Long-term working memory in text production

RONALD T. KELLOGG
University of Missouri, Rolla, Missouri

In reading and other high-level cognitive tasks, Ericsson and Kintsch (1995) proposed that the limited capacity of short-term working memory (STWM) is supplemented by long-term working memory (LTWM) for individuals with a high degree of domain-specific knowledge. In Experiment 1, college students ($N = 80$) wrote persuasive and narrative texts concerning baseball; domain-specific knowledge about baseball and verbal ability was assessed. The results showed that verbal ability and domain-specific knowledge independently affected writing skill, supporting the view that literacy depends on both knowledge sources and refuting one argument raised in support of the LTWM hypothesis. Experiment 2 ($N = 42$) replicated this outcome and tested the prediction that a high degree of domain-specific knowledge would lessen interference on a secondary task. The data supported the interference prediction, offering evidence that LTWM plays a role in the production of text.

Many kinds of writing tasks impose considerable demands on working memory, the system responsible for processing and storing information on a short-term basis (McCutchen, 1996; Ransdell & Levy, 1996; Torrance & Jeffery, 1999). One way to index the cognitive effort invested in writing processes is to compare baseline reaction time (RT) to an auditory probe presented in isolation against secondary RT to the same probe presented during composition. The more RT interference caused by writing processes, the more momentary cognitive effort they consume. The secondary RT task as well as writing demand limited capacity executive functions, such as focusing attention and scheduling a response (Jonides & Smith, 1997). The results from such comparisons show that planning ideas, linguistically translating ideas or generating sentences, and reviewing ideas and text are all effortful, although the pattern of differences among these processes varies with the task (Kellogg, 1988, 1994; Levy & Ransdell, 1996; Piolat, Roussey, Olive, & Farioli, 1996). In comparison with learning and reading tasks, writing places substantial demands on executive functioning.

Managing the demands of writing on short-term working memory (STWM) may take many forms, such as strategies for allocating resources to planning, translating, and reviewing over time and automating the motor execution requirements (Fayol, 1998; McCutchen, 1996). Another possibility is that skilled writers may draw on an alternative system of working memory. Ericsson and Kintsch (1995) proposed that the classic distinction between long-term memory and STWM is inadequate to explain a large body of evidence on reading and other complex cognitive skills. They suggested that individuals with a high degree of domain-specific knowledge process information held in long-term working memory (LTWM), whereas those with less knowledge are bound by the capacity limitations of STWM. The capacity of STWM may be equally limited for both skilled and unskilled readers, but they differ on a reading span test because the skilled reader can use extensive retrieval structures that aid in comprehension and memory. In short, the effective capacity of working memory is enlarged through the use of LTWM (see Spilich, Vesonder, Chiesi, & Voss, 1979).

One line of evidence cited in support of their theory concerned the different effects observed with domain knowledge versus verbal ability. Individuals with a high rather than a low degree of baseball knowledge wrote more elaborate narratives about the game and recalled more relevant propositions after reading such texts (Voss, Vesonder, & Spilich, 1980). In contrast, null effects for verbal ability have been reported in comparable reading and memory tasks (Recht & Leslie, 1988; Schneider, Körkel, & Weinert, 1989; Walker, 1987). College students with a high degree of verbal ability and general aptitude did no better than those with low ability; only their score on tests of domain-specific knowledge predicted their performance.

Ericsson and Kintsch (1995) argued that the null outcome for verbal ability in reading tasks makes sense if STWM has little to do with individual differences in reading ability. Instead of attributing such differences to STWM capacity (Just & Carpenter, 1992), they suggested that retrieval of knowledge from long-term memory was the critical factor. Implicit in their argument is that verbal ability is directly related to STWM capacity differences (Daneman & Merikle, 1996; Engle, Cantor, & Carullo, 1992; Waters & Caplan, 1996).

Another line of evidence advanced by Ericsson and Kintsch (1995) supporting the construct of LTWM is that individuals with a high degree of domain knowledge can perform a current task with little interference. For example, counting from 1 to 10 failed to disrupt an expert waiter's ability to learn dinner orders, but did interfere with the ability of nonwaiters (Ericsson & Polson, 1988). This suggests that the strategies used to encode and retrieve the dinner orders were not dependent on STWM in the expert. Task interference, therefore, also tests whether long-term memory can boost the effective working memory capacity of individuals with a high degree of domain knowledge.

The present research examined both lines of evidence advanced by Ericsson and Kintsch (1995) to determine whether LTWM may be used by writers with a high degree of domain knowledge. Experiment 1 sought to determine whether domain-specific knowledge, but not verbal ability, would affect writing skill. Experiment 2 examined both this issue and the question of interference with a concurrent task. Specifically, measurements were taken of the degree of interference in secondary probe RT during text production. The LTWM hypothesis predicts that writers with a high degree of domain-specific knowledge should show less interference than low-knowledge writers.

A second motivation for the research was to examine further the effects of verbal and domain knowledge on writing performance. The domain-knowledge effects are well established for text comprehension, but in production the evidence is less compelling (Benton, Corkill, Sharp, Downey, & Khramtsova, 1995; Rowan, 1990). For instance, Benton et al. found that baseball knowledge only weakly correlated with the number of game actions ($r = .24$) and relevant-nongame actions ($r = .11$) included in written stories. The evidence in production tasks is actually stronger for verbal ability than for domain knowledge. Standardized tests of verbal ability correlate moderately to strongly with reader judgments of the quality of written texts (Huot, 1990).

Theoretically, one can argue that writers use domain knowledge about the topic to generate ideas and discourse knowledge to organize relevant ideas into a coherent text (Bereiter & Scardamalia, 1987; McCutchen, 1986). For example, applying Kintsch's (1988) construction-integration model to production, writers formulate both a text base and a situation model in creating the text. The text base forms a coherent network of propositions expressed in a text, whereas the situation model forms a more complex representation that integrates the text base with prior knowledge. Coherence of the text and the prior domain knowledge of the reader both contribute to text comprehension (McNamara, Kintsch, Songer, & Kintsch, 1996). A well-written text facilitates the reader's ability to construct the text base, whereas domain knowledge contributes to the situation model. Only readers with a high degree of domain knowledge can provide correct elaborations of the text in their recall (Moravcsik & Kintsch, 1993). Similarly, individual differences in reading comprehension ability are related chiefly to the development of the text base, in contrast to the benefits of prior domain knowledge for the development of the situation model (Voss & Silfies, 1996).

Presumably in text production, the writer's verbal ability and domain knowledge should also independently contribute to an effective text that a reader can comprehend and remember. A high degree of verbal ability should aid in fashioning a syntactically correct coherent text base for the reader. On the other hand, a high degree of domain knowledge should be primarily helpful in formulating a situation model that integrates prior knowledge with the text base.

EXPERIMENT 1

Differences in domain knowledge were assessed using the baseball quiz of Voss et al. (1980) and differences in verbal ability were based on standardized test scores. Participants wrote a narrative account of a half-inning of a baseball game. The propositions expressed were categorized according to whether they were game actions, relevant-nongame actions, or irrelevant-nongame actions, as in previous work (Benton et al., 1995; Voss et al., 1980). It was anticipated that a high degree of domain knowledge would increase the proportion of game-relevant actions. Grammatical errors were tabulated with the expectation that a high degree of verbal ability would decrease their number. Finally, an analytic rating procedure involving separate judgments about the content and style of writing was employed as well as a summed overall measure of text quality (Huot, 1990; Kellogg, 1994). It was expected that content judgments would improve with more domain knowledge, whereas style judgments would vary with verbal ability.

Besides the narrative text studied in past research, the participants also wrote a persuasive text in an effort to establish the generality of any knowledge effects observed. The grammatical error and analytic rating measures were again used, but the analysis of propositions as game actions and so on was not applicable.

Method

The writers were 80 college students who received credit in their general psychology class for participating. They wrote both a narrative and a persuasive text concerning the topic of baseball. Half the participants wrote the narrative first and the other half the persuasive. For the narrative task, the instructions read as follows:

In this writing assignment you should write a narrative account of one-half inning of a fictitious baseball game involving major league teams. Imagine that you are writing a play-by-play account in a newspaper sports story. You may make up the teams, players, and events or draw on a game that you actually saw or heard about. You should make the story as interesting as possible for the readers.

For the persuasive task, the instructions read as follows:

In this writing assignment, you should write a persuasive editorial regarding the skyrocketing salaries of major league baseball players. The average salary of a major league player is $750,000, with more than 40 players earning in excess of $3,000,000 per year. Your task is to write a newspaper editorial calling for a cap on salaries at $1,000,000, a level that only the very best players should receive. Your task is to persuade the public that the current excessive salaries will ruin the national pastime. You must formulate convincing arguments for the salary cap regardless of how you personally feel about the matter.

The English score of the ACT indexed verbal ability, but this score was unavailable for 19 of the 80 students tested. The median score for the remaining 61 equaled 21. The writers with scores of 21 or less were assigned to the low knowledge condition ($n = 30$) and those with scores of 22 or more were assigned to the high knowledge condition ($n = 31$).

Voss et al.'s (1980) baseball quiz provided the index of domain knowledge. The 80 students were divided into two groups based on their score on the 45-item quiz. A score of 26 or less led to assignment to the low-knowledge condition ($n = 41$), whereas 27 or more led to the high-knowledge condition ($n = 39$).

Testing was done in groups of 8–15 students. They took the baseball knowledge quiz and then received the first writing task instructions and studied them for 2 min. The instructions specified the approximate length of the text (350 words), allowed the use of paper to plan, and emphasized that the final text must be legible and coherent. The experimenter then timed the 30-min writing session, indicating on the chalkboard when 15, 10, 5, and 1 min remained. When the session ended, the experimenter collected the task instructions and the texts, and repeated the procedure for the second writing assignment.

The propositional content of the narrative texts was analyzed using the procedure described by Benton et al. (1995). As in the Voss et al. (1980) analysis, the ideas expressed in the text were categorized as game actions, relevant-nongame actions, and irrelevant-nongame actions. The narratives were segmented into minimally parsable units of information in the text and then categorized in relation to the conceptual framework of the half-inning narrative. Game actions were behaviors that caused a change in the state of the game that moved closer to the goal of winning (e.g., "the batter hit the pitch"). Relevant-nongame actions were events that did not change the state of the game, but that were relevant to the game (e.g., "the catcher gave the sign"). Irrelevant-nongame actions were events that were unrelated to the game actions, including descriptions of the setting (e.g., "it was a beautiful day at the ballpark"). After all 80 narratives were segmented and categorized by a research assistant, a random sample of 20 were rescored by a second evaluator. The proportion of agreement in categorizations was .91.

Two paid judges rated each of the documents for quality. One of these judges also marked and tallied grammatical errors as they read and judged the texts. The grammatical errors marked were run-on sentences, subject–verb disagreements, errors in pronoun usage, and noun–modifier disagreements. Both judges scored at the 90th percentile or better on the ACT English test and had taken two college composition classes. They judged the content and style of typed transcriptions of the texts handwritten by the participants. Content was defined in terms of how well ideas were developed, how coherently organized the text was, and how effectively the text communicated its message. Style was defined in terms of the quality of word choice, sentence structure, spelling, and grammar. The judges were trained (see Kellogg, 1987) to rate content and style on 7-point scales ranging from 1 (*poor*) to 7 (*excellent*).

The subjective ratings of quality were averaged for the two judges after it was determined that the interjudge reliability coefficients were statistically significant at $p < .001$. For the persuasive texts, these coefficients were $r = .73$ for content ratings and $r = .63$ for style ratings. The narrative texts showed slightly less agreement between the judges, with $r = .59$ for content ratings and $r = .59$ for style ratings.

Results

To assess the effects of domain knowledge and verbal ability, the 61 participants with a ACT English test score as well as the baseball quiz score were divided into four conditions. These were defined by writers who scored low on both domain and verbal knowledge ($n = 20$), high on both ($n = 18$), low on domain but high on verbal ($n = 13$), and high on domain but low on verbal ($n = 10$).

The proportions of narrative events categorized as game actions, relevant-nongame actions, and irrelevant-nongame actions are shown in Table 1 for the four conditions. An analysis of variance (ANOVA) revealed a reliable interaction of domain knowledge and action type [$F(2,114) = 6.59, MS_e = .054, p < .05$]. A high degree of domain knowledge resulted in a higher proportion of game actions ($M = .38$) included in the narratives relative to low-domain knowledge ($M = .25$). Although the knowledge conditions differed little with respect to relevant-nongame actions, the high-knowledge writers ($M = .23$) included fewer irrelevant-nongame actions relative to low-knowledge writers ($M = .41$). The kinds of propositions expressed in the narratives were not influenced reliably by verbal ability. No other source of variance was reliable.

In contrast, verbal ability but not domain knowledge affected the frequency of grammatical errors detected in the narrative and persuasive texts, as shown in Table 2. Although the individual cell means are noisy, collapsing over domain knowledge and text type reveals that low-verbal writers ($M = 3.27$) generated more grammatical errors than did high-verbal writers ($M = 1.84$). An ANOVA revealed a reliable main effect of verbal knowledge [$F(1,57) = 6.79, MS_e = 7.97, p < .05$]. The only other reliable source of variance was a main effect of text type [$F(1,57) = 6.83, MS_e = 2.99, p < .05$]. Persuasive texts ($M = 2.08$) included fewer errors overall than did narratives ($M = 3.02$). This difference probably reflected the reliably shorter length of the persuasive texts ($M =$

Table 1
Mean Proportions of and Standard Errors for Game, Relevant-Nongame, and Irrelevant-Nongame Actions in Experiment 1

Domain Knowledge	Action Type		Relevant Nongame		Irrelevant Nongame	
	M	SE	M	SE	M	SE
Low Verbal Ability						
Low	.24	.03	.37	.09	.39	.04
High	.37	.03	.41	.08	.22	.08
High Verbal Ability						
Low	.25	.05	.32	.05	.43	.05
High	.39	.02	.37	.06	.24	.05

Table 2
Mean Number of and Standard Errors
for Grammatical Errors in Experiment 1

| | Type of Text | | | |
| | Narrative | | Persuasive | |
Domain Knowledge	*M*	*SE*	*M*	*SE*
Low Verbal Ability				
Low	4.10	.66	2.70	.55
High	1.95	.62	2.60	1.03
High Verbal Ability				
Low	1.84	.49	1.53	.46
High	2.39	.59	1.50	.33

296 words) compared with the narratives [$M = 356$ words, $F(1,57) = 11.01$, $MS_e = 7,311$, $p < .05$]. The length of texts was not reliably affected by either verbal ability or domain-specific knowledge.[1]

It was anticipated that the ratings of content would be influenced by domain knowledge, whereas ratings of style would reveal differences in verbal ability. However, it turned out that content and style ratings were strongly correlated with the content ratings for both the persuasive ($r = .82$, $p < .001$) and narrative ($r = .90$, $p < .001$) assignments. Further preliminary analyses showed that domain knowledge and verbal ability each affect content and style ratings in the same fashion. Consequently, these were summed to provide an overall measure of text quality.

As shown in Figure 1, the mean overall quality judgments, averaged across narrative and persuasive texts, showed that domain knowledge and verbal ability exert independent effects on overall quality. An ANOVA showed no interaction but did show reliable main effects of domain knowledge [$F(1,57) = 14.04$, $p < .01$] and verbal knowledge [$F(1,57) = 9.06$, $MS_e = 6.47$, $p < .001$]. The narrative and persuasive texts were combined here, because each revealed this pattern. An ANOVA of the ratings of narratives revealed a main effect favoring high-domain knowledge writers [$F(1,57) = 14.70$, $MS_e = 4.63$, $p < .001$] and a main effect favoring high-verbal writers [$F(1,57) = 7.57$, $p < .01$]. Similarly, for the persuasive texts, main effects were obtained for domain knowledge [$F(1,57) = 6.55$, $MS_e = 4.27$, $p < .05$] and verbal ability [$F(1,57) = 17.81$, $p < .001$].[2]

Discussion

Variations in writers' domain knowledge and verbal ability independently affected quality judgments of both narrative and persuasive texts. Contrary to some findings on text comprehension and memory, verbal ability had a decisive influence on how well writers produced texts. High verbal ability decreased the number of gram-
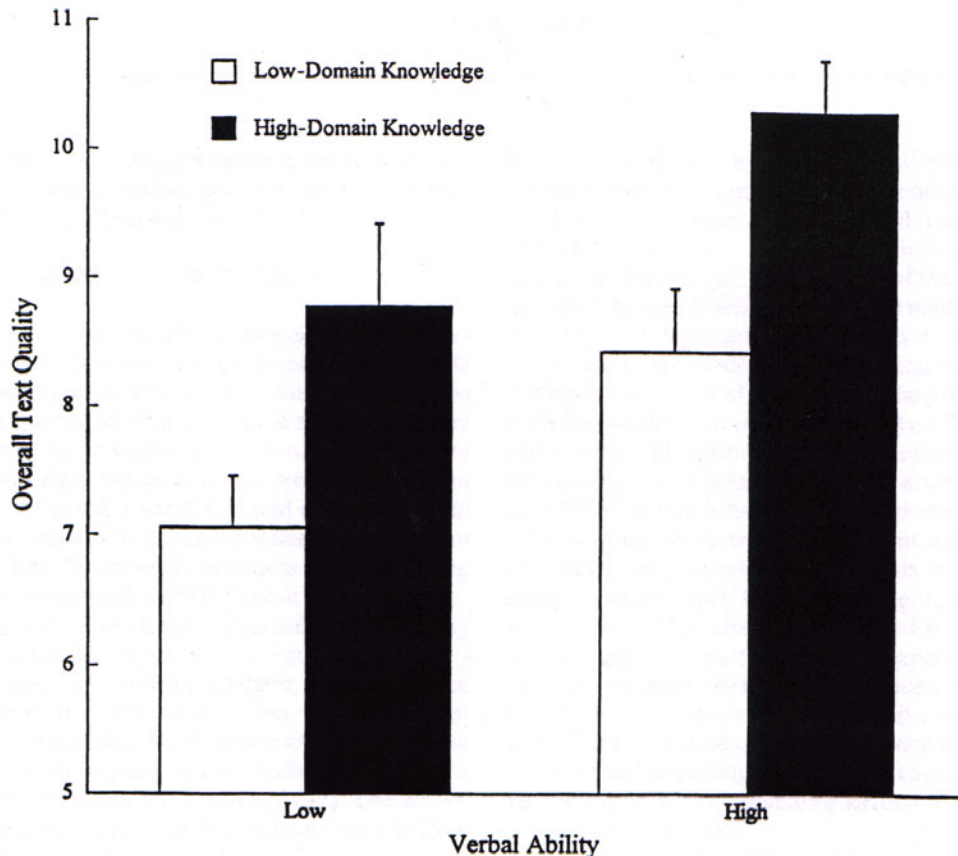

Figure 1. Mean overall quality in Experiment 1 as a function of domain knowledge and verbal ability.

matical errors made in the text and increased its subjectively rated quality relative to low verbal ability. Ericsson and Kintsch (1995) contended that working memory capacity varies depending on the domain tapped by the language task given that domain knowledge, but not verbal ability, affected text recall. The present research shows that with text production both factors can have a clear impact. Thus, one line of evidence cited by Ericsson and Kintsch in support of LTWM is not sustained when the task is one of production rather than comprehension.

Putting aside the LTWM hypothesis, the results are compatible with comprehension theories that assign roles to both verbal ability and other forms of discourse knowledge, on the one hand, and domain-specific knowledge, on the other. For example, the results can be interpreted as supporting the construction–integration model in that writers needed to construct both a text base and a situation model. Verbal ability helped them to develop a text base relatively free of grammatical errors for both the narrative and persuasive texts. Domain expertise allowed the writers to include a relatively high proportion of propositions concerning game actions in their narrative accounts of a half-inning of a baseball game. The subjective ratings of quality, for both persuasive and narrative texts, were highest when writers were apparently successful in constructing both text base and situation models effectively.

EXPERIMENT 2

A second argument for LTWM is that experts escape the usual interference observed when doing two tasks simultaneously that should overload the limited capacity of STWM. In Experiment 2, interference in the time needed to detect an auditory probe presented at random during text production was measured. Secondary task RT has commonly been used to assess spare attentional capacity (Kahneman, 1973); it thus provides an index of the engagement of the executive functions of working memory in the primary task (Jonides & Smith, 1997). Again, individual differences in domain knowledge and verbal ability were examined. According to Ericsson and Kintsch's (1995) model, a high degree of expertise should lessen the degree of interference observed.

Method

The design of Experiment 1 was replicated with the addition of a RT task (Kellogg, 1987, 1994). When writers detected a tone on a variable-interval schedule, they said "Stop" as rapidly as possible and then indicated whether they were planning ideas, translating ideas into sentences, reviewing previous work, or were engaged in other activities when the tone occurred. Increases in RT over baseline provide estimates of the cognitive effort devoted to these three writing processes. The amount of interference observed depends on the frequency of tones and task tradeoffs (Piolat et al., 1996). But on the same schedule (on average, once every 30 sec), the relative differences between knowledge conditions can be fruitfully compared.

College students ($N = 42$) were tested individually and screened on the basis of having an ACT English score available. The median score for dividing into high- and low-verbal groups was 2 points higher (23) than in Experiment 1. The same cutoff on the baseball quiz was used (26). The procedure resulted in these cell sizes: low verbal–low domain ($n = 11$), low verbal–high domain ($n = 10$), high verbal–low domain ($n = 11$), and high verbal–high domain ($n = 10$).

The 42 narrative texts were again segmented and categorized as game actions, relevant-nongame actions, and irrelevant-nongame actions by a research assistant not involved in the Experiment 1 analysis. A random sample of 10 texts were rescored by the same individual who checked for agreement in the earlier analysis, and the proportion of agreement in categorizations this time was .82. Subjective quality was again rated by the same two judges employed in Experiment 1 (one of whom again tallied grammatical errors), and the interjudge reliability coefficients were again statistically significant ($p < .001$). The style ($r = .79$) and content ratings ($r = .57$) for the narrative texts and the style ($r = .71$) and content ratings ($r = .62$) for the persuasive texts were again averaged across the two judges. The style and content ratings were reliably correlated for both the narrative ($r = .64$) and persuasive ($r = .61$) texts. Overall quality scores were again derived by summing the content and style ratings.

Results

The proportions of narrative events categorized as game actions, relevant-nongame actions, and irrelevant-nongame actions are shown in Table 3 for the four conditions. An ANOVA revealed a reliable interaction of domain knowledge and action type [$F(2,76) = 3.39, MS_e = .04, p < .05$]. A high degree of domain knowledge resulted in a higher proportion of game actions ($M = .33$) included in the narratives relative to low-domain knowledge ($M = .22$). Although the knowledge conditions differed little with respect to relevant-nongame actions, the high-knowledge writers included fewer irrelevant-nongame ($M = .33$) actions relative to low-knowledge writers ($M = .45$). The kinds of propositions expressed in the narra-

Table 3
Mean Proportion of and Standard Errors for Game, Relevant-Nongame,
and Irrelevant-Nongame Actions in Experiment 2

Domain Knowledge	Action Type		Relevant Nongame		Irrelevant Nongame	
	M	SE	M	SE	M	SE
Low Verbal Ability						
Low	.18	.03	.32	.06	.50	.06
High	.31	.03	.34	.05	.35	.05
High Verbal Ability						
Low	.25	.04	.37	.06	.38	.08
High	.34	.04	.37	.05	.30	.05

tives were not influenced reliably by verbal ability. The only other reliable source of variance was a main effect of action type [$F(2,76) = 3.57, p < .05$], reflecting that game actions were less frequent than relevant and irrelevant nongame actions overall.

As in Experiment 1, verbal ability but not domain knowledge affected the frequency of grammatical errors detected in the narrative and persuasive texts, as shown in Table 4. Low-verbal writers ($M = 2.40$) generated more grammatical errors than did high-verbal writers [$M = 1.61; F(1,38) = 5.52, MS_e = 2.36, p < .05$]. No other sources of variance were significant. The narrative texts did not contain reliably more grammatical errors and they were not on average 60 words longer than the persuasive texts, as found in Experiment 1. Nonetheless, the difference in length again reliably favored the narrative ($M = 281$ words) over the persuasive [$M = 248$ words; $F(1,38) = 10.87, p < .01$] texts, and this was the only significant source of variance in the analysis of text length.

Plotted in Figure 2 are the mean overall quality scores.[3] The ANOVA revealed main effects of domain knowledge [$F(1,38) = 8.49, MS_e = 3.83, p < .01$] and of verbal ability [$F(1,38) = 16.96, p < .001$]. Replicating the findings of Experiment 1, the effects of domain knowledge and verbal ability were independent.

The median baseline RT was subtracted from the medians associated with planning, translating, and reviewing responses in the introspection task; this interference metric indicated the degree to which writing required a momentary allocation of limited working memory capacity. The greater the RT interference, the more momentary effort required. A preliminary analysis of the median RT data showed that the baseline group means for the low domain–low verbal ($M = 301$ msec, $SE = 13.2$), high domain–low verbal ($M = 292$ msec, $SE = 17.8$), low domain–high verbal ($M = 278$ msec, $SE = 28.4$), and high domain–high verbal ($M = 290$ msec, $SE = 64.2$) conditions were not reliably different.

The RT interference analysis revealed one reliable source of variance, a main effect of domain knowledge. The degree of interference in milliseconds was roughly the same for planning, translating, and reviewing responses, and did not vary reliably by the kind of text written. However, the writers with a high degree of domain knowledge showed less interference, and therefore greater

spare capacity, than those with a low degree of domain knowledge [$F(1,38) = 5.72, MS_e = 137,006, p < .05$]. This advantage for the high-domain knowledge condition was slightly greater for planning ($M = 233$ vs. $M = 351$ msec) and reviewing ($M = 224$ vs. $M = 350$ msec) than for translating ($M = 225$ vs. $M = 315$ msec). However, the interaction was unreliable, so the data are averaged across process type in Figure 3. Neither the main effect of verbal ability nor any other source of variance was reliable. The correlation of RT interference with the scores on the baseball quiz was reliable ($r = -.48; p < .01$), but not with the ACT English scores ($r = -.11$).

Discussion

The data replicated the findings of Experiment 1 for analyses of text content, grammatical errors, and subjective quality ratings. Of key interest, interference in RT on the secondary probe detection task was reliably lower for high-domain knowledge writers compared with those with low-domain knowledge. As in two previously reported experiments, domain knowledge lowered the momentary effort required by planning, translating, and reviewing (Kellogg, 1987). These RT findings are consistent with Ericsson and Kintsch's (1995) view that LTWM reduces the interference between two tasks that both demand the resources of STWM. The secondary RT task requires executive functions that are limited in capacity, as does the planning, sentence generation, and other cognitive operations involved in writing. The amount of interference between the two tasks was lessened for writers with a high degree of domain-specific knowledge presumably because only they could readily retrieve information from LTWM during the composition task.

More recently, Ericsson and Delaney (1998) suggested the same pattern might be obtained for verbal ability, in light of the growing evidence that it independently affects reading and writing performance. This was not found in Experiment 2. Data from two previous experiments reported by Kellogg and Mueller (1993) for which ACT English scores were available and that showed a wider range were pooled ($N = 68$) and analyzed as reported here. In those data, the RT interference negatively correlated with verbal ability at a modest but reliable level $r = -.31$. Data from another recent writing experiment were analyzed with respect to this issue, and the correlation between verbal ability and RT interference was also reliably negative ($r = -.48$; Kellogg, in press). The null effect reported here may, therefore, reflect a Type II statistical error that warrants further investigation.

However, even if a negative relationship between verbal ability and RT interference is established, its interpretation in this context is unclear. Verbal ability and other measures of general aptitude are correlated with individual differences in STWM span (Daneman & Merikle, 1996; Engle et al., 1992; Waters & Caplan, 1996). A negative relationship between verbal ability and RT interference could be just as readily attributed to variations in STWM instead of the use of LTWM. This line of reason-

Table 4
Mean Number of and Standard Errors
for Grammatical Errors in Experiment 2

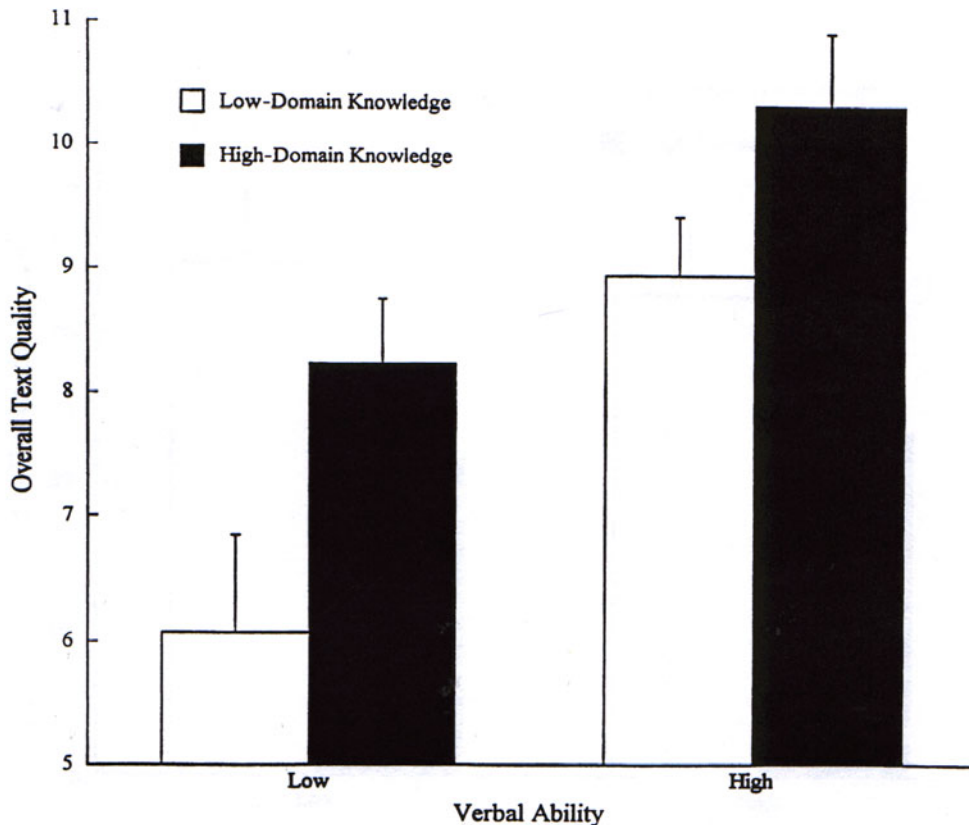| | Type of Text | | | |
| | Narrative | | Persuasive | |
Domain Knowledge	M	SE	M	SE
Low Verbal Ability				
Low	1.72	.30	2.36	.28
High	2.90	.43	2.60	.40
High Verbal Knowledge				
Low	2.00	.45	1.63	.39
High	1.50	.34	1.30	.30

Figure 2. Mean overall quality in Experiment 2 as a function of domain knowledge and verbal ability.

ing fails in the case of domain-specific knowledge, because such knowledge differences are not consistently related to capacity differences in STWM.

GENERAL DISCUSSION

A high degree of verbal ability enhanced writing performance in both narrative and persuasive assignments concerned with baseball. Domain-specific knowledge also improved the quality of the texts written and these factors did not interact. Only verbal ability was reliably related to the number of grammatical errors produced. Domain knowledge, on the other hand, resulted in differences in the degree to which writers incorporated game-relevant propositions into the narrative accounts of a half-inning of baseball (Benton et al., 1995; Voss et al., 1980).

A growing body of evidence suggests that both reading and writing skills are dependent on both verbal ability and domain-specific knowledge (Adams, Bell, & Perfetti, 1995; Benton et al., 1995; McCutchen, 1986; Singer & Ritchot, 1996). Kintsch's (1988) construction–integration model has been chiefly applied to reading, but it is also relevant to text production (Nicolich, 1997). The writer, as well as the reader, presumably must construct a situation model that integrates prior domain knowledge with the text base. The present results suggest that verbal abil-

ity aids in developing a text base free of grammatical ambiguities, whereas domain knowledge helps the writer to construct the necessary situation model from prior knowledge and the text base. Each of these influences contributed to the overall quality judgment of the text. Detailed testing of the construction–integration model in text composition is an important area for further research.

One argument made by Ericsson and Kintsch (1995) in support of a LTWM started with the premise that domain knowledge, but not verbal ability, affected literacy performance. This argument must be rejected on the basis of evidence now available. However, a second argument for LTWM is that a high degree of domain knowledge reduces interference with a concurrent task. The results of Experiment 2 clearly supported this in a writing task using secondary task RT. A high degree of knowledge about baseball lessened the degree of interference observed, suggesting that the baseball experts were relying on long-term memory to solve the content and rhetorical problems of composing both narrative and persuasive texts. The low degree of RT interference observed in the high baseball knowledge condition is consistent with two earlier experiments preceding Ericsson and Kintsch's theory (Kellogg, 1987).

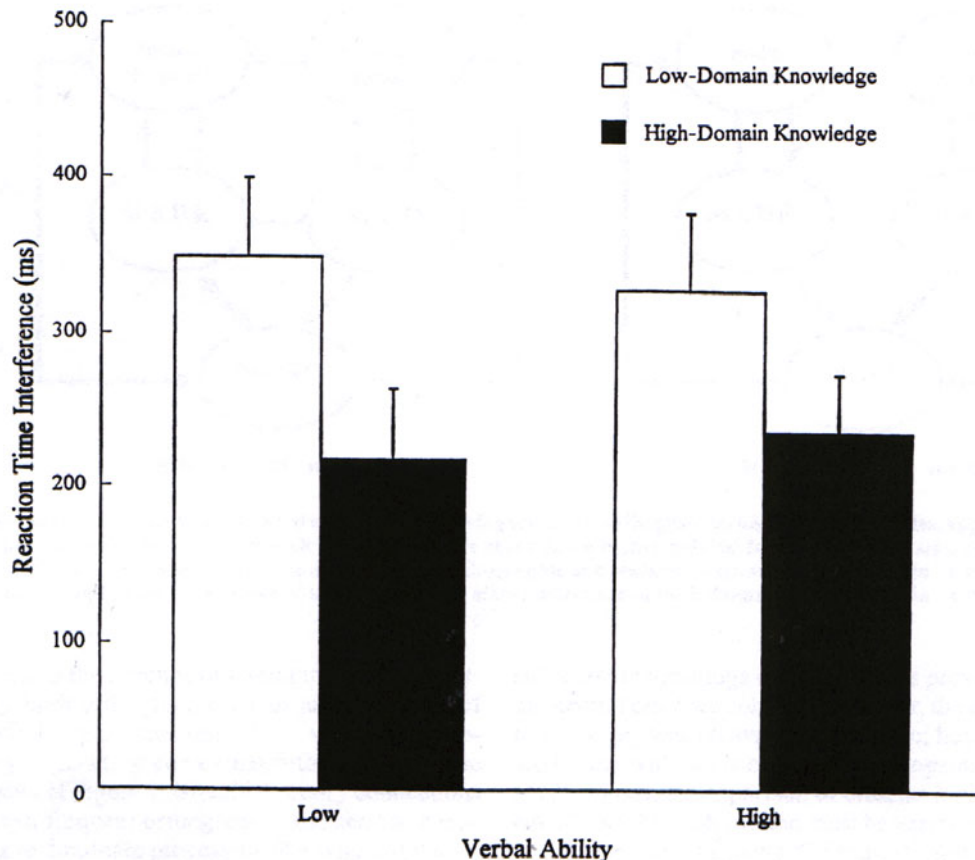One interpretation of the domain knowledge results is that the highly knowledgeable writer can rely on a sys-

Figure 3. Mean reaction time interference (in milliseconds) in Experiment 2 as a function of domain knowledge and verbal ability.

tem of LTWM to solve the rhetorical and content problems not available to the less knowledgeable writer (Ericsson & Kintsch, 1995; Spilich et al., 1979). The results do not rule out the possibility that individual differences in the capacity of STWM also affect reading and writing performance in addition to relevant domain expertise (Singer & Ritchot, 1996).

An alternative interpretation of the interference data is that the experts automatically wrote their texts. If one refers only to the relatively automatic retrieval of well-learned content knowledge from long-term memory, then one is making the same claim as the hypothesis of LTWM. But beyond the retrieval of ideas, it seems unlikely that automaticity can account for the present data. First, it is unclear why any interference was observed in the high domain knowledge condition if writing was automatic. Second, the extensive interaction among the planning, translating, and reviewing processes observed in skilled writers is inconsistent with automaticity and modularity (McCutchen, 1988). For example, in planning a text, an experienced writer sets rhetorical goals that are shaped by the intended audience. Such high-level planning has

consequences for the selection of ideas to be included and even the choice of words in a given sentence.

Verbal ability was unrelated to secondary RT interference in Experiment 2, but reliable negative correlations have been observed in other experiments. The findings of Voss and Silfies (1996) suggest one explanation for the varying outcomes. Using fictional narrative accounts, they contrasted the reading of an unexpanded text with the reading of an expanded version that spelled out the causal relationships of importance. For the unexpanded version, they found that prior knowledge about history predicted performance on memory completion and written essay tests. Reading comprehension ability was significantly related to performance when participants read the expanded text. For the sparse, unexpanded text, the reader's prior knowledge of history allowed the drawing of inferences needed to develop an adequate situation model. Prior knowledge was less important when the text was already unpacked in terms of causal relations. In this case the development of the text base was the limiting factor, and so reading ability explained variance in performance.

Here the baseball writing tasks presumably required high levels of prior knowledge for the successful construction of a situation model. Perhaps RT interference was sensitive to the effort required to construct the situation model rather than the text base in the present task. Where reliable negative correlations have been obtained between verbal ability and RT interference, construction of the situation model was arguably less dependent on high levels of knowledge. For example, Kellogg and Mueller (1993) designed their tasks so they could be composed by undergraduates without specialized knowledge (e.g., argue whether an attorney should defend a murderer whom he/she knows to be guilty). Kellogg (in press) designed the tasks to be highly familiar to all students in the experiment (e.g., create a narrative about taking a test). In these cases, verbal ability may have correlated with RT interference because construction of the text base was the limiting factor.

REFERENCES

ADAMS, B. C., BELL, L. C., & PERFETTI, C. A. (1995). A trading relationship between reading skill and domain knowledge in children's text comprehension. *Discourse Processes*, **20**, 307-323.

BENTON, S. L., CORKILL, A. J., SHARP, J. M., DOWNEY, R. G., & KHRAMTSOVA, I. (1995). Knowledge, interest, and narrative writing. *Journal of Educational Psychology*, **87**, 66-79.

BEREITER, C., & SCARDAMALIA, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Erlbaum.

DANEMAN, M., & MERIKLE, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, **3**, 422-434.

ENGLE, R. W., CANTOR, J., & CARULLO, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 972-992.

ERICSSON, K. A., & DELANEY, P. F. (1998). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 93-114). Cambridge: Cambridge University Press.

ERICSSON, K. A., & KINTSCH, W. (1995). Long-term working memory. *Psychological Review*, **102**, 211-245.

ERICSSON, K. A., & POLSON, P. G. (1988). An experimental analysis of the mechanisms of a memory skill. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 305-316.

FAYOL, M. (1998). From on-line management problems to strategies in written composition. In M. Torrance & G. C. Jeffery (Eds.), *The cognitive demands of writing: Processing capacity and working memory in text production* (pp. 13-23). Amsterdam: Amsterdam University Press.

HUOT, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational Research*, **60**, 237-263.

JONIDES, J., & SMITH, E. E. (1997). The architecture of working memory. In M. D. Rugg (Ed.), *Cognitive neuroscience* (pp. 243-276). Cambridge, MA: MIT Press.

JUST, M. A., & CARPENTER, P. A. (1992). A capacity theory of comprehension. *Psychological Review*, **99**, 122-149.

KAHNEMAN, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

KELLOGG, R. T. (1987). Effects of topic knowledge on the allocation of processing time and cognitive effort to writing processes. *Memory & Cognition*, **15**, 256-266.

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

KELLOGG, R. T. (1994). *The psychology of writing.* New York: Oxford University Press.

KELLOGG, R. T. (in press). Competition for working memory among writing processes. *American Journal of Psychology*.

KELLOGG, R. T., & MUELLER, S. (1993). Performance amplification and process restructuring in computer-based writing. *International Journal of Man-Machine Studies*, **39**, 33-49.

KINTSCH, W. (1988). The use of knowledge in discourse processing: A construction–integration model. *Psychological Review*, **95**, 163-192.

LEVY, C. M., & RANSDELL, S. (1996). Is writing as difficult as it seems? *Memory & Cognition*, **23**, 767-779.

McCUTCHEN, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory & Language*, **25**, 431-444.

McCUTCHEN, D. (1988). Functional automaticity in children's writing. *Written Communication*, **5**, 306-324.

McCUTCHEN, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, **8**, 299-325.

McNAMARA, D. S., KINTSCH, E., SONGER, N. M., & KINTSCH, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition & Instruction*, **14**, 1-43.

MORAVCSIK, J. E., & KINTSCH, W. (1993). Writing quality, reading skills, and domain knowledge as factors in text comprehension. *Canadian Journal of Experimental Psychology*, **47**, 360-374.

NICOLICH, R. M. (1997). The design of writing: Rhetorical and content interactions and epistemic change. *Dissertation Abstracts International*, **58**, 5665B. (UMI Microform No. 9812908)

PIOLAT, A., ROUSSEY, J.-Y., OLIVE, T., & FARIOLI, F. (1996). Charge mentale et mobilisation des processus rédactionnels: Examen de la procédure de Kellogg. *Psychologie Française*, **41**, 339-354.

RANSDELL, S., & LEVY, C. M. (1996). Working memory constraints on writing quality and fluency. In C. M. Levy & S. Ransdell (Ed.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 93-101). Mahwah, NJ: Erlbaum.

RECHT, D. R., & LESLIE, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, **80**, 16-20.

ROWAN, K. E. (1990). Cognitive correlates of explanatory writing skill: An analysis of individual differences. *Written Communication*, **7**, 316-341.

SCHNEIDER, W., KÖRKEL, J., & WEINERT, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology*, **81**, 306-312.

SINGER, M., & RITCHOT, K. F. M. (1996). The role of working memory and knowledge access in text inference processing. *Memory & Cognition*, **24**, 733-743.

SPILICH, G. J., VESONDER, G. T., CHIESI, H. L., & VOSS, J. F. (1979). Text processing of domain-related knowledge for individuals with high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior*, **18**, 275-290.

THIESMEYER, E. C., & THIESMEYER, J. E. (1990). *Editor: A system for checking usage, mechanics, vocabulary, and structure.* New York: Modern Language Association.

TORRANCE, M., & JEFFERY, G. C. (1999). Writing processes and cognitive demands. In M. Torrance & G. C. Jeffery (Eds.), *The cognitive demands of writing: Processing capacity and working memory in text production* (pp. 1-11). Amsterdam: Amsterdam University Press.

VOSS, J. F., & SILFIES, L. N. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition & Instruction*, **14**, 45-68.

VOSS, J. F., VESONDER, G. T., & SPILICH, G. J. (1980). Text generation and recall by high-knowledge and low-knowledge individuals. *Journal of Verbal Learning & Verbal Behavior*, **19**, 651-667.

WALKER, C. H. (1987). Relative importance of domain knowledge and overall aptitude on acquisition of domain-related information. *Cognition & Instruction*, **4**, 25-42.

WATERS, G. S., & CAPLAN D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology*, **49A**, 51-79.

NOTES

1. The texts were also analyzed to determine the use of cohesive ties between clauses, following the procedures of McCutchen (1986). On the basis of her findings with younger writers, it was anticipated that verbal ability and even domain knowledge would increase the degree of local coherence in the text. In the present experiment, the use of different kinds of linguistic ties was unrelated to individual differences in verbal ability and domain knowledge, although some differences between persuasive and narrative texts were found. Further, an automated analysis of the texts using the Editor software (Thiesmeyer & Thiesmeyer, 1990) assessed problems in usage, mechanics, and vocabulary. Although differences were again observed between the two types of texts, the key variables of domain knowledge and, particularly, verbal ability showed no effects in these analyses.

2. The data of all 80 participants were first analyzed ignoring the effects of verbal ability (because of missing ACT scores). For both narrative and persuasive texts, reliable main effects were observed for domain knowledge in terms of both content and style ratings.

3. The style and content ratings were also analyzed separately in Experiment 2. As before, they each revealed main effects and no interaction of domain knowledge and verbal ability. These effects were again found for both narrative and persuasive texts.