed to represent the extent to which writers develop the global structure of their text and their plans during writing, the other – planned sentence production – which is assumed to represent the extent to which writers plan their sentences before writing them down or, instead, write them down spontaneously and then modify them once they have been produced. In the course of

the analysis, we then, in addition, identified the reduction in length of writers' original ideas after writing as an indicator of the extent to which a mental outline had been maintained during writing (memorisation). The most important finding was that increases in understanding were related to an interaction between all three of these processes, and the key question for the dual process model is how well this corresponds with the process it characterises as dispositional spelling out.

The fact that increased understanding is related to unplanned text production appears at first sight to provide strong support for the dual process model since this is compatible with its main claim that discovery depends on the spontaneous synthesis of ideas rather than on explicit planning to satisfy rhetorical goals. Furthermore, the fact that it also depends on the extent to which writing is not mentally outlined supports the claim that in order to lead to development of understanding, sentence production should be relatively free of top-down control, and should instead be guided by the implicit organisation of semantic memory. Finally, the fact that it is depends also on the extent to which the global structure of the text is revised in the course of writing is compatible with the claim that the global structure of the text emerges in the course of writing rather than being imposed beforehand.

The model also receives some support from the fact that, although the difference was not significant, it was the low self-monitors' synthetically planned texts that produced the least planned sentence production and which were most non-linearly produced. Furthermore, it was precisely when the high self-monitors wrote most like low self-monitors in the synthetically planned condition – less planned sentence production and less linearly produced text that they experienced increases in understanding. Clearly, though, these results need to be replicated in further research with a larger sample, in order to establish whether there is a reliable difference.

In broad outline, then, these results appear to support the dual process model. However, this depends to a large extent on what the precise processes are that are involved in unplanned and planned sentence production. The above account assumes that less planned sentence production is equivalent to what the dual process model characterises as dispositionally guided text production. Although the fact that this process involves relatively brief pauses before writing suggests that this is not an explicitly planned process, it should be remembered that this process also includes non-linear global structuring during the course of sentence production. These could, in principle, involve explicit rhetorical evaluation of the sentence as it is produced, and modifications of the sentence could be designed to satisfy rhetorical goals. If this were the case, then this would provide support for a central tenet of the problem solving models. By contrast, the dual–process model claims that the revision of sentences as they are produced is a consequence of a spontaneous response of the

writer's implicit disposition towards the topic. Further research is clearly needed, perhaps using think aloud protocols, or content analysis of the revisions, in order to decide this question.

Overall, we think that these results provide broad support for the dual process model's claim that discovery through writing depends on the processes by which ideas are produced rather than the amount of change in content that occurs, and that this involves an interaction between explicit organising processes and more spontaneous sentence production processes.

### *5.4.4 Relationships between processes and idea change*

These relationships were complex and are hard to interpret, particularly because it is impossible to be sure of the causal relationships between variables. Our first general observation is simply that there are a range of strong relationships between the three process measures and the idea change measures. This confirms that the process measures are indeed related to the way in which content is managed during writing. Furthermore, the fact that there are complex interactions between the process measures suggests that the functions the processes serve vary depending on the context in which they occur.

Our second observation is that the major preserved and new ideas are related to the memorisation and global linearity measures rather than to the planned sentence production measure. This suggests that the major ideas are primarily affected by more explicit global planning processes, either directly, or in response to the outcome of sentence production processes.

We will now consider the effects on major preserved and new ideas. The first finding here was that the high self-monitors retained more major preserved ideas when they maintained a mental outline during writing, whereas the low self-monitors did not. This is compatible with the assumption that high self-monitors impose more top-down control on writing than low self-monitors. The second finding was that synthetically planned texts retained less major preserved ideas than outline planned texts, unless the engaged in non-linear revision of earlier text. This suggests that, in the absence of an outline, there is a tendency for writers to drift away from their original ideas, unless they actively re-read and revise previous text in order to reconstruct their original intentions. We use the term "drift" to emphasise that this loss of major preserved ideas in the course of linear text production need not reflect a systematic change in the writer's ideas about the topic, and hence is not necessarily associated with the development of understanding.

For the number of major new ideas we found that low self-monitors' outline planned texts and high self-monitors' synthetically planned texts produced more new ideas when their texts were linearly produced. By contrast, high self-monitors' outlined planned texts and low self-monitors' synthetically planned texts produced more new ideas when their texts were non-linearly produced. Although this provides an explanation for the differences in the amount of new ideas produced in these conditions in terms of a difference in the way linear text production is carried out, this does not in itself explain why global linearity should have this different relationship with major new ideas in the different conditions. The best explanation that we can come up with is that this is a consequence of a combination of the top-down control and "drift" processes that we suggested for the major preserved ideas. According to this explanation, the high self-monitors' outline planned texts are the most top-down controlled texts. They therefore involve actively maintaining the writers' original ideas and using these to guide text production. So long as text production proceeds smoothly the linear production of text will therefore not involve creating new content, and new ideas will only be created when the writer is unable to maintain the forward progression of text production. For the low self-monitors' outline planned texts and the high self-monitors' synthetically planned, top-down control is less strong (because of the low self-monitors' different priorities, and because of the absence of an outline for the high self-monitors) as a result there is a greater tendency to drift during the course of linear text production and the writer hence writes about different ideas to those that they originally intended. The puzzling feature is why the low self-monitors' synthetically planned texts, which would be expected to be the least top-down controlled, do not show the same effect. Our suggestion is that this is because these writers are now predominantly controlled by sentence production processes, and hence their writing does not involve forming a global representation at all. They only actively reconstruct this when the local sentence production processes run into trouble. Although, this is clearly speculation, it does have the virtue that the same concepts of "top-down control" and "drift " are used to explain the effects for both the new and preserved ideas.

Overall, then, this explanation suggests that differences in the number of major new and preserved ideas that are produced after writing are partly a consequence of the different balance between top-down control processes and the drifting effect of the forward progression of text. In themselves, these differences are not assumed to lead to developments of understanding. Hence there is no direct relationship between the number of major preserved and new ideas and increases in understanding.

### *5.4.5 General conclusion*

This experiment has provided strong support for the main features of the dual process model. It has shown that the processes associated with the development of understanding are as predicted by the model. In particular it has demonstrated that change in understanding occur as a consequence of an interaction between explicit organising processes and more spontaneous sentence production processes. Furthermore, it has extended previous research by showing that the account it has developed to explain the development of understanding can also be applied to explaining differences in text quality. A key feature is that it both predicts, and explains why, the development of understanding is not necessarily associated with the production of high quality text. The production of the high quality text does not necessarily involve a knowledge transformation process.

That said, there is a need for further research into a number of features. First, more detailed investigation is needed into the processes involved in sentence production, particularly into the nature of revision during unplanned text production. Second, the assumptions that have been made about the role of a mental outline in writing need to be confirmed. But, most importantly, further research is needed into how the different processes lead to idea change and into how different configurations of ideas correspond to different levels of understanding. This could, in part, be achieved by collecting data from a larger sample of participants so that a reliable path analysis could be carried out to establish the interrelationships between variables. However, this can only address issues at an aggregated level. It seems, therefore, more adequate to turn to multilevel analysis. Hence, a better understanding could be achieved by investigating how ideas are used and changed in the course of text production itself, and how this in turn relates to developments in understanding. How do writers build their understanding as they write their texts?

# Chapter 6

*Using and Changing Ideas in Writing*

## *Abstract*

A common feature of the cognitive models in the field of writing is that discovery through writing depends on the extent to which writers change their ideas (Flower & Hayes, 1980; Bereiter & Scardamalia, 1987; Galbraith, 2009). However, as far as we know, there have not been any studies that examined the distribution of old and new ideas during the process of writing and, more importantly, which assessed the relationship between the use of ideas during writing and the development of understanding through writing. The aim of the present paper is to investigate how ideas are created in the course of text production itself and how this, in turn, relates to the development of understanding through writing. The results reveal that participants who experience development of understanding through writing show distinctly different patterns of idea use than participants who did not experience development of understanding through writing. Furthermore, these results clearly show that two different processes are associated with development of understanding through writing. These two processes directly relate to the way in which old and new ideas are introduced during writing in different planning conditions.


*Keywords*; development of understanding, process of writing, idea change, generalized additive modelling, self-monitoring

## 6.1 INTRODUCTION

The development by Hayes and Flower (1980) of a model of the cognitive processes involved in writing brought about a shift in research on writing from an almost exclusive focus on the written product towards a focus on the processes lying behind the creation of the product. Where before, the main concern had been with the way that thought was presented in language, and with the more or less effective use of language, after Hayes and Flower, research on writing shifted towards characterising the processes involved, and to identifying differences in the processes and strategies employed by expert and novice writers. Initially, this research focussed on the higher level thinking processes involved. More recently, however, greater attention has been paid to text production processes and to the way in which these are deployed, on-line, as the text is created (see Galbraith & Torrance, 1999; Torrance, van Waes & Galbraith, 2007, for reviews). An important step in this has been the introduction of statistical methods for examining how different processes are deployed during the course of text production itself (van den Bergh & Rijlaarsdam, 1996; Rijlaarsdam & van den Bergh, 1996; van den Bergh, Rijlaarsdam, Janssen, Braaksma, Weijen, van, & Tillema, 2009). This research has revealed, for example, that differences in the quality of the text that writers produce depend not just on the overall frequency with which different kinds of process are carried out, but also on when during writing the processes have been carried out.

Although this shift towards process and, then, towards the temporal characteristics of processes, has been very fruitful, it has not been without its critics. Schilperoord and Sanders (1999), for example, have suggested that in emphasizing on the process, writing research has lost sight of the written product, and have suggested that the text itself should be placed in a more central position. They advocated, in particular, that process research should be combined with text analysis methods. In this paper, we explore the potential benefits of incorporating the product in process research. In order to do so, we use measures of idea change that we have used in previous research to analyse the way in which writers move between old and new ideas during writing.

This approach evolved out of a study investigating the development of understanding through writing (Baaijen, Galbraith & de Glopper, in preparation). Low and high self-monitors were asked to write an article for a university newspaper, under two different planning conditions: conventional outlining or a non-outline condition that we called synthetic planning. The participants were asked to list their ideas before and after writing and then to rate the similarity of the ideas contained on the two lists. This enabled us to identify ideas that had been generated prior to

writing and then preserved after writing, as well as new ideas that had only been produced after writing. In addition, writers were asked to rate how much they felt they knew about the topic before and after writing. This enabled us to assess the extent to which writers experienced increases in understanding under different conditions and to relate these increases to changes in ideas that had occurred after writing. We also collected keystroke logs to assess how these changes in ideas and understanding were related to writing processes.

This experiment revealed a number of important features of the development of understanding through writing. In brief summary, it showed that writers were more likely to experience increases in understanding after synthetically planned writing than after outline planned writing and, that these increases in understanding were associated with unplanned sentence production combined with less linear production of text. It also showed clear differences between low and high self-monitors in the way these processes related to text quality. High self-monitors produced better text when they produced linear text, whereas low self-monitors produced their best texts when planned sentence production was combined with non-linear text production. In general, text quality was not related to the development of understanding.

Our overall conclusion was that these results could be explained by a contrast between a top-down and a bottom-up approach to writing proposed by Galbraith's (2009) dual process model of writing. According to the model, these approaches involve different ways of combining an explicit organising process with an implicitly organised text production process. High self-monitors are assumed to prioritise the explicit organising process and to use this to control text production in a top-down manner. Low self-monitors are assumed to prioritise the implicit text production process, and to reorganise the product of this process after it has been produced. Thus, both groups are assumed to use the two processes but to combine them in different ways. These different strategies are then moderated by whether outline planning or synthetic planning is carried out before writing. Outline planning is assumed to facilitate top-down processing, and to reduce or enhance the extent to which low and high self-monitors are able to implement their preferred strategy. Thus, the model assumes that writing is at its most top-down when high self-monitors produce outline planned text and at its most bottom-up when low self-monitors produce synthetically planned text. Intermediate combinations are assumed to occur when high self-monitors write synthetically planned texts, and when low self-monitors write outline planned texts.

A key further assumption of the dual process model is that, while the explicit organising process plays an important role in producing well-formed text, it only

involves the re-organisation of existing content and does not lead to the development of the writer's understanding. By contrast, although the implicit text production process disrupts the production of well organised text (which is why it requires further non-linear revision if it is to produce higher quality text), it does lead to the development of the writer's understanding. When this assumption was combined with the assumption about top-down and bottom-up strategies, the dual process model appeared to be able to account for many of our results. It explained why, for example, more increases in understanding occurred after synthetically planned writing than after outline planned writing, and why these increases in understanding were associated with unplanned sentence production combined with non-linear organisation. It also explained why text quality was not correlated with increases in understanding and why high quality text was associated with different combinations of processes for low and high self-monitors.

However, one puzzling feature of this experiment was that the relationship between idea change and increases in understanding was rather unclear. This is partly what the dual process model would predict. According to the dual process model explicit organising processes can produce different ideas after writing compared to before writing, but that these are simply pre-existing ideas that have been retrieved as more relevant to the rhetorical context than the ideas that were initially retrieved before writing. Hence, to the extent that idea change is a consequence of explicit organising processes it will not be associated with increases in understanding. This accounts for the fact that, although outline planning conditions do typically produce a high number of new ideas, these are not typically associated with increases in understanding (Galbraith, 1999; Galbraith, Torrance & Hallam, 2006). Furthermore, this explains why the relationship between idea change and development of understanding is further obscured when ideas produced by rhetorical planning processes are mixed in with ideas produced by implicit text production processes in other writing conditions. In general, then, the dual process model claims that the extent to which new ideas will be associated with increased understanding depends on the process by which these ideas are produced.

A more general problem with this approach to assessing the role of idea change in discovery is that the ideas that are assessed are produced after writing has taken place. Furthermore, the measures of processes that are related to these ideas are aggregated across the whole writing process. Although this does relate to questions about the effect that writing has on a writer's ideas and hence to issues related to writing-to-learn, it does not allow one to assess how ideas are created in the course of text production. And this is in fact the central question for models of discovery through writing. The main claim of different writing models is not so much about

writing-to-learn as about the process of writing itself as a process of discovery. The fundamental notion is that discovery through writing is not about translating preconceived ideas into text but rather about generating ideas in the course of producing the text itself. It is then assumed that when this content is either expressed in relation to a rhetorical context or is produced by dispositionally guided text production processes this change in ideas will lead to the development of understanding. But even aside from how these ideas are associated with increases in understanding, a key question for models of writing in general is how new ideas occur during the process of text production.

### *6.1.1 Aim of the present experiment*

To our understanding, there are no studies that have investigated directly how ideas are used and produced in the course of writing and, furthermore, how this relates to the development of understanding. Our aim in this exploratory study was to shift the focus of attention to a direct examination of how new and old ideas emerge in the text in order to investigate the relationship between idea change and development of understanding more directly. We therefore examined whether we could use the lists of ideas produced before and after writing to identify new and old ideas as they were produced in the text. In doing so we hoped to be able to incorporate the product into a process analysis, as recommended by Schilperoord and Sanders (1999).

We set out with three broad questions in mind. First of all, can lists of ideas produced before and after writing be used to identify when old and new ideas are produced during writing? Secondly, given the results of the aggregate level analysis that we have already carried out on this data, will there be differences in the way that new and old ideas are produced that correspond to the contrast between top-down and bottom up processes? Specifically, will the pattern of the production of old and new ideas of the high self-monitors in the outline planning condition correspond to a top-down process, and will the pattern of the production of old and new ideas of low self-monitors in the synthetic planning condition correspond with a more bottom-up process of writing? Thirdly, will the identification of the point during writing at which new and old ideas are produced enable us to assess more directly what the relationship is between idea change and increases in understanding? For this first attempt at this approach, we restricted ourselves to analysing a basic process feature of the texts, namely the location of new and old ideas within the linear sequence of bursts that make up writing.

Therefore, in this paper, we will investigate the occurrences of old and new ideas during the process of writing and we will explore whether these can be analysed

using generalised mixed effect regression modelling. Furthermore, we will assess how the production of new ideas and the use of old ideas during the process of writing are related to the development of understanding through writing.

## 6.2 METHOD

### *6.2.1 Participants*

84 participants from the Faculty of Arts at the University of Groningen were recruited to participate in the experiment. They were all native Dutch speakers, average age 22.2 years (*SD*=3.8), and were pre-selected using Snyder's revised 18 item self-monitoring scale (Snyder & Gangestad, 1986). Participants could only take part when they were classified either as a high or a low self-monitor. They were classified as high self-monitors (HSM, *N*=42) if they scored 11-18 on the scale and as low self-monitors (LSM, *N*=42) if they scored 0-8 on the scale. Due to technical problems we lost the process data from one of the participants, therefore this paper reports about the results of 83 participants.

### *6.2.2 Design and Procedure*

High and low self- monitors were randomly allocated to two planning conditions resulting in the four experimental groups: (i) High self-monitors outline planning (HO, *N*=21); (ii) High self- monitors, synthetic planning (HS, *N*=21); (iii) Low self-monitors outline planning (LO, *N*=20), and; (iv) Low self-monitors synthetic planning (LS, *N*=21).

In all four writing conditions, participants were asked to plan and write an article for the university newspaper discussing whether "*our growing dependence on computers and the Internet is a good development or not*". The writing task was divided into three components.

- **Before writing**

Before writing, participants were first given 10 minutes to list all the ideas they could think of relevant to the topic. It was stressed that each idea should be no longer than a sentence in length. They were then asked to rate how much they felt they know about the topic.

- **Writing task**

During the writing assignment, participants were given 5 minutes to either write down a single sentence summing up their overall opinion (synthetic planning) or to

construct a structured outline (outline planning). They were then given 30 minutes to write a well-structured article for the university newspaper. It was stressed that they had to produce a reasoned argument reflecting their own opinion about the matter. Participants were allowed to consult their written outlines during writing. During writing, keystrokes were logged using Inputlog (Leijten & van Waes, 2006).

- **After writing**

Immediately after writing, participants were asked again to rate how much they felt they knew about the topic. They were then given ten minutes to list all the ideas they could think of relevant to the topic. When the ten minutes had expired, they were asked to compare the lists produced before and after writing, and to rate the correspondence between ideas produced on list 2 with ideas that they had originally produced on list 1 on a 6-point scale ranging from 1=identical point and 6= no correspondence. Participants were given both lists and a separate sheet with three columns to do so. In the first column they found a numbered list of their ideas produced on the second list. In the second column of the sheet they were asked to note down ideas from the first list that could be related to the idea on the second list. Finally, in the third column they were asked to rate the correspondence between the ideas produced on the second and on first list.

### *6.2.3 Measures*

#### *6.2.3.1 Development of understanding*

The ratings of understanding were used to assess subjective changes in understanding as a consequence of writing. On a 7-point scale they had to indicate how much they felt they know about the topic ranging from 1= extremely little, to 7= a great deal. We established increases in understanding when participants had higher post understanding rating than their prior understanding rating.

For present analyse purposes, we have distinguished two groups of writers in terms of development of understanding. We have divided the data set into participants who reported that they had experienced increases in understanding and labelled these writers with 'Knowledge Change' (KC, *N*=22) and we have identified a group of writers who did not report increases in understanding and we labelled these writers with 'No Knowledge Change' (No KC, *N*=57).

#### *6.2.3.2 Development of ideas*

This was assessed using a procedure used in previous research (Galbraith, 1992). Ideas produced in list 1, before writing, were defined as old ideas. Furthermore, ideas

in list 2, produced after writing, which received ratings from 1 to 3 on the correspondence measure with ideas in list 1 were also defined as old ideas. New ideas were defined as ideas on list 2 that received rating from 4 to 6 for their correspondence with ideas on the first list. For present purposes, we have coded the occurrences of all of these ideas in the text so that we could analyse the relation between old and new ideas during the construction of the text.

### *6.2.4 Idea coding of old and new ideas in the written output and process files*

In order to assess (i) whether ideas occurred during writing; and if so, to assess (ii) the exact location of the occurrence of ideas in the text, we have developed files in which we mapped the process data onto of the final product. This log therefore shows in one eye-side which text is introduced at a specific moment during the process, whether it is remained in the end product and whether the text is produced as a part of text production at the leading edge or during revisions. Therefore, for each burst of language production we know at what time during the process it was produced. This log of the combined process and product features enabled us to code when and where ideas were introduced and to analyse the relationships between ideas and the process of writing. The coding of ideas is done as described below.

#### *6.2.4.1 Idea coding in the text*

On the basis of a sample of 10 texts, we developed a detailed coding scheme that enabled us to code the occurrences of ideas produced in list 1 and 2 in the text. It was decided that ideas could either be expressed, or they could be 'implied' or expressed not in separate sentence parts, but rather as a set of sentences. For the coding of the ideas in the text is was decided not to allow the rater to interpret what a writer had possibly meant with an expression in the text, but rather to take the expression of the ideas on the lists as literally as possible. Therefore, ideas had to be expressed objectively and certain in order to be located. However, sometimes, allowing a bit more interpretation from the rater, expressions in the text could be related to ideas produced on the lists. To not miss out on these more implicitly expressed ideas, we have coded these ideas as 'implied', because the link between the expression in the text and the expression on the list was more implicitly established.

In order to minimalize the overlap of ideas on the lists produced before and after writing, the lists produced before and after writing were prepared prior to the coding of the ideas in the text. This procedure involved crossing out ideas that were identically expressed on both list, so that just one of these ideas had to be located in the text. When the expression of that idea was located in the text it was then

established that both the idea from list 1 as the idea from list 2 occurred in the text. Furthermore, it involved separating out different concepts or propositions within one idea, so that these could be located in the text separately from each other. An example of such an idea would be an idea like "Blogs, You Tube". It was established that both of these concepts could be located in the text separately from each other, but that the original idea was located in the text when one of these concepts was expressed. So, it wasn't necessary that both of these concepts of one idea occurred in the text in order to locate the idea in the text. Finally, in the cases that expressions of ideas in list 1 and 2 were similar to each other but one idea involved a more detailed description of the idea or included examples of the concepts expressed in the idea, it was decided that these more detailed descriptions could be located separately from the more general expression of the idea. This most frequently occurred when the original ideas in list 1 were just briefly mentioned in list 2. This way of preparing the lists in advance of the actual coding of the ideas in the text enabled us to more precisely locate the expressions of ideas in the text. The research assistant prepared all lists after an intensive training. The premier researcher checked all prepared lists before the start of the actual coding of the ideas in the start.

#### *6.2.4.2 Agreement about the coding of ideas in the text*

The final coding was done on the basis of 20 other randomly selected texts from the sample. This meant that the two judges rated 25% of all data independently from each other. Interrater agreement was established on the basis of (i) whether ideas occurred in the text (ii) and what the location of those ideas is. We coded the ideas in the files that combined the process and product data. We have chosen the sentences in those files as the units to assign ideas to. Just for clarification reasons it is stressed once more that the sentences in these files displayed both the text that was retained in the end product as well as the processes that were executed at that location but were not necessary retained in the final product such as insertions in the sentence or text which was deleted at the leading edge of that sentence.

Overall, the two independent judges reached 92.8% agreement about whether ideas occurred in the text regardless of the location of those ideas. Furthermore, the independent judges reached 86.2% agreement when the location of the ideas was also taken into account.

The trained research assistant then completed rating the rest of the sample. However, the premier researcher did check all the rated ideas and decided, in correspondence with the second rater, which ideas occurred in the text and what the location of the ideas was.

### *6.2.5 Analysis of ideas coded in the text*

This procedure of coding the ideas in the text allows us to investigate whether the ideas produced on the lists before and after writing were used and or introduced during the process of writing or whether they are just produced before or after writing without being expressed in the text. Furthermore, it allows us to examine the distribution of old and new ideas during the process of writing. For all written output in the process we have coded whether the burst of language production executed at a certain moment in time contains new ideas, old ideas, or whether no ideas were expressed in that specific burst of language production. To investigate how both old and new ideas are produced during the process of writing and whether the distribution of old and new ideas in the text varies depending on the different writing conditions and writer types we have turned to generalized linear models and we will analyse our date with R (Baayen, 2008).

The analysis of our data involves binary dependent variables. To make this more concrete in our first analysis, that investigates whether ideas are introduced during the process of writing, we will use a variable labelled "IdeaInText". This variable contains zeros for all bursts that do not contain ideas produced on the lists and this contains ones for bursts that do contain ideas that were produced on one of the lists. In later analyses, where we will assess the distribution of old and new ideas in the text, we have two binary dependent variables that are coded as 1 (yes old/ new idea) or 0 (no old or new idea). Therefore, the analyses of these data require logistic regression. Logistic regression does not model the dependent variable directly, but instead attempts to model the probability of a given output ("yes old/ new idea") given the predictors. This probability is estimated in terms of the logits, the natural logarithm of the odds of observing a certain value (in our case, for the first analysis, the probability that text produced contains an idea) (Baayen, 2008; Wieling, 2012).

Generalized logistic mixed effects regression models are derived from linear regression models. Linear regression models are statistical models in which a univariate response is modelled as the sum of a 'linear predictor' and a zero mean random error term. The same counts for multiple linear regression models in which more than one predictor can be included. For present purposes the idea behind this analysis is that the linear predictor depends on the predictor variable and the unknown parameter that needs to be estimated. A key feature of these models is that the linear predictor depends linearly on the parameter. In logistic regression models, the probability of occurrence of a certain outcome value is predicted on the basis of the known values of the predictor variable. Generalized linear models (GLM) apply a less strict assumption of linearity assumption, by allowing the expected value of the

response to depend on a smooth monotonic function on the linear predictor. Furthermore, these models allow response variables from the exponential family (for instance binomial). This is requested for our analysis since we have shown above that our data has binary dependent variables.

#### *6.2.5.1 Relative burst order*

To analyse the process by which old and new ideas are produced during writing we have taken the process burst order of the participants as a sequence measurement. The occurrences of ideas are coded on the basis of sentences in the combined process and product documents described above. These sentences are built up of different production bursts during the process of writing. The later on in the process an idea is introduced, the higher the number of the burst. In this context, it should be noted that each participant has a different total number of process bursts, the maximum being 209 for one participant and the minimum being 62 for another participant. This makes it difficult to compare the moment in time during which an ideas is introduced across participants. Burst 30 of the participant with a total of 62 bursts is happening approximately half way through the process whereas for the participant with 209 bursts, burst 30 is relatively early on during the process. To compare these bursts orders across participants we, therefore, rescaled their bursts sequence as a proportion of their total number of bursts ranging from 0 to 1 where 1 is reflecting their total number of bursts produced. This was done by dividing the burst order data by the maximum number of bursts observed for each individual participant. This allowed us to investigate whether including a time measure would explain the variance in the data better than without taking time into account.

Another thing to note about the coding of ideas on the basis of bursts nested in sentences is that a sentence in the written output can contain both new and old ideas. These sentences are built up of different bursts and hence, bursts can also contain both new and old ideas. This means that when a new idea occurs in the text it does not necessarily mean that no old idea can be produced there. The distribution of old and new ideas during the process of writing is therefore conceptually independent of each other. Locations that contain old ideas can also contain new idea.

#### *6.2.5.2 Assessing non-linear trends in the data*

In order to analysis non-linear trends in our data we have turned to two different functions in R (Baayen, 2008). A common way to test non-linearity is to apply growth models to the data. These growth models assess which trend of the rate of change over time best describes the data (Field, 2009). A disadvantage of growth curve fitting is that they impose a very specific functional form onto a curve (Baayen, 2008). To

estimate the distribution of old and new idea in the text we have used the *rcs*-function in R which refers to *Restricted Cubic Splines* (Harrell, 2001). This function allows more flexible non-linearity estimation. The name refers to splines in the context of construction where these flexible strips of metal or rubber are used for the drawing of curved parts of objects. For the statistical purposes of this paper, spline is a function that is used for the modelling of non-linear relationships. The function constructs a curve from series of sections of cubic polynomial joined together so that the curve is continuous. The points at which these sections are joined together are known as the knots of the spline (Wood, 2006; Baayen, 2008). The number of knots determines the number of sections and hence the number of knots determines how smooth the curve will be. The more knots are added to the model, the more subtle the non-linearity that the models allows and the wigglier the curve can be. The minimal number of knots is three, hence two cubic spline sections. Restricted cubic splines are cubic splines that are adjusted to avoid over fitting for the more extreme values of the predictor. In our model we will allow 5 knots, and hence four sections will be estimated. We have analysed these with the *rms-package* of R (Harrell, 2001).

In order to analyse the relationships between the distribution of old and new ideas over time and the development of understanding through writing we have turned to a generalized additive mixed-effect regression model (GAM). A general additive model (GAM) provides flexible tools for modelling complex interactions describing wiggly surfaces (Wieling, 2012). This can be seen as a general linear model in which a part of the linear predictor is specified in terms of a sum of smooth functions of predictor variables (Wood, 2006; Baayen, 2008). A generalized additive model consists of two parts, a parametric part which is basically identical to that of standard linear models, and a non-parametric part that provides non-parametric functions for modelling wiggly surfaces. In this study, we have used a tensor product to model thin plate regression splines, since our predictor variable 'relative burst order' is a continuous variable. When fitting a tensor smoother to the data, it is important to avoid under smoothing or over smoothing of the data. We have used the default of the GAM implementation of the *mgcv-package* of Wood (2006) (version 1.7-18) to estimates the optimal smoothness from the data using generalized cross validation. The significance of the regression spline is further assessed with a Chi-square test evaluating whether the invested estimate degrees of freedom entail an improved fit to the model. When the estimated degrees of freedom (edf) are close to 1, the line can basically be interpreted as a linear relationship, whereas more estimate degrees of freedom indicate more bends in the constructed curves (Wood, 2006; Baayen 2010; Tremblay & Baayen, 2010).

**Table 6-1** Different levels included in our generalized additive regression analysis

| Level | Knowledge change | Idea type | Writing condition | N |
|---|---|---|---|---|
| NewCond16 1-New-HO | KC | New | High self-monitor outline planning | 5 |
| NewCond16 1-New-HS | KC | New | High self-monitor synthetic planning | 6 |
| NewCond16 1-New-LO | KC | New | Low self-monitor outline planning | 3 |
| NewCond16 1-New-LS | KC | New | Low self-monitor synthetic planning | 8 |
| NewCond16 1-Old-HO | KC | Old | High self-monitor outline planning | 5 |
| NewCond16 1-Old-HS | KC | Old | High self-monitor synthetic planning | 6 |
| NewCond16 1-Old-LO | KC | Old | Low self-monitor outline planning | 3 |
| NewCond16 1-Old-LS | KC | Old | Low self-monitor synthetic planning | 8 |
| NewCond16 0-New-H0 | No KC | New | High self-monitor outline planning | 15 |
| NewCond16 0-New-HS | No KC | New | High self-monitor synthetic planning | 13 |
| NewCond16 0-New-LO | No KC | New | Low self-monitor outline planning | 16 |
| NewCond16 0-New-LS | No KC | New | Low self-monitor synthetic planning | 13 |
| NewCond16 0-Old-HO | No KC | Old | High self-monitor outline planning | 15 |
| NewCond16 0-Old-HS | No KC | Old | High self-monitor synthetic planning | 13 |
| NewCond16 0-Old-LO | No KC | Old | Low self-monitor outline planning | 16 |
| NewCond16 0-Old-LS | No KC | Old | Low self-monitor synthetic planning | 13 |

In order to fit a generalized additive mixed effect regression model we have prepared the data in such a way that each writing group that we included in the model represented a context for which we were interested in the pattern of ideas over time. For this analysis we have four writing conditions (HO, LO, HS and LS), we distinguished between old and new ideas and we want to fit these distributions for participants who reported development of understanding/ knowledge change (KC) and participants who did report knowledge change (No KC). Therefore, we wanted to fit thin plate regression splines for 16 different sub groups. We created a new variable

(NewCond16) that had each of these sub groups as a level. The different levels are shown below in table 6-1.

## 6.3 RESULTS AND PRELIMINARY DISCUSSION

First of all, we set out to test the validity of the idea coding in the text by comparing the number of new and old ideas in the list produced after writing with the number of new and old ideas that were located in the text. Furthermore, we set out to test whether there were effects of self-monitoring and type of planning on the extent to which new and old ideas were included in the text.

### *6.3.1 Number of new and old ideas produced on the lists and coded in the text*

A repeated measures ANOVA with a Greenhouse-Geisser correction determined that there was a significant difference in the number of new ideas in the list produced after writing and in the text ($F$ (1, 75) = 89.57, $p$< .001). Overall, significantly more new ideas were produced in list 2 ($M$=6.22, $SD$=3.76) than there were identified in the text ($M$=3.54, $SD$=2.62). However, this did not vary as a function of self-monitoring or type of planning, suggesting that this occurred to the same extent across the conditions, and that the conditions did not vary in the number of new ideas produced in the text.

A repeated measures ANOVA with a Greenhouse-Geisser correction showed that there was a significant difference in the number of old ideas in the lists produced before and after writing and in the text ($F$(1, 75)=93.77, $p$<.001). Overall, significantly more old ideas were produced in the lists ($M$= 8.34, $SD$=3.65) than were coded in the text ($M$=6.03, $SD$=3.27). Again, this difference did not vary as a function of self-monitoring and type of planning.

In general, then, significantly fewer old and new ideas were identified in the text than appeared in the lists produced before and after writing. This may in part be an indication of the fact that our coding system was rather conservative, since it focussed on the explicit expression of ideas in the text. This may mean that we only coded a subset of the most explicit ideas from the lists in the texts. It could, however, also indicate that writers add new ideas to their lists after writing the text itself, or reproduce old ideas in the list that they haven't actually written about. To check whether this was associated with differences between writers we correlated the number of new and old ideas that were coded in the text with the number of new and old ideas listed. This showed strong positive correlations for both new ($r$=.77, $p$<.001) and old ideas ($r$=.83, $p$<.001). This suggests that, overall, the number of new and old ideas that the participants listed before and after writing correlated reasonably well

with the number that was identified in their texts, and that the coding system had been applied in an unbiased way across the conditions.

### *6.3.2 The extent to which ideas from the lists were written about*

The first analysis will focus on whether text produced during the process contains ideas (yes or no), regardless of their being produced before, during or after writing, and examine whether this differs across the different writing conditions.

In order to investigate this we compared a simple baseline model which just included participants as a random factor and the occurrence of ideas as a fixed factor with a more specific model that included participants as a random factor and the four writing conditions as a fixed factor. To analyse these data we created the variable NewCond which consisted of the four different writing conditions. NewCond**HO** refers to **H**igh self-monitors **O**utline planning condition (and is included as the intercept, and reference category, in the regression table below), NewCond**HS** refers to **H**igh self-monitors **S**ynthetic planning, NewCond**LO** refers to **L**ow self-monitors **O**utline planning and finally, NewCond**LS** refers to **L**ow self-monitors **S**ynthetic planning. These groups will be used for all the remaining analyses. This analysis showed that the more specific model that included the main effects of the four writing conditions was a better fitting model ($\chi$2 (3) = 8.0869, $p$<.044) than the baseline model. The full model is shown in table 6-2.

**Table 6-2** Generalized linear mixed effects model for occurrence of ideas during the process of writing

| **Random effects:** | **Variance** | | |
|---|---|---|---|
| Participant (Intercept) | 0.22 | | |
| **Fixed effects:** | **Estimate** | **Std. Error** | ***z* value** |
| (Intercept) | 0.66 | 0.11 | 5.96*** |
| NewCondHS | 0.15 | 0.16 | 0.92 |
| NewCondLO | 0.10 | 0.16 | 0.65 |
| NewCondLS | 0.44 | 0.16 | 2.78** |
| *Formula*: IdeaInText ~ NewCond + (1 \| Participant) | | | |

***$p$<.001, **$p$<.01, *$p$<.05

As can be seen in table 6-2, the intercept for the fixed effect component is significantly different from zero ($\beta$ =0.66, $z$= 5.96 $p$<.001). This indicates that during the process of writing more content is associated with ideas that were contained on the lists produced before or after writing than with ideas that were not referred to on the list.

Furthermore, since all the coefficients are positive, this holds for all the different writing conditions. Note, however, that the low self-monitors synthetic planned texts contained significantly more text corresponding to ideas from the lists produced before and after writing than the high self-monitors' outline condition ($\beta$ =.0.44, $z$= 2.78. $p$<.001). By contrast, there were no differences for the other two conditions. If we translate these estimates (which are expressed as logits) back to probabilities, this reveals that, for low self-monitors synthetically planned texts, about 75% of their text is likely to be associated with the content of the ideas produced on the lists. For the other three conditions, the probability of occurrences of ideas on the list is a bit less and varies between 66% and 69%.

Overall, then, this analysis shows that for all writing conditions a reasonable amount of the text produced is captured by the ideas produced on the lists. Furthermore, it shows that the ideas produced by the low self-monitors in synthetic planning capture a higher percentage of the text than the other writing conditions.

### *6.3.3 The distribution of old and new ideas during the writing process as a function of self-monitoring and type of planning*

The next step in the analysis was to investigate when both old and new ideas are produced during the process of writing and whether the distribution of old and new ideas in the text varies depending on the different writing conditions.

#### *6.3.3.1 Old ideas*

In order to analyse the point at which old ideas were produced during writing we compared four different generalized logistic regression models. The first model included the linear trend of the relative burst order and the main effects for the writing conditions as fixed factors. We compared this with a model that included interactions between the linear trend of the relative burst order and the writing conditions. In both these models participants were included as a random factor. Comparison of these models revealed that the model with the interaction terms explained the data better ($\chi$2 (3) = 56.64, $p$<.001) than the baseline model. The next step in the analysis was to allow non-linear relationships between the relative burst order and the old ideas. In this model we therefore included the non-linear terms (allowing 4 sections) and the main effects of writing conditions as fixed factors, and the participants as random factors. We again compared this model with the baseline model. Model comparison revealed that the more complex model which allowed non-linear relationships fitted the data better ($\chi$2 (3) = 156.86, $p$<.001). Finally, we compared the latter model with a model that included the non-linear terms, the main

effects and the interaction terms of relative burst order and writing conditions as fixed factors and participants as random factor. This proved to be the best fitting model ($\chi$2 (12) = 118.82, $p$<.001, see table 6-3).

**Table 6-3** Generalized linear mixed effect model for old ideas as a function of processing time and writing condition

| **Random effects:** | **Variance** | | |
|---|---|---|---|
| Participant (Intercept) | 0.26 | | |
| **Fixed effects:** | **Estimate** | **Std. Error** | ***z*-value** |
| (Intercept) | -1.04 | 0.20 | -5.14*** |
| NewCondHS | 0.93 | 0.30 | 3.18** |
| NewCondLO | 0.61 | 0.28 | 2.18* |
| NewCondLS | 2.29 | 0.30 | 7.65*** |
| Main effect non-linear function S1 | 4.89 | 0.98 | 4.98*** |
| Main effect non-linear function S2 | -10.66 | 5.42 | -1.97* |
| Main effect non-linear function S3 | 22.61 | 16.99 | 1.33 |
| Main effect non-linear function S4 | -26.88 | 23.15 | -1.16 |
| S1*NewCondHS | -3.28 | 1.45 | -2.26* |
| S2*NewCondHS | 10.78 | 8.04 | 1.34 |
| S3*NewCondHS | -29.47 | 25.23 | -1.17 |
| S4*NewCondHS | 36.11 | 34.41 | 1.05 |
| S1*NewCondLO | 0.10 | 1.38 | 0.07 |
| S2*NewCondLO | -11.11 | 7.74 | -1.44 |
| S3*NewCondLO | 37.85 | 24.36 | 1.55 |
| S4*NewCondLO | -46.07 | 33.34 | -1.38 |
| S1*NewCondLS | -8.60 | 1.47 | -5.86*** |
| S2*NewCondLS | 40.24 | 8.12 | 4.95*** |
| S3*NewCondLS | -128.51 | 25.44 | -5.05*** |
| S4*NewCondLS | 178.90 | 34.46 | 5.19*** |

*Formula*: HasOldIdea ~ rcs(Burstorder.rel, 5) * NewCond + (1 | Participant)

***$p$<.001, **$p$<.01, *$p$<.05

The full-interaction model for this best fitting model is shown in table 6-3. For ease of interpretation, we have changed the names of the *rcs*-function which refer to different parts of the cubic spline to the equivalent sections that they refer to. S1 (section 1) refers to the first part of the cubic spline, section 2 refers to the second part of the cubic spline etc. The first four rows of the table show the main effects of the writing conditions, and the next four rows show the main effects of the non-linear function. Note that the reference category for this analysis is again the high self-monitor outline

planning condition (NewCondHO). This means that the main effects of the non-linear function therefore also need to be interpreted as the probability that occurrence of old ideas over time are significantly non-linear for the reference condition and this is estimated for 4 different parts of the cubic spline, since we have allowed 5 knots. The interaction terms in the table then indicate how the different writing conditions deviate from this non-linear trend of the reference condition.

In figure 6-1, we have plotted the fixed effects of the above interaction model. On the basis of this figure, we will describe the results.

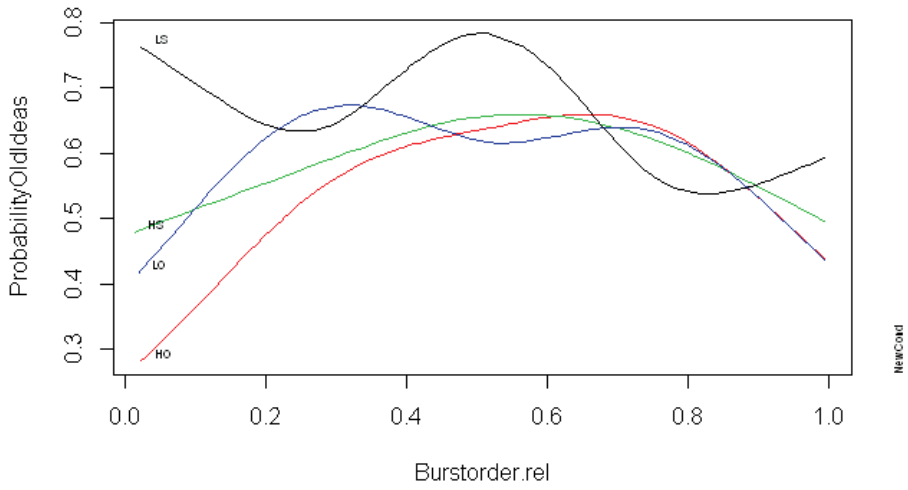


**Figure 6-1** Probability of occurrence of old ideas over time (relative burst order) plotted for the four different writing conditions (HO =red line, LO=blue line, HS=green line and LS = black line).

As can be seen in figure 6-1 the intercepts for the four different writing conditions are very different. At the beginning of the writing process, the probability of occurrences of old ideas is highest for the low self-monitors in the synthetic planning condition (NewCondLS, black line). This is significantly different ($\beta$ =2.29, $z$= 7.65, $p$<.001) from the probability of occurrence of old ideas for the high self-monitors in the outline planning condition (intercept/reference condition, red line). The high self-monitors in outline planning have the lowest probability of occurrences of old ideas at the beginning of the process. The low self-monitors in the outline planning condition (NewCondLO, blue line) ($\beta$ =0.61, $z$= 2.18 $p$<.05) and the high self-monitors in the synthetic planning condition (NewCondHS, green line) ($\beta$ =0.93, $z$= 3.18 $p$<.01) also show a significantly higher probability of producing old ideas at the beginning of the

writing process in comparison with the high self-monitors in the outline planning condition.

Overall, then, this component of the regression model suggests that the four conditions differ strongly in the likelihood that they begin with an old idea, with low self-monitors' synthetically planned texts having a 75% probability of beginning with an old idea, and high self-monitors' outline planned texts having about a 30% likelihood of doing so. Note, in addition, that the order of these differences corresponds very directly with the extent to which it is assumed that writing is a top-down or a bottom-up process. The most top-down condition (high self-monitors outline) is least likely to start with an old idea; the most bottom-up condition (low self-monitors' synthetic) is most likely to start with an old idea; with the two conditions which are assumed to involve a combination of top-down and bottom-up processing lying in between.

The next four terms in the model, correspond to the non-linear function over time for the high self-monitors' outline condition. These show that the distribution of old ideas is significantly different from linear at the beginning of the writing process (S1) ($\beta$ =4.89, $z$= 4.98, $p$<.001) for the high self-monitors in the outline planning condition. After this initial steep increase in the probability of old ideas occurring during the process of writing the line flattens out ($\beta$ =-10.66, $z$= -1.97, $p$<.05), and then remains flat, with no significant non-linearities.

The final set of 12 interaction terms in the model shows the extent to which the different writing conditions vary from this pattern. For the low self-monitors' outline condition and the high self-monitors' synthetically planned condition, where most of the coefficient are non-significant, the patterns are essentially the same as for the high self-monitors' outline condition. The one exception being the initial part of the curve for the high self-monitors' synthetically planned condition which does not rise as steeply at the beginning of the writing process as the high self-monitors in outline planning condition (a negative coefficient ($\beta$ =-3.27, $z$= -2.26 $p$<.05). Overall, although these conditions differ in how likely they are to start with an old idea, after about 30% of the bursts have been produced, they have a similar likelihood of producing an old idea (about 60%) and this remains constant for the reminder of the time.

By contrast, the low self-monitors' synthetically planned texts show a very different pattern, with all four of the coefficients differing significantly from the high self-monitors' outline planning condition. At the beginning of the writing process the low self-monitors in the synthetic planning condition are very likely to produce an old idea. The likelihood then decreases during the first part of the writing process ($\beta$ =-8.60, $z$=-5.86, $p$<.001), increases again during the second part of the writing process ($\beta$

=40.24, $z$= 4.95 $p$<.001), significantly decreases again ($\beta$ =-128.51, $z$= -5.05 $p$<.001) and finally, at the end of the writing process, increases again ($\beta$ =178.90, $z$= 5.19, $p$<.001).

Overall, then, this analysis shows a clear difference between the high self-monitors' outline condition (the condition assumed to be most top-down) and the low self-monitors synthetically planned condition (the condition assumed to be the most bottom-up). The low self-monitors' synthetically planned condition shows a distinctive oscillation in the likelihood of producing an old idea across time, starting with a high probability, but then varying regularly up and down across the writing process.

#### *6.3.3.2 New ideas*

We conducted the same stepwise comparison of different models for the new ideas as described above for the occurrence of old ideas. In all the models that we compared, participants were included as random factors. The model comparisons revealed that, in order to explain the probability of occurrence of new ideas over time, the interactions between the non-linear terms and the writing conditions needed to be included ($\chi$2 (12) = 54.42, $p$<.001). The full interaction model of this regression analysis is shown in table 6-4.

Table 6-4 can be interpreted in the same way as table for the regression model of the old ideas. So, first we will present the main effects of the writing conditions and the non-linear terms and then we will present the interaction terms. It should be noted that the high self-monitors in the outline planning condition are again the reference condition. Figure 6-2 shows the fixed effects of the interaction model.

First of all, it is important to note that the overall probability of occurrences of new ideas is much lower than the probability of occurrences of old ideas. That said, we will describe the results from the regression table on the basis of figure 6-2.

First, the intercept is significantly different from zero ($\beta$ =-1.60, $z$= -4.56 $p$<.001. The intercept in this model represents the high self-monitors in the outline planning condition. The intercepts of the other writing conditions do not significantly differ from the reference condition (HO) expect for the low self-monitors in the synthetic planning condition ($\beta$ =-1.24, $z$= -2.38 $p$<.05). This negative estimate reveals that low self-monitors in synthetic planning show a lower probability of starting writing process with a new idea.

**Table 6-4** Generalized linear mixed effect model for new ideas as a function of processing time and writing condition

| **Random effects:** | **Variance** | | |
|---|---|---|---|
| Participant (Intercept) | 1.64 | | |
| **Fixed effects:** | **Estimate** | **Std. Error** | ***z* value** |
| (Intercept) | -1.60 | 0.35 | -4.56*** |
| NewCondHS | -0.54 | 0.52 | -1.04 |
| NewCondLO | 0.10 | 0.49 | 0.21 |
| NewCondLS | -1.24 | 0.52 | -2.38* |
| Main effect non-linear function S1 | -2.48 | 1.24 | -2.00* |
| Main effect non-linear function S2 | 7.42 | 7.03 | 1.06 |
| Main effect non-linear function S3 | -7.13 | 22.05 | -0.32 |
| Main effect non-linear function S4 | -19.23 | 29.86 | -0.64 |
| S1*NewCondHS | 3.36 | 1.92 | 1.75 |
| S2* NewCondHS | 16.07 | 10.69 | -1.50 |
| S3*NewCondHS | 42.89 | 33.39 | 1.29 |
| S4*NewCondHS | -28.53 | 44.75 | -0.64 |
| S1*NewCondLO | 1.30 | 1.75 | 0.74 |
| S2 *NewCondLO | -7.28 | 9.96 | -0.73 |
| S3*NewCondLO | 17.02 | 31.38 | 0.54 |
| S4*NewCondLO | -4.89 | 42.62 | -0.12 |
| S1*NewCondLS | 8.48 | 1.88 | 4.51*** |
| S2*NewCondLS | -36.30 | 10.30 | -3.52*** |
| S3*NewCondLS | 91.80 | 32.08 | 2.86** |
| S4*NewCondLS | -76.09 | 43.04 | -1.77 |

*Formula*: HasNewIdea ~ rcs(Burstorder.rel, 5) * NewCond + (1 | Participant)

***$p$<.001, **$p$<.01, *$p$<.05

The results for the non-linear function over time (S1-S4) show that, at the beginning of the writing process, there is a marginally significant decrease in the likelihood of the high self-monitors in the outline planning condition producing a new idea ($\beta$ =-2.48, $z$= -2.00 $p$<.05). Overall, though, this is essentially a linear function. Since none of the coefficients for the interaction terms for low self-monitors' outline planned texts and high self-monitors' synthetically planned texts are significant, these show essentially the same pattern as the high self-monitors' outline planned texts.

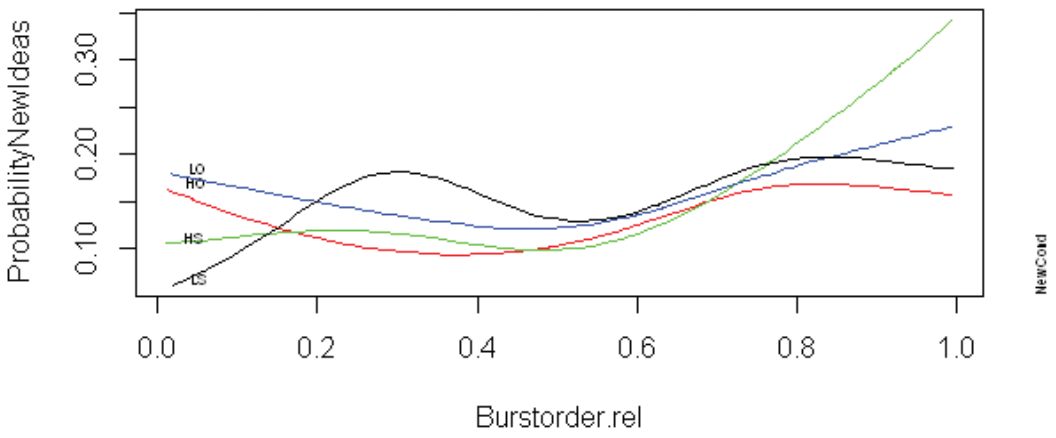


**Figure 6-2** Probability of occurrence of new ideas over time plotted for the four different writing conditions (HO =red line, LO=blue line, HS=green line and LS = black line).

The exception is, again, the low self-monitors in the synthetic planning condition. Their distribution of new ideas over time is more wiggly, significantly so for the first part ($\beta$ =8.48, $z$= 4.51 $p$<.001) the second part ($\beta$ =-36.30, $z$= -3.52 $p$<.001) and the third part of the writing process ($\beta$ =-91.80, $z$= -2.86 $p$<.001), but not for the fourth part. These writers start the writing process with a lower probability of producing a new idea. The probability then rises to a peak at about 30% of the way into the writing process and then oscillates regularly.

These results mirror those found for the old ideas. The low self-monitors in the synthetically planned condition are less likely to produce a new idea at the start of writing than the other groups, and then show a distinctively different, oscillating pattern. This is suggesting that as the probability of an old idea decreases the probability of a new idea increases, and vice versa. It is important to note here that the old and new ideas are independently measured, and could in principle co-occur in the same burst. Furthermore, there are also bursts that are not associated with the occurrence of ideas on the list. This means that at different points during writing, writers may not be writing about either an old or a new idea, but instead about content that does not appear in the lists produced before or after writing. At the beginning of writing in particular, where, apart from the low self-monitors' synthetically planned condition, the groups have a probability less than 50% of producing an old idea (only about 30% for the high self-monitors' outline planned condition) and less than 20 % of producing a new idea, it appears that there is a relatively high probability that writers will produce content that is not included in the lists of ideas. This presumably reflects general introductory content.

Overall, then, this analysis suggest a clear difference between the low self-monitors' synthetically planned writing and the other conditions. This difference is compatible with the contrast between a top-down and bottom-up process. For the high self-monitors' outline planned texts, in which an initial introduction is created, and there is a low probability of an old idea and a higher probability of a new idea at the start of writing, the process appears to reflect explicit organisation prior to writing the text. For the low self-monitors' synthetically planned text, in which there is a high probability of starting with a pre-existing idea, and in which old and new ideas oscillate across the course of text production, the process appears to reflect a gradual change from old to new and back again. This is compatible with new ideas emerging in the course of text production.

### *6.3.4 Relationships between patterns of ideas and development of understanding through writing*

The final analysis investigated how these patterns of ideas were related to increases in understanding. In order to investigate this we fitted a generalized additive mixed effect regression model (GAM) to the data. In order to fit such a model we prepared the data as described in the method section.

For this analysis, we had four writing conditions (HO, LO, HS and LS), and we distinguished between old and new ideas, and between participants who reported increases in knowledge/ understanding (KC) and who did not report increases in knowledge/ understanding (No KC). Therefore, we wanted to fit the thin plate regression splines for 16 different groups. However, since the previous analyses indicated that the main difference was between the low self-monitors in the synthetic planning condition and the other writing conditions, we started off with model comparisons between a model that only distinguished between the LS condition (Low self-monitors, synthetic planning condition) and the Non–LS conditions (HO, HS and LO conditions). When old and new ideas, and participants who improved their understanding or did not improve their understanding, were included, this resulted in 8 different writing groups. To test whether this simpler model with 8 groups explained the data better than the more complex model with 16 groups we used an analysis of deviance. The results of this comparison are shown below in table 6-5.

**Table 6-5** Analysis of deviance for model comparison

| | Resid. Df | Resid. Dev | Df | Deviance |
|---|---|---|---|---|
| 1 | 21713 | 24252 | | |
| 2 | 21684 | 24038 | 28.78 | 213.3*** |

*Formula Model 1*: HasIdea ~ te(Burstorder.rel, by = NewCond8, bs = "tp") + NewCond8 + s(Participant, bs = "re")

*Formula Model 2*: HasIdea ~ te(Burstorder.rel, by = NewCond16, bs = "tp") + NewCond16 + s(Participant, bs = "re")

****p*<.001, ***p*<.01, **p*<.05*

This generalized likelihood test to compare the two models indicates that model 1 (the simpler model with 8 writing groups) should be rejected in favour of model 2 (the more complex model with 16 groups). Overall, model 1 explained 17.3% of the variance whereas model 2 explained 18.1% of the variance.

**Table 6-6** Generalized additive mixed-effect model

| **Parametric coefficients**: | **Estimate** | **Std. Error** | ***z* value** |
|---|---|---|---|
| Intercept) | -1.46 | 0.14 | -10.85*** |
| NewCond16 0-New-HS | -0.01 | 0.20 | -0.06 |
| NewCond16 0-New-LO | -0.15 | 0.17 | -0.81 |
| NewCond16 0-New-LS | 0.17 | 0.20 | 0.85 |
| NewCond16 0-Old-HO | 1.67 | 0.07 | 23.48*** |
| NewCond16 0-Old-HS | 1.20 | 0.19 | 10.36*** |
| NewCond16 0-Old-LO | 1.76 | 0.18 | 9.63*** |
| NewCond16 0-Old-LS | 2.04 | 0.20 | 10.41*** |
| NewCond16 1-New-HO | -0.45 | 0.26 | -1.72 |
| NewCond16 1-New-HS | 0.20 | 0.25 | 0.79 |
| NewCond16 1-New-LO | 0.42 | 0.29 | 1.44 |
| NewCond16 1-New-LS | 0.05 | 0.23 | 0.23 |
| NewCond16 1-Old-HO | 1.66 | 0.25 | 6.66*** |
| NewCond16 1-Old-HS | 1.50 | 0.25 | 6.04*** |
| NewCond16 1-Old-LO | 2.34 | 0.29 | 8.01*** |
| NewCond16 1-Old-LS | 2.34 | 0.23 | 10.31*** |

HasIdea ~ te(Burstorder.rel, by = NewCond16, bs = "tp") + NewCond16 + s(Participant, bs = "re")

****p*<.001, ***p*<.01, **p*<.05

We therefore created a generalized additive mixed-effect regression model capturing the 16 different writing groups. This varied type of idea (old and new), development of understanding (KC =1 versus No KC =0) and the different writing conditions (HO, HS, LO, and LS). The full regression model of this analysis is shown in table 6-6 and 6-7. Note that the 'No Knowledge Change (0), New ideas (New), High self-monitors in the Outline planning condition (HO)' is taken as the reference condition (NewCond 0-New-HO) in both tables.

The parametric part of the model shows the intercepts of the different groups. This shows that the intercept (0-New-HO) is significantly different from zero ($\beta$ =-1.46, $z$= -10.85 $p$<.001). The other coefficients in this table indicate whether the intercepts for the other writing groups are significantly different from the reference condition. As can be seen in table 6-6, some of the intercepts from different writing groups differ from the reference condition and others do not. Overall, it should be noted that all intercepts related to the production of old ideas are significantly different from the reference condition and furthermore, that these intercepts are higher. This basically represents the fact that the probability of occurrences of old ideas is more likely than the occurrences of new ideas. This is expected and also confirms our previous finding that overall probability of occurrences of new ideas is much lower than the probability of occurrences of old ideas.

Table 6-7 then presents the non-parametric part of the model which shows the significance of the smooth terms for the patterns of idea distributions. These patterns, that show how the distributions for old and new ideas vary over time for the 16 distinguished writing groups, are presented in figure 6-3. It should be noted that in these plots the intercepts are not taken into account. As said, the intercepts of the parametric model presented in table 6-6 report the likelihood of ideas occurring in a specific condition. If these intercepts would have been taken into account in the plots, the dotted lines in the figure would not be centred at zero by at the intercept level. These intercepts can be calculated by adding up the intercept of the reference condition and the intercept of the writing group of interest.

**Table 6-7** Approximate significance of smooth terms of the generative mixed effect regression model

| Approximate significance of smooth terms | edf | Ref.df | Chi Square |
|---|---|---|---|
| Burstorder: 0-New-HO | 2.76 | 3.27 | 17.99*** |
| Burstorder: 0-New-HS | 2.43 | 2.94 | 63.40*** |
| Burstorder: 0-New-LO | 1.01 | 1.03 | 8.56** |
| Burstorder: 0-New-LS | 2.95 | 3.45 | 30.37*** |
| Burstorder: 0-Old-HO | 3.34 | 3.76 | 52.72*** |
| Burstorder: 0-Old-HS | 2.87 | 3.38 | 33.31*** |
| Burstorder: 0-Old-LO | 3.58 | 3.90 | 34.33*** |
| Burstorder: 0-Old-LS | 3.20 | 3.66 | 12.2* |
| Burstorder: 1-New-HO | 3.59 | 3.90 | 70.97*** |
| Burstorder: 1-New-HS | 3.19 | 3.66 | 27.65*** |
| Burstorder: 1-New-LO | 3.19 | 3.65 | 39.70*** |
| Burstorder: 1-New-LS | 3.76 | 3.96 | 15.73** |
| Burstorder: 1-Old-HO | 3.36 | 3.77 | 66.58*** |
| Burstorder: 1-Old-HS | 3.08 | 3.57 | 64.64*** |
| Burstorder: 1-Old-LO | 3.66 | 3.93 | 40.70*** |
| Burstorder: 1-Old-LS | 3.91 | 3.96 | 78.93*** |
| s(Participant) | 68.41 | 75.00 | 825.84*** |

***$p$<.001, **$p$<.01, *$p$<.05

Taken that these intercepts are different for the different groups we were then interested in the patterns of idea distributions over time. These patterns do not change when the intercepts are taken into account. If the intercepts are taken into account the whole pattern would only move up or down in relation with the intercept estimate. The patterns of idea distributions over time are presented in figure 6-3. The non-parametric part of the model shows whether these patterns are significantly different from zero. As can be seen in table 6-7, the plotted smooth terms are significant for all the plotted patterns. GAM's basically measures whether the invested degrees of freedom (edf, which roughly represents the number of curves that are needed to fit each pattern; 1 being a linear line and more edf indicating more wiggly curves) is allowed. This means that all the patterns plotted are significantly different from 0 and furthermore, that the invested degrees of freedom significantly explain this pattern of idea distribution over time.

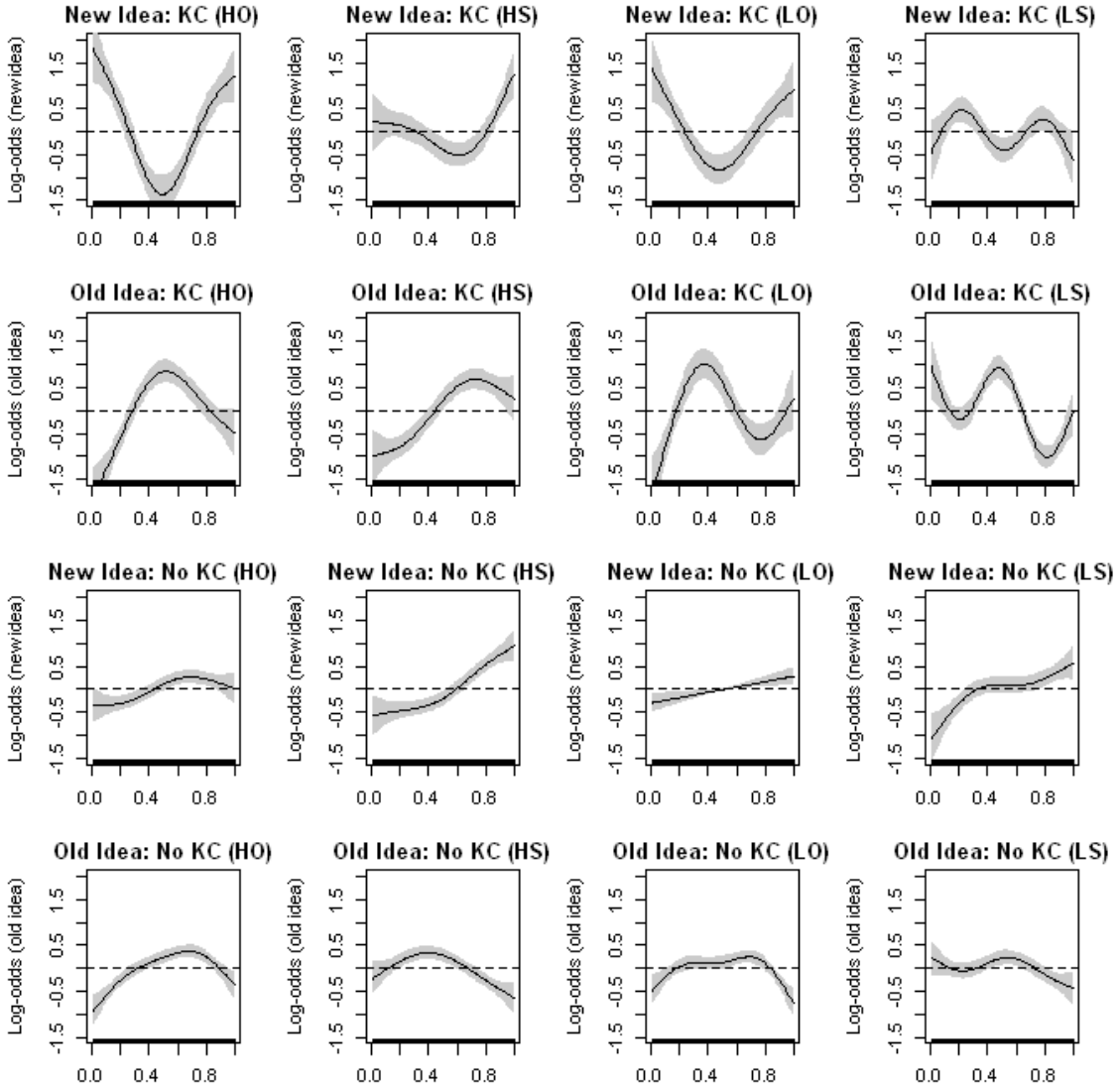


**Figure 6-3** Log odds for the distribution of new and old ideas across time (x-axis relative burst order) plotted separately for participants who increase in understanding (first two rows KC) and who do not increase their understanding (last two rows, No KC) in the different writing conditions (HO, HS, LO and LS).

Then, in figure 6-3 the estimated effects of the relative burst order on the distributions of old and new ideas for participants who showed knowledge change (KC) and did not show knowledge change (No KC) are plotted separately for the different writing conditions (HO, LO, HS and LS). In these plots, the log probability of old /new ideas occurring at a certain moment in time is presented by the thin spline plate. The log odds represent the log probability of having a new idea divided by the log probability not to have a new idea (Field, 2009). The grey shades show the 95% confidence bands for the regression lines. For the 'New Ideas-No KC-HO condition, the intervals are so

small that the area coincides with the data points (Baayen, 2008). The dotted line in the middle of the figures indicates that above the line there is more than 50% change that an idea occurred, whereas below the line there is less than 50% change that an idea occurred.

The four plots on the top row of the figure show the non-linear patterns of log odds of new ideas occurring over time for participants who increase in understanding (New idea: KC) in the four different writing conditions (HO, HS, LO and LS). The second row shows the log odds of old ideas occurring over time for participants who increase in understanding (Old idea: KC) in the four different writing conditions (HO, HS, LO and LS). The final two rows show the patterns of the log odds of new (third row) and old ideas (fourth) occurring over time for the participants who do not increase in understanding (New Idea: No KC; Old Idea: No KC) for the different writing conditions (HO, HS, LO and LS). The interpretation of these plots will be given below.

A first general look at the figure immediately shows that the patterns of the lines in the top two rows (corresponding to new and old ideas for participants who reported development of understanding) are very different from the bottom two rows (participants who did not report development of understanding). In what follows, we will describe the results for participants who did not report development of understanding separately from those whose understanding increased.

#### *6.3.4.1 Participants who did not report development of understanding*

In general, the trends for the participants who did not experience development of understanding through writing seems to suggest that there is a low log probability that they start the writing process with either new or old ideas. This may, as we mentioned earlier, reflect the fact that writers generally, at least in these conditions, where there were no increases in understanding, tend to start with general introductory material. This is similar across all the conditions, with the exception, perhaps, of the old ideas in the synthetic conditions, which may be slightly higher. But given the width of the confidence intervals, it is safest to conclude that there is little evidence of a difference as a function of self-monitoring and type of planning.

For the new ideas (in the third row), there is some evidence that the probability of new ideas being produced is higher in all conditions towards the end of the writing process than at the beginning, with the possible exception of the high self-monitors' outline condition, where the confidence intervals overlap. This may reflect a tendency to introduce a new idea as an overall conclusion to the text. If it does, then the fact that these participants did not experience increases in understanding suggests that this is not associated with a change in understanding. It is noticeable also that there is a

corresponding drop in the likelihood of producing an old idea at the end of writing, and in this case the confidence intervals clearly do not overlap with the middle portions of the curves. This is compatible with the inference that this final section is concerned with drawing an overall conclusion rather than with producing ideas related to the onward progression of the argument.

The general pattern for this data is of a lack of variation across the writing process, and the main interest of what variation there is, is as a baseline against which to compare the patterns for participants whose understanding increased.

#### *6.3.4.2 Participants who reported development of understanding*

The first feature of these data that we will consider is the pattern in the outline conditions (first column = high self-monitors; third column = low self-monitors; row 1 = new ideas; row 2 = old ideas). As can be seen in the figure, these are very similar for the low and high self-monitors, suggesting that this is a general effect of outlining, and are markedly different to the patterns for participants who did not increase their understanding. Both low and high self-monitors show a high probability of producing a new idea at the beginning and end of writing – much higher than when understanding did not increase – and a correspondingly very low probability of producing a new idea in the middle section of writing. Similarly, for the old ideas, both groups show a low probability of producing an old idea at the beginning of writing, followed by a pronounced increase in probability during the middle section. The one area where there is less evidence of a difference is at the end of writing, where there is little difference between participants whose understanding increased and those whose understanding did not increase. The low self-monitors also show some evidence of non-linearity here, but since the confidence interval is very wide, it is questionable whether this is of any significance.

Overall these data suggest a clear difference between the participants whose understanding increased and those whose understanding did not increase. We draw two main conclusions from this. First, the pattern is consistent with the change in understanding being a consequence of top-down processing, involving explicit organisation during pre-planning. The fact that the new ideas have a high probability at the beginning of writing suggests that they have not been produced as a consequence of text production, but rather as a consequence of the writer formulating an overall goal for the text during pre-planning. Similarly, the fact that new ideas also have a higher probability at the end of writing is compatible with the formulation of a new overall conclusion for the text at the end of writing. It appears, however, since the main body of the text consists of old rather than new ideas, that this is a matter of creating a new global structure for the writer's existing ideas rather than a widespread

change in ideas across the writing process. The writers' old ideas appear to have been contextualised within a new overall framework. Furthermore, the clear distinction between different phases of writing suggests that the process is more controlled than when understanding does not increase: the occurrence of new and old ideas is highly structured. The second conclusion is that, although this is compatible with the dual process model's claim that outline planning is associated with an explicit organising process, it contradicts the claim that this is not associated with increased understanding.

The second feature of these data is the contrasting patterns for the synthetic planning conditions. This is most pronounced for the low self-monitors (the fourth column, top two rows). The first distinctive feature is at the start of the writing process, where the low self-monitors show a higher probability of starting with an old idea, and a correspondingly lower probability of starting with a new idea, than when their understanding did not increase, and then all the other conditions where understanding increased. This suggests less explicit organisation prior to writing, and a greater tendency to start by formulating a pre-existing idea in text. The second distinctive feature is that these "wiggly" interactions between old and new ideas as text production proceeds look as if the new ideas are emerging during the process of writing, in a way that is consistent with the dual process model's characterisation of the knowledge-constituting process. The writer starts with an existing idea and then, as successive utterances are produced, finds a new idea gradually emerging. Once this idea has been formulated, the writer then appears to retrieve a further existing idea from their plan, which as it is formulated in text, again leads to the emergence of a new idea. If this characterisation is correct, then it provides strong support for the dual process model, as it claims that this kind of dispositionally guided text production will be at a maximum for the low self-monitors in the synthetically planned condition. Even if the precise details of the characterization are not correct, the fact that the pattern for low self-monitors in the synthetic planning condition is so markedly different to the pattern in the outline planning conditions provides strong support for the general contrast between a top-down, explicit planning process and a bottom-up, text production process.

If this reasoning is valid, and the pattern for the low self-monitors in the synthetic planning condition can be taken as evidence for the knowledge constituting process, this would also explain why the pattern for the high self-monitors in synthetic planning is more moderate. The pattern for these writers looks like a hybrid of the outline planning and the low self-monitors' synthetically planned conditions. Thus, unlike the outline planning conditions, there is a relatively low probability that writers begin with a new idea at the beginning of writing. This is compatible with the

assumption that this depends on the writer being able to outline before writing. However, during the middle section of writing there is a much less marked contrast between the probability of an old and a new idea being produced, suggesting that writing is less top-down controlled, and that new ideas may be emerging in the course of text production. The reason why the pattern for this group is less distinctive, both compared to the other writing conditions, and compared to participants within the same condition whose understanding did not increase, may be because a mixture of processes is involved.

## 6.4 GENERAL CONCLUSIONS

Both models of writing make claims about the process by which ideas are produced during writing, but research within this context has only assessed indirectly how ideas are produced or distributed during the process of writing. In this paper, we set out to examine the distribution of ideas in the course of text production and how this relates to development of understanding through writing.

First of all, we evaluated whether ideas produced on the lists before and after writing express similar content as the ideas expressed during the process of writing. Even though our coding scheme seems somewhat conservative we were able to locate ideas in the text and have shown that these ideas and the patterns of their distributions can adequately be modelled using mixed effect logistic regression models.

Secondly, we have shown that the patterns of idea use across time seem to be compatible with the contrast between a top-down and bottom-up process in different writing conditions. For the high self-monitors' outline planned texts there is a low probability of an old idea and a higher probability of a new idea occurring at the start of writing and hence the process appears to reflect explicit organisation prior to writing the text. For the low self-monitors' synthetically planned text there is a high probability of starting with a pre-existing idea and then evidence of a process in which old and new ideas oscillate across the course of text production. This is compatible with new ideas emerging in the course of text production.

The final analysis investigated how these patterns of ideas related to the development of understanding. There, we found that the patterns observed for the participants who did experience development of understanding through writing, provide strong support for the assumption that two distinctive kinds of processes are involved. At one extreme, in the high self-monitors' outline planned condition, this seems to involve a top-down, explicit organising process. At the other extreme, in the low self-monitors' synthetically planned condition, this seems to involve a bottom-up,

text production process. The fact that there is a clear relationship between the pattern of ideas in the high self-monitors' outline condition and increased understanding suggests that, contrary to the dual process model, explicit organising processes can lead to developments of understanding.

Future research should focus on how these patterns of ideas are related to different writing processes. For instance, are the new ideas introduced in the outline planning conditions associated with brief pauses or with long pauses? This could shed more light on the question whether these ideas are really generated during the pre-planning phase. Furthermore, are the ideas that emerge in the synthetic planning condition associated with revision processes during writing? In addition, it would also be interesting to find out how these patterns of ideas related to the end quality of the writing product.

Of course, not all has been said about these patterns. One can speculate about the precise interpretation of a curve going more up at the writing process or going more down for another group of writing. Future research should investigate these patterns more closely and preferably on the basis of a larger number of participants. The simple fact, however, that we have shown that these patterns are different for participants who develop their understanding and who do not develop their understanding, contributes to our understanding about discovery through writing. It also shows that there is a clear and straightforward relationship between idea change and development of understanding. This research has suggested that in order to look at these relations one needs to look more closely at how ideas are produced in the course of text production. Finally, in this paper we have provided some evidence that two processes are important for the development of understanding through writing: one which involves explicit organisation of knowledge and one which involves the implicit organisation of knowledge. This result suggests that Galbraith's dual process model should take into account that also more explicit planning processes are associated with the development of understanding through writing.

# Chapter 7

*Conclusion and Discussion*

## 7.1 INTRODUCTION

The aim of this thesis was to test how processes of writing relate to the development of understanding through writing. In order to do so, both high and low self-monitors wrote an article for the university newspaper in which they discussed our dependence on the computer and internet. They were asked to list, and rate the importance of, their ideas and to rate how much they felt they know about this topic before and after writing. Before writing the text itself, they were asked either to make an outline or to write down a single sentence summing up their overall opinion. During writing, keystrokes were collected using Inputlog (Leijten & van Waes, 2006). **Chapter 1** briefly discussed the rationale for the study and presented the outline of the thesis. **Chapter 2** reviewed the literature on cognitive models of writing as they relate to the question of how writing contributes to the development of understanding. **Chapter 3** investigated how implicit beliefs about writing affect text quality and development of understanding. The chapter evaluated the Writing Beliefs Inventory developed by White and Bruning (2005) and concluded that the scales need to be further developed in order to be used in future research. These scales were, therefore, not used in the main analyses of the thesis. **Chapter 4** evaluated the procedures and measures that were used in order to align keystroke logging data with underlying cognitive processes. These measures were then used in **Chapter 5** to assess how development of understanding and text quality vary depending on type of planning and self-monitoring. This was followed up by a more detailed analysis of the distribution of ideas during the process of writing in **Chapter 6**. Finally, in **Chapter 7**, the current chapter, a summary of the results of the research is presented. In addition, this chapter will address some methodological issues, suggestions for future research and implications for educational practices and well as for Galbraith's dual process model of writing.

## 7.2 RESEARCH PROBLEM

The series of studies presented in this thesis were carried out in order to get more insight in the effect of writing on the development of the writer's understanding. Over the last thirty years, the dominant problem solving models in research on writing have treated expert writing as involving knowledge-transforming and hence, as a process of discovery. Typically, this research has focused its attention on the reflective processes involved in writing, rather than on the processes involved in translating ideas into written output.

More recently, the dual process model of writing (Galbraith, 2009) has shifted the view of writing from a problem solving approach towards a view of writing as text production. Galbraith (2009) claims that discovery through writing occurs during text production and hence, treats text production not as a passive process of translating preconceived ideas into written output, but as an active knowledge constituting process which itself involves the generation of ideas. This process focuses on the implicit organisation of knowledge during writing. The dual process model acknowledges the existence of an explicit rhetorical component within writing, labelled as the knowledge retrieval process, but claims that this component only involves the retrieval of existing knowledge from long term memory and, therefore, does not lead to changes in understanding. According to Galbraith this process, which involves the explicit organisation of knowledge, can only account for the re-organisation of existing knowledge.

The crucial feature of Galbraith's dual process model is that writers need both these processes during writing. The knowledge retrieval process is required for the production of coherent text and the knowledge constituting process is important in order to develop one's understanding during writing. The problem, then, is that these two processes operate best under different writing conditions and hence, are in potential conflict with each other. Galbraith has suggested, on the basis of previous research (Galbraith, 1992, 1999; Galbraith et al., 2006), that high and low self-monitors prioritise the explicit organisation and the implicit organisation processes in different ways.

## 7.3 WRITING BELIEFS

In Galbraith's research self-monitoring is used as a general individual difference measure. In chapter 3 of this thesis we evaluated the Writing Beliefs Inventory designed by White and Bruning (2005) as a more specific writing oriented individual difference measure. Similarly to White and Bruning, we identified two independent sets of writing beliefs - transactional beliefs and transmissional beliefs - and we have suggested that they are associated with important effects on the development of understanding through writing as well as on text quality. The results of this study showed that the effects of writing beliefs vary depending on type of planning. Within the synthetic planning condition, high transactional and low transmissional writers produced better quality text, and high transactional writers experienced greater increases in understanding than low transactional writers. By contrast, in the outline planning condition, low transactional writers produced text similar in quality to the high transactional writers, and high transactional writers were less likely to

experience increases in understanding. Further analysis suggested that, for writers with low transactional beliefs, text modification was associated with the production of lower quality text and with moderate increases in understanding; for writers with high transactional beliefs, text modification was strongly related to increased understanding but unrelated to text quality. On the basis of these results, and in response to the unsatisfactory interpretation of two sub-configurations of the Writing Beliefs Inventory, we have suggested a different interpretation of the transmissional and transactional scale. We suggested that transactional beliefs are essentially beliefs about the process of writing whereas transmissional beliefs are essentially about the source of content in writing. We concluded that these scales need to be further developed in order to use them as a valid and reliable measure. In the subsequent chapters we therefore only used the self-monitoring scale as an individual difference measure.

## 7.4 KEYSTROKE LOGGING

In general, keystroke logging software enables researchers to investigate the online writing processes in far more detail than was previously possible (see Sullivan & Lindgren, 2006, for an overview of recent research). However, it is difficult to relate the automatic output of keystroke logging programs to the underlying processes. In its raw form, keystroke logs only provide information about empty time (when no keys are being pressed) and filled time (when, and which, keys are being pressed). To make sense of these logs, the researcher has to relate this log to a model of different types of processes involved in writing. In addition, due to the complex and recursive nature of the writing process the undifferentiated measures of global properties of keystroke logs, such as mean pause length, are likely to be extremely insensitive measures of the underlying writing process. Some authors have even claimed that this might be a fruitless enterprise (Schilperoord, 2001). In chapter 4, the procedures and measures that were carried out in order to align our keystroke logging data with the underlying cognitive processes were described.

Some of the results obtained in chapter 4 will be highlighted here and we will address some methodological issues. First of all, we have shown how text production can be isolated from other activities during writing. This enabled us to distinguish text production as a part of the forward progression of the text from revision activities. Secondly, on the basis of an evaluation of the language production model from Hayes (Chenoweth & Hayes, 2003) and the speech production model from Levelt (1989) we suggested that pauses during writing may reflect the mental revision of already planned language as much as forward planning of the next linguistic unit. Therefore,

we proposed a more differentiated classification system of burst types with the view of improving the mapping between bursts and different types of process for keystroke logging data. Furthermore, we turned to mixture models to identify subcomponents of the normally much skewed distributions of pauses. The assessment of the relationship between pause length and burst length has shown that long pauses between sentences are associated with shorter clean bursts rather than with longer bursts. This finding implies that bursts in keystroke logging data also reflect the reviewer component of the Chenoweth and Hayes's text production model rather than only the planning of the next linguistic unit. Hence, we have suggested that the length of P-bursts might, and relative number of P- and R- bursts, may, in part, also reflect the strategy of the writer. It could well be that some writers rehearse several bursts of language production before producing a clean P-burst while others do not. Thirdly, we used principal component analysis to identify groupings of interrelated measures. The components derived from theses groupings of measures were used to capture underlying writing processes. In the subsequent chapters we have used components retained from a principal component analysis to assess whether these characteristics vary as a function of self-monitoring and planning as well as how they related to the development of understanding and text quality.

This last feature of the results is therefore of particular importance for the analysis in the other chapters in this thesis. In chapter 5 we reported about two components retained from a somewhat different principal component analysis. We have again labelled these components as planned sentence production and global linearity. The planned sentence production component reflects long pauses and clean bursts versus short pauses and more within sentence revision. This component therefore captures how different writers plan and revise at sentence level. The second component represents global linearity and reflects linear production of the text with little insertions elsewhere in the text and linear transitions between global boundaries such as sentences and paragraphs versus more insertions elsewhere and more revision between global text boundaries. This component captures the extent to which different writers move back and forth in the text, altering the linear ordering of sentences. An important feature of these components is that they are orthogonal to one another. In other words, writers who score highly on planned sentence production do not necessarily score higher or lower on the global linearity component. These measures, therefore, enable us to identify writers who combine planned/ unplanned sentence production with non-linear global structuring of the text. It is an important claim from the dual process model from Galbraith (2009) that high and low self-monitors combine these two processes differently. These measures therefore help to interpret the writing processes in a way that would not be possible

with single measures alone, let alone with global or undifferentiated measures of process.

Some methodological issues should be addressed. A particularly interesting possibility raised by the analysis of keystrokes is how pauses relate to bursts lengths and potential differences between speech and writing. This was in part a conceptual question and in part an empirical question. The conceptual question is whether pauses reflect the same thing in spoken language production and written language production given the possibility that monitoring of inner speech is more feasible in writing than in speaking. This raises questions about the validity of thinking aloud data as an indicator of sentence production processes. If these differ in the way that proposed language is monitored then it is possible that pauses may be related differently to burst lengths in speech and writing. An important area for future research is to compare spoken and written production to assess how pauses, burst lengths and percentages of P- and R-bursts are related with one another in spoken and written language. This should establish whether the same negative relationship holds between pauses and burst length in speech and writing, or whether pauses are more directly tied to conceptual planning in speech. The point here is that it should not be assumed that spoken protocols necessarily reflect the same processes as occur during writing. This needs to be empirically tested by explicitly comparing these relationships in the two contexts.

A second issue this raises is about how one can find out what is going on in the pauses before bursts. A natural assumption is that verbal protocols would reveal the hidden thoughts within the pauses, and that one could therefore simply examine protocols to see whether individuals are proposing language parts during lengthy pauses. The difficulty with this is that, whether or not this is the case, one cannot be sure that the same processes are in operation and hence that verbal protocols give a valid picture of written language production.

Related to this issue, it should be noted that the principal component analysis of the main study of this thesis (chapter 5) did not include measures about the length of P-bursts. Initial screening of the data showed that the burst length measures were highly correlated with the measures of pause duration. However, the results discussed in chapter 4 have shown that length of burst can also cast valuable light on the processes strategies of writers. Future research should focus on developing more detailed writing process components which also include information about bursts length. Potentially, these more fully developed components might prove useful to disentangle different sentence production strategies of writers.

## 7.5 TESTING THE DIFFERENT CLAIMS

Chapter 2 discussed the two different approaches within writing research to the question of how writing facilitates learning. On the one hand, the problem solving models (Hayes & Flower, 1980; Bereiter & Scardamalia, 1987) focus on the processes involved in writing and on how these differ between experts and novices. They describe expert writing as a knowledge-transforming process and contrast it with the knowledge-telling process employed by novices. The problem solving models (Hayes & Flower, 1980; Bereiter & Scardamalia, 1987) focus on writing as a deliberate problem solving activity and suggest that when writers direct their writing towards rhetorical goals, it then becomes a means of developing understanding.

An alternative approach is pursued by Galbraith (1992; 2009) who focusses on the effect of writing on discovery by explicitly testing the development of understanding through writing under different writing conditions. Galbraith's results contradicted important features of the problem solving models since he showed that development of understanding was greater for low self-monitors (who hold dispositional goals) writing full text (assumed to involve greater cognitive load) rather than high self-monitors (which reflects rhetorical goals). These results led Galbraith to develop a dual process model that claims that, during writing, the rhetorical planning process and the implicit text production process occur simultaneously. Furthermore, Galbraith claims that high and low self-monitors prioritise these two processes in different ways. High self-monitors are assumed to engage in top-down processing – focusing first on explicit organisation and then on translating content specified by this process into text; low self-monitors are assumed to engage in bottom-up processing – focusing first on formulating their implicit disposition in connected sentences, and then on adapting the global organisation of the text to the ideas that are constituted in the text.

The results of chapter 5 showed that, overall, participants reported more development of understanding in synthetic planning than in outline planning. Furthermore, we showed that development of understanding through writing is associated with unplanned text production, combined with non-linear global structuring of the text. On the assumption that unplanned sentence production can be equated with dispositionally spelling out, these results provide strong evidence for the dual process model. However, further research is necessary to establish more precisely what the exact nature of planned and unplanned sentence production is. This issue will be addressed more fully later on.

The results also showed that text quality depends on the joint effects of high level planning processes and more local sentence production processes, and suggest

that these have conflicting effects. They confirm the prediction of Galbraith's dual process model that high self-monitors' better quality writing is associated with a top-down process, whereas low self-monitors' is associated with a bottom-up relationship between sentence production and global planning. An important point to note also is that these findings are based on measures that are independent of self-monitoring. These results are therefore providing objective evidence that there are genuine differences between low and high self-monitors that cannot be attributed to a correlated difference in how they carry out the measurement tasks. This provides important support for the claim of the dual process model that hypothesizes that high and low self-monitors differ in how they combine the explicit planning and implicit text production processes.

## 7.6 MAIN INTERPRETATION OF THE RESULTS

In the following section we will relate the results of chapter 5 to the claims of the different writing models. The broad claims of the dual process model are that quality and development of understanding are not necessarily related. Galbraith's dual process model explains this by claiming that deliberate rhetorical planning involves explicit organising processes, and that these play an important role in organising the text and in tailoring it to the needs of readers, but that these processes do not lead to increases in understanding. By contrast, dispositionally guided text production involves implicit organising processes, which do lead to the development of understanding, but which disrupt the production of well-organised text.

In this research we presented results that support this broader claim of the dual process model that development of understanding and text quality are unrelated. In our results this was most pronounced for the high self-monitors. When they imposed linear organisation on their writing, their texts showed the highest text quality whereas they had to organise their writing non-linearly and combine this with unplanned text production in order to experience development of understanding through writing.

Even though Hayes and colleagues and Bereiter and Scardamalia do not explicitly claim that there is a direct relationship between knowledge-transforming processes and better text quality, this is a common assumption within this field (Klein et al., 2011). Therefore, our results have shown that the term "knowledge-transforming" should no longer be used to characterise high quality text. This further suggests that, at the very least, the problem solving models of writing need to develop a differential account for text quality and development of understanding.

Furthermore, our results also support the broad claim of Galbraith's dual process model (2009) that development of understanding is associated with the interaction between implicit organisation processes and explicit organisation processes. Overall, development of understanding was associated with unplanned text production in combination with non-linear global structuring of the text. At the very least, this research supports that a dual process model is required in order to explain development of understanding through writing. For the problem solving models of writing this means that they cannot fully account for the findings of this research. In order to do so, the problem solving models of writing need to take into account that more locally organised sentence production processes are associated with the development of understanding through writing.

That said, we need to consider the precise detail of these results. One issue here is that the interpretation of the process data depends on inferences on our part. On the assumption that dispositionally guided text production is a bottom-up process we have suggested that unplanned text production invites subsequent non-linear global structuring of the text. This assumption is in line with the claims of the dual process model, but this is not directly observable in our data. Given that the data are correlational in nature, this inference about the causal relation between these two process measures should be further tested in future research and needs to be confirmed.

Another issue is the precise nature of the processes involved during planned and unplanned sentence production. First of all, it should be noted that in itself the measure of planned and unplanned sentence production does not reveal whether text production is dispositionally driven. Secondly, on the assumption expressed above that the organisation of sentence production leads to non-linear global structuring of the text, it should be noted that for text quality this was related to planned sentence production whereas for development of understanding this was related to unplanned sentence production. This might suggest that it is the general focus on sentence production processes combined with explicit planning processes, rather than the precise nature of planned and unplanned sentence production, that leads to the development of understanding. Future research should focus on defining more precisely what planned and unplanned sentence production are and should investigate how planned and unplanned sentence production correspond with the knowledge constituting process as described by Galbraith's dual process model. The present research has not gone down to this level, but showed that local organisation and global organisation do conflict with each other and that the way in which these processes are prioritised by high and low self-monitors lead to different writing outcomes. The issue about the precise nature of text production processes relates to

the issues addressed about keystroke logging analysis. Establishing what happens during pauses is a complex issue that needs to be further addressed in future research.

Having stressed the importance of research establishing what happens in pauses, our results showed that unplanned sentence production combined with non-linear organisation leads to the development of understanding. On the assumption that this captures what Galbraith defines as dispositionally driven text production, some further issues should be raised. One of which is how this text production process relates to idea change. One less straightforward result from our experiment was that it produced a different pattern of idea change to previous research. In this context it was remarkable that none of the idea change measures were related to the planned sentence production process whereas other processes were related to idea change. This is particularly important for the claim of the dual process model that ideas produced by dispositionally driven text production processes are related to development of understanding whereas ideas produced by explicit planning processes are not related to the development of understanding.

A possible explanation for the lack of a relationship between planned sentence production and idea change might be related to the fact that the precise nature of these processes still needs to be further established. Planned and unplanned sentence production could in principle be different ways of accomplishing the same goals, with some writers doing this during the pauses before they produce sentences and other writers doing this as they revise sentences. In particular, the problem solving account could claim that the difference is essentially between rhetorical evaluation of proposed bursts before they are written down and rhetorical evaluation of bursts after they have been written down. Equally, the dual process model could claim that what matters is whether pauses and revisions are concerned with dispositional goals, rather than whether this is carried out before or after a burst is transcribed. One explanation, therefore, for the lack of relationship between the planned sentence production measure and idea change could be that it is not this aspect of the sentence production process that is responsible for discovery but rather some other difference in what goes on during pauses or revisions. Therefore, one of the more important questions for future research is disentangling the complex and dynamic interactions and interrelationships between processes measures and idea change measures.

## 7.7 IDEA CHANGE

Concerning the evaluation of the claims of the two models it should be noted that, although they attribute the development of understanding to different processes, they are compatible with one another in the claim that when writers change their ideas

during writing this leads to the development of understanding. In this context, one of the more striking results of the previous experiment was that idea change seemed to be unrelated with development of understanding through writing. The dual process model explains this by claiming that the process by which ideas are produced accounts for whether ideas are associated with the development of understanding. According to the dual process model, ideas produced by rhetorical planning processes are unrelated to the development of understanding whereas ideas produced by dispositionally driven text production are assumed to be related to the development of understanding.

We assume that in our writing conditions ideas produced by dispositionally guided text production processes were mixed in with ideas produced by rhetorical planning processes. Therefore, we were unable to detect direct relationships between idea change, processes and the development of understanding. First of all, this implies that more sophisticated measures of change in understanding need to be developed. Secondly, it suggests that it should be investigated in more detail how writers build their understanding as they write their text.

In the final empirical chapter of this thesis it was therefore investigated how writers use their ideas during the process of writing and how the distribution of these ideas relates to the development of understanding through writing. The results of this chapter provide evidence that suggests that increases in understanding do depend on how ideas are produced during the process of writing. The patterns of idea distribution that we observed for participants who reported development of understanding were distinctly different from the patterns observed from participants that did not report development of understanding. Furthermore, the results suggested that there are clear differences between the outline planned and synthetically planned texts. This suggests that there are clear relationships between patterns of idea change and increases in understanding, but that in order to establish these one needs to look more closely at ideas as they are created in the course of text production rather than on the basis of average measures of processes across the whole text.

We suggested that within the outline planning conditions, participants generate their new ideas during pre-planning, and then, during the writing process, focus on the alignment of their original ideas within the context of these new ideas, which occur at the beginning and the end. Basically, this pattern of idea use looks as if the writer writes out their original ideas but create and introductory overview and a final summary. This relates to the problem solving models in that is suggests that writers are creating a new global structure and a new context for their original ideas. This is a distinctly different kind of change of understanding driven by the global structuring of the text. This result therefore suggests that Galbraith's dual process

model needs to take into account that also the explicit planning processes can be related to the development of understanding in writing.

In contrast, for the synthetic planning conditions, especially in the low self-monitor synthetic planning condition, we have shown that their pattern of idea use relates more to the hypothesized features of dispositionally guided text production. The pattern observed in that condition seems to suggest that writing is indeed a dialectical interaction between the text written so far and the writer's disposition towards the topic, and hence ideas emerge in the course of text production. Future research should establish in more detail what the exact pattern of this idea use is in order to establish that it does relate to the knowledge constituting process as described by Galbraith.

To the extent that these results show that different organisation is imposed on the use of ideas during the process of writing in the synthetically planned and the outline planned text, these results support the dual process model. The dual process model is right in claiming that new ideas can be produced without any relationship with development of understanding. We have shown this in chapter 5 and provided the explanation that according to the dual process model this implies that these ideas merely involve rhetorical changes. However, the results of chapter 6 have shown that the development of understanding through writing has at least two accounts and that also planned idea change can lead to the development of understanding. This underlines the need for the dual process model to refine the nature of explicit planning processes. In order to define in more detail how these explicit planning processes relate to the development of understanding further research is needed to investigate how the use of ideas in the course of text production is related to different writing processes. Future research should also assess how this affects text quality.

## 7.8 GENERAL ISSUES

In this section we will discuss some general methodological issues related to the findings of the results presented in this thesis. In light of these considerations we will give some possible suggestions for further research.

### *7.8.1 Individual difference measures*

Individual differences between high and low self-monitors have played a central role in this and previous research. This raises two methodological issues. One issue is that research into self-monitoring and writing is based on the assumption that low self-monitors' writing is directed towards dispositional goals whereas high self-monitors'

writing is directed towards rhetorical goals. Despite the fact that previous research has provided convincing evidence that high and low self-monitors show different patterns of idea and knowledge change the self-monitoring measure in itself does not guarantee that high self-monitors prioritise rhetorical goals and that low self-monitors do prioritise dispositional goals. This could be examined by letting high and low self-monitors think aloud when composing an outline before writing and examining whether these show evidence of a difference in the goals towards which they direct their writing. This could also be done by using the triple task during writing (Kellogg, 1988). In that case, high and low self-monitors would be presented with an intermittent tone during writing. When the tone is presented, they should be prompted to indicate which writing activity they are currently engaged in.

Another, more general, issue with individual difference measures is that observed differences may be due to some other confounded variable rather than the assumed individual difference. In the case of self-monitoring, one cannot rule out that the differences observed in previous research could be a consequence of differences in the way that high and low self-monitors list their ideas before and after writing, or of differences in the way they subjectively rate their understanding. In this context it was reassuring that we did not find differences in the number of ideas, the importance of ideas or the length of ideas produced by high and low self-monitors before writing. Furthermore, we showed that objective measures such as our keystroke logging data and the externally rated text quality were able to pick up genuine differences between high and low self-monitors. However, in order to establish more directly whether there are no differences in the way high and low self-monitors approach the different tasks the issue of measurement invariance needs to be addressed, i.e. the extent to which the factorial structure of the measure is equal for the different types of writers (Meredith, 1993; Vandenberg & Lance, 2000).

Our results also indicate that future research should look at how writing beliefs and self-monitoring interact. Although there was no direct correlation between self-monitoring and writing beliefs, these may interact with each other in their effects. In this study we did not pursue this avenue any further because the scales from the writing beliefs inventory were difficult to interpret, and because in addition the sample size made it impossible to examine high level interactions between so many variables in any reliable way. For future research, the first step here is to improve the reliability and interpretability of the writing belief scales. This will open the way to more detailed analysis of writing beliefs, their relationship to more detail process analysis and other measures such as self-monitoring.

### *7.8.2 Writing task*

In our experiment participants were given 30 minutes to complete a coherent article for the university newspaper. Therefore, it should be noted that the writing task in this experiment involved a single task. Moreover, even though writers were allowed to revise their text, and also showed evidence of such processes, the writing assignment merely consisted of a single draft. The limited amount of time available for writing task might have limited the extent to which writers were able to experience development of understanding through writing. The fact that the writing assignment merely consisted of a single draft implies that writers were not given the opportunity to revise their text over an extended period of time. This is an important thing to note in the context of the claims of the dual process model of writing. In the light of research from Elbow (1981), the dual process model advocates a dual drafting strategy whereby the first draft focusses on getting ideas on paper and the second draft is about trying to organise the ideas produced in the first draft in a coherent manner. Due to the instructions in this experiment writers are unlikely to have been able to employ a full dual drafting strategy. Note though that there were at least one or two writers who showed evidence of writing several drafts. Future research should focus on writing tasks in which writers are given the opportunity to compose their texts in multiple drafts.

In light of this, future research should also focus on the relationship between text quality and development of understanding. The current experiment showed evidence that this might not be interrelated, but this should be further assessed using a bigger sample of participants. Although we recruited a relatively large sample of 84 participants for this experiment, a much larger sample would be needed to carry out a reliable factor analysis of the keystroke measures, and to carry out a full path analysis of the interrelationships between self-monitoring, type of planning, the process measures, idea change, changes in understanding and text quality. Keystroke analysis is a promising method in principle because it enables one to collect a large amount of data. However it is important to stress that to do this future research is needed to develop more automatic methods for coding process logs. The data preparations needed to derive these measures from our keystroke logs have involved a huge amount of time.

## 7.9 EDUCATIONAL AND THEORETICAL IMPLICATIONS

As mentioned in the introduction an important goal of this thesis was to test the predictions of cognitive models of how writing facilitates the development of

understanding. Even though this research was closely related to writing-to-learn research we addressed a conceptually different question. Therefore, this research does not intend to address issues in the teaching of writing in schools. Nevertheless, the results from this thesis are relevant for the teaching, learning and use of writing. Despite the fact that it is difficult to provide hard advice for the educational context, since our results are correlational in nature, we feel the need to address some implications for educational settings below. Furthermore, this thesis presented results that have implications for the development of the cognitive models. Taken that a dual process is required to explain the processes involved in effective writing we will focus on implications for Galbraith's dual process model of writing.

### *7.9.1 Educational implications*

First of all, the results have shown that individual difference measures, such as self-monitoring and implicit writing beliefs, are able to pick up genuine differences between writers. Even aside from what the precise nature of these differences is, this result has important implications for educational practice. It implies that different types of writers may need different kinds of instruction (see for example, Kieft, Rijlaarsdam & van den Bergh, 2006; Kieft, Rijlaarsdam, Galbraith & van den Bergh, 2007).

This was best illustrated by our text quality result that showed that high and low self-monitors produce their best text quality under different circumstances. This result showed that writing quality was related to a complex and dynamic relation between types of writers and writing processes. It is difficult to extract hard advice for educational settings from this result. But even if the nature of our data did allow us to give advice in the way which is common within the field of aptitude treatment research (Snow, 1989) whereby different forms of instruction are offered to different types of learners, we would refrain from doing so. Our data do suggest that different approaches have differential effects for different types of writers, but they also show that all writers engage in all types of processes, be this at different stages of their writing process. We therefore restrict ourselves to express some advice that is useful for all writers. Then, one issue would be that this result implies that focussing on the fact that both higher order planning and more local text production processes are responsible for the production of high quality text might be a good start for educational practices. Subsequently, this implies that writing instructions should take into account that the interrelations between different processes matter. Hence, educational practices would gain from a focus on the function of the two processes rather than necessarily on a global drafting strategy. This awareness of the different

functions of writing processes might provide students with useful tools to organise their writing processes in more flexible manners.

Also our study into writing beliefs has shown that different conceptions about the nature of writing can influence the performance of the writing task. In order to develop one's understanding through writing it was shown that writers need to hold the belief that writing can be used as a tool to do so. When writers hold the belief that writing merely involves the transmission of sources they are most likely not going to develop their understanding through writing. This has consequences for educational settings and suggests that an important focus for writing instruction should be the often tacit beliefs that students have about writing.

In the context of educational settings the most common advice is to let students make an outline plan before writing. Various different aspects of our data have suggested that outlining was not uniformly beneficial for everybody. First of all, our research failed to replicate the general beneficial effect that outlining is assumed to have on text quality. Our results showed that for high transactional writers synthetic planning can produce texts similar in quality perhaps even of higher quality, to outline planned text. Since Kellogg's experiments have typically compared outline planning with conditions in which writers have to combine planning with text production in a single draft without pre-planning, this result suggests that synthetic planning – in which the writer has 5 minutes pre-writing time but is not required to explicitly organise their ideas – may be more effective than writing a single draft. Further research should test the role of different types of pre-planning more explicitly. It should be noted though, that our study into writing beliefs showed that low transactional writers wrote better texts in the outline planning condition than in the synthetic planning condition. This result does fit well with the assumption that separating the generation and organisation of ideas from text production enables writers to focus on text production during writing. Combined with the result that high self-monitors produce the highest quality texts when they produce their text linearly this suggests that text quality is associated with the extent to which, at least certain types of, writers are able to work out what they want to say before they write it down.

Furthermore our results showed that outlining has a negative effect on the development of understanding. For educational settings this means that one should carefully consider what the goal of writing is when instructing writers to start with an outline plan before writing. If the goal of writing is to develop one's understanding through writing, it might be better to start with a spontaneous draft rather than an explicit plan for the text.

When this result is related to the writing-to-learn context it might partly explain why writing-to-learn assignments are not always effective. Our results have shown that not all writing assignments enhance the development of understanding. This suggests that real thought need to go into how these writing-to-learn tasks are designed. Future research should investigate what the optimum conditions are for writing-to-learn assignments. That goes beyond this research, because we have only looked at the subjective development of understanding. A second point in this context is, then, how subjective understanding relates to more objective learning outcomes. At the very least, this research has suggested that writing-to- learn research needs to take into account that learning might not be the same as producing good quality writing.

### *7.9.2 Implications for the dual process model of writing*

Galbraith's dual process model claims that dispositionally guided text production involves implicit organising processes, which do lead to the development of understanding, but which disrupt the production of well-organised text. The key feature of the dual process model is that this dispositionally driven text production process needs to be combined with explicit organisation processes. In order to develop their understanding writers need to, first, focus on formulating their implicit disposition in connected sentences, and then on adapting the global organisation of the text to the ideas that are constituted in the text. This interaction between these two processes is crucial for the development of understanding through writing. However, it is also a subtle interaction. When the explicit planning processes are introduced to early on in the process it will limit the extent to which writers are able to use the dispositionally guided text production process. The sketch of the dual process model that is put forward by Galbraith at the moment only presents the knowledge constituting part of the model (see figure 7-1). Since our research has provided evidence that supports the claim from the dual process model that writing involves the complex interaction between explicit planning processes and sentence production processes, we believe the time has come to develop the dual process model further. The model needs to specify more precisely how the two processes are interrelated. For example, how does the model account for the fact that, when rhetorical processes are imposed early on in the writing process by means of outlining or high self-monitors, writers will still be able to constitute their disposition towards the topic during writing? Our results have shown that when the high self-monitors wrote more like low self-monitors in the synthetic planning conditions – less planned sentence production and less linearly produced text – they experienced increases in understanding. Synthetic planning might enable them to focus more on the local

organisation of the text production, but the simple fact that they are high self-monitors still implies that from early on in the writing process these writers have to handle the conflicting features of the two processes. The sketch of the current dual process model does not account for this complex interaction between the two processes. It is therefore stressed that the dual process model would gain when it presents a more detailed sketch of the interrelationships between the two processes or when it provides a more detailed description of the set of explicit processing stages during writing.

Related to this issue are our results about memorisation. Our results have suggested that memorisation might be an indicator of the extent to which writers store their ideas during writing. According to the process model, this process will, similarly to outlining, limit the extent to which writers are able to focus on the spontaneous unfolding of sentences during text production. Memorisation and outlining are assumed to impose a linear plan to the synthesis process which will stop the extent to which the next utterance is produced by dispositionally spelling out. On the assumption that ideas that are maintained during writing are stored in working memory, this implies that the processes in semantic memory are interlinked with working memory processes. It is therefore stressed that the more elaborate sketch of the dual process should also indicate how working memory and semantic memory are linked.

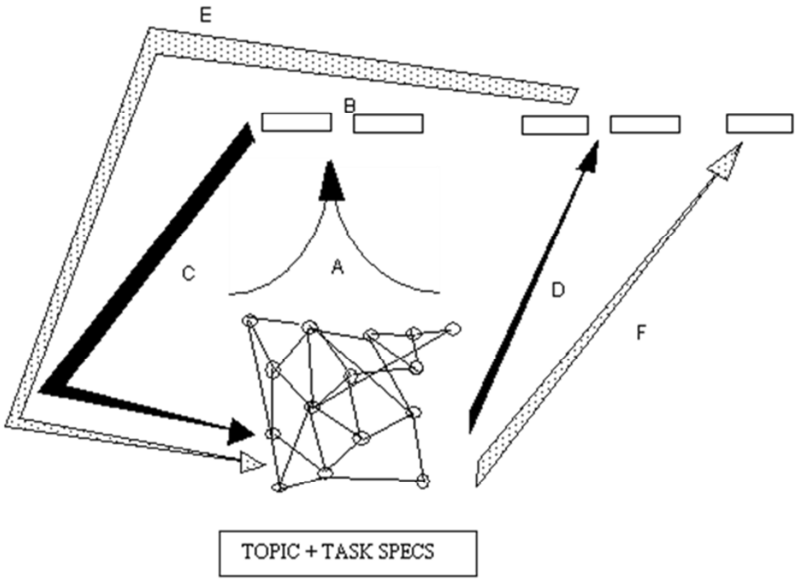


**Figure 7-1** Visual representation of writing as a knowledge constituting process (Galbraith, 1999).

Furthermore, the dual process model would gain from a more detailed description of the language processes involved. The dual process model claims that the dispositionally guided text production process focusses on the formulation of the writer's implicit disposition in connected processes. We have already raised that further research is needed to investigate in more detail how, for instance, the planned sentence component corresponds with the process Galbraith characterizes as dispositionally spelling out. Having said that, some additional comments are needed to explain the request for a more detailed description of the language processes involved. Galbraith (personal communication) interprets the language production component of his model in the context of Slobin's (1987) "thinking for speaking". The "thinking for speaking" hypothesis refers to how speakers organise their thinking to meet the demands of online encoding during the act of speaking. "Thinking before speaking" relates to the claim from Galbraith that ideas put forward by the knowledge constituting process are inherently related to the writer's disposition towards the topic. However, dispositionally guided text production is often interpreted as if the language processes itself are dispositionally spelled out whereas the close reader can observe that the dual process model does not make any explicit claims about the way in which messages are converted into linguistic strings or written text (see arrow labelled A, in figure 7-1).

To relate this claim to the text production model from Chenoweth and Hayes (2003), it basically means that the proposer and the translator of Hayes' text production model are included within the network of interconnected units in the middle of the diagram (see figure 7-1). This synthesis process is therefore responsible for the production of the idea package and for converting this idea package into a linguistic string. In Hayes' text production model the transcriber is then responsible for converting this linguistic string into written text and in Galbraith's dual process model this step is represented by arrow A. It should therefore be noted that the dual process model of writing does not make any specific claims about the nature of the process in which linguistic strings are put down in words on paper. This suggests that the languages processes involved in this step are similar to the language processes proposed by other models such as the text production model from Hayes (Chenoweth & Hayes, 2003). To take this even one step further, this suggests that, similar to the problem solving models, Galbraith's dual process model provides explanations for processes that are responsible for the production of an utterance and not for how these utterances are transcribed or, to say it in Galbraith's terms, constituted in written language. Nevertheless, because the content that is produced by this synthesis process is unpredictable and it is assumed to relate directly to the writer's disposition towards to topic, the knowledge constituting process from Galbraith still provides an

account of how the dialectical interaction between the writer's disposition and the emerging text can lead to the development of understanding through writing. However, the dual process model would gain by a more detailed description of the language processes involved.

This research has already suggested that further studies are required to test the more detailed claims about the knowledge constituting process. It should be investigated what the precise nature of planned and unplanned text production is and how this relates to the knowledge constituting process as described by Galbraith. Furthermore, it should be examined what the direction is of the interaction between explicit planning processes and local sentence production processes. These future studies should be able to confirm the claim from the dual process model that dispositionally guided text production leads to the non-linear global structuring of the text. Yet another way to examine more precisely how writers build their understanding as they write their text is by looking at the patterns of ideas in the course of text production. All these suggestions for future research can give more insight in the broader claims of the nature of the knowledge constituting process. However, it should be noted that the knowledge constituting process in itself is based on theoretical assumptions about the nature of knowledge representations in semantic memory. In that sense, the knowledge constituting process is a theoretical model trying to explain how content that is produced can be dispositionally organised. In order to test these more fundamental theoretical claims from Galbraith's dual process model different kinds of future experiments are required. Perhaps simulation studies could get some way at answering how writers constitute their disposition into language. However, we will leave that puzzle for future research to unravel.

## 7.10 GENERAL CONCLUSION

The final point to be made is that the approach used in chapter 6 represents a potentially new avenue of research. In the past, writing was initially viewed as being about the final product. This then shifted to a focus on the cognitive processes involved in the production of the text, and latterly towards research investigating writing processes as they occur online. In the exploratory analysis that we carried out in chapter 6, we combined the general approach embodied in Galbraith's research into idea change and applied it to the development of ideas during writing itself, rather than before and after writing. This raises an important alternative approach to the writing process. For the writer, it may be the development of content during writing that is important. Questions like *What do I want to say*? *What have I said so far*? *Does it make sense to me*? *Will it make sense to my readers*? *What else do I need to say*? are familiar

to all writers. However, in past research these questions have been addressed in the context of rhetorical goals for the text. In contrast, this research suggests that these are judgments about the content rather than evaluations about the process of writing. Perhaps writing processes are driven by the writer's judgments about the content of the text under construction and not only by what the overall goals of the writing task are. Future research should focus on what is driving text production, on the extent to which ideas drive the creation of text.

# Bibliography

Ackerman, J.M. (1993). The promise of writing to learn. *Written Communication, 10*, 334-370.

Applebee, A.N. (1984). Writing and reasoning. *Review of Educational Research, 54* (4), 577-596.

Baaijen, V.M., Galbraith, D., & de Glopper, K. (2010). Writing: The process of discovery. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society,* (pp. 1774-1779). Austin, TX: Cognitive Science Society.

Baaijen, V.M., Galbraith. D., & de Glopper, K. (2012). Keystroke analysis: Reflections on procedures and measures. *Written Communication, Special Issue on Writing and Cognition: In Honor of John R. Hayes*, *29* (3), 246-277.

Baayen, R.H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R.* Cambridge: University Press.

Baayen, R.H. (2010). The directed compound graph of English. An exploration of lexical connectivity and it processing consequences. In S. Olsen (Ed.), *New Impulses in Word-Formation (Linguistic Berichte Sonderhelf 17*), (pp. 383-402). Hamburg: Buske.

Bergh, H, van den., & Rijlaarsdam, G. (1996). The dynamics of composing: modeling writing process data. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences & applications,* (pp.207-232). Mahwah, NJ: Lawrence Erlbaum Associated.

Bergh, H, van den., & Rijlaarsdam, G. (2001). Changes in cognitive activities during the writing process and relationships with text quality. *Educational Psychology*, *21*(4), 373-385.

Bergh, H, van den., Rijlaarsdam, G., Janssen, T., Braaksma, M., Weijen, D, van., & Tillema, M. (2009). Process execution of writing and reading: Considering text quality, learner and task characteristics. In M. Shelly, B. Hand, & L. Yore (Eds.), *Quality research in literacy and science education: International perspectives and global standards,* (pp.399-426). Spinger.

Berninger, V.W. (2012). *Past, present, and future contributions of cognitive writing research to cognitive psychology.* New York: Psychology Press.

Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.

Bangert-Drowns, R.L., Hurley, M.M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, *74* (1), 29-58.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Britton, J. (1982). Writing-to-learn and learning to write. In G.M. Pradl, (Ed.), *Prospect and retrospect: Selected essays of James Britton,* (pp. 94-111.) Boynton/Cook Publishers, Inc., Montclair, NJ, (Reprinted from *The Humanity of English: NCTE Distinguished Lectures 1972).*

Bochardt, I. (1984). Het schrijfproces: Cognitief-psychologisch onderzoek van Flower en Hayes. *Tijdschrift voor Taalbeheersing, 6* (1), 23-42.

Bruning, R., Dempsey, M., Kauffman, D., & Zumbrunn, S. (2011, April). *Writing viewed as transmission and transaction: An empirical study of implicit writing*

*beliefs*. Poster session presented at the annual meeting of the American Educational Research Association, New \Orleans.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication, 18* (1), 80-98.

Chenoweth, N. A., & Hayes, J. R. (2003). The inner voice in writing. *Written Communication 20 (*1), 99-118.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. A. (2003). *Applied multiple regression/correlation analysis for the behavior sciences.* Mahwah, NJ: Erlbaum.

Collins, A., & Gentner, D. (1980). A framework for a cognitive model of writing. In L. W. Gregg & Steinberg, E.R. (Eds.), *Cognitive processes in writing,* (pp.51-72). Hillsdale, NJ: Lawrence Erlbaum Associates.

Crawley, M.J. (2005). *Statistics: an introduction using R.* Chichester, West-Sussex: Wiley.

Elbow, P. (1981). *Writing with power. Techniques for mastering the writing process.* Oxford: University Press.

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication, 28*, 122-128.

Field, A. (2009). *Discovering statistics using SPSS (and sex and drugs and rock 'n'roll.* Londen: Sage Publications.

Flower, L.S., & Hayes, J.R. (1980a). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, *31*(1), 21-32.

Flower, L.S., Hayes, J.R. (1980b). The dynamics of composing: Making plans and juggling constraints. In L.W. Gregg & E.R. Steinberg (Eds.), *Cognitive processes in writing,*(pp.31-50). Hillsdale, N.J.: Lawrence Erlbaum Associates Publishers.

Flower, L., & Hayes, J.R. (1984). Images, plans, and prose: The representation of meaning in writing. *Written Communication*, *1*(1), 120-160.

Friedenberg, J., & Silverman, G. (2006). *Cognitive science. An introduction to the study of mind.* London: Sage Publications.

Galbraith, D. (1992). Conditions for discovery through writing. *Instructional Science*, *21*, 45-72.

Galbraith, D. (1996). Self-monitoring, discovery through writing and individual differences in drafting strategy. In G. Rijlaarsdam, H. van den Berg & M. Couzijn (Eds.), *Theories, models and methodology in writing research,* (pp.121-141). Amsterdam: University Press.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production,* (pp. 139-160). Amsterdam: University Press.

Galbraith, D., & Rijlaarsdam, G. (1999). Effective strategies for the teaching and the learning of writing. *Learning and Instruction, 9*, 93-108.

Galbraith, D., & Torrance, M. (1999). *Knowing what to write: Conceptual processes in text production.* Amsterdam: University Press.

Galbraith, D., & Torrance, M. (1999a). Conceptual processes in writing: From problem solving to text production. In D. Galbraith & M. Torrance (Eds.), *Knowing what to write: Conceptual processes in text production,* (pp. 1-12). Amsterdam: University Press.

Galbraith, D., & Torrance, M. (2004). Revision in the context of different drafting strategies. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes,* (pp.63-86). Dordrecht, NL: Kluwer Academic Publishers.

Galbraith, D., Ford, S., Walker, G., & Ford, J. (2005). The contribution of different components of working memory to knowledge transformation through writing. *L1-Educational Studies in Language and Literature*, *5*, 113-145.

Galbraith, D., Torrance, M., & Hallam, J. (2006). Effects of writing on conceptual coherence. *Proceedings of the 28th Annual Conference of the Cognitive Science Society,* 1340–1345.

Galbraith, D., Waes, L, van., & Torrance, M. (2007). Introduction. In G. Rijlaarsdam (Series Ed.) and M. Torrance, L. van Waes & D. Galbraith (Volume Eds.), *Writing and cognition: Research and applications* (Vol. 20, pp. 1-10).Amsterdam: Elsevier.

Galbraith, D. (2009). Writing as discovery. *British Journal of Educational Psychology*, *00*, 1-23.

Gregg, L.W., & Steinberg, E.R. (1980). *Cognitive processes in writing*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Harrell, F.E. (2001). *Regression modelling strategies*. Berlin: Spinger.Hayes, J.R., & Flower, L.S. (1980). Identifying the organization of writing processes. In L.W. Gregg & E.R. Steinberg (Eds.), *Cognitive processes in writing,* (pp.3-30). Hillsdale, N.J.: Lawrence Erlbaum Associates Publishers.

Hayes, J.R., & Flower, L.S. (1986). Writing research and the writer. *American Psychologist*, *1*(10), 1106-1113.

Hayes, J.R., Flower, L.S., Schriver, K.A., Stratman,J.F., & Carey, L. (1987). Cognitive processes in revision. In S Rosenberg (Ed.), *Advantages in applied psycholinguistics, Volume 2: Reading, writing and language learning,* (pp.176-240). Cambridge, NY: Cambridge University Press.

Hayes, J.R., & Nash. J.G. (1996). On the nature of planning in writing. In C.M. Levy & S. Ransdell (Eds.), *The Science of writing: Theories, methods, individual differences and Applications,* (pp.29-55). Mahwah, NJ: Lawrence Erlbaum Associates.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications,* (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum Associates.

Hayes, J.R. (2009). From idea to text. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds.), *The SAGE handbook of writing development*, (pp.65-79). London: SAGE Publication Ltd.

Kass, R.E., & Raftery, A.E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 775-795.

Kaufer, D., Hayes, J.R., & Flower, L.S. (1986). Composing written sentences. *Research in the Teaching of English, 20*, 121-140.

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* (2), 355-365.

Kellogg, R.T. (1990). Effectiveness of prewriting strategies as a function of task demands. *American Journal of Psychology*, *103*, 327-342.

Kellogg, R.T. (1994). *The psychology of writing* Oxford, UK: Oxford University Press.

Keys, C.W. (2000). Investigating the thinking processes of eighth grade writers during the composition of a scientific laboratory report. *Journal of Research in Science Teaching, 37,* 676-690.

Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2006). Writing as a learning tool: Testing the role of students' writing strategies. *European Journal of Psychology of Education*, *XXI* (1), 17-34.

Kieft, M., Rijlaarsdam, G., Galbraith, D., & van den Bergh, H.. (2007). The effects of adapting a writing course to students' writing strategies. *British Journal of Educational; Psychology, 77*, 565-578.

Kirsner, K., Hird, K., & Dunn, J.C. (2005). Communication disorders following stroke: First step towards a new fluency protocol. *Brain and Language*, *95* (1), 165-166.

Klein, P.D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review*, *11*(3), 203-270.

Klein, P.D. (2004). Constructing scientific explanations through writing. *Instructional Science*, 32, 191-231.

Klein, P.D., & Kirkpatrick, L.C. (2010). A framework for content area writing: Mediators and moderators. *Journal of Writing Research*, *2*(1), 1-46.

Langer, J.A., & Applebee, A.N. (1987). *How writing shapes thinking: A study of teaching and learning*. National council of teachers of english. Urbana, IL.

Lavelle, E., Smith, J., & O' Ryan, L. (2002). The writing approaches of secondary students. *British Journal of Educational Psychology, 72,* 399-418.

Leijten, M., & Waes, L, van. (2006). Inputlog: New perspectives on the logging of online writing. In K.P.H. Sullivan & E. Lindgren (Eds.), Studies in Writing: Vol. 18, *Computer keystroke logging and writing: Methods and applications,* (pp.73-93). Oxford: Elsevier.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.

Levelt, W. (1999). Producing spoken language: A Blueprint of the speaker. In C. Brown & P. Hagoort (Eds.), *The neuro-cognition of language*, (pp. 83-122). New York: Oxford University Press.

Levy, C.M., & Ransdell, S. (1996). *The science of writing. Theories, methods, individual differences and applications*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Levy, C.M., & Ransdell, S (1996a). Working memory constraints on writing quality and fluency. In C.M. Levy, & S. Ransdell (Eds.), *The science of writing. Theories, methods, individual differences and applications*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Lindgren, E., & Sullivan, K.P.H. (2006). Writing and the analysis of revision: An overview. In K.P.H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing: Methods and applications,* (pp.31-44). Oxford: Elsevier.

MacArthur, C.A., Graham, S., & Fitzgerald, F. (2006). *The handbook of writing research,* (pp.386 – 402). NY: Guilford Publications.

McLachlan, G.J., & Peel D. (2000). *Finite mixture models*. New York: Wiley.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58* (4), 525-543.

Murray, D.M. (1978). Internal revision: A process of discovery. In C.R. Cooper & L. Odell (Eds.), *Research on composing: Points of departure*, (pp. 85-104). National council of teachers of english, Urbana, IL.

O' Grady, W., Dobrovolsky, M., & Katamba, F. (1996). *Contemporary linguistics*. London: Longman.

Rijlaarsdam, G., Bergh, H, van den., & Couzijn, C. (1996). *Theories, models and methodology in writing research*. Amsterdam: University Press.

Rijlaarsdam, G. & Bergh, H, van den. (1996). The dynamics of composing – an agenda for research into an interactive compensatory model of writing: many questions, some answers. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences & applications*, (pp.107-125). Mahwah, NJ: Lawrence Erlbaum Associated.

Sanders, T., & Wijk, C, van. (1996). Text analysis as a research tool: how hierarchical text structure contributes to the understanding of conceptual processes in writing. In C.M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences & applications*, (pp.252-269). Mahwah, NJ: Lawrence Erlbaum Associated.

Sanders, T. & Schilperoord, J. (2006). Text structure as a window on the cognition of writing: How text analysis provides insights in writing products and writing processes. In C.A. MacArthur, S.Graham & J.Fitzgerald (Eds.), *The handbook of writing research*, (pp. 386 - 402). NY: Guilford Publications.

Slobin, D. (1987). Thinking for speaking. In J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Papers from the 13th Annual Meeting at the Berkeley Linguistic Society*. Berkeley, CA: BLS.

Snow, R.E. (1989). Aptitude –treatment interaction as a framework for research on individual differences in learning. In P.L. Ackerman, R.J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research*, (pp. 13-59). New York: W.H. Freeman Publishers.

Rosenberg, S. (1987). *Advances in applied psycholinguistics, volume 2: Reading, writing, and language teaching*. Cambridge, NY: Cambridge University press.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science*, *8*(2), 173-190.

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written communication. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics, volume 2: Reading, writing, and language teaching*, (pp.142-1745). Cambridge, NY: Cambridge University press.

Schilperoord, J., & Sanders, T. (1999). How hierarchical text structure affects retrieval processes: implications of pause and text analysis. In D. Galbraith & M. Torrance (Eds.), *Knowing what to write: Conceptual processes in text production*, (pp. 13-30). Amsterdam: University Press.

Schilperoord, J. (2001). On the cognitive status of pauses in discourse production. In T. Olive & C.M. Levy (Eds.), *Contemporary Tools and Techniques for Studying Writing*, (pp.61-87). Dordrecht: Kluwer Academic Publishers.

Schraw, G., & Bruning, R. (1996). Readers' implicit models of reading. *Reading research Quarterly, 31*, 290-305.

Schraw, G., & Bruning, R. (1999). How implicit models of reading affect motivation to read and reading engagement. *Scientific Studies of Reading*, *3* (3), 281- 302.

Spelman Miller, K., Lindgren, E., & Sullivan, K.P.H. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, *42* (3), 433-454.

Snyder, M. (1974). Self-monitoring of expressive self-behaviour. *Journal of Personality and Social Psychology, 30*, 526-537.

Snyder, M. (1986). *Public appearances, private readings: The psychology of self-monitoring.* New York: W.H. Freeman and Company.

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

Stevenson, M., Schoonen, R., & Glopper, K, de. (2006). Revising in two languages: A multidimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing, 15* (3), 201-233.

Sullivan, K.P.H., & Lindgren, E. (2006). *Computer keystroke logging and writing: methods and applications*. Oxford: Elsevier.

Torrance, M. (1996). Is writing expertise like other kinds of expertise? In G.Rijlaarsdam & E. Espéret (Series Ed.), G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Vol Eds.), *Studies in Writing vol.2, Theories, models and methodology in writing research* (pp. 3-9). Amsterdam: Amsterdam University Press.

Tremblay, A., & Baayen, R.H. (2010). Holistic processing of regular four-word sequences: A behavioural and ERP study of the effects of structure, frequency, and probability on immediate free call. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication*, (pp. 151-173). London: The continuum international publishing group.

Vandenberg, R.J. & Lance, C.E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3* (1), 4-70.

Vigliocco, G., & Hartsuiker, R.J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin 128* (3), 442-477.

Waes, L, van, & P.J. Schellens. (2003). Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics*, *35* (6), 829-853.

Waes, L, van, Leijten, M., Wengelin, A., & Lindgren, E. (2012). Logging tools to study digital writing processes. In V.W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology,* (pp.507-533). New York: Psychology Press.

Wallace, D.L., & Hayes, J.R. (1991). Redefining revision for freshmen. *Research in the Teaching of English*, 25, 54-66.

Wason, P.C. (1980). Specific thoughts on the writing process. In L.W. Gregg, &,E.R. Steinberg (Eds.), *Cognitive processes in writing*, (pp.129-137). Lawrence Erlbaum Associates, Hillsdale, NJ.

Wengelin, A. (2006). Examining pauses in writing: Theory, methods and empirical data. In K.H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing: Methods and application,* (pp.107-130). Oxford: Elsevier.

White, M.J., & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educational Psychology, 30*, 166-189.

Witte, S., & Cherry, R. (1986). Writing processes and written products in composition research. In C. Cooper, & S. Greenbaum (Eds.), *Studying writing: Linguistic approaches,* (pp.112-153). Newbury Park, CA: Sage.

Wood, S. N. (2006). *Generalized additive models. An introduction with R.* Boca Raton:Taylor & Francis Group.

Wieling, M. B. (2012*). A quantitative approach to social and geographical dialect variation*. Dissertation, University of Groningen.

# Appendices

## Contents

## Self-monitoring scale (Dutch)

*Instructies*

De onderstaande uitspraken hebben te maken met persoonlijke reacties op een aantal verschillende situaties. Er zijn geen uitspraken die hetzelfde zijn, dus het is de bedoeling dat je elke uitspraak zorgvuldig bekijkt voordat je antwoord geeft. Als een uitspraak voor jou waar is of meestal waar is, kruis dan het blokje voor waar aan. Als een uitspraak onwaar is of voor jou meestal niet waar is, kruis dan het blokje voor onwaar aan. Het is belangrijk dat je alle vragen zo eerlijk mogelijk probeert te antwoorden. Je antwoorden zullen natuurlijk vertrouwelijk behandeld worden.

| | | |
|---|---|---|
| 1 | Ik vind het moeilijk om het gedrag van andere mensen te imiteren. | ☐ Waar ☐ Onwaar |
| 2 | Op feestjes en in groepen doe ik geen moeite om die dingen te doen of te zeggen die andere mensen leuk vinden. | ☐ Waar ☐ Onwaar |
| 3 | Ik kan alleen achter ideeën staan waar ik al in geloof. | ☐ Waar ☐ Onwaar |
| 4 | Ik kan goed geïmproviseerde praatjes houden, zelfs over onderwerpen waar ik niet zoveel van weet. | ☐ Waar ☐ Onwaar |
| 5 | Soms probeer ik indruk te maken op mensen of mensen te vermaken door een show op te voeren. | ☐ Waar ☐ Onwaar |
| 6 | Ik zou waarschijnlijk een goede acteur/actrice zijn. | ☐ Waar ☐ Onwaar |
| 7 | In een groep sta ik bijna nooit in het middelpunt van de belangstelling. | ☐ Waar ☐ Onwaar |
| 8 | In verschillende situaties en met verschillende mensen, gedraag ik me vaak als heel verschillende personen. | ☐ Waar ☐ Onwaar |
| 9 | Ik ben er niet zo goed in ervoor te zorgen dat andere mensen me aardig vinden. | ☐ Waar ☐ Onwaar |
| 10 | Ik ben niet altijd de persoon die ik lijk te zijn. | ☐ Waar ☐ Onwaar |
| 11 | Ik zou mijn ideeën, of de manier waarop ik dingen doe, niet veranderen om bij iemand in de smaak te vallen. | ☐ Waar ☐ Onwaar |
| 12 | Ik denk er wel eens over om later iets te gaan doen als entertainer. | ☐ Waar ☐ Onwaar |
| 13 | Ik ben nooit goed geweest in spelletjes waarbij je moet improviseren. | ☐ Waar ☐ Onwaar |
| 14 | Ik heb er moeite mee me aan te passen aan verschillende mensen en verschillende situaties. | ☐ Waar ☐ Onwaar |
| 15 | Op een feestje laat ik het aan anderen over om grappen te maken en verhalen te vertellen. | ☐ Waar ☐ Onwaar |
| 16 | In een groep mensen voel ik me niet erg op mijn gemak: ik kom dan niet zo goed over als ik zou willen. | ☐ Waar ☐ Onwaar |
| 17 | Ik kan iedereen aankijken en met een uitgestreken gezicht een leugen vertellen. | ☐ Waar ☐ Onwaar |
| 18 | Ik kan mensen misleiden door te doen alsof ik ze aardig vind, terwijl dat helemaal niet zo is. | ☐ Waar ☐ Onwaar |

## Self-monitoring scale (English)

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.

1. I find it hard to imitate the behaviour of other people (F)
2. At parties and social gatherings, I do not attempt to say things that others will like (F)
3. I can only argue for ideas which I already believe (F)
4. I can make impromptu speeches on topics about which I have almost no information (T)
5. I guess I put on a show to impress or entertain others (T);
6. I would probably make a good actor (T)
7. In a group of people I am rarely the center of attention (F)
8. In different situations and with different people, I act like very different persons (T)
9. I am not particularly good at making other people like me (F)
10. I'm not always the person I appear to be (T)
11. I would not change my opinions (or the way I do things) in order to please someone else or win their favor (F)
12. I have considered being an entertainer (T)
13. I have never been good at games like charades or improvisational acting (F)
14. I have trouble changing my behavior to suit different people and different situations (F)
15. At parties I let others keep the jokes and stories going (F)
16. l feel a bit awkward in company and do not show up as well as I should (F)
17. I can look anyone in the face and tell a lie with a straight face (if for a right end) (T)
18. I may deceive people by being friendly when I really dislike them (T)

## Writing Beliefs Inventory (Dutch)

*Instructies* Vul hier je studentnummer in……..

De onderstaande uitspraken hebben betrekking op opvatting over schrijven. Er zijn geen uitspraken die hetzelfde zijn, dus het is de bedoeling dat je elke uitspraak zorgvuldig bekijkt voordat je aangeeft in welke mate je het eens bent met de uitspraak. Als je het heel erg met de uitspraak eens bent geef je dit aan door een kruisje te zetten in het meest rechter blokje (het laatste blokje). Als je het juist helemaal niet eens bent met een uitspraak kun je dit aangeven door een kruisje te zetten in het meest linker blokje (het eerste blokje).

| | | Helemaal mee oneens | | | | Helemaal mee eens |
|---|---|---|---|---|---|---|
| 1. | Goede schrijvers gebruiken veel citaten van deskundigen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 2. | Het belangrijkste doel van schrijven is het geven van informatie aan anderen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 3. | Een hoofddoelstelling van schrijven zou moeten zijn dat je zo weinig mogelijk aanpassingen hoeft te maken. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 4. | Schrijven zou zich moeten richten op de informatie uit boeken en artikelen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 5. | De sleutel tot goed schrijven is het accuraat rapporteren over wat deskundigen denken. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 6. | De belangrijkste reden om te schrijven is om te rapporteren over wat deskundigen denken over een onderwerp. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 7. | Schrijven vereist het steeds herzien van de tekst om te verbeteren wat is geschreven. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 8. | Schrijven is een proces waarbij veel verschillende emoties een rol spelen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 9. | Het is belangrijk om een karakteristieke schrijfstijl te ontwikkelen. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 10. | Goede schrijvers wijken weinig af van de informatie die zij nodig hebben voor een onderwerp. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 11. | Goed schrijven houdt in dat je veelvuldig aan het reviseren bent. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 12. | Schrijven brengt vaak intense ervaringen met zich mee. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 13. | Schrijven helpt me bij het beter begrijpen van de dingen waar ik over na denk. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 14. | Ik heb altijd het gevoel dat één extra revisie mijn tekst nog beter zal maken. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 15. | Schrijven helpt me om inzicht te krijgen in de complexiteit van ideeën. | ☐ | ☐ | ☐ | ☐ | ☐ |

16. Mijn gedachten en ideeën worden helderder wanneer ik schrijf en herschrijf. ☐ ☐ ☐ ☐ ☐
17. De gezichtspunten van schrijvers zouden duidelijk moeten worden uit hun schrijven. ☐ ☐ ☐ ☐ ☐
18. Schrijven is vaak een emotionele gebeurtenis. ☐ ☐ ☐ ☐ ☐
19. Schrijvers moeten zich onderdompelen in hun schrijven. ☐ ☐ ☐ ☐ ☐

## Writing Beliefs Inventory (English)

White, M.J. & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educational Psychology (30)*, 166-189.

1. Good writers include a lot of quotes from authorities in their writing
2. Writing's main purpose is to given other people information
3. A primary goal of writing should be to have to make as few changes as possible
4. Writing should focus around the information in books and articles
5. The key to successful writing is accurately reporting what authorities think
6. The most important reason to write is to report what authorities think about a subject
7. Writing requires going back over it to improve what has been written
8. Writing is a process involving a lot of emotion
9. It is important to develop a distinctive writing style
10. Good writers stick closely to the information they have about a topic
11. Good writing involves editing it many times
12. Writing often involves peak experiences
13. Writing helps me understand better what I am thinking about
14. I always feel that just one more revision will improve my writing
15. Writing helps me to see the complexity of ideas
16. My thoughts and ideas become more clear to me as I write and rewrite
17. Writers views should show through their writing
18. Writing is often an emotional experience
19. Writers need to immerse themselves in their writing

## Brief summary of the procedure of the experiment

- Administer WBI
- Assign topic: *"We are getting more and more dependent on the Computer and Internet. Is this a good development yes or no?"*
- List ideas about topic (10 minutes)
- Rate importance of ideas (as long as it takes).
- Rate how much they feel they know about the topic.
- Give instructions for writing task. (2 different planning tasks)

    - **Synthetic planning**: You get 5 minutes to prepare the writing task. At the end of the 5 minutes I want you to have written one sentence summing up your overall opinion about the topic. You can have more than one attempt to write this single sentence down, but you are not allowed to write anything else than this single sentence.
        - Then 30 minutes to write the article for the university newspaper on the computer. You have to write a well-structured article in the time available. During writing participants are allowed to consult their plan. Keystrokes are logged with Inputlog (Leijten & van Waes, 2006).

    - **Outline planning**: You get 5 minutes to prepare the writing task. At the end of the 5 minutes I want you to have made a structured outline for the text. In the outline I want you to order your ideas and to indicate which ideas are important and in which order they should appear in the text that you are about to write. You may start with making some notes, but at the end of the 5 minutes I want you to have made a structured outline.
        - Then 30 minutes to write the article for the university newspaper on the computer. You have to write a well-structured article in the time available. During writing participants are allowed to consult their plan. Keystrokes are logged with Inputlog (Leijte & van Waes, 2006).

- Ask to rate 'how much they feel they know about the topic now'.
- List ideas as they occur now. (10 minutes)
- Rate importance of ideas now (as long as it takes)
- Compare first list of ideas with second list
    - Take each idea in second list
    - Search for corresponding idea/ideas in first list
    - If present write down, otherwise leave blank
    - Rate similarity of idea/ideas (as group) to the corresponding idea in the second list.
- Debriefing

## Lijst met ideeën 1

*Instructies*
Schrijf al je ideeën op die te maken hebben met het onderwerp *"We worden steeds afhankelijker van de computer en van het internet. Is dit een goede ontwikkeling of niet?"* Je hoeft je daarbij geen zorgen te maken over hoe je deze ideeën opschrijft. Je mag ze in één zin formuleren, maar één enkel woord is ook voldoende. De enige beperking is dat de ideeën niet langer mogen zijn dan één zin. Wanneer één punt meerdere aspecten heeft en je kunt die verschillende aspecten niet meer in één zin formuleren, schrijf die aspecten dan op als aparte ideeën.

Je krijgt straks 10 minuten van me om deze lijst te maken. Het is de bedoeling dat je in die 10 minuten probeert om zo veel mogelijk relevante ideeën op te schrijven.

**Lijst met ideeën**

1.

2.

3.

{…}

24.

25.

26..

## Meting van belang 1

*Instructies*

Neem de lijst met ideeën door die je net opgesteld hebt en geef op dit formulier aan hoe belangrijk elk idee is voor het artikel dat je straks gaat schrijven. Gebruik de volgende schaal (1 = minst belangrijke punt, 5 = meest belangrijke punt).

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Minst belangrijke punt | | | | Meest belangrijkste punt |

**Idee nummer**

| | | |
|---|---|---|
| 1. | 11. | 21. |
| 2. | 12. | 22. |
| 3. | 13. | 23. |
| 4. | 14. | 24. |
| 5. | 15. | 25. |
| 6. | 16. | 26. |
| 7. | 17. | 27. |
| 8. | 18. | 28. |
| 9. | 19. | 29. |
| 10. | 20. | 30. |

## Meting onderwerp 1

*Instructies*
Geef aan hoeveel je weet over het volgende onderwerp:

*We worden steeds afhankelijker van de computer en van het internet. Is dit een goede ontwikkeling of niet?*

Geef aan hoeveel je weet over dit onderwerp op een schaal van 1 tot 7 (omcirkel het juiste antwoord: 1 = heel erg weinig, 7 = heel erg veel).

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Heel erg weinig | | | Niet veel, maar ook niet weinig | | | Heel erg veel |

## Meting onderwerp 2

*Instructies*
Geef aan hoeveel je NU weet over het volgende onderwerp:

*We worden steeds afhankelijker van de computer en van het internet. Is dit een goede ontwikkeling of niet?*

Geef aan hoeveel je nu weet over dit onderwerp op een schaal van 1 tot 7 (omcirkel het juiste antwoord: 1 = heel erg weinig, 7 = heel erg veel).

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Heel erg weinig | | | Niet veel, maar ook niet weinig | | | Heel erg veel |

## Lijst met ideeën 2

*Instructies*
Nu wil ik graag dat je alle ideeën opschrijft die NU relevant zijn voor de vraag of het een goede ontwikkeling is dat wij steeds afhankelijker worden van de computer en van het internet. Je hoeft weer geen zorgen te maken over hoe je deze ideeën opschrijft. Je mag ze in een zin formuleren, maar één enkel woord is ook voldoende. De enige beperking is dat de ideeën niet langer mogen zijn dan één zin. Wanneer één punt meerdere aspecten heeft en je kunt die verschillende aspecten niet meer in één zin formulieren, schrijf die aspecten dan op als aparte ideeën.

Je krijgt 10 minuten om deze lijst met ideeën te maken. Probeer in deze 10 minuten zoveel mogelijk relevante ideeën op te schrijven.

**Lijst met ideeën**

1.

2.

3.

{…}

24.

25.

26.

## Meting van belang 2

*Instructies*

Neem de lijst met ideeën door die je net gemaakt hebt en geef op dit formulier aan hoe belangrijk elk idee is. Gebruik hiervoor de volgende schaal (1 = minst belangrijke punt, 5 = meest belangrijke punt).

1_____2_____3_____4_____5

Minst belangrijke punt Meest belangrijkste punt

**Idee nummer**

| | | |
|---|---|---|
| 1. | 11. | 21. |
| 2. | 12. | 22. |
| 3. | 13. | 23. |
| 4. | 14. | 24. |
| 5. | 15. | 25. |
| 6. | 16. | 26. |
| 7. | 17. | 27. |
| 8. | 18. | 28. |
| 9. | 19. | 29. |
| 10. | 20. | 30. |

## Mate van overeenstemming

*Instructies*

Nu ga je de ideeën uit de lijst die je voor het schrijven hebt opgesteld vergelijken met de ideeën uit de lijst die je na het schrijven hebt gemaakt. De nummers van de ideeën uit de tweede lijst staan weergegeven in de linker kolom. Het is bedoeling dat je het idee dat correspondeert met dit nummer bekijkt en dat je daarna je eerste lijst erbij pakt en kijkt of er ideeën zijn die overeenkomen met dit idee uit de tweede lijst. Als je in de eerste lijst corresponderende ideeën vindt, schrijf dan het nummer(s) van dit corresponderende punt(en) op in de daarvoor bestemde tweede kolom.

Van alle punten die overeenstemming vertonen, schrijf je op hoeveel ze met elkaar overeenkomen door gebruik te maken van de onderstaande schaal. Je kunt hiervoor de derde kolom gebruiken met de titel *'mate van overeenstemming'*. Als er meer dan 1 punt correspondeert met een punt uit lijst 2 beoordeel dan de overeenstemming van de groep ideeën. Als er geen corresponderende ideeën te vinden zijn in de eerste lijst laat dan de tweede kolom leeg en vul in de derde kolom schaal 6 in (geen overeenstemming).

1_____ ___2____ ___3____ ___4____ ___5____ _____6

**Identiek punt** **Geen overeenstemming**

| **Nummer van het idee in lijst 2** | **Vergelijkbare ideeën in lijst 1** | **Mate van overeenstemming** |
|---|---|---|
| 1. | | |
| 2. | | |
| 3. | | |
| {..} | | |
| 24. | | |
| 25. | | |
| 26. | | |

## Listing of ideas 1

*Instructions*
Write all your ideas down related to the topic: *"We are getting more and more dependent on the computer and the Internet. Is this a good development yes or no?"* Don't worry about how well expressed these are, or whether anyone else could understand what you mean, just jot down a sentence, word or phrase representing the idea for yourself. A point can be as short as a single word but no longer than sentence. If a point has several different aspects which you can't write as a single sentence then put these down as separate points.

You will have ten minutes to do this in. The aim is to write down as many relevant ideas as you can think of in the time available.

**List of points**

1.

2.

3.

{..}

24.

25

26.

## Rating of importance 1

*Instructions*

Read through the list of points you have just made and rate how important each idea is on the form below. Use the following scale (note: 5 = more important ideas, 1 = less important ideas)

1_____2_____3_____4_____5

**Least important points** **Most important points**

**Point number**

| | | |
|---|---|---|
| 1. | 11. | 21. |
| 2. | 12. | 22. |
| 3. | 13. | 23. |
| 4. | 14. | 24. |
| 5. | 15. | 25. |
| 6. | 16. | 26. |
| 7. | 17. | 27. |
| 8. | 18. | 28. |
| 9. | 19. | 29. |
| 10. | 20. | 30. |

## Rating of prior understanding

*Instructions*
Please indicate how much you feel you know about the following topic:

*We are becoming more and more dependent on the computer and Internet. Is this a good development yes or no?*

How much do you feel you know about this topic on a scale from 1 (extremely little) to 7 (a great deal)?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Extremely little | | | A moderate amount | | | A great deal |

## Rating of post understanding

*Instructions*
Please indicate how much you feel you know about the following topic NOW:

*We are getting more and more dependent on the computer and Internet. Is this a good development yes or no?*

How much do you feel you know now about this topic on a scale from 1 (extremely little) to 7 (a great deal)?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Extremely little | | | A moderate amount | | | A great deal |

## Listing ideas 2

*Instructions*
I want you to write down a list of all the ideas you can think of *NOW* relevant to the question: *We are getting more and more dependent on the computer and Internet. Is this a good development yes or no?* Don't worry about how well expressed these are, or whether anyone else could understand what you mean, just jot down a sentence, word or phrase representing the idea for yourself. A point can be as short as a single word but no longer than sentence. If a point has several different aspects which you can't write as a single sentence then put these down as separate points.

You will have ten minutes to do this in. The aim is to write down as many relevant ideas as you can think of in the time available.

**List of points**

1.

2.

3.

{…}

24.

25.

26.

## Rating of importance 2

*Instructions*

Read through the list of points you have just made and rate how important each idea is on the form below. Use the following scale (note: 5 = more important ideas, 1 = less important ideas)

1_____2_____3_____4_____5

**Least important points** **Most important points**

**Point number**

| | | |
|---|---|---|
| 1. | 11. | 21. |
| 2. | 12. | 22. |
| 3. | 13. | 23. |
| 4. | 14. | 24. |
| 5. | 15. | 25. |
| 6. | 16. | 26. |
| 7. | 17. | 27. |
| 8. | 18. | 28. |
| 9. | 19. | 29. |
| 10. | 20. | 30. |

## Degree of correspondence

*Instructions*
I want you to compare the ideas in the lists you produced before and after writing. The numbers of the points in list 2 are written in the left hand column. I want you to take each of the points in this list in turn, and read through list 1 to see if there are any corresponding point(s). If there are, write the number(s) of the corresponding point (s) in the second column.

If there are any corresponding points(s) please rate how similar they are to the point in list 2, using the scale below. Put this rating in the third column headed "degree of correspondence".

If there is more than one point in list 1 corresponding to a point in list 2, then rate the similarity of the group as a whole to that point (i.e. give it a single rating). If there are no corresponding points in list 1 then leave the second column blank and put a rating of 6 in the third column.

1_____ ___2____ ___3____ ___4____ ___5____ _____6
**Identical point** **No corresponding point**

| **Number of point in list 2** | **Corresponding point (s) in list 1** | **Degree of correspondence** |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| {…} | | |
| 24 | | |
| 25 | | |
| 26 | | |

## Nederlandse samenvatting

In dit proefschrift staat de vraag centraal op welke wijze schrijven bijdraagt aan leereffecten. Om meer inzicht te krijgen in deze kwestie worden de twee belangrijkste stromingen binnen het cognitieve procesonderzoek naar schrijven geëvalueerd (Flower & Hayes, 1980; Bereiter & Scardamalia, 1987; Galbraith, 2009). Om de claims over leereffecten door schrijven van de concurrerende modellen uit deze benaderingen te kunnen vergelijken, zijn binnen dit onderzoek naar de ontwikkeling van begrip door schrijven steeds zowel de achterliggende cognitieve processen als de subjectieve beoordelingen van studenten over hun begrip als uitgangspunt genomen voor analyses. Om de schrijfprocessen te kunnen onderzoeken, is gebruik gemaakt van toetsaanslagregistratiedata met behulp van Inputlog (Leijten & Van Waes, 2006). Om het begrip vast te stellen, is gebruik gemaakt van een subjectieve beoordeling door de schrijvers zelf. Hieronder zal eerst het theoretische kader van dit onderzoek uiteengezet worden. Daarna wordt kort ingegaan op de vier verschillende empirische studies die centraal staan in dit proefschrift. Ten slotte zal er een samenvatting worden gegeven van de belangrijkste bevindingen van het onderzoek.

*Het probleemoplossingsparadigma*

Binnen de cognitieve psychologie is 'probleemoplossen' een centraal begrip. Men gaat ervan uit dat individuen bij het oplossen van problemen op een strategische wijze gebruikmaken van de informatie die zij tot hun beschikking hebben (Bochardt, 1984). Flower en Hayes (1980) hebben dit probleemoplossingsparadigma toegepast op het schrijven van teksten en hebben schrijven hierbij getypeerd als een cognitief probleem dat de schrijver op de een of andere manier moet oplossen. Door dit en aanverwant onderzoek is de visie op het schrijfgedrag vanaf de jaren '80 veranderd. Schrijven wordt niet langer gezien als een lineaire omzetting van gedachten in geschreven taal, maar als een recursief proces waarin verschillende factoren een rol spelen en waarin planning, formuleren en revisie elkaar afwisselen (Flower & Hayes, 1980). Tijdens een schrijftaak moet een schrijver verschillende schrijfprocessen organiseren. Zo moet een schrijver tijdens het schrijven nadenken over problemen die hij tegenkomt en moet hij nadenken over mogelijke oplossingen voor die problemen. Ook moet een schrijver tijdens het schrijven nadenken over ideeën, over de opbouw van de tekst en over de wijze van formuleren. Verder moet de schrijver rekening houden met de opdracht en met het publiek waarvoor de tekst geschreven wordt. Als laatste moet de schrijver schrijfdoelen bepalen en plannen ontwikkelen om die schrijfdoelen te bereiken. Dit alles maakt schrijven een complex proces (Bochardt, 1984).

Cognitieve procesmodellen, zoals dat van Hayes en Flower (1980), maar zeker ook dat van Bereiter en Scardamalia (1987), stellen dat schrijven kan bijdragen aan leren door expliciete denkprocessen, zoals plannen en reviseren. Het meest aangehaalde cognitieve schrijfproces van Hayes en Flower (1980) bestaat uit drie deelprocessen: 'plannen', 'formuleren' en 'reviseren'. Binnen de processen 'plannen' en 'reviseren' zijn verschillende subprocessen te onderscheiden. Hayes en Flower onderscheiden in het planningsproces de subprocessen genereren, organiseren en het bepalen van doelen. In het reviseerproces worden verder de subprocessen lezen en redigeren doorlopen. Al deze processen staan onder de controle van een monitor. Hayes en Flower benadrukken vooral de dynamische structuur van het schrijfproces. De schrijver doorloopt de verschillende cognitieve processen meerdere malen. Elk cognitief proces kan op elk willekeurig moment tijdens het schrijven optreden en alle processen kunnen door elkaar opgevolgd worden. Het schrijfproces is daardoor zo complex dat schrijvers nooit aan alle aspecten tegelijkertijd aandacht kunnen besteden. De beperkte capaciteit van het werkgeheugen maakt dat voor hen onmogelijk (Kellogg, 1994).

Om te voorkomen dat het werkgeheugen van een schrijver overbelast raakt, maken schrijvers gebruik van schrijfstrategieën. Bij strategieën kan bijvoorbeeld gedacht worden aan het maken van een gestructureerd plan voor de te schrijven tekst. Het maken van zo'n plan stelt de schrijver in staat om ideeën voor de tekst te genereren alvorens deze uit te moeten schrijven. Concreet betekent dit dat de planningsactiviteiten gescheiden worden van de formuleeractiviteiten. Het voordeel van het scheiden van deze twee cognitief belastende taken is dat de schrijver zich tijdens het schrijven geheel kan concentreren op het formuleren, omdat hij zich niet meer hoeft bezig te houden met het bedenken van het onderwerp en de wijze waarop hij zijn informatie zal ordenen. Een bijkomend en belangrijk voordeel is dat deze methode van schrijven kan bijdragen aan de kwaliteit van de geschreven tekst. In een serie experimenten heeft Kellogg (1999) aangetoond dat het maken van een schrijfplan voor aanvang van de schrijftaak leidt tot een betere tekstkwaliteit.

Het onderzoek dat is verricht binnen het kader van het model van Hayes en Flower (1980) is over het algemeen beschrijvend van aard. Het wordt overwegend uitgevoerd volgens het zogenaamde 'expert- versus-beginners-paradigma', waarbij de schrijfprocessen van goede of ervaren schrijvers worden vergeleken met de schrijfprocessen van slechte of onervaren schrijvers. Bij de bestudering van de cognitieve schrijfprocessen wordt doorgaans gebruik gemaakt van hardopdenkprotocollen. Wanneer er tijdens de bestudering van de hardopdenk-protocollen bewijs wordt gevonden dat schrijvers de ideeën voor de tekst ontlenen aan retorische doelen (bijvoorbeeld het informeren of overtuigen van je publiek)

wordt aangenomen dat de schrijver probleemoplossend bezig is. Hierdoor ontstaat een situatie waarin de schrijver mogelijkerwijs kan leren van het schrijven. Aangezien onervaren schrijvers vaak niet in staat zijn om rekening te houden met retorische doelen van de tekst, wordt aangenomen dat de schrijfprocessen van onervaren schrijvers niet bijdragen aan de ontwikkeling van het begrip van de schrijver. Onervaren schrijvers richten zich tijdens het schrijven meer op de inhoud van de tekst. Het schrijfproces bestaat dan uit het ophalen en meteen opschrijven van bestaande kennis uit het werkgeheugen zonder dat deze informatie wordt geëvalueerd of verwerkt aan de hand van de tekstdoelen. Het is belangrijk om hier op te merken dat het onderzoek dat binnen dit paradigma wordt uitgevoerd meestal niet expliciet meet of het schrijven daadwerkelijk heeft bijgedragen aan de begripsontwikkeling van de schrijver. Vaak wordt aangenomen dat schrijvers hun begrip ontwikkeld hebben, wanneer probleemoplossingsprocessen worden geobserveerd.

*De procestheorie van Bereiter en Scardamalia*

Ook Bereiter en Scardamalia (1987) benadrukken in hun procestheorie dat schrijven een complexe taak is. Zij geven aan dat de moeilijkheid er vooral in schuilt dat de communicatieve partner ontbreekt. Uit het onderzoek van Flower en Hayes (1980) is al naar voren gekomen dat vooral onervaren schrijvers het lastig vinden om rekening te houden met retorische doelen van schrijven. De theorie van Bereiter en Scardamalia beschrijft de ontwikkeling van het schrijfproces aan de hand van de termen *knowledge-telling* en *knowledge-transforming* (Bereiter & Scardamalia, 1987).

*Knowledge-telling* verwijst naar een strategie die toegepast wordt door onervaren schrijvers. Deze strategie wordt gebruikt om te kunnen omgaan met de beperkingen van het werkgeheugen. Onervaren schrijvers halen tijdens het schrijven een idee op uit het geheugen en schrijven dit vervolgens op. Op basis van dit idee wordt er een nieuw idee voortgebracht en dit proces wordt telkens herhaald tot de ideeën van de schrijver uitgeput zijn. Deze manier van schrijven wordt zodoende gekarakteriseerd door het ontbreken van plannings-, structurerings-, en revisieprocessen (Bereiter & Scardamalia, 1987).

Tegenover *knowledge-telling* staat *knowledge-transforming*. Schrijvers die gebruik maken van deze strategie verwerken gedachten en ideeën verschillende malen voordat ze op papier gezet worden. Bij de verwerking van de informatie wordt zowel aandacht besteed aan de inhoudelijke kant als aan de retorische kant van de tekst. In tegenstelling tot *knowledge-telling* kenmerkt *knowledge-transforming* zich door de aanwezigheid van plannings-, structurerings-, en revisieprocessen. Aangezien dit het schrijfgedrag is dat van ervaren schrijvers verwacht wordt, wordt *knowledge-*

*transforming* ook wel gezien als het standaardmodel. Het *knowledge-transforming*-model kan, net als het procesmodel van Flower en Hayes, gezien worden als een probleemoplossingsmodel en wordt door de naam van het model bij uitstek gebruikt om te illustreren hoe schrijven kan bijdragen aan leren.

*Het dual process-model van Galbraith*

Tegenover deze twee gevestigde procesmodellen uit de jaren '80 staat het cognitieve *dual process*-model van Galbraith (1992; 2009). Ook Galbraith claimt dat schrijven kan bijdragen aan leren, maar het *dual process*-model verklaart de leereffecten van schrijven uit impliciete taalproductieprocessen. De naam van dit cognitieve model geeft aan dat Galbraith vindt dat bij schrijven twee processen van belang zijn. Galbraith erkent, net als Hayes en Flower (1980) en Bereiter en Scardamalia (1987), dat probleemoplossingsprocessen een belangrijke rol spelen bij schrijven. Galbraith claimt echter dat het bij probleemoplossingsprocessen alleen gaat om het *ophalen* van al bestaande kennis en dat het eigenlijke schrijven in die gevallen bestaat uit het uitwerken van een vooraf bedacht schema. Door te denken alvorens ideeën op papier te zetten, wordt de mogelijkheid om tijdens het schrijven ideeën op te doen onbenut gelaten, aldus Galbraith. Om die reden zal dit proces volgens Galbraith niet bijdragen aan de ontwikkeling van begrip door de schrijver.

Tegenover het strategische planningsproces waarbij bestaande kennis wordt opgehaald uit het werkgeheugen, zet Galbraith een impliciet taalproductieproces waarbij het gaat om een synthese van ideeën in het semantische geheugen. Ideeën ontstaan zogezegd tijdens de taalproductie: pas *tijdens* het schrijven bedenken schrijvers wat voor informatie ze willen opnemen in de tekst en hoe zij die informatie willen orderenen. Zodoende ontdekt de schrijver ook zelf pas tijdens het schrijven wat voor ideeën hij heeft over een bepaald onderwerp. Het is dit proces van synthese dat Galbraith verantwoordelijk acht voor de ontwikkeling van begrip tijdens het schrijven.

In tegenstelling tot het onderzoek binnen het probleemoplossingsparadigma heeft Galbraith wel expliciet getest of schrijven bijdraagt aan de begripsontwikkeling van schrijvers. Aan de hand van een subjectieve begripsmeting heeft Galbraith in een serie experimenten laten zien dat schrijven voor verschillende typen schrijvers – *high* en *low self-monitors* – niet op gelijke wijze bijdraagt aan leereffecten. *Self-monitoring* is een persoonlijkheidsconstruct dat is ontwikkeld door Gangestad en Snyder (1986). *High self-monitors* zijn mensen die zich bij uitstek bewust zijn van de sociale context waarin zij zich bevinden. Van dit type mensen wordt aangenomen dat zij hun gedrag aanpassen aan de situatie waarin zij zich bevinden. *Low self-monitors* zijn mensen die zich meer oriënteren op hun eigen opvattingen en overtuigingen. Hierdoor zijn zij

minder geneigd om hun gedrag aan te passen aan de sociale situatie waarin zij zich bevinden. Galbraith heeft dit *self-monitoring* construct gebruikt om onderscheid te kunnen maken in de mate waarin schrijvers rekening houden met de retorische context van de schrijftaak. Galbraith neemt aan dat *high self-monitors* eerder geneigd zijn om hun tekstdoelen te ontlenen aan de hand van de retorische context, terwijl *low self-monitors* eerder geneigd zijn prioriteit te geven aan het opschrijven van hun persoonlijke opvatting en eerder al schrijvend zullen denken.

Naast het onderscheiden van verschillende schrijverstypen heeft Galbraith in zijn onderzoek ook veelvuldig verschillende schrijftaken gemanipuleerd (Galbraith, 1992; 2006; 2009). Hierbij kan onder andere gedacht worden aan het schrijven van een volwaardig artikel (*full prose*) tegenover het maken van notities voor een artikel (*notes*) (Galbraith, 1992) of aan verschillende voorbereidende planningstaken, zoals het maken van een gestructureerd plan voor aanvang van de schrijftaak (*outline planning*) versus het opschrijven van één samenvattende zin over je ideeën (*synthetic planning*) (Galbraith, Torrance, Hallam, 2006). Het idee hierachter is dat deze schrijftaken andere taalproductieprocessen zullen uitlokken. Wanneer een schrijver eerst nadenkt alvorens dingen op te schrijven (*outline planning*), zal hij tijdens het schrijven meer aandacht kunnen besteden aan het schrijven zelf. Dit proces beperkt echter de mogelijkheid om nog nieuwe ideeën op te doen tijdens het schrijven. Wanneer schrijvers, aan de andere kant, voor aanvang van de schrijftaak alleen één samenvattende zin opschrijven, is het mogelijk om nog tijdens het schrijven te bedenken welke ideeën ze in hun teksten willen behandelen. Zoals al eerder aangestipt, veronderstelt Galbraith dat alleen deze laatste manier van schrijven bijdraagt aan de ontwikkeling van begrip door schrijven.

In een serie experimenten heeft Galbraith empirisch getest hoe schrijven bijdraagt aan leren door aan *high* en *low self-monitors* te vragen om zowel voor aanvang als na afloop van de schrijftaak een lijst met ideeën op te stellen en om aan te geven hoeveel zij denken te weten over het onderwerp waarover zij moeten schrijven. Deze methode stelt Galbraith in staat om te onderzoeken of schrijvers ideeën ontwikkelen door schrijven en in welke mate zij leren van schrijven. De resultaten van dit onderzoek hebben aan het licht gebracht dat *high self-monitors*, de mensen die rekening houden met de retorische context, meer nieuwe ideeën ontwikkelen na het schrijven van korte notities en expliciete planningstaken. Deze resultaten sluiten aan bij het probleemoplossingsparadigma dat claimt dat leren door schrijven verklaard kan worden door expliciete planningsprocessen. Galbraith heeft echter ook laten zien dat deze nieuwe ideeën van *high self-monitors* niet samengaan met de ontwikkeling van begrip. Verder is uit zijn onderzoek naar voren gekomen dat *low self-monitors*, de mensen die georiënteerd zijn op de persoonlijke expressie van hun ideeën, meer

nieuwe ideeën ontwikkelen na het schrijven van volwaardige artikelen en dat deze nieuwe ideeën wél samengaan met de begripsontwikkeling van de schrijvers.

Op basis van deze empirische resultaten heeft Galbraith (2009) het *dual process*-model ontwikkeld, dat claimt dat zowel expliciete planningsprocessen als impliciete taalproductieprocessen een rol spelen bij schrijven, maar dat alleen impliciete taalproductieprocessen leiden tot de leereffecten van schrijven. Door expliciet te meten hoeveel schrijvers denken te weten voor en na aanvang van een schrijftaak, geeft het onderzoek van Galbraith meer inzicht in de wijze waarop schrijven bij kan dragen aan begripsontwikkeling. Het is echter belangrijk om op te merken dat Galbraith de achterliggende cognitieve schrijfprocessen niet onderzoekt.

*Doelstellingen voor dit promotieonderzoek*

Samenvattend kan gesteld worden dat de probleemoplossingsmodellen veronderstellen dat schrijven kan bijdragen aan begripsontwikkeling door expliciete denkprocessen. Het *dual process*-model van Galbraith (2009) verklaart de leereffecten van schrijven uit impliciete taalproductieprocessen. Daarnaast verschillen de beide benaderingen qua methode. Waar binnen de ene benadering gedetailleerd inzicht wordt gegeven in de achterliggende cognitieve processen, maar een begripsmeting ontbreekt, wordt binnen de andere benadering een begripsmeting afgenomen, maar ontbreekt het vastleggen van de achterliggende cognitieve schrijfprocessen. In de studies in dit proefschrift wordt in kaart gebracht hoe schrijven kan bijdragen aan de ontwikkeling van begrip door zowel te kijken naar schrijfprocessen als naar de manier waarop schrijvers hun ideeën ontwikkelen en leren door schrijven.

*Korte schets van de methodologische opzet van het onderzoek*

Om inzicht te verkrijgen in de vraag hoe schrijven bijdraagt aan leren en om daarbij de assumpties van beide cognitieve benaderingen te kunnen evalueren was het noodzakelijk om het ontwerp van de studies van Galbraith te repliceren. Zodoende is er ook in dit onderzoek gewerkt met twee typen schrijvers – *high* en *low self-monitors* – en is er ook zowel voor aanvang als na afloop van de schrijftaak aan de schrijvers gevraagd om hun kennis over het onderwerp te beoordelen en om een lijst met hun ideeën over het onderwerp van de schrijftaak op te stellen. Het onderwerp van de schrijftaak was *onze groeiende afhankelijkheid van de computer en het internet*. Aan 42 *high self-monitors* (schrijvers waarvan aangenomen wordt dat ze rekening houden met retorische doelen) en 42 *low self-monitors* (schrijvers waarvan aangenomen wordt dat ze meer belang hechten aan de expressie van hun eigen ideeën) is gevraagd om een artikel voor de universiteitskrant te schrijven over dit onderwerp. De schrijftaak is uitgevoerd onder twee verschillende planningscondities: het maken van een

gestructureerde opzet (*outline planning*) versus het vooraf formuleren van een samenvattende zin (*synthetic planning*). Er is gekozen voor deze twee planningstaken om zo nauw mogelijk aan te sluiten bij de experimentele opzet van eerder onderzoek van Galbraith, maar ook om de uitvoering van en de eisen aan de te schrijven tekst hetzelfde te kunnen houden over de verschillende schrijfcondities. Alleen met één en dezelfde schrijftaak was het mogelijk om de achterliggende schrijfprocessen te onderzoeken *en* de tekstkwaliteit van de geschreven teksten te beoordelen. Om toch de verschillende schrijfprocessen uit te lokken die centraal staan in het onderzoek van Galbraith – dus aan de ene kant 'de eerst denken, dan schrijven aanpak' en aan de andere kant 'de al schrijvend denken aanpak' – is er voor twee verschillende planningstaken gekozen. Er wordt aangenomen dat de *outline planningstaak* ervoor zal zorgen dat schrijvers van tevoren vaststellen wat ze willen zeggen en in welke volgorde ze deze informatie willen presenteren. Schrijvers denken dan over de retorische doelen van de tekst na zonder dat er tegelijkertijd werkgeheugencapaciteit wordt opgeëist voor het formuleren van ideeën in volwaardige zinnen. Het generen van ideeën wordt zo gescheiden van het formuleren. Tijdens het schrijven kunnen schrijvers zich zodoende concentreren op het formuleren. De *synthetic planningstaak* stelt schrijvers niet in staat om van tevoren een gestructureerd plan op te stellen voor de te schrijven tekst. Deze schrijvers zullen dus tijdens het schrijven nog moeten bedenken wat voor informatie zij willen opnemen in hun tekst. In deze schrijfconditie is het daardoor mogelijk om de impliciete taalproductieprocessen, die volgens Galbraith zo belangrijk zijn voor de leereffecten van schrijven, hun werk te laten doen. Voor de registratie van de cognitieve schrijfprocessen is tijdens het schrijven van het artikel voor de universiteitskrant gebruik gemaakt van het toetsaanslag-registratieprogramma Inputlog (Leijten & Van Waes, 2006).

De keuze om te werken met toetsaanslagregistratiedata is belangrijk geweest in de opzet van dit promotieonderzoek. Deze methode biedt de mogelijkheid om schrijfprocessen digitaal te registreren zonder dat de denkprocessen van de proefpersonen hier hinder van ondervinden. Ongemerkt wordt informatie verzameld over de duur van pauzes voor aanvang van elke toetsaanslag. Tevens wordt geregistreerd waar proefpersonen tekst wissen, herschrijven en wanneer ze op een eerdere locatie in de tekst een revisie uitvoeren. Na het voltooien van de schrijftaak kan heel gedetailleerd worden bekeken hoe de schrijfopdracht tot stand is gekomen. Deze methode en de opzet van het experiment maken het zodoende mogelijk om meer inzicht te verkrijgen in de vraag hoe schrijven bijdraagt aan leren en welke cognitieve processen hierbij een rol spelen. In de hierop volgende paragrafen zal achtereenvolgens uiteengezet worden hoe de verschillende studies in dit proefschrift bijdragen aan de beantwoording van dit vraagstuk.

*Schrijfopvattingen en hun effect op tekstkwaliteit en begripsontwikkeling*

Het *self-monitoring* construct dat centraal staat in de studies van Galbraith is niet specifiek gericht op schrijven. Om die reden is er binnen dit promotieonderzoek gezocht naar een vragenlijst die de opvattingen van schrijvers meer specifiek gericht op schrijven in kaart kon brengen. Daarvoor is gebruik gemaakt van de *Writing Beliefs Inventory* (White & Bruning, 2005). Hoofdstuk 3 evalueert de *Writing Beliefs Inventory* zoals die ontwikkeld is door White en Bruning (2005). In het hoofdstuk wordt gerapporteerd over een experiment dat test welke invloed impliciete opvattingen over schrijven hebben op schrijfprocessen, tekstkwaliteit en het ontwikkelen van begrip door schrijven.

De *Writing Beliefs Inventory* meet de impliciete schrijfopvattingen van schrijvers. Deze inventaris maakt onderscheid tussen *transmissional beliefs* en *transactional beliefs*. In het eerste geval wordt schrijven gekarakteriseerd als het overbrengen van informatie uit bronnen en vindt er geen interactie plaats tussen de kennis van de schrijver en de informatie in de bronnen. Een typisch voorbeeld van een *transmissional belief* is: "Goede schrijvers gebruiken veel citaten van deskundigen". De schrijver heeft de opvatting dat informatie uit bronnen zo correct mogelijk op papier gezet dient te worden. In het tweede geval, bij een *transactional belief* dus, vindt er meer interactie plaats tussen de informatie uit de bronnen en de persoonlijke kennis en opvattingen van de schrijver. In dit geval ziet de schrijver het als zijn taak om de informatie uit bronnen te evalueren en te verwerken alvorens de informatie in de tekst op te nemen. Een typisch voorbeeld van een *transactional belief* is: "Goed schrijven houdt in dat je veelvuldig aan het reviseren bent".

De resultaten van deze studie laten allereerst zien dat de interpretatie van de aard van *transmissional* en *transactional beliefs* door White en Bruning zelf niet geheel doorzichtig is. In de resultatensectie worden daarom suggesties gedaan voor de verbetering van de vragenlijst. Wel ondersteunen de resultaten de bevindingen van White en Bruning (2005) dat impliciete opvattingen over schrijven nauw samenhangen met tekstkwaliteit en de ontwikkeling van begrip door schrijvers. Evenals bij White en Bruning (2005) laten de resultaten zien dat schrijvers die de opvatting hebben dat schrijven meer is dan het rapporteren over bronnen betere tekstkwaliteit produceren dan schrijvers die deze opvatting niet hebben. Bovendien laten de resultaten zien dat het effect van de voorbereidingstaak beïnvloed wordt door de impliciete opvattingen van schrijvers. Schrijvers met *transmissional beliefs* zijn gebaat bij het maken van een gestructureerd plan van aanpak voor aanvang van de schrijftaak en produceren betere kwaliteit teksten in de *outline planning* conditie dan in de *synthetic planning* conditie. Voor schrijvers met *transactional beliefs* heeft het maken van een gestructureerd plan voor aanvang van de schrijftaak geen meerwaarde voor

de kwaliteit van de tekst. Daarnaast komt uit de resultaten naar voren dat schrijvers moeten geloven dat schrijven meer is dan het overbrengen van informatie uit bronnen om het daadwerkelijk te kunnen gebruiken als een instrument om het eigen begrip te ontwikkelen. Dit resultaat heeft implicaties voor het onderwijs: het lijkt erop te wijzen dat er binnen het schrijfonderwijs rekening gehouden dient te worden met de impliciete opvattingen van schrijvers.

*Reflecties op procedures en metingen*

In de studie die gepresenteerd wordt in hoofdstuk 4 wordt gereflecteerd op het gebruik en de toepassing van toetsaanslagregistratiedata. Hoewel toetsaanslag-registratie een veelbelovende methode is voor het bestuderen van cognitieve schrijfprocessen, is het doorgaans moeilijk om uit de ruwe data van toetsaanslagen betekenisvolle metingen af te leiden die iets zeggen over de achterliggende cognitieve processen. Deze studie beschrijft de procedures en metingen die binnen het kader van dit promotieonderzoek ontwikkeld zijn voor de analyses van de toetsaanslag-registratiedata. Het doel van de ontwikkeling van deze procedures was om een betere verbinding te bewerkstelligen tussen de geregistreerde toetsaanslagen en de achterliggende cognitieve processen. Deze metingen en procedures worden in de studie vervolgens gebruikt om pauzegedrag te analyseren, om *language bursts* te evalueren en om naar het revisiegedrag van schrijvers te kijken. De term *language burst* verwijst hier naar een aaneengesloten stuk geschreven taalproductie. Mensen produceren taal over het algemeen in 'horten en stoten'. Aaneengesloten stukken taalproducties worden bijvoorbeeld onderbroken met pauzes om te bedenken hoe de zin verder geformuleerd moet worden (*Pause-burst*, of *P-burst*). Daarnaast kan de taalproductie onderbroken worden door een revisie (*Revisie-burst*, of *R-burst*).

De resultaten van deze studie laten zien dat de verschillende typen *bursts* bij het schrijven vragen om een ander classificatiesysteem dan de gangbare *bursts* die onderscheiden worden in hardopdenkprotocollen. In het gangbare classificatie-systeem wordt alleen gekeken naar de manier waarop een *burst* afgesloten wordt. Voor toetsaanslagregistratiedata is het echter van belang om rekening te houden met zowel de manier waarop een *burst* geïnitieerd wordt als de manier waarop de *burst* beëindigd wordt. Daarnaast wordt in deze studie aangetoond dat *mixture models* beter in staat zijn om de aard van de scheve verdeling van pauzetijden te karakteriseren. Vervolgens blijkt dat schrijvers die langer pauzeren alvorens een nieuwe zin te formuleren kortere, maar meer vloeiende *bursts* produceren. Dit druist in tegen de gangbare observatie dat de lengte van de pauze de lengte van de *burst* kan voorspellen. Als laatste is er een *principal component analysis* uitgevoerd om de verschillende schrijfprocesmetingen te reduceren tot drie globale componenten. Eén

component karakteriseert geplande tekstproductie, één component representeert revisie binnen zinnen en als laatste wordt een component onderscheiden die staat voor revisies van de globale tekststructuur. Soortgelijke componenten worden in de volgende studie gebruikt om de schrijfprocessen van schrijvers te analysen.

*Schrijven: een proces van ontdekken*

In hoofdstuk 5 staat de hoofdstudie van dit promotieonderzoek centraal. De probleemoplossingsmodellen van het schrijfproces, zoals dat van Bereiter en Scardamalia (1987), stellen dat schrijven aan begripsontwikkeling kan bijdragen door expliciete denkprocessen. Het *dual process*-model van Galbraith (2009) claimt dat niet alleen expliciete denkprocessen, maar ook impliciete taalproductieprocessen bij leren door schrijven een rol spelen. In deze studie wordt in kaart gebracht hoe schrijven kan bijdragen aan leren door zowel te kijken naar schrijfprocessen als naar het ontwikkelen van begrip door schrijven. Daarnaast wordt er gekeken naar tekstkwaliteit.

De resultaten van deze studie laten zien dat zowel expliciete denkprocessen als meer impliciete taalproductieprocessen een belangrijke rol spelen bij schrijven. Deze resultaten ondersteunen de claims van het *dual process*-model (Galbraith, 2009). Bovendien laten de resultaten zien dat leren door schrijven en het produceren van een goede tekst niet hetzelfde zijn. De analyses van de schrijfprocessen laten zien dat de verschillende manieren waarop *low* en *high self-monitors* hun schrijfprocessen organiseren verschillende effecten hebben op het ontwikkelen van begrip en op het produceren van een goede tekst. Voor *high self-monitors* geldt dat geplande zinsproductie in combinatie met een lineaire organisatie van de tekst bijdraagt aan het produceren van een kwalitatief goede tekst. Om te leren van schrijven is het voor deze schrijvers echter van belang dat zij ongeplande zinsproductie combineren met een non-lineaire organisatie van de tekst als geheel. Voor *low self-monitors* geldt over het algemeen dat zij een betere tekstkwaliteit bereiken *en* leren van schrijven wanneer zij ongeplande zinsproductie combineren met een non-lineaire organisatie van de tekst als geheel. Deze resultaten tonen aan dat tekstkwaliteit en leren door schrijven om een andere aanpak van schrijven vragen.

*Het gebruiken en veranderen van ideeën*

In de laatste empirische studie uit dit proefschrift staan de ideeën die de proefpersonen geproduceerd hebben voor aanvang van en na afloop van de schrijftaak centraal. Eerder onderzoek (Galbraith 1992, 2006) richtte zich bij de analyse van ideeën vooral op de vraag of er meer of minder nieuwe ideeën geproduceerd worden door *high* en *low self-monitors* onder bepaalde schrijfcondities. Deze studie

brengt in kaart hoe de geproduceerde ideeën fungeren tijdens het schrijfproces zelf en hoe het gebruik en het ontwikkelen van ideeën samenhangt met het ontwikkelen van begrip door schrijvers.

De resultaten van deze studie laten allereerst zien dat ongeveer 70 procent van de tekst die geproduceerd wordt tijdens het schrijfproces bestaat uit ideeën die ook op de voor en na het schrijven gemaakte lijsten met ideeën worden gezet. Zoals te verwachten valt, wordt er tijdens het schrijfproces meer gebruik gemaakt van oude dan van nieuwe ideeën. Vervolgens wordt in kaart gebracht hoe zowel oude als nieuwe ideeën gedurende het schrijfproces geïntroduceerd worden in de verschillende planningscondities. Hieruit komt naar voren dat de verschillende planningsstaken een verschillend patroon van ideeëngebruik laten zien. De *outline* planningstaken laten een *top-down* patroon zien waarbij de kans op een nieuw idee aan het begin van het proces hoog is en de kans op een oud idee laag. Dit lijkt te wijzen op de aanwezigheid van expliciete planningsprocessen voorafgaand aan de schrijftaak. De *synthetic* planningscondities laten een tegenovergesteld, meer *bottom-up* beeld zien met een lage kans op een nieuw idee aan het begin van het proces en een hogere kans op een oud idee aan het begin van het schrijfproces. Voor deze planningsconditie geldt dat, naarmate het schrijfproces vordert, het gebruik van oude en nieuwe ideeën elkaar afwisselt. Dit lijkt erop te wijzen dat ideeën ontstaan tijdens het schrijven. Een laatste stap in de analyses brengt in kaart hoe deze patronen van ideeëngebruik samenhangen met het ontwikkelen van begrip door de schrijvers. Daarbij komt naar voren dat schrijvers die iets geleerd hebben van schrijven meer gebruik maken van hun ideeën dan de schrijvers die aangeven niet geleerd te hebben van schrijven. Bovendien laten de resultaten ook hier een verschillend patroon zien voor de verschillende schrijfcondities. Voor schrijvers die aangeven iets te leren van schrijven in de *outline* planningsconditie lijkt het patroon wederom een *top-down* proces te representeren. Voor schrijvers die aangeven geleerd te hebben van schrijven in de *synthetic* planning condities representeert het patroon van ideeën wederom een meer dynamisch proces waarbij nieuwe en oude ideeën gedurende het hele proces afwisselend geïntroduceerd worden. Dit patroon heeft veel weg van wat Galbraith typeert als *dispositionally spelling out* oftewel een manier van al denkend schrijven waarbij nieuwe ideeën ontwikkeld worden tijdens de productie van de tekst. Deze verkennende studie laat zien dat het voor toekomstig procesonderzoek niet alleen van belang is om naar schrijfprocessen te kijken, maar ook om meer specifiek te kijken naar hoe schrijvers hun ideeën gebruiken tijdens het schrijven.

*Conclusies en discussie*

In dit promotieonderzoek draaide het om de vraag hoe schrijven bij kan dragen aan het ontwikkelen van begrip. De belangrijkste resultaten van dit onderzoek hebben laten zien dat twee verschillende typen processen een belangrijke rol spelen bij schrijven: expliciete denkprocessen en impliciete taalproductieprocessen. Bovendien hebben de resultaten laten zien dat leren door schrijven en het produceren van een goede tekst niet noodzakelijk samengaan: het produceren van een goede kwaliteit tekst vergt een andere organisatie van de schrijfprocessen dan het ontwikkelen van begrip. Dit onderzoek ondersteunt daarmee de belangrijkste claims van het *dual process*-model. Daarmee is echter nog niet alles gezegd. Uit de laatste studie bleek dat schrijvers die tijdens het schrijven werken aan hun ideeën meer van de schrijftaak leren dan schrijvers die dat niet doen. Daarbij kwam naar voren dat het patroon van het gebruik van ideeën in de *outline* planning condities erg verschilt van het patroon van het gebruik van ideeën in de *synthetic* planning condities. Deze resultaten hebben zodoende implicaties voor beide cognitieve modellen.

Klassieke probleemoplossingsmodellen kunnen impliciete taalproductie-processen niet meer zonder meer links laten liggen. Dit onderzoek toont immers aan dat taalproductieprocessen een actieve rol spelen bij het ontwikkelen van begrip door schrijven. Bovendien lijkt dit onderzoek erop te wijzen dat onderzoek binnen het probleemoplossingsparadigma de assumptie moet loslaten dat een goed geproduceerde tekst gelijk staat aan *knowledge-transforming*.

Aan de andere kant moet Galbraith erkennen dat ook strategische planningsprocessen kunnen bijdragen aan het ontwikkelen van begrip. Deze planningsprocessen dienen dus niet alleen, zoals verondersteld door het *dual process*-model, voor het organiseren en structureren van al bestaande ideeën.

De resultaten brengen verder naar voren dat vervolgonderzoek noodzakelijk is naar de precieze werking en interactie van impliciete taalproductieprocessen en expliciete denkprocessen. Dit onderzoek dient zich met name te richten op de aard van impliciete tekstproductieprocessen. Wat is ongeplande zinsproductie precies? Maar ook: hoe verhouden geplande en ongeplande zinsproductie zich tot de lineaire organisatie van de tekst als geheel? Er moet nagegaan worden of de veronderstelling van het *dual process*-model dat ongeplande zinsproductie een non-lineaire organisatie van de tekst uitlokt, houdbaar is. De resultaten van dit promotieonderzoek maken bovendien duidelijk dat het *dual process*-model van Galbraith verder ontwikkeld moet worden. De huidige versie van het *dual process*-model toont alleen het proces van de impliciete taalproductieprocessen, terwijl dit onderzoek heeft laten zien dat het bij het verklaren van begripsontwikkeling door schrijven juist gaat om de interactie tussen expliciete denkprocessen en meer impliciete taalproductieprocessen.

## About the author

Veerle Baaijen (1982) studied Dutch Language and Literature at the University of Groningen. In 2007 she graduated with a Master of Arts, specialising in *Speech Communication and Discourse Analysis*. In 2008 she graduated with distinction (cum laude) from the University of Groningen with a Research Master in *Applied Linguistics*. During her two year research master program she spent one year as a research apprentice at Staffordshire University (UK) in Stoke -on-Trent to conduct her thesis research. In 2008 Veerle started her PhD project at the Center for Language and Cognition Groningen at the University of Groningen under the supervision of Kees de Glopper and David Galbraith. The work presented in this thesis is the result of this PhD project. Veerle Baaijen is currently employed as a lecturer at the Department of Communication and Information Sciences and at the Department of Speech Communication and Discourse Analysis of the University of Groningen.

## Acknowledgements

When you look at this finished thesis the creation may seem little more than the translation of preconceived ideas into words. Nothing could be further from the truth. We all know that writing is a complicated and effortful activity. And similar to the research data of this thesis - keystroke logs - the hidden traces of the creation of the text are often as interesting, if not more interesting, than the finished product. The last four years - which constitute the creation of this thesis - have been an invaluable experience which I would not have completed without the help of many people. Therefore, I would like to spend some words to make the hidden traces of writing this thesis visible and to thank the people who have supported me along the way.

First of all, I would like to thank my supervisors. Kees, thank you for being my promotor and for taking me up in the department. I truly appreciate the freedom that you gave me to develop this project to my own insights. I particularly appreciated your efforts towards the end of the project when you helped me get through the times when the deadline seemed impossible to achieve. Moreover, I am very grateful to you for getting me involved in teaching.

David, thank you for being such a good teacher. Doing my PhD with you as my supervisor has been the most adventurous roller-coaster ride imaginable. First, me trying to get up to your speed, learning about all the ins and outs of the not always easy to understand dual process model. Then, us gradually learning to agree to disagree. Then, more naturally, enabling each other to take the pilot and co-pilot seat when necessary. And finally, truly and intrinsically enjoying the pleasure of getting our story out and discussing possible explanations of the data. *Poef*, we have done it and I have to sincerely acknowledge that I would have been lost without you. Not only have you been an excellent teacher, you have also become a dear friend. I have truly enjoyed our Skype calls, our visits, meetings at conferences and all your literate and, from time to time, very funny e-mails. More than once you have cheered me up and you always knew how to comfort me in times of pressure. Your contribution to this thesis is enormous and I am immensely grateful for that.

In addition, I would like to express my gratitude to the members of my reading committee for their critical evaluations of my thesis: John Hayes, Carel Jansen and Gert Rijlaarsdam.

With great pleasure I would like to thank the people who have made an intrinsic contribution to my project. Thanks go out to Luca Nanetti and Rick Scholtes who

occasionally provided me with helpful tips when my own Excel knowledge was lacking. I would like to thank Erik Tjong Kim Sang for helping me to re-write the Inputlog xml-files. Next, I would like to thank Jacolien van Rij for writing an R-script that enabled me to quantify my data more efficiently. Furthermore, I would like to express my gratitude to Jacolien van Rij and Martijn Wieling for their help with R. Martijn, thank you for showing me the possibilities to analyse my data with R and to introduce me to the *rcs*-function and the GAM-function in R. Your help with these analyses is greatly appreciated. Jacolien, thank you for taking the time for my follow-up questions and for helping me with making the plots in R. Both of you, thank you for answering numerous statistical questions throughout the years and for being so passionate about what you do. It was contagious. Finally, special thanks goes to Huub van den Bergh. Thank you for helping me explore my data for multilevel analysis. I have learned a great deal from those two days in Utrecht and I truly appreciate your effort.

Moreover, there have been a lot of people on the fourth floor (or occasional other floors or buildings) from who I received general support. I would like to thank the secretaries of the Dutch Department, BCN, Wyke van de Meer and Karin Mast. I would also to express my thanks to all CLCG PhD's and the members of the department of *Taalbeheersing*. Thank you all for making the CLCG such a fun place to do research.

There are a number of people within the CLCG who I would like to mention specifically. First of all, I would like to thank my office mates through the years: Aletta, Jacolien and especially Myrte. Thank you for sharing your experiences with me and for being so extremely helpful at times. Myrte, the last four years there has been nobody with whom I have spent so much time in one room and I am happy that you survived. A frown of the brow was enough to know what we thought. I think we really did well and I want to thank you for sharing all our ups and downs!

I would like to thank Harm Brouwer, John Hoeks and Bart Hollebrandse for answering my intermediate (statistical) questions. Furthermore, I would like to thank Harrie Mazeland for his understanding when this book needed to be sent to the printer. Moreover, I am especially grateful to Kees de Bot for his help and comments on my CogSci proceedings paper. Thanks also go to Jan Berenst, Femke Kramer, Jacqueline van Kruiningen and Henrike Padmos for their support, for words of encouragement and for their interest in my work. I also want to thank my fellow board members from GRASP and the board members from the TABU Dag 2010 committee for valuable and fun organising experiences. Thank you Ildikó Berzlánovich, Tim Kallenborn and Ruggero Montalto for organising TABU Dag with

me. I learned a lot from working with you and organising a conference together. Finally, thank you Ildikó Berzlánovich, Anne van Buul, Gisi Cannizzaro, Marjolein Deunk, Agnes Engbersen, Myrte Gosen, Sanne Kuijper and Jacolien van Rij for regular lunch meetings, words of advice, providing me with chocolate (Kees has been very generous in this aspect as well) and just for always being there when I knocked on your doors. Finally, Anne and Myrte, I am very happy to have the two of you at my side as *paranimfen* at the defense.

Important other contributors to my project were all the participants who were willing to take part in my experiment. Thank you all. In addition, I would like to thank the colleagues who helped me with the recruitment of my participants. And I would to thank my student assistants throughout the years for their good work and help. Charlotte de Glopper, Lidewij Hamstra, Noortje Hemmen, and Sjors van Ooij, I truly enjoyed working together with you.

Thanks goes to the COST-network without whose financial support my collaboration with David Galbraith would not have been possible. I am grateful to all members of the COST-network and especially to the members who have organised relevant summer and training schools.

During my PhD project and during my year as a research apprentice at Staffordshire University I had the great privilege to meet numerous writing colleagues. I appreciate the way in which you have welcomed me within the writing research community and I value your interest in my work. I would like to extend my gratitude to Rui Alves, Martine Braaksma, Victoria Johansson, Joachim Grabowski, Eva Lindgren, Guido Nottbusch, Thierry Olive, Mark Torrance, and Asa Wengelin. Furthermore, I have truly enjoyed meeting people at conferences. Special thanks go to Catherine van Beuningen, Mariëlle Leijten, Theresa Limpo, Elke van Steendam, Marion Tillema, Lieve Verheyden, Ruth Villalon and Daphne van Weijen. I would also like to thank Gert Rijlaarsdam for the feedback that he has given me during various conferences.

Finally, I would like to thank my friends for their general support, their patience and understanding during the last couple of years and especially during the final year of my project during which I may have been a bit preoccupied at times. I truly appreciated your effort to catch me outside the office. In particular, I would like to thank Janneke and Marjolein. Thank you both for forcing me to relax from time to time, for your friendship and for your support: I wouldn't have coped without the two of you. In addition, I would like to especially thank Jenny, Chris, Sarah and Carla

for always welcoming me warmly on my numerous trips to England. Furthermore, I would like to thank Daan and the gentlemen from Kras - Gerard, Gosse and Wander – for taking me out on beautiful cycling trips. Cycling has been *the* way to relax during my PhD and it was great fun doing this with all of you.

This leaves my family. Mum, dad, Jurjen, Rozanne and Femke, thank you for being there for me. In this context, a special thanks goes to my mother. Nelly, I am grateful for your general interest and your active involvement in my project. You have proved to be an excellent 'mother-assistant'. With great pleasure I observed the precision with which you counted the yellow words for me. Now I know from whom I have my eye for detail. Also Leo and Roz, thank you so much for proofreading bits of the thesis for me.

I saved this final paragraph for Onno. *Otto Nico Norbert Ollie*, finishing this thesis without you at my side would have been a lot less fun. Thank you for helping me with the lay-out of the thesis and for designing the cover. Moreover, thank you for your support, your endless energy, for sharing your enchanted view on the world with me and for inspiring me to excel. Thank you for being you and for letting me shine.

## Groningen dissertations in linguistics (GRODIL)

1. Henriëtte de Swart (1991). *Adverbs of Quantification: A Generalized Quantifier Approach.*
2. Eric Hoekstra (1991). *Licensing Conditions on Phrase Structure.*
3. Dicky Gilbers (1992). *Phonological Networks. A Theory of Segment Representation.*
4. Helen de Hoop (1992). *Case Configuration and Noun Phrase Interpretation.*
5. Gosse Bouma (1993). *Nonmonotonicity and Categorial Unification Grammar.*
6. Peter I. Blok (1993). *The Interpretation of Focus.*
7. Roelien Bastiaanse (1993). *Studies in Aphasia.*
8. Bert Bos (1993). *Rapid User Interface Development with the Script Language Gist.*
9. Wim Kosmeijer (1993). *Barriers and Licensing.*
10. Jan-Wouter Zwart (1993*). Dutch Syntax: A Minimalist Approach.*
11. Mark Kas (1993*). Essays on Boolean Functions and Negative Polarity.*
12. Ton van der Wouden (1994). *Negative Contexts.*
13. Joop Houtman (1994). *Coordination and Constituency: A Study in Categorial Grammar.*
14. Petra Hendriks (1995). *Comparatives and Categorial Grammar.*
15. Maarten de Wind (1995). *Inversion in French.*
16. Jelly Julia de Jong (1996). *The Case of Bound Pronouns in Peripheral Romance.*
17. Sjoukje van der Wal (1996). *Negative Polarity Items and Negation: Tandem Acquisition.*
18. Anastasia Giannakidou (1997). *The Landscape of Polarity Items.*
19. Karen Lattewitz (1997). *Adjacency in Dutch and German.*
20. Edith Kaan (1997). *Processing Subject-Object Ambiguities in Dutch.*
21. Henny Klein (1997). *Adverbs of Degree in Dutch.*
22. Leonie Bosveld-de Smet (1998). *On Mass and Plural Quantification: The case of French 'des'/'du'-NPs.*
23. Rita Landeweerd (1998). *Discourse semantics of perspective and temporal structure.*
24. Mettina Veenstra (1998). *Formalizing the Minimalist Program.*
25. Roel Jonkers (1998). *Comprehension and Production of Verbs in aphasic Speakers.*
26. Erik F. Tjong Kim Sang (1998). *Machine Learning of Phonotactics.*
27. Paulien Rijkhoek (1998). *On Degree Phrases and Result Clauses.*
28. Jan de Jong (1999). *Specific Language Impairment in Dutch: Inflectional Morphology and Argument Structure.*
29. H. Wee (1999). *Definite Focus.*
30. Eun-Hee Lee (2000). *Dynamic and Stative Information in Temporal Reasoning: Korean tense and aspect in discourse.*
31. Ivilin P. Stoianov (2001). *Connectionist Lexical Processing.*
32. Klarien van der Linde (2001*). Sonority substitutions.*
33. Monique Lamers (2001). *Sentence processing: using syntactic, semantic, and thematic information.*
34. Shalom Zuckerman (2001). *The Acquisition of "Optional" Movement.*
35. Rob Koeling (2001). *Dialogue-Based Disambiguation: Using Dialogue Status to Improve Speech Understanding.*
36. Esther Ruigendijk (2002). *Case assignment in Agrammatism: a cross-linguistic study.*

37. Tony Mullen (2002). *An Investigation into Compositional Features and Feature Merging for Maximum Entropy-Based Parse Selection*.
38. Nanette Bienfait (2002). *Grammatica-onderwijs aan allochtone jongeren.*
39. Dirk-Bart den Ouden (2002). *Phonology in Aphasia: Syllables and segments in level-specific deficits.*
40. Rienk Withaar (2002). *The Role of the Phonological Loop in Sentence Comprehension.*
41. Kim Sauter (2002). *Transfer and Access to Universal Grammar in Adult Second Language Acquisition.*
42. Laura Sabourin (2003). *Grammatical Gender and Second Language Processing: An ERP Study.*
43. Hein van Schie (2003). *Visual Semantics*.
44. Lilia Schürcks-Grozeva (2003). *Binding and Bulgarian.*
45. Stasinos Konstantopoulos (2003). *Using ILP to Learn Local Linguistic Structures*.
46. Wilbert Heeringa (2004). *Measuring Dialect Pronunciation Differences using Levenshtein Distance.*
47. Wouter Jansen (2004). *Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology.*
48. Judith Rispens (2004). *Syntactic and phonological processing in developmental dyslexia.*
49. Danielle Bougaïré (2004). *L'approche communicative des campagnes de sensibilisation en santé publique au Burkina Faso: Les cas de la planification familiale, du sida et de l'excision.*
50. Tanja Gaustad (2004). *Linguistic Knowledge and Word Sense Disambiguation.*
51. Susanne Schoof (2004). *An HPSG Account of Nonfinite Verbal Complements in Latin.*
52. M. Begoña Villada Moirón (2005). *Data-driven identification of fixed expressions and their modifiability.*
53. Robbert Prins (2005). *Finite-State Pre-Processing for Natural Language Analysis*.
54. Leonoor van der Beek (2005) *Topics in Corpus-Based Dutch Syntax*
55. Keiko Yoshioka (2005). *Linguistic and gestural introduction and tracking of referents in L1 and L2 discourse.*
56. Sible Andringa (2005). *Form-focused instruction and the development of second language proficiency.*
57. Joanneke Prenger (2005). *Taal telt! Een onderzoek naar de rol van taalvaardigheid en tekstbegrip in het realistisch wiskundeonderwijs.*
58. Neslihan Kansu-Yetkiner (2006). *Blood, Shame and Fear: Self-Presentation Strategies of Turkish Women's Talk about their Health and Sexuality*.
59. Mónika Z. Zempléni (2006). *Functional imaging of the hemispheric contribution to language processing.*
60. Maartje Schreuder (2006). *Prosodic Processes in Language and Music*.
61. Hidetoshi Shiraishi (2006). *Topics in Nivkh Phonology*.
62. Tamás Biró (2006). *Finding the Right Words: Implementing Optimality Theory with Simulated Annealing.*
63. Dieuwke de Goede (2006). *Verbs in Spoken Sentence Processing: Unraveling the Activation Pattern of the Matrix Verb.*
64. Eleonora Rossi (2007). *Clitic production in Italian agrammatism*.

65. Holger Hopp (2007). *Ultimate Attainment at the Interfaces in Second Language Acquisition: Grammar and Processing.*
66. Gerlof Bouma (2008*). Starting a Sentence in Dutch: A corpus study of subject- and object-fronting.*
67. Julia Klitsch (2008). *Open your eyes and listen carefully. Auditory and audiovisual speech perception and the McGurk effect in Dutch speakers with and without aphasia.*
68. Janneke ter Beek (2008). *Restructuring and Infinitival Complements in Dutch.*
69. Jori Mur (2008). *Off-line Answer Extraction for Question Answering.*
70. Lonneke van der Plas (2008). *Automatic Lexico-Semantic Acquisition for Question Answering.*
71. Arjen Versloot (2008). *Mechanisms of Language Change: Vowel reduction in 15th century West Frisian.*
72. Ismail Fahmi (2009). *Automatic term and Relation Extraction for Medical Question Answering System.*
73. Tuba Yarbay Duman (2009). *Turkish Agrammatic Aphasia: Word Order, Time Reference and Case.*
74. Maria Trofimova (2009). *Case Assignment by Prepositions in Russian Aphasia.*
75. Rasmus Steinkrauss (2009). *Frequency and Function in WH Question Acquisition. A Usage-Based Case Study of German L1 Acquisition.*
76. Marjolein Deunk (2009). *Discourse Practices in Preschool. Young Children's Participation in Everyday Classroom Activities.*
77. Sake Jager (2009). *Towards ICT-Integrated Language Learning: Developing an Implementation Framework in terms of Pedagogy, Technology and Environment.*
78. Francisco Dellatorre Borges (2010). *Parse Selection with Support Vector Machines.*
79. Geoffrey Andogah (2010). *Geographically Constrained Information Retrieval.*
80. Jacqueline van Kruiningen (2010). *Onderwijsontwerp als conversatie. Probleemoplossing in interprofessioneel overleg.*
81. Robert G. Shackleton (2010). *Quantitative Assessment of English-American Speech Relationships.*
82. Tim Van de Cruys (2010). *Mining for Meaning: The Extraction of Lexico-semantic Knowledge from Text.*
83. Therese Leinonen (2010). *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects.*
84. Erik-Jan Smits (2010). *Acquiring Quantification. How Children Use Semantics and Pragmatics to Constrain Meaning.*
85. Tal Caspi (2010). *A Dynamic Perspective on Second Language Development.*
86. Teodora Mehotcheva (2010). *After the fiesta is over. Foreign language attrition of Spanish in Dutch and German Erasmus Student.*
87. Xiaoyan Xu (2010). *English language attrition and retention in Chinese and Dutch university students.*
88. Jelena Prokić (2010). *Families and Resemblances.*
89. Radek Šimík (2011*). Modal existential wh-constructions.*
90. Katrien Colman (2011). *Behavioral and neuroimaging studies on language processing*

*in Dutch speakers with Parkinson's disease*.

91. Siti Mina Tamah (2011). *A Study on Student Interaction in the Implementation of the Jigsaw Technique in Language Teaching*.
92. Aletta Kwant (2011). *Geraakt door prentenboeken. Effecten van het gebruik van prentenboeken op de sociaal-emotionele ontwikkeling van kleuters.*
93. Marlies Kluck (2011). *Sentence amalgamation*.
94. Anja Schüppert (2011). *Origin of asymmetry: Mutual intelligibility of spoken Danish and Swedish.*
95. Peter Nabende (2011). *Applying Dynamic Bayesian Networks in Transliteration Detection and Generation.*
96. Barbara Plank (2011). *Domain Adaptation for Parsing.*
97. Cagri Coltekin (2011).*Catching Words in a Stream of Speech: Computational simulations of segmenting transcribed child-directed speech.*
98. Dörte Hessler (2011). *Audiovisual Processing in Aphasic and Non-Brain-Damaged Listeners: The Whole is More than the Sum of its Parts*.
99. Herman Heringa (2012). *Appositional constructions*.
100. Diana Dimitrova (2012). *Neural Correlates of Prosody and Information Structure*.
101. Harwintha Anjarningsih (2012). *Time Reference in Standard Indonesian Agrammatic Aphasia.*
102. Myrte Gosen (2012). *Tracing learning in interaction. An analysis of shared reading of picture books at kindergarten.*
103. Martijn Wieling (2012). *A Quantitative Approach to Social and Geographical Dialect Variation.*
104. Gisi Cannizzaro (2012). *Early word order and animacy*.
105. Kostadin Cholakov (2012). *Lexical Acquisition for Computational Grammars. A Unified Model.*
106. Karin Beijering (2012). *Expressions of epistemic modality in Mainland Scandinavian. A study into the lexicalization-grammaticalization-pragmaticalization interface*.
107. Veerle Baaijen (2012). *The development of understanding through writing*.

GRODIL
Center for Language and Cognition Groningen (CLCG)
P.O. Box 716
9700 AS Groningen
The Netherlands