Available online at www.sciencedirect.com
SCIENCE DIRECT®

Contemporary Educational Psychology 30 (2005) 242–262

Contemporary Educational Psychology

www.elsevier.com/locate/cedpsych

What limits the encoding effect of note-taking? A meta-analytic examination

Keiichi Kobayashi

Faculty of Education, Shizuoka University, Shizuoka, 422-8529, Japan

Available online 20 January 2005

Abstract

Previous meta-analyses indicate that the overall encoding effect of note-taking is positive but modest. This meta-analysis of 57 note-taking versus no note-taking comparison studies explored what limits the encoding effect by examining the moderating influence of seven variables: intervention, schooling level, presentation mode and length, test mode, and publication year and source. It was found that (a) either positive interventions or rise in schooling level did not enhance the benefits of note-taking; (b) visual presentation of learning material interfered with the note-taking process, whereas longer presentation did not; (c) recall test detected the encoding effect more than recognition and higher-order performance tests; and (d) publication year and source contributed to the variation in effect sizes. These results suggest that the modest encoding effect is not due to the incompleteness of students' spontaneous note-taking procedures, but mechanical demands of note-taking, type of learning outcome measure, and publication characteristics.

Keywords: Note-taking; Encoding effect; Meta-analysis; Generative processing; Quality of note-taking procedure; Mechanical demands; Learning outcome measure; Publication characteristics

E-mail address: ekkobay@ipc.shizuoka.ac.jp.

0361-476X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.cedpsych.2004.10.001

1. Introduction

The present meta-analysis focused on the encoding hypothesis in note-taking studies. This hypothesis is primarily concerned with the questions of whether the process of taking notes promotes the encoding of lecture or text information, and if so, how much and why (Di Vesta & Gray, 1972). The encoding effect of note-taking can be estimated by comparing learning outcomes of note-takers with those of non-note-takers under the condition that they are given no opportunity of reviewing their own or provided notes afterward.

Examination of the encoding effect is important at least in two ways. First, note-taking is generally considered as an effective learning strategy. Many students believe in the positive effect of note-taking process itself on the learning performance. According to Van Meter, Yokoi, and Pressley (1994), for example, college students shared beliefs that the act of taking notes facilitates attending to the lecture, comprehension of the material to be learned, and the subsequent recall. Note that such beliefs probably affect their study behavior during or after class (Ryan, 2001). As a practical issue of school learning, it is important to clarify whether the students' beliefs about the encoding effect are well founded.

Second, note-taking is one of the commonest study activities in school settings (e.g., Carrier & Newell, 1984; Hartley & Davies, 1978; Lahtinen, Lonka, & Lindblom-Ylanne, 1997). During a lecture or text-learning, students are required to incessantly coordinate taking notes and other ongoing processes (e.g., listening, reading). This may have some influence on the lecture or text processing, whether note-taking directly promotes encoding or not. For example, students might give up grasping the point of a lecture while they are busy writing down what the lecturer said. Closer examination of the encoding effect should be informative to reveal what and how students (can) learn while being in class. Of course, in real academic situations, students often benefit from reviewing their notes after class more than taking them during class. It must be cautious to judge the practical value of note-taking only by the encoding effect (Kiewra, 1989; Van Meter et al., 1994).

Since 1920s, the field of educational psychology has garnered a large number of findings about the encoding effect. It is expected that the synthesis of those findings may provide a wealth of information about the encoding hypothesis. Some researchers (e.g., Bligh, 2000; Carrier & Titus, 1979; Kiewra, 1985, 1989; Ladas, 1980b; Suritsky & Hughes, 1991) have attempted to organize the divergent findings and to resolve the inconsistency through narrative and vote-counting reviews. For example, Hartley (1983) gathered 57 note-taking studies and found that 34 studies favored the positive effect of note-taking, 19 studies yielded nonsignificant results, and 4 studies revealed the negative effect. He concluded from this result of vote-counting that note-taking enhances learning in certain conditions. Despite the usefulness in presenting theoretical arguments and cataloguing research findings, however, there are limits for these approaches to rigorously summarize data across numerous studies (see Cooper, 1998). A meta-analytic approach is more appropriate for the comprehensive and quantitative synthesis.

In the past, two meta-analyses were conducted to aggregate findings of note-taking studies. Ryan (1982) reviewed 19 studies examining the effect of note-taking compared with reading on text-learning and indicated that the mean effect size was 0.22. Henk and Stahl (1985) conducted a meta-analysis of 14 studies comparing recall performances of note-taking groups with those of listening groups in lecture settings. They found the mean effect size to be 0.34. Taken together, the encoding effect of note-taking seems to be positive but modest. Previous meta-analyses, however, did not go beyond estimating the advantage of note-taking over no note-taking. These studies were silent on whether the variance in effect sizes across studies was homogeneous, and if not so, what moderator variables influenced the encoding effect. It would be too hasty to infer the value of the note-taking process from a mere estimate of the overall effect.

A variety of explanations have been put forward regarding why the encoding effect is not as high as one might expect (for reviews, see Carrier & Titus, 1979; Hartley & Davies, 1978; Kiewra, 1989; Ladas, 1980b). In the present meta-analysis, I addressed some of the most prominent explanations through the inclusion of moderator variables in coding. First, the encoding effect may be limited by the quality of note-taking procedures students employ. Proponents of the encoding hypothesis (e.g., Bretzing & Kulhavy, 1979; Di Vesta & Gray, 1972; Einstein, Morris, & Smith, 1985; Peper & Mayer, 1978, 1986) state that note-taking enhances learning by stimulating the note-takers to actively process the material and to relate it to their existing knowledge. However, this is not to suggest that the act of taking notes always activates generative processing (Kiewra, 1989). Different note-taking procedures seem to vary in terms of how much generative processing is activated. According to Bretzing and Kulhavy (1979), for example, paraphrasing and summarizing serve as catalysts to stimulate deeper semantic processing, whereas transcribing verbatim does not. It may be that the effect of note-taking process is modest because students' note-taking procedures are less generative.

Second, mechanical demands of taking notes with paper and pen may put restrictions on the encoding effect. These restrictions include the attentional demands of monitoring one's own hand-movements and the products, the labor involved, and the time required for writing. Owing to these demands, the process of taking notes may be affected by presentation variables such as speed, mode, length, and information density (Ladas, 1980a; Suritsky & Hughes, 1991). For example, students often complain of the rapid presentation of lecture information for the reason that it reduces time to take notes in a generative way (Van Meter et al., 1994). The modest encoding effect may stem from the interaction between mechanical demands of note-taking and presentation variables.

Finally, some types of learning outcome measures may be ill-suited to detect the advantage of note-taking over no note-taking. Whether a certain type of outcome measure is suitable or not depends on the quality of mental representation students construct from the presented material. For example, Peper and Mayer (1986) argue that the generative effect of note-taking is undetectable on verbatim recognition and fact recall tests because generative processing makes it difficult to discriminate externally presented information from internally generated information. Many note-taking researchers, nevertheless, have used a variety of tests as learning outcome measures without considering the quality of mental representation constructed (Lonka, Lindblom-Ylanne, & Maury, 1994). This may distort the estimate of the encoding effect.

To test these explanations, I examined the influence of five possible moderator variables on the encoding effect: intervention, schooling level, presentation mode and length, and test mode. Additionally, for explanatory purposes, I included two publication characteristics (i.e., publication year and source) in the present meta-analysis.

1.1. Quality of note-taking procedure

If students do not gain much from the note-taking process because their note-taking procedures are less generative, the encoding effect would increase as they become more skillful in taking generative notes. To test this possibility, I identified two factors that affect the quality of note-taking procedure: intervention and schooling level.

1.1.1. Intervention

According to Kiewra (1989), students often produce poor notes (e.g., verbatim transcript, omission of important ideas), suggesting that their spontaneous note-taking procedures are defective in processing of lecture or text information. To compensate for such a defect, positive interventions in the note-taking procedure would be helpful. Positive interventions include pre-training of note-taking skills or techniques, providing framework notes, and giving verbal instructions to employ an effective note-taking procedure. If the use of less generative note-taking procedures lowers the benefits of note-taking, the encoding effect would be greater for note-takers who received any positive intervention than for those who did not. Although this idea is not new (e.g., Jonassen, 1984; Kiewra, 1989; Robin, Martello, Foxx, & Archable, 1977; Spires, 1993), prior empirical findings were inconsistent and inconclusive. A meta-analysis based on numerous research findings would provide a rigorous test of whether positive interventions improve the encoding effect.

1.1.2. Schooling level

I also examined the influence of schooling level on the encoding effect, which has rarely been the object of empirical research. Several researchers (Faber, Morris, & Lieberman, 2000; Rinehart & Thomas, 1993; Simbo, 1988) argue that students at lower schooling level do not benefit much from the note-taking process because they are lacking in cognitive abilities and skills necessary for taking generative notes. Probably those abilities and skills gradually develop as one grows up and gains educational experience in schools. Indeed, there is some evidence that students at higher schooling level take notes more skillfully than students at lower schooling level (Brown & Smiley, 1978; Hidi & Klaiman, 1983), though this does not imply that the former note-taking procedures are complete. Thus, if the encoding effect depends on the quality of note-taking procedure, the advantage of note-taking over no note-taking would be greater for higher schooling level than for lower schooling level.

1.2. Mechanical demands of note-taking

The second possibility is that mechanical demands of note-taking may restrict the encoding effect. Presentation mode and length were used to test this possibility. Despite the theoretical and practical importance (Carrier & Titus, 1979; Ladas, 1980a; Weener, 1974), these factors have as yet been scarcely dealt with in empirical research.

1.2.1. Presentation mode

Taking notes with paper and pen often requires the note-takers to visually monitor their hand-movements and the products. For this reason, what type of mode information is presented through may affect the note-taking process (Greene, 1934; Howe & Singer, 1975; Kiewra, DuBois, Christian, Kim, & Lindberg, 1989). In school settings, learning material is usually presented to the students through one among audio, visual, and text modes, or their combinations. Above all, note-taking researchers have used audio (e.g., taped lecture), audio-visual (e.g., live lecture, videotaped lecture), or text modes for the presentation of the material in their experiments.

Of these presentation modes, audio-visual and text modes require note-takers to direct their visual attention to the material, and therefore the two modes may be incompatible with the act of taking notes (Ash & Carlton, 1953; Ladas, 1980a). In contrast to audio-visual presentation which is in the control of a lecturer, however, text presentation allows note-takers to voluntarily alternate between reading and taking notes. Thus, reading a text may not seriously interfere with the note-taking process (Kiewra et al., 1989; Lahtinen et al., 1997). If this were the case, the encoding effect would be weaker when the material is presented through the audio-visual mode than when presented through the audio or text modes.

1.2.2. Presentation length

Listening or reading with note-taking needs physical and mental effort more than without note-taking. Consequently, the fatigue effect on the note-takers may increase as the length of a lecture or the time given to read a text gets longer (Locke, 1977; Scerbo, Warm, Dember, & Grasha, 1992). Owing to the negative influence of fatigue on the cognitive processing (e.g., attention decrement), longer presentation of learning material may reduce the advantage of note-taking over no note-taking compared with shorter presentation.

1.3. Type of learning outcome measure

The encoding effect may be dependent on the modes of test used as a learning outcome measure. For the reason mentioned earlier, if the primary function of note-taking is to promote generative processing, the advantage of note-taking over no note-taking would be detectable on higher-order performance (e.g., inference, comprehension, transfer, and synthesis) tests rather than on fact retention tests. In addition, the note-taking effects may differ depending on types of fact retention test: free-recall, cued-recall, and recognition tests. In his narrative review, for example, Weener (1974) suggests that the benefits of note-taking are greater on recall tests than on recognition tests. However, there is no attempt to comprehensively and quantitatively synthesize prior findings about the influence of test mode. To examine this factor, I estimated the encoding effect separately for free-recall, cued-recall, recognition, and higher-order performance tests.

1.4. Publication characteristics

Publication year and source were included in the present meta-analysis. Focus, theoretical background, and practical value of note-taking studies have changed in the course of years. Such historical changes may bias research findings through the methodological aspects of research or the selection criteria for publication. In addition, the publication source was examined to identify publication bias (Begg, 1994; Lipsey & Wilson, 1993). Owing to publication bias, journal articles may be apt to show greater effect sizes than dissertations or ERIC reports.

2. Method

2.1. Literature search

Two sources to locate studies for inclusion were used in the present meta-analysis. First, I searched the major computerized databases, including PsycINFO, ERIC, and Dissertation Abstracts International. Key words used to identify relevant articles were *note-taking*, *summarization*, *study technique*, *learning strategy*, and variants on these terms. Second, I used references listed in review articles, book chapters, and empirical articles.

To be included in the meta-analysis, a study had to (a) focus on note-taking with paper and pen in academic situations (e.g., lecture, text-learning, and entrance examination), (b) include comparison between outcomes of note-taking groups that were allowed or instructed to take notes but not to review the notes afterward and those of no note-taking groups that were not allowed to take notes, (c) measure how much the participants learned newly presented material, and (d) report outcome data amenable to the calculation of an effect size. Although the synthesis of correlational findings based on the association between content of notes and learning performance might provide another estimate of the encoding effect (e.g., Ganske, 1981; Hartley & Davies, 1978; Kiewra, 1985), the correlational studies were not included in the present meta-analysis. This was because the number of correlational studies reporting the independent effect of note-taking process is sparse, and because the correlation does not indicate a direct casual relationship between note-taking and learning performance.

2.2. Dependent learning measures

To identify independent comparisons, I employed the independent sample as the unit of analysis (Cooper, 1998). In the primary analyses, effect sizes of all dependent learning measures within an independent note-taking group compared with no note-taking group were averaged into one estimate. Additionally, I conducted the secondary analyses in which dependent learning measures were classified into four constructs of test mode: (a) free-recall test (e.g., writing down everything one can remember from the lecture, writing down all the passage), (b) cued-recall test (e.g., fill-in-the-blanks test measuring factual knowledge), (c) recognition test (e.g., multiple-choice test of factual information, sentence-verification test of explicit statements presented), and (d) higher-order performance test (e.g., application test, sentence-verification test of implicit statements presented, and transfer test). Effect sizes of multiple dependent learning measures within an independent sample were separately averaged into one estimate for each construct.

An independent coder was asked to code about 20% of the studies included in the present meta-analysis. The percentage of agreement with the author was 84.2%. Disagreements, most of which were due to the coder's failing to locate crucial information that was present in the original studies, were resolved by discussion.

2.3. Moderator variables

For the primary analyses, each study was coded according to a set of continuous or discrete variables: (a) intervention (behavioral positive, verbal positive, neutral, or verbal negative); (b) schooling level (lower [e.g., primary school students, junior high school students, and senior high school students] or higher [e.g., college students, graduate students, and applicants to university medical faculty]); (c) presentation mode (audio-visual [e.g., film, videotaped lecture, and live lecture], audio [e.g., taped lecture, oral presentation of passage], or text); (d) presentation length; (e) publication year; and (f) publication source (journal article or dissertation/ERIC report). Intervention was coded as *behavioral positive* when pre-training sessions of note-taking skills or techniques were prepared for the treatment groups, or when framework notes (e.g., matrix notes, skeleton notes, and outline notes) were provided; as *verbal positive* when only verbal instructions (e.g., instructions to outline, instructions to summarize, and instructions to take brief notes) that would facilitate generative processing were given; as *neutral* when any special intervention in the note-taking procedure was not conducted (e.g., instructions to take notes as usual, mere instructions to take notes), or when equivocal verbal instructions were given; and as *verbal negative* when only verbal instructions (e.g., instructions to take verbatim notes, instructions to search some letters while reading a text) that would deteriorate generative processing were given. With regard to presentation length, the length of a lecture or the time given to read a text was coded.

An independent coder was asked to code about 20% of the studies included in the present meta-analysis. The percentage of agreement with the author was 95.4%. Disagreements were resolved by discussion.

2.4. Meta-analytic procedures

I used Cohen's d as the measure of effect size. When only a test static such as t and F values was reported, each effect size was calculated using a variety of formulas provided by Lipsey and Wilson (2001). For seven studies that did not report the number of participants per group, the total number of participants was divided by the number of groups, and then effect sizes were estimated from the means, standard deviations, and averaged number of participants assigned to each group.

Each effect size from individual independent samples was weighted by its sample size and combined into a composite mean effect size. Effect sizes which fell outside three interquartile rage below the first quartile or above the third quartile were detected as outliers. The values of those outliers were winsorized to the values of the next maximum effect size in the respective distributions. A 95% confidence interval (CI) was also computed for each mean weighted effect size. Homogeneity test statistic (Q) was used to test whether a set of effect sizes was homogeneous. When the value of Q was significant at $p < .05$, a between-groups statistic (Q_B) and a within-groups statistic (Q_W) were conducted to analyze the relation between possible moderator variables and variation in effect sizes. Finally, I conducted a multiple regression analysis in which each effect size was weighted by its inverse variance weight to estimate the independent effects of possible moderator variables.

3. Results

3.1. Primary analyses

A sample of 57 studies met the inclusion criteria and was included in the primary analyses. In total, these studies yielded 131 independent samples and 306 effect sizes of note-taking groups versus no note-taking groups comparison. Information (i.e., sample characteristics and effect sizes) concerning studies included in the primary analyses can be obtained from the author on request. The mean weighted effect size across 131 independent samples was 0.22 (the mean unweighted $ES = .29$), with 95% confidence interval ranging from 0.17 to 0.27. According to Cohen's (1988) criteria, this effect size was in the range of small to medium magnitude. The number of additional studies with a null result needed to negate this finding (fail-safe N) was 15,134, indicating high tolerance for future null results. Of the 131 effect size values, 6 were large negative (−1.08 to −0.86, including two winsorized effect sizes); 5 were medium to large negative (−0.72 to −0.51); 13 were small to medium negative (−0.44 to −0.22); 19 were small negative (−0.19 to −0.00); 16 were small positive (0.01–0.16); 25 were small to medium positive (0.20–0.47); 20 were medium to large positive (0.51–0.76); and 27 were large positive (0.81–1.40, including nine winsorized effect sizes).

Test of homogeneity indicated that the effect sizes were not significantly homogeneous, $Q(130) = 426.84$, $p < .001$. To account for the variation in effect sizes, moderator analyses were conducted separately for six possible moderator variables: intervention, schooling level, presentation mode and length, and publication year and source. Results of the moderator analyses are shown in Table 1. For five of the six analyses, between-groups statistics were significant, indicating that the subgroups were heterogeneous. It should be noted, however, that within-groups statistics were

Table 1
Moderator analyses for note-taking versus no note-taking comparison studies

Variable and category	Weighted					
	Q_B	*k*	Mean Es	95% CI		Q_W
Intervention	35.16***					
Behavioral positive		19	.53	.38	−.67	40.21***
Verbal positive		37	.14	.06	−.22	98.70***
Neutral		66	.26	.19	−.32	204.35***
Verbal negative		9	−.11	−.29	−.07	48.43***
Schooling level	29.82***					
Higher		97	.14	.09	−.20	305.34***
Lower		30	.43	.34	−.52	89.80***
Mixed		4	.08	−.14	−.30	1.88
Presentation mode	39.02***					
Audio-visual		31	−.02	−.11	−.07	63.71***
Text		73	.27	.20	−.33	216.37***
Audio		27	.43	.31	−.55	107.73***
Presentation length	1.22					
<30 min		64	.20	.14	−.27	275.09***
30–59 min		28	.26	.14	−.39	43.99*
>59 min		13	.20	.07	−.33	23.80*
Unspecified		23	.26	.14	−.37	82.75***
Publication year	46.94***					
Before 1970		12	−.04	−.13	−.05	50.78***
1970s		37	.36	.27	−.46	133.23***
1980s		46	.33	.24	−.42	120.87***
1990–2003		36	.27	.18	−.37	75.02***
Publication source	24.84***					
Journal article		103	.16	.11	−.21	337.15***
Dissertation/ERIC report		28	.47	.36	−.58	64.85***

Note. Q_B, between-groups homogeneity statistic. *k*, number of independent samples. CI, confidence interval. Q_W, within-groups homogeneity statistic.
* $p < .05$.
*** $p < .001$.

significant for most of the subgroups in these analyses. Therefore, the contrasts between subgroups must be interpreted with caution.

3.1.1. Intervention

Interventions in the note-taking procedure were classified into four: behavioral positive, verbal positive, neutral, and verbal negative. Four samples of participants from Bretzing and Kulhavy (1981), who received verbal instructions to take notes from the perspective of a teacher or expert, were coded as neutral. On the surface, the perspective-taking instructions appear to promote generative processing. According to Bretzing and Kulhavy, however, whether the influence of the perspective-taking on the note-taking procedure is positive or negative depends on the nature of material to be learned. Therefore, I considered the perspective-taking instructions equivocal. There was a significant effect for intervention on the variation in effect sizes, $Q_B(3) = 35.16$, $p < .001$. The mean weighted effect sizes were 0.53 ($k = 19$) for behavioral positive, 0.14 ($k = 37$) for verbal positive, and 0.26 ($k = 66$) for neutral, all of which were significantly greater than zero. On the other hand, the mean weighted effect size for verbal negative was −0.11 ($k = 9$), which was not significantly different from zero. This indicates that the outcomes of note-taking groups that received verbal negative interventions did not significantly differ from those of no note-taking groups.

3.1.2. Schooling level

Participants' schooling levels were divided into lower and higher levels. Lower schooling level ranged between 6th and12th graders. For higher schooling level, the majority of samples were undergraduate students. Four independent samples from Stordahl and Christensen (1956) were coded as *mixed* because the participants' years of education were in the range of 6–16 years, probably consisting of both schooling levels. There was a significant effect for schooling level on the variation in effect sizes, $Q_B(2) = 29.82$, $p < .001$. The mean weighted effect sizes were 0.43 ($k = 30$) for lower schooling level and 0.14 ($k = 97$) for higher schooling level, both of which were significantly greater than zero. This result indicates that the advantages of note-taking over no note-taking were greater for the students at lower schooling level than for the students at higher schooling level.

3.1.3. Presentation mode

The significant homogeneity test statistic for this variable, $Q_B(2) = 39.02$, $p < .001$, revealed that presentation mode of learning material was significantly associated with the variation in effect sizes. The mean weighted effect sizes were 0.43 ($k = 27$) for audio, 0.27 ($k = 73$) for text, and −0.02 ($k = 31$) for audio-visual, respectively. The former two effect sizes were significantly different from zero, whereas the latter effect size was not.

3.1.4. Presentation length

The median length of presentation was 20.5 min, with a range of 5–93. Studies were classified into three ranges of presentation length: <30 min, 30–59 min, and >59 min. Ten studies did not provide information on the length of presentation. Samples from these studies were coded as *unspecified.* Test of homogeneity revealed that presentation length was not significantly associated with the variation in effect sizes, $Q_B(3) = 1.22$. The mean weighted effect sizes were 0.20 ($k = 67$) for <30 min, 0.26 ($k = 28$) for 30–59 min, and 0.20 ($k = 13$) for >59 min, respectively.

3.1.5. Publication year

Studies were classified into four ranges of publication year: before 1970, 1970s, 1980s, and 1990–2003. The publication years of studies coded as before 1970 ranged between 1934 and 1956. There was a significant effect for publication year on the variation in effect sizes, $Q_B(3) = 46.94$, $p < .001$. For studies published before 1970, the mean weighted effect size was −0.04 ($k = 12$), which was not significantly different from zero. The mean weighted effect sizes for studies published in 1970s or after were statistically significant and in the range of small to medium: 0.36 ($k = 37$) for 1970s, 0.33 ($k = 46$) for 1980s, and 0.27 ($k = 36$) for 1990–2003.

3.1.6. Publication source

Forty-three of 57 studies included in the present meta-analysis were journal articles, 11 dissertations, and 3 ERIC reports. The significant homogeneity test statistic for this variable, $Q_B(1) = 24.84$, $p < .001$, revealed that publication source was significantly associated with the variation in effect sizes. The mean weighted effect sizes were 0.16 ($k = 103$) for journal article and 0.47 ($k = 28$) for dissertation/ERIC report, both of which were significantly greater than zero.

3.1.7. Multiple regression analysis

To test whether the interrelations among possible moderator variables confounded the results of each moderator analysis, a multiple regression analysis was conducted. The predictor variables were behavioral positive versus other (behavioral positive = 1, other = 0), positive or neutral versus verbal negative (positive or neutral = 1, verbal negative = 0), schooling level (higher = 1, lower = 0), audio or text versus audio-visual (audio or text = 1, audio-visual = 0), audio versus text or audio-visual (audio = 1, text or audio-visual = 0), presentation length (in minutes), publication year (in chronological year), and publication source (journal article = 1, dissertation/ERIC report = 0). Overall, these variables accounted for 38.4% (adjusted $R^2 = 0.33$) of the variation in effect sizes, $F(8, 95) = 7.40$, $p < .001$. Variables that had significant independent effects at $p < .05$ were positive or neutral versus negative ($b = .94$, $SE = .15$), schooling level ($b = -.20$, $SE = .08$), audio or text versus audio-visual ($b = .33$, $SE = .08$), audio versus text or audio-visual ($b = .26$, $SE = .01$), publication year ($b = .009$, $SE = .001$), and publication source ($b = -.15$, $SE = .08$). Behavioral positive versus other and presentation length did not significantly contribute to the regression ($b = .05$, $SE = .09$, and $b = -.003$, $SE = .002$, respectively).

For behavioral positive versus other, the result of moderator analysis using Q statistic was inconsistent with that of multiple regression analysis. That is, the mean weighted effect size for behavioral positive intervention was much greater than those for other interventions, whereas behavioral positive versus other was not a significant predictor. This may be due to the fact that the interrelations between intervention and other moderator variables confounded the former result. To identify possible interrelations, I conducted χ^2 tests for the discrete variables (i.e., schooling level, presentation mode, and publication source) and t tests for the continuous variables (i.e., presentation length, publication year). χ^2 of behavioral positive versus other crossed with schooling level and publication source were significant: $\chi^2(1, N = 127) = 8.62$, $p < .01$, and $\chi^2(1, N = 131) = 13.18$, $p < .001$, respectively. Also, there was a significant difference in publication year between behavioral positive and other, $t(129) = 2.74$, $p < .01$. These results suggest that the greater effect sizes for behavioral positive cannot be explained away as the independent effect of intervention.

Table 2
Effect sizes and confidence intervals for test mode

Test mode	Weighted				
	k	Mean *ES*	95% CI		Q_W
Free-recall	38	.55	.44	–.65	252.86***
Cued-recall	32	.47	.36	–.58	68.20***
Recognition	47	.18	.10	–.26	106.39***
Higher-order performance	54	.26	.18	–.33	133.96***
Unclassified	20	.02	−.07	–.10	70.25***

*** $p < .001$.

3.2. Secondary analyses

To examine whether the encoding effects differ depending on test mode, the meta-analyses for note-taking groups versus no note-taking groups comparison studies were conducted separately for free-recall, cued-recall, recognition, and higher-order performance tests. Of studies included in the secondary analyses, a total of 17 note-taking group versus no note-taking group comparison studies for free-recall test yielded 37 independent samples and 62 effect sizes; 12 studies for cued-recall test yielded 32 independent samples and 55 effect sizes; 23 studies for recognition test yielded 47 independent samples and 62 effect sizes; and 21 studies for higher-order performance test yielded 54 independent samples and 117 effect sizes. Samples from 20 studies were coded as *unclassified* because they did not provide sufficient information on test mode, or because they reported the data on learning outcomes without discriminating between test modes. Information (i.e., sample characteristics and effect sizes) concerning studies included in the secondary analyses can be obtained from the author on request.

The results for each test mode are shown in Table 2. When learning outcomes were measured by free-recall and cued-recall tests, the mean weighted effect sizes were in the range of medium magnitude (*ES* = .55 and .47, respectively). Conversely, the mean weighted effect sizes calculated for recognition and higher-order performance tests were small or small to medium in magnitude (*ES* = .18 and .26, respectively), though significantly greater than zero.

4. Discussion

The present meta-analysis of 57 studies demonstrated that the overall effect of note-taking compared with no note-taking was positive but modest (mean weighted *ES* = .22, and mean unweighted *ES* = .29). This result is consistent with the findings of prior meta-analyses conducted by Ryan (1982) and Henk and Stahl (1985). They obtained the mean unweighted effect sizes of 0.22 and 0.34, respectively. Homogeneity test static, however, revealed that the variation in effect sizes was significantly larger than that expected from sampling error, indicating the systematic influence of moderator variables on the encoding effect. To explore what limits the overall encoding effect, the present meta-analysis included seven possible moderator variables: intervention, schooling level, presentation mode and length, test mode, and publication year and source.

First, intervention and schooling level were examined to test a possibility that the encoding effect is limited because the students' spontaneous note-taking procedures are less generative. This assumes that both intervention and schooling level affect the quality of note-taking procedures they employ. With respect to intervention, the results of moderator analyses indicated that interventions were significantly associated with the variation in effect sizes. Note-taking groups that received verbal negative interventions did not significantly differ from no note-taking groups ($ES = -.11$), whereas note-taking groups that received behavioral positive, verbal positive, or neutral interventions significantly outperformed no note-taking groups ($ES = .53$, .14, and .26, respectively). It is not surprising that verbal negative interventions were inferior to other interventions in the enhancement of learning performance. Verbal negative interventions forced the participants to employ the poor note-taking procedures, such as copying and searching letters, whereby to hinder generative processing (Bretzing & Kulhavy, 1979). Consequently, the note-takers who received verbal negative interventions may have gained no profit from the process of taking notes.

The aforementioned finding is consistent with the idea that the use of less generative note-taking procedure lowers the encoding effect. On the other hand, however, I found no evidence that positive interventions heightened the benefits of note-taking. There was little difference in effect sizes between verbal positive and neutral interventions. Also, the multiple regression analysis indicated that the variable contrasting behavioral positive with other did not significantly contribute to the regression. Indeed, the effect sizes for behavioral positive were greater than those for neutral, but this was probably due to the confounding effect between intervention and other moderator variables. Unless positive interventions had no impact on the quality of the students' note-taking procedures, these findings throw doubt on the possibility that the defects in their spontaneous note-taking procedures reduce the encoding effect.

Another factor related to the quality of note-taking procedure was schooling level. Contrary to the prediction, the students at lower schooling level benefited from note-taking significantly greater than did the students at higher schooling level ($ES = .43$ and .14, respectively). It seems unlikely, however, that the former students surpassed the latter in the quality of note-taking procedure. Rather, the difference in effect sizes may be accounted for by the cognitive characteristics of students at each schooling level. In general, cognitive abilities and skills necessary for learning from a lecture or text develop with age and experience in schooling (e.g., Schneider & Bjorklund, 1998; Siegler, 1986). It may be that the students at higher schooling level had already acquired those abilities and skills to a certain degree. For this reason, they may have been able to perform successfully without the help of note-taking. On the other hand, for the students at lower schooling level who were less skilled learners, note-taking may have substantially aided them to process the lecture or text information.[1] Of course, this is not to suggest that it is useless for highly skilled learners to take notes while being in class. Remember that note-taking produces a written record for later review. Possibly they gain much from the activity of reviewing notes rather than from the process of taking notes.

Overall, the present findings did not support the idea that the limited effect of note-taking results from the incompleteness of students' spontaneous note-taking procedures. It should be noted, however, that few studies included in the present meta-analysis provided data adequate to estimate the quality of note-taking procedures the participants actually employed. Therefore, I could not confirm whether intervention and schooling level may have actually brought about changes in their note-taking procedures. Contrary to the present assumption, it is possible that neither positive interventions nor rise in schooling level improved the quality of note-taking procedures. For example, the imposed generative note-taking procedures might not easily become established in the students without an intensive training (Thornton, Bohlmeyer, Dickson, & Kulhavy, 1990). Of the studies coded as positive interventions in the present meta-analysis, nevertheless, only a few prepared the training sessions for the treatment groups. Also, generative note-taking usually requires time and effort much more than less generative note-taking (Kiewra et al., 1991). Accordingly, students might be unwilling to take generative notes even if they can. Future studies should measure the quality of note-taking procedure as well as the learning outcome.

Second, through moderator analyses for presentation mode and length, I tested the possibility that mechanical demands of taking notes with paper and pen may limit the encoding effect. With respect to presentation mode, as predicted, the mean weighted effect sizes were greater for audio ($ES = .43$) than for text ($ES = .27$) and audio-visual ($ES = -.02$), and greater for text than for audio-visual. It is important to note that prior research has indicated that traditional presentation mode (e.g., video, oral, and written) by and in itself makes little difference in remembering and understanding of lecture material (e.g., Greene, 1934; Howe & Singer, 1975; see also Bligh, 2000). This suggests that the difference in effect sizes across presentation modes cannot be accounted for by different effects of each presentation mode on no note-taking groups. Given this consideration, it is reasonable to infer from the present results that visual attention to one's own hand-movements and the products while taking notes interferes with processing of information presented through visual mode, especially when the note-takers cannot voluntarily control the pace of presentation.

On the other hand, the result of moderator analysis using Q statistic indicated that presentation length was not a significant moderator variable. Similar result was also gained from the multiple regression analysis: presentation length did not significantly contribute

[1] Samples of participants coded as lower schooling level ranged between 6th and 12th graders, suggesting that their cognitive abilities and skills varied widely. If note-taking compensates for lack of cognitive abilities and skills, younger students at lower schooling level would benefit from note-taking greater than older students. To test this possibility, I estimated the encoding effects separately for 6–8th and 9–12th graders. The result showed that the mean weighted effect sizes were greater for 6–8th graders ($ES = .68$, $k = 18$, 95% CI = 0.54, 0.82) than for 9–12th graders ($ES = .23$, $k = 11$, 95% CI = 0.11, 0.36).

to the regression. These findings suggest that longer presentation of material does not cause fatigue effect to the note-taker greater than does shorter presentation. The labor involved in taking notes seems to have little impact on the lecture or text processing.

Third, I examined whether the benefits of note-taking vary according to test modes. To secure independence at the sample level, the encoding effects were estimated separately for test modes. The results showed a distinct difference in effect sizes across test modes. That is, the mean weighted effect sizes for free-recall and cued-recall tests were substantial ($ES = .55$ and .47, respectively), whereas those for recognition and higher-order performance tests were modest ($ES = .18$ and .26, respectively). The difference in effect sizes between recall and recognition tests can be understood in the light of the encoding specificity principle (Tulving, 1983). According to this principle, successful recollection depends on how much the information available at the time of retrieval coincides with the episodic information encoded at the time of learning. Retention tests differ in how many the test condition provides retrieval cues akin to the episodic information encoded. The recognition test condition provides lots of retrieval cues, whereas the recall test condition demands learners to create retrieval cues more autonomously. Thus, the greater effect sizes for recall test compared with those for recognition test suggest that the function of note-taking is to help note-takers create retrieval cues in the memory representation, whereby they can recollect information efficiently with the minimum retrieval cues the recall test condition provides (Santa, Abrams, & Santa, 1979). In other words, for recognition test, non-note-takers may have been able to perform little less than note-takers because the test condition provided abundant retrieval cues.

The result for higher-order performance test suggests that note-taking serves little generative function. It should be added, however, that the variation in effect sizes for each test was not homogeneous, indicating the influences of moderator variables. Indeed, the generative hypothesis of Mayer and his associates (Peper & Mayer, 1978, 1986; Shrager & Mayer, 1989) predicts the thee-way interaction between treatment (i.e., note-taking versus no note-taking), cognitive characteristics of the learner, and test mode. According to Shrager and Mayer (1989), for example, note-taking may facilitate the higher-order test performance only for less skilled learners, but not for highly skilled learners; and furthermore, the same pattern of interaction between note-taking and cognitive skills may not be observed on a retention test. Unfortunately, there were no consistent data on cognitive characteristics of the participants to examine this possibility in the present meta-analysis. Instead, I conducted interactive comparison between treatment, schooling level, and test mode. As mentioned above, it seems likely that the schooling level of students is roughly correspondent to the level of their cognitive abilities and skills. The results of analyses are shown in Table 3. The students at lower schooling level benefited from note-taking significantly greater than the students at higher schooling level, regardless of test mode. Although this pattern of results was not entirely congruous with the prediction of the generative hypothesis, the mean weighted effect size for the students at lower schooling level on higher-order performance tests was 0.68 ($k = 16$). This raises the possibility that for less skilled learners, note-taking may substantially facilitate deeper understanding as well as retention performance.

Table 3
Interactive comparison between note-taking, schooling level, and test mode

Variable and category	Weighted					
	Q_B	k	Mean ES	95% CI		Q_W
Free-recall	10.39^{***}					
Higher schooling level		34	.47	.35	−.59	238.56^{***}
Lower schooling level		4	.93	.68	−1.19	3.91^{*}
Cued-recall	5.79^{**}					
Higher schooling level		24	.37	.23	−.50	53.59^{***}
Lower schooling level		8	.65	.47	−.84	8.82
Recognition	14.71^{***}					
Higher schooling level		28	.02	−.10	−.13	23.16
Lower schooling level		19	.33	.22	−.44	68.52^{***}
Higher-order performance	43.70^{***}					
Higher schooling level		34	.10	.01	−.20	57.41^{**}
Lower schooling level		16	.68	.54	−.83	30.98^{**}
Mixed		4	.08	−.14	−.30	1.88

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

Finally, publication year and source were exploratory examined as possible moderator variables. The result of moderator analysis for publication year indicated that the mean weighted effect sizes were smaller for studies appearing before 1970 ($ES = -0.04$) than for those appearing in 1970s or after ($ES = .27$–$.36$). It is difficult to infer the reason for the difference in effect sizes because publication year covaries with other variables, such as focus of research, methodology, educational system, and teaching methods in schools. For example, the present finding might reflect a substantial shift in the focus of research that occurred early in the 1970s. Note-taking studies appearing before 1970 were generally product-oriented. The focus was to determine whether a note-taking procedure is superior to other study techniques or which one is the most effective among a variety of note-taking procedures. In contrast, studies appearing in 1970s or after were more process-oriented, reflecting the cognitive revolution in educational psychology (Mayer, 1992; Walberg & Haertel, 1992). Based on cognitive theory, many researchers took interest in the factors related to the process of taking notes. Such a change might improve the quality of research, thereby helping detect the encoding effect.

Publication source was also a significant moderator variable. Contrary to publication bias assumed commonly (Begg, 1994; Lipsey & Wilson, 1993), however, dissertations and ERIC reports had greater benefits of note-taking than did journal articles ($ES = .47$ and $.16$, respectively). It might reflect the difference in quality of research methodology rather than publication bias. To test this possibility, microanalytic approach (Ladas, 1980b) that critically examines methodological characteristics of individual study is needed.

Although the present findings contribute to some advances in understanding the encoding function of note-taking, two limitations should be noted. One is that the present meta-analysis could address only a few of the explanations regarding why the encoding effect is limited. Other potential explanations still remain to be examined. In real academic situations, for example, students usually take notes in expectation of being able to review their notes afterward (Van Meter et al., 1994). Reviewing notes would provide another opportunity of generative processing (Kiewra et al., 1991). Accordingly, they might devote little effort to study the material while taking notes, thereby reducing the encoding effect. I could not test this explanation, however, because few studies reported whether the participants expected to review their notes. Future researchers should incorporate this variable into their studies. Another limitation is that a number of studies included in the present meta-analysis used laboratory settings that were quite different from the ordinary learning situations (Hartley & Davis, 1980; Locke, 1977; Nye, Crooks, Powley, & Tripp, 1984). For example, the participants attending experiments often knew that their performance on post-tests was irrelevant to the final grades in the regular course, and therefore they may have not been highly motivated to take notes in a generative way (Slotte & Lonka, 2001). This might threaten the external validity of the present meta-analytic findings. Thus, extensive generalization must be cautioned.

References

Ash, P., & Carlton, J. (1953). The value of note-taking during film learning. *British Journal of Educational Psychology, 23*, 121–125.

Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–409). New York: Russell Sage Foundation.

Bligh, D. A. (2000). *What's the use of lectures?* San Francisco, CA: Jossey-Bass.

Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology, 4*, 145–153.

Bretzing, B. H., & Kulhavy, R. W. (1981). Note-taking and passage style. *Journal of Educational Psychology, 73*, 242–250.

Brown, A. L., & Smiley, S. S. (1978). The development of strategies for studying texts. *Child Development, 49*, 1076–1088.

Carrier, C. A., & Newell, K. J. (1984). What dental hygiene students think about notetaking in lectures. *Canadian Dental Hygienist, 18*, 13–17.

Carrier, C. A., & Titus, A. (1979). The effects of notetaking: A review of studies. *Contemporary Educational Psychology, 4*, 299–314.

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

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews*. Thousand Oaks, CA: Sage.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63*, 8–14.

Einstein, G. O., Morris, J., & Smith, S. (1985). Note-taking, individual differences, and memory for lecture information. *Journal of Educational Psychology, 77*, 522–532.

Faber, J. E., Morris, J. D., & Lieberman, M. G. (2000). The effect of note taking on ninth grade students' comprehension. *Reading Psychology, 21*, 257–270.

Ganske, L. (1981). Note-taking: A significant and integral part of learning environments. *Educational Communication and Technology Journal, 29*, 155–175.

Greene, E. B. (1934). Certain aspects of lecture, reading and guided reading. *School and Society, 39*, 619–624.

Hartley, J. (1983). Note-taking research: Resetting the scoreboard. *Bulletin of British Psychological Society, 36*, 13–14.

Hartley, J., & Davies, I. K. (1978). Note-taking: A critical review. *Programmed Learning and Educational Technology, 15*, 207–224.

Henk, W. A., & Stahl, N. A. (1985). A meta-analysis of the effect of notetaking on learning from lecture. Paper presented at the 34th Annual Meeting of the National Reading Conference, St. Petersburg Beach, Florida (ERIC Document Reproduction Service No. ED258533).

Hidi, S., & Klaiman, R. (1983). Notetaking by experts and novices: An attempt to identify teachable strategies. *Curriculum Inquiry, 13*, 377–395.

Howe, M. J. A., & Singer, L. (1975). Presentation variables and student activities in meaningful learning. *British Journal of Educational Psychology, 45*, 52–61.

Jonassen, D. H. (1984). Effects of generative text processing strategies on recall and retention. *Human Learning, 3*, 241–256.

Kiewra, K. A. (1985). Investigating notetaking and review: A depth of processing alternative. *Educational Psychologist, 20*, 23–32.

Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review, 1*, 147–172.

Kiewra, K. A., DuBois, N. F., Christian, D., Kim, S., & Lindberg, N. (1989). A more equitable account of the note-taking functions in learning from lecture and from text. *Instructional Science, 18*, 217–232.

Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology, 83*, 240–245.

Ladas, H. S. (1980a). Note taking on lectures: An information-processing approach. *Educational Psychologist, 15*, 44–53.

Ladas, H. S. (1980b). Summarizing research: A case study. *Review of Educational Research, 50*, 597–624.

Lahtinen, V., Lonka, K., & Lindblom-Ylanne, S. (1997). Spontaneous study strategies and the quality of knowledge construction. *British Journal of Educational Psychology, 67*, 13–24.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*, 1181–1209.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Locke, E. A. (1977). An empirical study of lecture note-taking among college students. *Journal of Educational Research, 77*, 93–99.

Lonka, K., Lindblom-Ylanne, S., & Maury, S. (1994). The effect of study strategies on learning from text. *Learning and Instruction, 4*, 253–271.

Mayer, R. E. (1992). Cognition and instruction: Their historical meeting within educational psychology. *Journal of Educational Psychology, 84*, 405–412.

Nye, P. A., Crooks, T. J., Powley, M., & Tripp, G. (1984). Student note-taking related to university examination performance. *Higher Education, 13*, 85–97.

Peper, R. J., & Mayer, R. E. (1978). Note taking as a generative activity. *Journal of Educational Psychology, 70*, 514–522.

Peper, R. J., & Mayer, R. E. (1986). Generative effects of note-taking during science lectures. *Journal of Educational Psychology, 78*, 34–38.

Rinehart, S. D., & Thomas, K. F. (1993). Summarization ability and text recall by novice studiers. *Reading Research and Instruction, 32*, 24–32.

Robin, A. L., Martello, J., Foxx, R. M., & Archable, C. (1977). Teaching note-taking skills to underachieving college students. *Journal of Educational Research, 71*, 81–85.

Ryan, M. P. (2001). Conceptual models of lecture learning: Guiding metaphors and model-appropriate notetaking practices. *Reading Psychology, 22*, 289–312.

Ryan, M. T. (1982). Effects of paraphrase notetaking on prose learning. (Doctoral dissertation, University of Connecticut, 1981. *Dissertation Abstracts International, 42* (9-A), 3921 (UMI No. 820356).

Santa, C. M., Abrams, L., & Santa, J. L. (1979). Effects of notetaking and studying on the retention of prose. *Journal of Reading Behavior, 11*, 247–260.

Scerbo, M. W., Warm, J. S., Dember, W. N., & Grasha, A. F. (1992). The role of time and cuing in a college lecture. *Contemporary Educational Psychology, 17*, 312–328.

Schneider, W., & Bjorklund, D. F. (1998). Memory. In W. Damon (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 467–521). New York: Wiley.
Shrager, L., & Mayer, R. E. (1989). Note-taking fosters generative learning strategies in novices. *Journal of Educational Psychology, 81*, 263–264.
Siegler, R. S. (1986). *Children's thinking*. New Jersey: Prentice-Hall.
Simbo, F. K. (1988). The effects of notetaking approaches on student achievement in secondary school geography. *Journal of Educational Research, 81*, 377–381.
Slotte, V., & Lonka, K. (2001). Note-taking and essay writing. In P. Tynjälä, L. Mason, & K. Lonka (Eds.), *Writing as a learning tool: Integrating theory and practice* (pp. 131–143). Dordrecht, Netherlands: Kluwer.
Spires, H. A. (1993). Learning from a lecture: Effects of comprehension monitoring. *Reading Research and Instruction, 32*, 19–30.
Stordahl, K. E., & Christensen, C. M. (1956). The effect of study techniques on comprehension and retention. *Journal of Educational Research, 49*, 561–570.
Suritsky, S. K., & Hughes, C. A. (1991). Benefits of note taking: Implications for secondary and postsecondary students with learning disabilities. *Learning Disabilities Quarterly, 14*, 7–18.
Thornton, N. E., Bohlmeyer, E. M., Dickson, L. A., & Kulhavy, R. W. (1990). Spontaneous and imposed study tactics in learning prose. *Journal of Experimental Education, 58*, 111–124.
Tulving, E. (1983). *Elements of episodic memory*. New York: Academic Press.
Van Meter, P., Yokoi, L., & Pressley, M. (1994). College students' theory of note-taking derived from their perceptions of note-taking. *Journal of Educational Psychology, 86*, 323–338.
Walberg, H. J., & Haertel, G. D. (1992). Educational psychology's first century. *Journal of Educational Psychology, 84*, 6–19.
Weener, P. (1974). Notetaking and student verbalization as instructional learning activities. *Instructional Science, 3*, 51–74.

Appendix: Studies Included in the Meta-Analysis

Anderson, W. A. (1977). The relationship between student note-taking during an instructional television lesson and student learning. (Doctoral dissertation, Michigan State University, 1976. *Dissertation Abstracts International, 37* (9-A), 5540. (UMI No. 7705766).
Annis, L. F. (1985). Student-generated paragraph summaries and the information-processing of prose learning. *Journal of Experimental Education, 47*, 92–96.
Annis, L. F., & Davis, J. K. (1975). The effect of encoding and an external memory device on note taking. *Journal of Experimental Education, 44*, 44–46.
Ash, P., & Carlton, J. (1953). The value of note-taking during film learning. *British Journal of Educational Psychology, 23*, 121–125.
Bentley, D. A., & Blount, P. (1980). Testing the spaced lecture for the college classroom. Paper presented at The Georgia Psychological Association Annual Meeting, Macon, Georgia. (ERIC Document Reproduction Service No. ED188559).
Benton, S. L., Kiewra, K. A., Whitfill, J. M., & Dennison, R. (1993). Encoding and external-storage effects on writing processes. *Journal of Educational Psychology, 85*, 267–280.
Berliner, D. C. (1972). *The generalization of aptitude-treatment interactions across subject matter*. Paper presented at the Meeting of The American Educational Research Association, Chicago, Illinois. (ERIC Document Reproduction Service No. ED062642).
Bigelow, M. L. (1993). The effects of information processing strategies and cognitive style on achievement of selected educational objectives. (Doctoral dissertation, Pennsylvania State University, 1992. *Dissertation Abstracts International, 53* (7-A), 2337 (UMI No. 9236787).
Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology, 4*, 145–153.
Bretzing, B. H., & Kulhavy, R. W. (1981). Note-taking and passage style. *Journal of Educational Psychology, 73*, 242–250.

Bretzing, B. H., Kulhavy, R. W., & Caterino, L. C. (1987). Notetaking by junior high students. *Journal of Educational Research, 80*, 359–362.

Davis, J. K., & Annis, L. (1978). The effect of study techniques, study preferences, and familiarity on later recall. *Journal of Experimental Education, 47*, 92–96.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63*, 8–14.

Doctorow, M., Wittrock, M. C., & Marks, C. (1978). Generative processes in reading comprehension. *Journal of Educational Psychology, 70*, 109–118.

Dunkel, P., Mishra, S., & Berliner, D. (1989). Effects of note taking, memory, and language proficiency on lecture learning for native and nonnative speakers of English. *TESOL Quarterly, 23*, 543–549.

Dyer, J. W., Riley, J., & Yekovich, F. R. (1979). An analysis of three study skills: Notetaking, summarizing, and rereading. *Journal of Educational Research, 73*, 3–7.

Einstein, G. O., Morris, J., & Smith, S. (1985). Note-taking, individual differences, and memory for lecture information. *Journal of Educational Psychology, 77*, 522–532.

Fisher, J. L., & Harris, M. (1973). Effect of note taking and review on recall. *Journal of Educational Psychology, 65*, 321–325.

Foos, P. W. (1995). The effect of variation in text summarization opportunities on test performance. *Journal of Experimental Education, 63*, 89–95.

Freedman, C. R. (1979). The effects of behavioral objectives, note taking and note review on recall and criterion test performance: A test of the encoding versus external storage hypothesis. (Doctoral dissertation, Indiana University, 1978. *Dissertation Abstracts International, 39* (9-A), 5408. (UMI No. 7905999).

Freyberg, P. S. (1956). The effectiveness of notetaking. *Education for Teaching, 39*, 17–24.

Greene, E. B. (1934). Certain aspects of lecture, reading and guided reading. *School and Society, 39*, 619–624.

Haenggi, D., & Perfetti, C. A. (1992). Individual differences in reprocessing of text. *Journal of Educational Psychology, 84*, 182–192.

Hertzbach, A. (1981). The effects of prior knowledge, note taking, and note review on free recall and comprehension performance from an oral presentation. (Doctoral dissertation, Indiana University, 1980. *Dissertation Abstracts International, 41* (12-A), 5540. (UMI No. 8112489).

Hooper, S., Sales, G., & Rysavy, S. D. M. (1994). Generating summaries and analogies alone and in pairs. *Contemporary Educational Psychology, 19*, 53–62.

Howe, M. J. A. (1970). Notetaking strategy, review and long-term retention of verbal information. *Journal of Educational Research, 64*, 285.

Howe, M. J. A., & Singer, L. (1975). Presentation variables and student activities in meaningful learning. *British Journal of Educational Psychology, 45*, 52–61.

Jonassen, D. H. (1984). Effects of generative text processing strategies on recall and retention. *Human Learning, 3*, 241–256.

Kiewra, K. A. (1985a). Students' note-taking behaviors and the efficacy of providing the instructor's notes for review. *Contemporary Educational Psychology, 10*, 378–386.

Kiewra, K. A. (1985b). Learning from a lecture: An investigation of notetaking, review and attendance at a lecture. *Human Learning, 4*, 73–77.

Kiewra, K. A. (1985c). Examination of the encoding and external storage functions of notetaking for factual and higher-order performance. *College Student Journal, 19*, 394–397.

Knight, L. J., & McKelvie, S. J. (1986). Effects of attendance, note-taking, and review on memory for a lecture: Encoding vs. external storage functions of notes. *Canadian Journal of Behavioral Science, 18*, 52–61.

Kulhavy, R. W., Dyer, J. W., & Silver, L. (1975). The effects of notetaking and test expectancy on the learning of text material. *Journal of Educational Research, 68*, 363–365.

Kung, C. C. (1991). The effects of different notetaking strategies on learning achievement by tenth-grade earth science students. (Doctoral dissertation, Arizona State University, 1990. *Dissertation Abstracts International, 52* (3-A), 792-793 (UMI No. 9116757).

Lahtinen, V., Lonka, K., & Lindblom-Ylanne, S. (1997). Spontaneous study strategies and the quality of knowledge construction. *British Journal of Educational Psychology, 67*, 13–24.

Maqsud, M. (1980). Effects of personal lectures notes and teacher-notes on recall of university students. *British Journal of Educational Psychology, 50*, 289–294.
Mathews, C. O. (1938). Comparison of methods of study for immediate and delayed recall. *Journal of Educational Psychology, 29*, 101–106.
McIntyre, S. R. (1991). The relationship between information processing and notetaking effectiveness as demonstrated by university undergraduates. Unpublished doctoral dissertation, Indiana University.
Peper, R. J., & Mayer, R. E. (1978). Note taking as a generative activity. *Journal of Educational Psychology, 70*, 514–522.
Peverly, S. T., Brobst, K. E., Graham, M., & Shaw, R. (2003). College adults are not good at self-regulation: A study on the relationship of self-regulation, note taking, and test taking. *Journal of Educational Psychology, 95*, 335–346.
Reese, J. L. (1984). The effect of notetaking, lecture structure, and ability level of performance. (Doctoral dissertation, University of Iowa, 1984. *Dissertation Abstracts International, 45* (7-A), 2046 (UMI No. 8423592).
Ryan, M. T. (1982). Effects of paraphrase notetaking on prose learning. (Doctoral dissertation, University of Connecticut, 1981. *Dissertation Abstracts International, 42* (9-A), 3921 (UMI No. 820356).
Shikanai, N. (1980). Youyaku, nouto oyobi andarainkatsudou ga sanbunkyouzai no gakushuu ni oyobosu kouka. [The effects of summarizing, note-taking and underlining activities on the learning of prose material]. *The Reading Science, 24*, 1–9.
Shrager, L., & Mayer, R. E. (1989). Note-taking fosters generative learning strategies in novices. *Journal of Educational Psychology, 81*, 263–264.
Slater, W. H., Graves, M. F., & Piche, G. L. (1985). Effects of structural organizers on ninth-grade students' comprehension and recall of four patterns of expository text. *Reading Research Quarterly, 20*, 189–202.
Stordahl, K. E., & Christensen, C. M. (1956). The effect of study techniques on comprehension and retention. *Journal of Educational Research, 49*, 561–570.
Thiede, K. W., & Anderson, M. C. M. (2003). Summarizing can improve metacomprehension accuracy. *Contemporary Educational Psychology, 28*, 129–160.
Thornton, N. E., Bohlmeyer, E. M., Dickson, L. A., & Kulhavy, R. W. (1990). Spontaneous and imposed study tactics in learning prose. *Journal of Experimental Education, 58*, 111–124.
Wellington, K. R. (1981). A study of the encoding vs. memory storage function in college notetaking. (Doctoral dissertation, University of Minnesota, 1980. *Dissertation Abstracts International, 42* (1-A), 158. (UMI No. 8109526).
Wittrock, M. C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal, 27*, 489–502.
Yu, H. L. Berliner, D. C. (1981). *Encoding and retrieval of information from lecture*. Paper presented at the 65th Annual Meeting of American Educational Research Association, Los Angeles, CA. (ERIC Document Reproduction Service No. ED206738).