Concreteness and Imagery Effects in the Written Composition of Definitions

By: Mark Sadoski, William A. Kealy, Ernest T. Goetz, and Allan Paivio

Sadoski, M., Kealy, W. A., Goetz, E. T., & Paivio, A. (1997). Concreteness and imagery effects in the written composition of definitions. *Journal of Educational Psychology*, 89, 518-526. Digital Object Identifier: 10.1037/0022-0663.89.3.518

Abstract:

Concreteness and imagery effects have been found to be among the most powerful in explaining performance on a variety of language tasks. Concreteness and imagery effects involve the capacity of concrete language to evoke sensory images in the mind (e.g., *juicy watermelon*), whereas abstract language has relatively less capacity to do so (e.g., *agriculturalproduce*). The effects of concreteness and imagery on reading and text recall have been well-established (e.g., Goetz, Sadoski, Fatemi, & Bush, 1994; Paivio, 1971, 1986; Paivio, Walsh, & Bons, 1994; Sadoski, Goetz, & Avila, 1995; Sadoski, Goetz, & Fritz, 1993a, 1993b). Concrete words, phrases, sentences, and texts have been found to be more imageable, comprehensible, memorable, and interesting than abstract language units even when other relevant contextual variables are carefully controlled. These results can be consistently interpreted by dual coding theory (Paivio, 1971, 1986, 1991), which maintains that cognition involves the operation of two separate but interconnected systems, one for verbal representations and processes and one for nonverbal (imagery) representations and processes.

Article:

Although the operation of nonverbal imagery processes in language reception (e.g., reading) is well-established, little research has explored their function in language production (i.e., speech and writing). Some theorists have suggested that imagery as well as verbal processes might be involved in the composition of text, implicating concreteness as a semantic factor in composing (Flower & Hayes, 1984; Sadoski, 1992). In an empirical study, Tirre, Manelis, and Leicht (1979)assigned undergraduates to use either a verbal or an imagery strategy in reading concrete and abstract passages matched for comprehensibility. When students were asked to explain in writing how a set of content words from each passage were related, passage concreteness had a significant effect on quality, although strategy assignment did not. In other studies, imagery instructions produced improvements in the originality of the compositions of gifted elementary students (Jampole, Konopak, Readence, & Moser, 1991; Long & Hiebert, 1985).

The most in-depth study of the effect of concreteness on composition was conducted by Reynolds and Paivio (1968), who investigated the production of oral definitions of concrete and abstract words. That study involved 48 undergraduates, identified as either high or low in verbal-associative productivity (i.e., the number of words produced as associates of a given list of words), who were given 30 s per word to orally define five concrete and five abstract nouns equated for frequency of usage and meaningfulness. Definitions of concrete words had shorter latencies, contained more words, included shorter words, and were rated higher in quality than definitions of abstract words. High associative productivity students produced shorter latencies, more words per definition, and definitions rated higher in quality.

The present study was an effort to investigate the effects of concreteness and imagery on the processes and products of composing written language, extending the work of Reynolds and Paivio (1968)to writing.

Although there are obvious differences between oral and written composition, both are instances of language production and may share common cognitive mechanisms. Two experiments were used. The first experiment investigated the production of definitions for the full set of 10 words used by Reynolds and Paivio under a brief time limit, and the second experiment used a longer time limit. The process and product measures noted above were used along with additional variables from the literature on writing quality and strategy self-reports.

The most extensive available review of literature on written composition indicated that very few syntactic variables are consistent predictors of writing quality as determined by reliable ratings on some quality scale (Hillocks, 1986). Chief among these are the overall length of the composition and the use of final free modifiers in T-units. A T-unit is defined as an independent clause with all its modifiers, including dependent clauses (Hunt, 1965). A final free modifier is a modifier placed at the end of a T-unit (Christensen, 1978). For example, the modifier "quickly" could be placed before, within, or after the T-unit "The cat chased the mouse." T-unit length was not predictive of quality by itself, suggesting that the use of the modifier in the final position was more relevant than elaboration in general.

The notion of the final free modifier resonates somewhat with the conceptual peg hypothesis of dual coding theory. The conceptual peg hypothesis maintains that concrete words serve as mental "pegs" for following words and help to unify them into an integrated and memorable whole. Another aspect of the conceptual peg hypothesis suggests that processing concrete material before abstract material may have some transfer effect that enhances the processing of the abstract material (see Paivio, 1986, pp. 166–171 and Sadoski et al., 1993b, for reviews).

For purposes of the present experiments, the percentage of T-units with final free modifiers and overall length in words were included as syntactic correlates of writing quality. The number of T-units was also included as an exploratory variable.

Strategy self-reports were also collected to determine whether verbal-associative thinking, mental imagery, or both strategies were used in composing each definition. Dual coding theory would predict that the verbal-associative strategy would be more likely to occur in composing definitions of abstract words, whereas the imagery strategy would be more likely to occur in composing definitions of concrete words (Paivio, 1986).

Experiment 1

Experiment 1 was designed to replicate and extend aspects of the study by Reynolds and Paivio (1968)except that written rather than oral definitions were elicited. Using dual coding theory and the results of Reynolds and Paivio, the definitions of concrete words were predicted to exhibit shorter latencies, greater length, higher quality content, a greater percentage of T-units with final free modifiers, and more self-reports of an imagery strategy. Definitions of abstract words were predicted to exhibit longer words and more self-reports of a verbal-associative strategy. Consistent with Reynolds and Paivio's findings, participants with higher associative productivity scores were predicted to exhibit shorter latencies, longer definitions, and higher quality ratings. In addition to ratings for the quality of the definitions, a rating for textuality was included in this experiment for exploratory purposes. Because in this experiment participants had only a brief time period to write each definition, they had little opportunity to attend closely to grammar, coherence, and mechanics. Definitions were rated on a 5-point scale in terms of their progression from a list of disconnected words or phrases to a series of grammatical sentences that constituted a cohesive and nonredundant text. Content and style have been treated traditionally as separate factors in scoring writing analytically (Cooper, 1977). Number of T-units was also included for exploratory purposes.

The multivariate nature of the study also afforded an opportunity to use correlational techniques to investigate relationships and detect any underlying structure in the data to provide further theoretical insights. In particular, Paivio (1986)has contended that there may be a relationship between word concreteness and the volubility of the definition because organized mental images activated by the concrete words might in turn activate organized, fluent language.

Method

Participants

Forty-eight undergraduates enrolled in an upperclass education course at Texas A & M University participated in the study for extra credit; 38 were female, 10 were male.

Materials

The five concrete and five abstract nouns used by Reynolds and Paivio (1968)and matched for frequency of usage (Thorndike & Lorge, 1944) and meaningfulness values (Noble, 1952) were used as stimuli. The concrete words were *library, prisoner, picture, hotel,*and *mother;*the matched abstract words were *crime, science, mind, fun,*and *death.*

Procedure

The stimulus words were presented on microcomputers. Students were tested in groups in a microcomputer laboratory. They were instructed to press a key that brought instructions to the screen explaining that they would be defining 10 common words, each within a 90-s time limit, and instructing them to write clearly at a normal rate and not to be excessively concerned with the vocabulary, grammatical correctness, spelling, or editing of their definitions.

The 10 words were then presented one at a time near the top left of the screen for 90 s in an individually randomized order. Through pilot testing, we had determined that 90 s was sufficient to produce reasonable, dictionary-style definitions for one usage of each word. Latencies from the appearance of the word until the first keystroke and the definitions produced were automatically recorded. After 90 s, the screen flashed blue, and the next word appeared.

Immediately after all 10 definitions were completed, a screen appeared with the instructions that the words just defined would be shown again together with rating tasks for the thinking strategies that might have occurred while defining the words. The words then appeared again individually in random order near the top left of the screen. The two rating tasks appeared beneath the words one at a time in a random sequence. The imagery rating task read “When I was defining this word, I thought of mental pictures of objects, scenes, or events.” The verbal-associative rating task read “When I was defining this word, I thought of other words, phrases, and related language.” Using the microcomputer's mouse, students were directed to click on one of four boxes to indicate how much a particular strategy was used: “not at all,” “very little,” “somewhat,” or “very much.” These choices were coded as 1–4, respectively.

When the ratings were completed, a sequence of screens directed students through the associative productivity test. After an introduction and two trial items, students viewed a sequence of 10 words presented in an individually randomized order. For each word they typed as many associates as possible in 60 s. Responses were automatically recorded. When the associative productivity test was completed, two screens of debriefing concluded the procedures.

Results

Scoring

Definitions were analyzed for the number of T-units per definition and the percentage of T-units with final free modifiers by a research assistant trained in grammatical analysis but naive to the purposes of the study. Definitions were also scored on 5-point scales for quality (content) and textuality (style). The rating scales are given in Appendix A. One of the researchers and a research assistant naive to the purposes of the study independently scored a randomly selected 10% of the definitions to assess reliability. Interrater correlations were .86 for content and .91 for style.

Strategy ratings exhibited alpha reliability coefficients of .63 for the imagery strategy and .67 for the verbal-associative strategy. Because these reliabilities were likely attenuated by the small number of items rated (i.e., 10), Spearman-Brown estimates were calculated for the respective reliabilities of rating 25 items. The results were .81 for the imagery strategy and .84 for the verbal-associative strategy. As a validity check on the strategy ratings, the mean strategy ratings for the 10 words were correlated with the concreteness norms for the respective words established in Paivio, Yuille, and Madigan (1968). In those norms, the higher the rating the more concrete the word, and the lower the rating the more abstract the word. The results were $r = .71$, $p < .05$, for the imagery strategy, and $r = -.73$, $p < .05$, for the verbal-associative strategy.

Associative productivity scores were determined by summing the number of nonredundant responses for each word. Replicating the study by Reynolds and Paivio (1968), a median split was used to discriminate high- and low-associative productivity scores. The median score was 7.2 ($M = 7.24$, $SD = 1.72$). Means and standard deviations for the dependent variables, aggregated across the five concrete and five abstract words for each participant, are given in Table 1.

Table 1
Means and Standard Deviations for Characteristics of Definitions in Experiment 1

Dependent variables	Independent variables											
	Concrete words						Abstract words					
	Low AP		High AP		Total		Low AP		High AP		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Latency	4.23	1.27	4.24	1.85	4.23	1.57	5.36	1.93	6.20	3.41	5.78	2.77
Number of words	26.25	7.99	26.26	7.14	26.25	7.49	20.38	6.32	21.81	7.71	21.09	7.01
Length of words	4.50	0.30	4.54	0.37	4.52	0.33	4.74	0.25	4.84	0.37	4.79	0.32
Number of T-units	0.69	0.67	1.05	1.00	0.87	0.86	0.66	0.64	1.02	1.12	0.84	0.92
% of T-units with FFM	40.98	34.47	47.61	33.08	44.29	33.59	37.30	32.50	37.35	36.04	37.33	33.95
Content score (0–4)	3.23	0.59	3.29	0.56	3.26	0.57	2.54	0.67	2.70	0.59	2.62	0.63
Style score (0–4)	2.48	0.46	2.71	0.56	2.60	0.52	2.43	0.47	2.69	0.66	2.56	0.58
Imagery strategy (1–4)	3.61	0.31	3.57	0.41	3.59	0.36	2.92	0.57	2.93	0.61	2.92	0.58
Verbal strategy (1–4)	2.79	0.33	2.78	0.52	2.78	0.43	3.03	0.54	3.12	0.47	3.07	0.50

Note. AP = associative productivity; FFM = final free modifier.

Means and Standard Deviations for Characteristics of Definitions in Experiment 1

In interpreting the style scores, it should be noted that although the means for the style scores appear to be low (i.e., the means in ranged from 2.4 to 2.7), students' responses were in most cases grammatical sentences or close approximations. A rating of 2 was given to phrases and clauses that could be easily expanded into grammatical, well-written sentences; a rating of 3 indicated one or more grammatical, well-written sentences. Most definitions rated 2 could be read as sentences because students used the word on the screen as the first word of the definition, leaving introductory determiners (e.g., *a*) and linking verbs (e.g., *is*) to be understood. For example, a typical definition read as follows, with the word in italics as given on the screen, words in parentheses inferred, and the rest of the definition as it was written by a student: "*Science*(is) Studying something, usually in an objective sense rather than a subjective one. *Science*(is) Concerned mainly with provable facts or statements, rather than values or opinions." Such a definition would receive a style score of 2, but this may largely be an artifact of having the word to be defined displayed on the screen and the students taking advantage of its presence in coping with the brief time limit.

Analysis of variance (ANOVA)

Each dependent variable in was subjected to a 2 (word concreteness: concrete or abstract) × 2 (associative productivity: high or low) mixed ANOVA with concreteness as a within-subject variable and associative productivity as a between-subjects variable. A preliminary multivariate analysis of variance (MANOVA) revealed that only the main effect of concreteness was significant, $F(9, 38) = 23.07$, $p < .0001$; the effects of associative productivity and the interaction also failed to reach significance in any of the univariate analyses. Therefore, only the tests of concreteness effects will be reported. Effect sizes (ES), determined by the difference between the means divided by the pooled standard deviation (Cohen, 1988), are also reported.

The analysis for latency revealed that the participants started typing their definitions sooner after the presentations of concrete words, $F(1, 46) = 22.16$, $p < .0001$, ES = .65. The definitions of concrete words contained significantly more words, $F(1, 46) = 44.95$, $p < .0001$, ES = .67. The definitions of abstract words had significantly longer words, $F(1, 46) = 37.89$, $p < .0001$, ES = .77. No significant effect was found for number of T-units in the definitions. The definitions of concrete words had a significantly greater percentage of T-units with final free modifiers, $F(1, 46) = 4.73$, $p < .03$, ES = .21.

The definitions of concrete words were rated significantly higher in quality, $F(1, 46) = 74.05$, $p < .0001$, ES = .94. No significant effect was found for style scores.

Reported use of an imagery strategy was significantly higher for concrete words, $F(1, 46) = 70.51$, $p < .0001$, ES = 1.14. Reported use of a verbal-associative strategy was significantly higher for abstract words, $F(1, 46) = 18.99$, $p < .0001$, ES = .60.

For each of the variables that exhibited a significant effect in the ANOVAs by participants, a corollary one-way ANOVA by items was performed to determine whether the effects were generalizable across the 10 words studied (Clark, 1973). Data were averaged across participants for each of the items, producing 10 observations, 5 concrete and 5 abstract. Significant concreteness effects ($p < .05$) were found for all variables except the percentage of T-units with final free modifiers.

Correlational analyses

Bivariate correlational analyses were performed on data aggregated across the concrete and abstract words for each participant, producing 96 observations. As shown in Table 2, tests of significance were conducted using 46 degrees of freedom, reflecting the fact that there were two nonindependent observations for each of the 48 participants. The associative productivity score for each student was added to these correlational analyses as a continuous variable.

Table 2
Correlation Matrix of Variables in Experiment 1

Variables	1	2	3	4	5	6	7	8	9	10
1. Latency	—	−.39	.17	−.21	−.15	−.34	−.15	−.18	.21	−.00
2. Number of words		—	−.16	.60	.49	.72	.56	.29	.07	.18
3. Length of words			—	.17	−.16	.11	−.18	−.14	.27	.14
4. Number of T-units				—	.77	.19	.90	−.02	.06	.22
5. % of T-units with FFM					—	.15	.82	.15	.07	.13
6. Content score						—	.16	.40	.04	.16
7. Style score							—	−.04	.05	.26
8. Imagery strategy								—	−.03	−.02
9. Verbal strategy									—	.07
10. AP										—

Note. With 46 degrees of freedom, $p < .05$ at $r = \pm.28$; $p < .01$ at $r = \pm.37$. FFM = final free modifiers; AP = associative productivity.

Correlation Matrix of Variables in Experiment 1

To investigate the underlying structure of the correlations, a principal components factor analysis was performed. Using the 1.0 eigenvalue criterion, three factors, accounting for 66% of the variance, were retained and rotated to an oblique (orthoblique) solution. The resulting factor loadings are displayed in Table 3. Factor 1 correlated with Factor 2 at $r = -.05$, Factor 1 correlated with Factor 3 at $r = -.08$, and Factor 2 correlated with Factor 3 at $r = .32$.

Table 3
Principal Components Structure for all Variables After Orthoblique Rotation in Experiment 1

Variables	Factor 1	Factor 2	Factor 3
Latency	.43	−.20	−.57
Number of words	−.06	.64	.81
Length of words	.77	−.22	−.03
Number of T-units	−.07	.95	.30
% of T-units with FFM	−.09	.88	.28
Content score	.12	.21	.89
Style score	−.06	.96	.25
Imagery strategy	−.20	−.02	.64
Verbal strategy	.69	.09	−.00
AP	.42	.30	.17

Note. FFM = final free modifiers; AP = associative productivity.

Principal Components Structure for all Variables After Orthoblique Rotation in Experiment 1

Discussion

Experiment 1 was an extension of Reynolds and Paivio (1968)using written rather than oral composition. In general, the predictions of dual coding theory were strongly confirmed.

As in the original study, the definitions of concrete words exhibited significantly shorter latencies, greater length in words, shorter words, and higher quality content. Unlike the original study, no significant main effects were found for associative productivity, although the means for the high associative productivity participants were virtually always higher than the means for the low associative productivity participants (i.e., the means were in the predicted direction).

Self-reports of the use of imagery and verbal-associative strategies also confirmed the predictions of dual coding theory. Reported use of an imagery strategy was significantly greater for concrete words, whereas reported use of the verbal-associative strategy was significantly greater for abstract words.

Syntactic variables generally favored the definitions of concrete words. Overall length in words and the percentage of T-units with final free modifiers are the two syntactic variables most consistently predictive of ratings of writing quality (Hillocks, 1986). As noted above, overall length in words was significantly greater for the definitions of concrete words. Although the number of T-units was equivalent for the definitions of concrete and abstract words, the percentage of T-units with final free modifiers was significantly greater for the definitions of concrete words for participants but not for items.

Theoretical support for dual coding theory can also be seen in the correlational evidence. Paivio and Begg (1981)and Paivio (1986)held that the results of Reynolds and Paivio (1968)were consistent with the interpretation that concrete words generated more voluble definitions because complex nonverbal images referentially activated extensive descriptions in the verbal system. The relative fluency and, to some extent, the grammatical ordering of the definitions could be the result of the availability of organized, complex images (cf. Segal, 1976, discussed in Paivio, 1975). The definitions of the abstract terms presumably relied to a greater degree on properties of the verbal system such as the availability of verbal-associative sequences as measured, for example, by the associative productivity test. This interpretation is supported by the factor-analytic results here.

Factor 3 received substantial loadings from the imagery strategy, number of words, and content ratings, suggesting the use of images in the evocation of verbal information to articulate a fluent, meaningful definition. The negative loading of latency on this factor (i.e., starting sooner) also supports the fluency interpretation. Factor 2 received substantial loadings from all of the grammatical variables. Its correlation with Factor 3 (.32) suggests that the image system plays some positive role in the generation of sentences describing concrete objects and events. Specifically, concrete words have more access to the imagery system and the organized scenes it can create, which in turn may be more readily described in organized, grammatical language.

Factor 1 is uncorrelated with the other factors and received substantial loadings from the verbal strategy and word length and moderate loadings from the associative productivity score and latency (i.e., starting later). This factor suggests a lexically based organization of associated words and phrases that differs from syntactic knowledge and may require time to sequence syntactically and fluently. This is consistent with findings that lexical and grammatical abilities are factorially separate, and that these linguistic abilities are factorially separate from imaginal abilities (Carroll, 1993; Paivio, 1986).

In a more general sense, the factors can be interpreted as a straightforward progression from the level of lexical knowledge (Factor 1) to the level of syntactic productivity (Factor 2) to the level of semantic productivity (Factor 3). This is consistent with a depth of processing interpretation that incorporates both verbal and nonverbal processes and relationships between them (Paivio, 1986).

Experiment 2

Experiment 2 was designed to replicate and extend certain aspects of Experiment 1 using a more ecologically valid time limit. Writing accurate definitions is a productive classroom task for secondary school students (Hillocks, 1995; Hillocks, Kahn, & Johannessen, 1983; Johannessen, 1989). For example, in an experimental investigation, students who were given sets of concrete data to analyze (e.g., newspaper reports of a variety of crimes) in a classroom setting learned to compose higher quality definitions of abstract concepts (e.g., accomplice) than students who were given model definitions to imitate (Hillocks et al., 1983). In the present experiment, participants were given more extended time limits to write full and polished paragraphs of definition for subsets of one concrete word and one abstract word from the set of 10 words used in Experiment 1.

Half of the participants defined a concrete word first and half defined an abstract word first. Besides its function as an experimental control, counterbalancing allowed for the investigation of order effects. Consistent with the conceptual peg hypothesis of dual coding theory, the effect of presenting a concrete language unit before an abstract language unit has been found to enhance the recall of the abstract language unit in a variety of studies of reading (Anderson, Goetz, Pichert, & Halff, 1977; Corkill, Glover, & Bruning, 1988; Paivio, 1965; Royer & Cable, 1975; Sadoski et al., 1993b). An extension of the investigation of this hypothesis to effects on composing was therefore included here.

Method

Participants

Fifty undergraduates enrolled in upperclass education courses at Texas A & M University participated in the study for extra credit; 29 were female, 21 were male.

Materials

The set of 10 words from Experiment 1 was used. Each student received 1 concrete word and 1 abstract word from that set. Each of the 25 possible pairs of words was used twice in counterbalanced order (i.e., concrete, abstract; abstract, concrete).

Procedure

The stimulus words were presented on microcomputers as they were in Experiment 1. Procedures were the same except for the following changes. Students were told that they would be asked to define two common words in one paragraph each with 15 min per definition. They were asked to write polished paragraphs, that is, to organize their ideas, use complete sentences, choose words carefully, and use correct mechanics (e.g., grammar, spelling, punctuation). The words were then presented near the top left of the screen one at a time and remained there for 15 min. Once the student began to type, a small clock displaying the remaining time appeared in the upper right corner of the screen. After the definitions were completed, the students were given the strategy rating task as in Experiment 1. The associative productivity test and the style score were deleted from this experiment because they failed to produce any significant effects in Experiment 1. Analysis by items was also eliminated because the design included different sets of participants per item mean.

Results

Scoring

Representative samples of definitions are exhibited in Appendix B. Definitions were analyzed for number of T-units, percentage of T-units with final free modifiers, and content scores by a trained research assistant naive to the purposes of the study. Because the amount of content was greater in the longer definitions in this experiment, the content scale from Experiment 1 was used but left open-ended. The range of content scores was 3–17. Interrater reliability for content scores between a research assistant naive to the purposes of the study and one of the researchers on a randomly selected 10% of the data was $r = .87$. Means and standard deviations for the concrete and abstract words by order are given in Table 4.

Table 4
Means and Standard Deviations for Characteristics of Definitions in Experiment 2

Dependent variables	Independent variables											
	Concrete words						Abstract words					
	C first		A first		Total		C first		A first		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Latency	13.58	10.69	10.67	6.69	12.13	8.95	16.08	14.57	17.20	11.85	16.64	13.16
Number of words	155.04	56.62	157.08	53.56	156.06	54.55	159.92	50.69	130.20	45.79	145.06	50.11
Length of words	5.59	.32	5.63	.29	5.61	.31	5.50	.30	5.76	.34	5.64	.35
Number of T-units	9.80	4.37	10.28	4.12	10.04	4.21	11.28	4.48	9.40	3.83	10.34	4.23
% of T-units with FFM	89.56	10.67	91.16	10.68	90.36	10.60	85.40	13.72	92.08	10.54	88.74	12.57
Content score	8.84	2.66	8.76	2.77	8.80	2.69	7.84	2.98	7.80	2.52	7.82	2.73
Imagery strategy (1–4)	3.44	.77	3.76	.52	3.60	.67	2.72	1.06	3.32	.80	3.02	.98
Verbal strategy (1–4)	3.12	.88	2.68	.63	2.90	.79	3.24	.72	3.28	.79	3.26	.75

Note. C = concrete word; A = abstract word; FFM = final free modifier.

Means and Standard Deviations for Characteristics of Definitions in Experiment 2

ANOVA

The dependent variables in were subjected to 2 (word concreteness: concrete or abstract) × 2 (order: concrete first and abstract second or abstract first and concrete second) mixed ANOVAs with concreteness as a within-subject variable and order as a between-subjects variable. A preliminary MANOVA revealed significant effects for concreteness, $F(8, 41) = 6.62$, $p < .0001$; order, $F(8, 41) = 2.57$, $p < .02$; and the Concreteness × Order interaction, $F(8, 41) = 2.32$, $p < .04$.

For latency, only a significant main effect for concreteness was found, $F(1, 48) = 4.76$, $p < .03$, ES = .39. Students started typing the definitions of concrete words sooner.

For number of words, a marginally significant main effect for concreteness was found, $F(1, 48) = 3.64$, $p < .06$, ES = .21. The main effect for order was not significant, but the Concreteness × Order interaction was significant, $F(1, 48) = 7.58$, $p < .008$, ES = .96. Definitions of abstract words tended to be shorter overall; when an abstract word was defined first, its definition was much shorter than in the other three conditions.

For length of words, the main effect for concreteness was not significant. The main effect for order was significant, $F(1, 48) = 4.84$, $p < .03$, ES = .45, as was the Concreteness × Order interaction, $F(1, 48) = 4.85$, $p < .03$, ES = .92. Students used longer words in both definitions when the abstract word was defined before the concrete word. However, when the abstract word was defined first, the words in that definition were longer than in the other three conditions.

For number of T-units, only the Concreteness × Order interaction was significant, $F(1, 48) = 6.57$, $p < .01$, ES = 1.03. When the abstract word was defined first, fewer T-units were used than for the other three conditions. For the percentage of T-units with final free modifiers, no effects were significant.

For content scores, only the main effect for concreteness was significant, $F(1, 48) = 7.56$, $p < .008$, ES = .36. The definitions of concrete words were rated higher in quality.

For the imagery strategy rating, the only significant effects were for concreteness, $F(1, 48) = 14.63$, $p < .0001$, ES = .66; and order, $F(1, 48) = 7.13$, $p < .01$, ES = .52. Reported use of an imagery strategy was greater for concrete words and for the order in which a concrete word was defined last.

For the verbal-associative strategy rating, only the main effect for concreteness was significant, $F(1, 48) = 5.13$, $p < .03$, ES = .46. Reported use of a verbal-associative strategy was greater for abstract words.

Discussion

Most of the findings for Experiment 1 were replicated with the longer time limit and instructions to write a polished paragraph. Definitions of concrete words exhibited shorter latencies, marginally greater length in words, higher quality content, and greater reported use of an imagery strategy. Definitions of abstract words exhibited greater use of a verbal-associative strategy. However, all effect sizes were smaller, and differences in word length and use of the final free modifier construction failed to reach significance.

Defining a concrete word first had a transfer effect on the definitions of abstract words. Abstract words defined after a concrete word had the most words and T-units; abstract words defined first had the fewest words and T-units. For word length, the opposite pattern occurred: the definitions of abstract words defined after a concrete word had the shortest words; the definitions of abstract words defined first had the longest words. Hence, the results for the abstract word defined first are generally consistent with the results for the definitions of the abstract words in Experiment 1 (i.e., fewer words and longer words). Evidence of such a transfer effect is consistent with the conceptual peg hypothesis of dual coding theory. However, there was no evidence of transfer on the quality of the definitions; they were uniformly higher for concrete words. The greater reported use of an imagery strategy when a concrete word was defined last was an unexpected finding and may suggest a recency effect in which the use of imagery with the last word defined generalized to the rating for the abstract word.

General Discussion

Across both experiments, when college students wrote definitions of concrete words, they began sooner, wrote more, produced definitions rated higher in quality, and reported using an imagery strategy more than when defining abstract words that were matched for frequency and meaningfulness. When defining abstract words, students used a verbal-associative strategy more. When the time limit was restricted, students also used a greater percentage of final free modifiers in defining concrete words, a construction consistently related to writing quality in previous research; they also used longer words in defining abstract words, a variable not consistently related to quality in previous research. As might be expected, effects were reduced when students were instructed to write polished paragraphs and given an ample time limit, especially for syntactic variables.

However, some differences remained stable. Under the more restricted time limit, students took 1.37 times longer to begin writing the definition of an abstract term than for a concrete term. Under the longer time limit, the proportional increase in latency was identical, 1.37 times longer. The number of T-units used did not significantly differ under either time limit, but this variable has not been consistently associated with writing quality in previous research. Overall, measures consistently associated with the quality of written composition were positively affected by the concreteness of the language to be defined and the use of mental imagery as a composing strategy. These results have both theoretical and practical implications.

An important theoretical implication is that common cognitive mechanisms appear to be involved in the production of both spoken and written language. These mechanisms appear to be those involved in the reception of spoken and written language as well (e.g., Paivio, 1971, 1986, 1991; Sadoski et al., 1993a, 1993b). The results of this study are consistent with general and specific predictions of dual-coding theory for verbal tasks.

The present study was an initial effort to bring predictions from an established general theory of cognition to bear on the empirical study of written composition. Other cognitive theories of composition have been specific to writing and have not been subjected to experimental test (e.g., Emig, 1971; Hayes & Flower, 1980). The present study derived and tested hypotheses from dual coding theory and extended empirical findings about language production to the investigation of written composition.

Although these results provide some new theoretical insights and directions, their limitations should be observed. Only the task of composing definitions was studied here. The time limits, while varying, were relatively brief and did not allow for external feedback. The scores for content and style should not be taken as adequate measures of the more sophisticated and extended texts that might be composed over several drafts with feedback. Studies of longer composing episodes in different genres and classroom situations are necessary.

Although the present research was theoretically motivated, it suggests some practical implications. The use of concrete concepts and the reported use of an imagery strategy in composing definitions was associated with both higher quality content and some effective stylistic devices. Simply being asked to define concrete concepts in finished paragraphs produced an effect size in quality ratings (.36, Experiment 2) over defining abstract concepts that was greater than the mean experimental group–control group effect size for 73 instructional interventions included in the most extensive meta-analysis of writing instruction to date (.28; Hillocks, 1986). This effect size is also larger than the mean experimental group–control group effect size for writing quality produced by using word processors in writing instruction in 20 studies (.27; Bangert-Drowns, 1993). Therefore, although effect sizes in Experiment 2 were attenuated relative to Experiment 1, they were still quite substantial from an instructional point of view. Stylistically, the use of mental imagery as a composing strategy and the use of concrete concepts as content may result in developing the few syntactic writing characteristics consistently associated with higher ratings for quality (i.e., length in words and the use of final free modifiers; Hillocks, 1986).

These findings also support the use of concrete data and imagination in composing definitions of concepts in classroom writing activities (Hillocks, 1995). Such activities have been associated with improved writing in both quantitative and qualitative studies (Hillocks et al., 1983; Johannessen, 1989). Hillocks (1995)contended that writing careful, analytical definitions is an important activity in improving writing "because they lie at the heart of analysis, argument, and the dialectical processes that drive inquiry in every field" (p. 58).

References

Anderson, R. C., Goetz, E. T., Pichert, H. M., & Halff, H. M. (1977). Two faces of the conceptual peg hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 142–149.

Bangert-Drowns, R. L. (1993). The word processor as an instructional tool: A meta-analysis of word processing in writing instruction. *Review of Educational Research*, *63*, 69–93.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Christensen, F. (1978). A generative rhetoric of the sentence. In B.Christensen (Ed.), *Notes toward a new rhetoric* (2nd ed., (pp. 23–44). New York: Harper & Row.

Clark, H. H. (1973). The language-as-a-fixed-effect-fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

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

Cooper, C. R. (1977). Holistic evaluation of writing. In C. R.Cooper & L.Odell (Eds.), *Evaluating writing: Describing, measuring, judging* (pp. 3–31). Urbana, IL: National Council of Teachers of English.

Corkill, A. J., Glover, J. A., & Bruning, R. H. (1988). Advance organizers: Concrete vs. abstract. *Journal of Educational Research*, *82*, 76–81.

Emig, J. (1971). *The composing processes of twelfth graders*. (Research Rep. No. 13). Urbana, IL: National Council of Teachers of English.

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

Goetz, E. T., Sadoski, M., Fatemi, Z., & Bush, R. (1994). That's news to me: Readers' responses to brief newspaper articles. *Journal of Reading Behavior*, *26*, 125–138.

Hayes, J. R., & Flower, L. S. (1980). Writing as problem solving. *Visible Language*, *14*, 388–399.

Hillocks, G., Jr. (1986). *Research on written composition: New directions for teaching*. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills and the National Conference on Research in English.

Hillocks, G., Jr. (1995). *Teaching writing as reflective practice*. New York: Teachers College Press.

Hillocks, G., Jr., Kahn, E. A., & Johannessen, L. R. (1983). Teaching defining strategies as a mode of inquiry: Some effects on student writing. *Research in the Teaching of English*, *17*, 275–284.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. (Research Rep. No. 3). Urbana, IL: National Council of Teachers of English.

Jampole, E. S., Konopak, B. C., Readence, J. E., & Moser, E. B. (1991). Using mental imagery to enhance gifted elementary students' creative writing. *Reading Psychology, 12*, 183–197.

Johannessen, L. R. (1989). Teaching writing: Motivating inquiry. *English Journal, 78*, 64–66.

Long, S., & Hiebert, E. H. (1985). Effects of awareness and practice in mental imagery on creative writing of gifted children. In J. A.Niles & R. V.Lalik (Eds.), *Issues in literacy: A research perspective. Thirty-fourth Yearbook of the National Reading Conference* (pp. 381–385). Rochester, NY: National Reading Conference.

Noble, C. E. (1952). An analysis of meaning. *Psychological Review*, *59*, 421–430.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associated learning. *Journal of Verbal Learning and Verbal Behavior*, *4*, 32–38.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

Paivio, A. (1975). Imagery and synchronic thinking. *Canadian Psychological Review*, *16*, 147–163.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.

Paivio, A., & Begg, I. (1981). *The psychology of language*. Englewood Cliffs, NJ: Prentice-Hall.

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why?*Journal of Experimental Psychology: Learning, Memory,and Cognition*, *20*, 1196–1204.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). *Concreteness, imagery, and meaningfulness values for 925 nouns*. *Journal of Experimental Psychology Monographs, 78* (1, Pt. 2).

Reynolds, A., & Paivio, A. (1968). Cognitive and emotional determinants of speech. *Canadian Journal of Psychology*, *22*, 164–175.

Royer, J. M., & Cable, G. W. (1975). Facilitated learning in connected discourse. *Journal of Educational Psychology*, *67*, 116–123.

Sadoski, M. (1992). Imagination, cognition, and persona. *Rhetoric Review*, *10*, 266–278.

Sadoski, M., Goetz, E. T., & Avila, E. (1995). Concreteness effects in text recall: Dual coding or context availability?*Reading Research Quarterly*, *30*, 278–288.

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993a). A causal model of sentence recall: Effects of familiarity, concreteness, comprehensibility, and interestingness. *Journal of Reading Behavior*, *25*, 5–16.

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993b). Impact of concreteness on comprehensibility, interest, and memory for text: Implications for dual coding theory and text design. *Journal of Educational Psychology*, *85*, 291–304.

Segal, A. U. (1976). *Verbal and nonverbal encoding and retrieval differences* Unpublished doctoral dissertation, University of Western Ontario, London, Canada.

Tirre, W. C., Manelis, L., & Leicht, K. L. (1979). The effects of imaginal and verbal strategies on prose comprehension by adults. *Journal of Reading Behavior, 11,* 99–106.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's book of 30, 000 words*. New York: Columbia University, Teachers College, Bureau of Publications.

APPENDICES

APPENDIX A: Rating Scales for Definitions: Experiment 1

	Score
0	No response or uninterpretable response.
1	One substantial, defining class or characteristic. Count all examples as one defining characteristic.
2	Two independent, substantial, defining classes or characteristics. Each characteristic offers new illumination or illustration.
3	Three independent, substantial, defining classes or characteristics.
4	More than three independent, substantial, defining classes or characteristics.

	Score
0	No response or uninterpretable response.
1	Lists of words or brief phrases that do not approximate grammatical, well-written sentences.>
2	Phrases and clauses that can be easily expanded into grammatical, well-written sentences.
3	One or more complete, grammatical, well-written sentences.
4	Complete, grammatical, well-written sentences that are organized into a cohesive and nonredundant text.>

APPENDIX B

A library is a building that houses books and documents that people may borrow to read. They serve many purposes whether it be academic or for entertainment. Libraries are essential to a literate society in order to spread the knowledge of the past and present. In addition to books, most libraries have periodicals and films. Libraries have books on so many topics it would be almost impossible to list them all. Most schools have a library to aid in the learning studying of its students. Most towns, small and large, have public libraries which give free access to books for all its inhabitants. Libraries allow people to take some books home for a period of time. Some books cannot be removed, so most libraries are equipped with facilities that allow them to use these materials on the premises. People of all occupations find libraries helpful for work related purposes. Libraries help the public find out about past events or historic figures.

Science consists of a wide variety of topics. In short, science explores problems and consists of experiments to help solve these problems. There are many different branches of science. These include chemistry, which deals with chemicals; biology, which deals with the living systems of animals and plants; botany, which deals with flowers and other plants; geology, which deals with minerals and rocks; and geography, which is concerned with the makeup of various features of the planet Earth. These are but a few of many branches of science. Science seeks to explore and investigate. It attempts to uncover previous unknown data and to further explain and develop already known information through experiments. Many experiments follow a procedure known as the scientific method, which through a series of well-defined steps, attempts to prove a hypothesis (a guess that is made about the experiment before the experiment is begun).