Computers & Education 123 (2018) 13–25

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Impact of slide-based lectures on undergraduate students' learning: Mixed effects of accessibility to slides, differences in note-taking, and memory term

Hyeyoun Kim[1]

Dongguk University–Seoul, South Korea

A R T I C L E I N F O

Keywords:
Authoring tools and methods
Improving classroom teaching
Interactive learning environments
Pedagogical issues
Teaching/learning strategies

A B S T R A C T

This paper addresses the effects of access to slide copies during lectures using PowerPoint® for undergraduate students on their learning outcomes depending on the quantity of notes they take and immediate vs. delayed testing. Seventy-one students repeatedly participated in the following six lecture conditions: accessibility to slides (full, partial, and no slide copy) × memory term (immediate and delayed test). Thus, the present study adopted a 3 × 2 within-subjects design with two note-taking covariates (the quantity of words and markers in notes). A mixed-effects model and counterbalancing method were applied to control idiosyncrasies and order effects caused by repeated measurement. The results revealed that accessibility to slide copies and students' note-taking predicted their learning outcomes. The effects of no slide copy were significant in both short- and long-term memory conditions as compared to those of access to full and partial copies. Access to full and partial slide copies did not have significantly different results. However, according to the interaction results between accessibility and memory term, the long-term encoding effect was assumed for the partial slide copy condition. Regarding note-taking variables, students' performance was considerably impacted by the number of markers but none of the number of words. The findings suggest educational implications for the way slides are prepared and provided and the way students take notes during slide-based lectures from a perspective of writing-to-learn.

1. Introduction

How instructors deliver content is a key question in education research (Stephenson, Brown, & Griffin, 2008). Many instructors have created electronic slides using computer presentation software such as PowerPoint® or Prezi®, which are widespread tools for delivering lectures to students (James, Burke, & Hutchins, 2006; Mantei, 2000; Peper & Mayer, 1986; Williams & Eggert, 2002). Numerous studies have mentioned the effects of electronic slides and yet have yielded contradictory results—skeptical or negative rather than positive—regarding their effectiveness on students' learning (Bartsch & Cobern, 2003; Buchko, A., Buchko, K., & Meyer, 2012; Daniels, 1999; Doumont, 2005; Lowry, 1999; Mantei, 2000; Savoy, Proctor, & Salvendy, 2009; Susskind, 2005, 2008; Szabo & Hastings, 2000; Tufte, 2003). Some studies have sought to solve this contradiction by testing more concrete and elaborate variables such as types of slide and/or access to them. Worthington and Levasseur (2015), for example, used the term instructor-provided slides (IP slides) to refer to focus more on providing slides or their copies. However, the results have also been inconsistent with regard to effectiveness: some have indicated positive effects when offering partial IP slides (Cornelius & Owen-DeSchryver, 2008; Katayama &

E-mail address: hyeyoun.kim@dongguk.edu.
[1] Full address: Dept. of Korean Language Education, Dongguk University, 30, Pildong-ro 1-gil, Jung-gu, 04620 Seoul, South Korea.

https://doi.org/10.1016/j.compedu.2018.04.004
Received 19 March 2017; Received in revised form 27 January 2018; Accepted 11 April 2018
Available online 14 April 2018
0360-1315/

Crooks, 2003) or access to slide copies (Chen & Lin, 2008), whereas others demonstrated negative effects for these factors (Babb & Ross, 2009; Grabe & Christopherson, 2005; Le, Joordens, Chrysostomou, & Grinnell, 2010; Worthington & Levasseur, 2015).

Moreover, the effects of electronic IP slides in relation to differences in students' note-taking capabilities have rarely been investigated. Few researchers have controlled for the variability of students' note-taking or have mentioned how they have handled this in their research design. Nevertheless, the issues of students' additional and voluntary note-taking may become a significant mediating variable since the variability of students' note-taking can be sufficiently high to impact the effectiveness of electronic IP slides. Many students in actual lecture rooms are free to take down their own notes in addition to what is already provided in IP slides/notes or what their instructors wanted them to write. Furthermore, types of electronic IP slides may influence any effect of students' note-taking (e.g., encoding effect). Thus, the present study seeks to apply experimental control to detect the effects of electronic IP slides on students' learning outcomes in consideration of their autonomous note-taking behaviors during lectures.

2. Theoretical backgrounds

2.1. Electronic IP slides

Using electronic IP slides has become ubiquitous in university lecture rooms and has been discussed by many researchers, particularly with regard to its positive effects (Daniels, 1999; Lowry, 1999; Mantei, 2000; Stephenson et al., 2008). Since electronic IP slides are presupposed as being presented using computers (Levasseur & Sawyer, 2006; Worthington & Levasseur, 2015), they have been compared to traditional lectures using chalk and boards for their effects and circumstances (Bartsch & Cobern, 2003; Debevec, Shih, & Kashyap, 2006; Savoy et al., 2009; Susskind, 2005, 2008; Szabo & Hastings, 2000). However, some researchers have demonstrated that there is no evidence that students who attend lectures that involve electronic IP slides outperform those who attend traditional lectures (Buchko, Buchko, & Meyer, 2012; Susskind, 2005, 2008; Szabo & Hastings, 2000). Other researchers have even insisted that there are negative effects associated with the in-class use of slides for students' learning (Bartsch & Cobern, 2003; Savoy et al., 2009). Despite researchers showing conflicting opinions regarding the use of electronic IP slides, surveys from students and lecturers tended to show positive responses for it considering their preference, paying attention, and their self-awareness regarding learning efficiency (Levasseur & Sawyer, 2006). Therefore, if electronic IP slides are preferred by students and lecturers in spite of their uncertain effects, we must investigate the specific conditions that affect their effectiveness to improve the circumstances in which slides are used.

As Doumont (2005) pointed out, the effectiveness of electronic IP slides is closely related to how instructors produce and provide them to students. In this regard, partial notes, which are a frequently mentioned method of enhancing effectiveness when giving lectures, may offer some examples of a better way to produce electronic IP slides. Some researchers insisted on the use of *partial notes*, which include intentionally missing several important parts of IP notes, by drawing a contrast and comparison with the full notes that will include all critical points (Cornelius & Owen-DeSchryver, 2008, p. 7). Many studies present significant positive effects for the use of partial notes (Annis, 1981; Austin, Lee, & Carr, 2004; Austin, Lee, Thibeault, Carr, & Bailey, 2002; Barbetta & Skaruppa, 1995; Cornelius & Owen-DeSchryver, 2008; Katayama & Crooks, 2003; Katayama & Robinson, 2000; Kiewra, 1985b; Konrad, Joseph, & Eveleigh, 2009; Williams, Weil, & Porter, 2012). However, as Larwin and Larwin (2013) pointed out, the effects can be moderate or invalid according to the research subject or situation. The effects of partial notes might differ according to school's level, whether students take notes, and whether they have copies of lecture notes (Larwin & Larwin, 2013). In this regard, as Worthington and Levasseur (2015, p. 16) assumed, it would be better to distinguish partial notes that emphasize instructors' handouts from *partial slides* that emphasize computer presentation. Interestingly, studies of partial slides showed both positive and negative results (Williams et al., 2012; Worthington & Levasseur, 2015), whereas studies of partial notes tended to show more positive results. In other words, it is not certain that students in lectures that employ a partial slides condition outperform those in the full slides condition. Therefore, more specific factors must be developed to investigate why partial slides have conflicting effects on students' learning.

Some researchers have emphasized the function of partial notes or slides as a lecturing tool by suggesting that they have a *cuing effect*. The cuing effect refers to any effect caused by a method "used to increase the salience of some feature of a stimulus" (Scerbo, Warm, Dember, & Grasha, 1992, p. 315). Titsworth and Kiewra (2004) divided the types of cuing techniques into four categories as determined by the interaction between two criteria, the forms (written vs. spoken) and functions (selection vs. organization) of lecture cues. For example, lecturers may use a spoken selection cue such as, "This is very important." Alternatively, they may use a written selection cue by putting additional information on the blackboard, or they may use a written organization cue such as partial notes with intentionally absent information. Since cuing techniques are one of the main lecturing methods that lecturers depend on, some researchers have also emphasized the use of partial notes or slides because they act as a cuing role to encourage students' note-taking (Scerbo et al., 1992; Titsworth & Kiewra, 2004).

How IP slides are produced and how they should be provided are two of the main factors that affect the effectiveness of IP slides. The in-class use of IP slides must be distinguished from access to them before and during lectures because the effects may "simply result from students having copies" (Levasseur & Sawyer, 2006, p. 112). Therefore, some studies have centered on the impact of students' access to IP slides. However, they have tended to also show conflicting results regarding providing slides on web prior to lectures. Chen and Lin (2008) reported a small positive effect, while Murphy and Cross (2002) suggested that students who were offered slide copies prior to lectures underperformed on exams compared to those who did not use the copies. In addition, some research studies have presented results that support negative views on providing instructor's slides or furnishing students with slide copies prior to lectures (Babb & Ross, 2009; Grabe & Christopherson, 2005; Grabe, 2005; Le et al., 2010; Worthington & Levasseur, 2015). However, it should be noted that such studies that focused on access to IP slides did not apply the same conditions. Some researchers requested that students download the slides prior to the lecture (Babb & Ross, 2009; Chen & Lin, 2008; Grabe & Christopherson, 2005; Grabe, 2005; Le et al., 2010), while others distinguished uploading slides to the web from bringing copies to the class (Murphy & Cross, 2002; Worthington & Levasseur, 2015). Since uploading itself cannot guarantee that students review them beforehand, it is necessary to confirm whether prior learning happened, for more elaborate experimental conditions. Furthermore, since uploading cannot also guarantee that students bring slide copies to the class, whether they bring copies should also be verified to ultimately understand the detailed conditions of access to IP slides. Therefore, it is necessary to reanalyze and reify some controversial conditions in which IP slides have been made or provided and examine why such contradictory results have been attained.

2.2. Note-taking

Students' note-taking during lectures has long been regarded as one of the most effective learning strategies (Carrier & Titus, 1979; Castelló & Monereo, 2005; Kobayashi, 2005, 2006; Lahtinen, Lonka, & Lindblom-Ylänne, 1997; Van Meter, Yokoi, & Pressley, 1994). Studies on note-taking generally emphasize its two representative functions: *encoding* and *external storage* (Di Vesta & Gray, 1972; Hartley, 1983; Kiewra, 1985a, 1989; Knight & McKelvie, 1986; Kobayashi, 2005, 2006; Peper & Mayer, 1978). The encoding effect means that note-taking behavior has a certain effect on students' learning regardless of whether they review what they have written, whereas the external storage effect focuses on the advantage of reviewing notes (Kiewra, 1985a, 1989; Kobayashi, 2005).

However, the advantage of the encoding effect has not been actively studied and it is regarded as less important than the external storage effect (Kobayashi, 2005, 2006). Some researchers have ascribed why the encoding effect has not been remarkable to the quality or method of note-taking (Bretzing & Kulhavy, 1979; Kiewra, 1989) or to the problem of how learning outcomes are measured (Lahtinen et al., 1997). In contrast, other researchers have argued that the encoding effect should be more heavily scrutinized since it relates to the core cognitive processes of note-taking and seems to be a prerequisite for the external storage function in most actual note-taking circumstances (Bui & Myerson, 2014; Kiewra, Dubois, & Christensen, 1989). Note-taking strategies, one of the representative aspects of the encoding effect, have been examined by some researchers. Jansen, Lakens, and IJsselsteijn (2017) insisted that existing note-taking strategies can be divided into two representative types: organizing and transcribing. Bretzing and Kulhavy (1979) presented four different types of students' note-taking and revealed that verbatim and letter search types were worse than not taking notes, whereas summary and paraphrasing were better. Similarly, Piolat, Olive, and Kellogg (2005) argued that note-taking should be differentiated from transcribing a text or lecture. As the analysis of Kobayashi (2005) revealed, note-taking strategies can yield a certain moderating influence on the encoding effect.

Aside from these strategies, some positive results to support the encoding effect were presented by researchers who were able to demonstrate that the quality and quantity of notes are predictors of test performance. Some researchers have tried to measure quality or efficiency of notes from a qualitative perspective (Brobst, 1996; Kirby & Pedwell, 1991; Peverly, Garner, & Vekaria, 2014; Williams & Eggert, 2002), and others have done so from a quantitative perspective (Norton & Hartley, 1986; Oakhill & Davies, 1991), or by combining both perspectives (Baker & Lombardi, 1985; Peverly et al., 2007). However, the reliability of the tools that measure quality of notes has been controversial (Williams & Eggert, 2002, p. 176). Moreover, the concepts of quality or efficiency of notes are not applicable to students' note-taking when they have free access to IP slides and are equipped with copies in class since they do not need to write down the essential points that IP slides already state. On the other hand, the quantity of notes, which has not drawn as much interest as the quality, has proven to be a predictor of the encoding effect (Aiken, Thomas, & Shennum, 1975; Bui, Myerson, & Hale, 2013). Bui et al. (2013) suggested that the quantity of notes taken during the lecture relates to better maintenance of the contents in memory. While the quality of notes can be controversial due to the ambiguous criteria of quality itself, determining the quantity of notes is likely to be considered an obvious and reliable factor. However, most studies adopting quantity-related variables only measured the amount of words in students' notes. Various aspects or ways to measure the quantity of notes must be developed to learn more about encoding processes.

Cognitive features during note-taking have more to do with the encoding effect than the external storage. As Bui and Myerson (2014) have pointed out, the most important cognitive features of note-taking have a close relation to working memory abilities. The term *working memory* can be used to refer to the manipulation of information as well as its temporary storage (Baddeley, 2012, p. 4). In addition to temporarily holding the memory of the lecture content, the working memory engages in manipulating and mediating the multitasking processes during note-taking (Bui & Myerson, 2014, p. 13). Piolat et al. (2005) conducted experiments to measure the cognitive efforts of note-taking, transcribing, and creating a text and thereby verified that note-taking places as much demand on the working memory as creating a text. Jansen et al. (2017) stated that some studies have revealed the mediating role of cognitive overload on the relationship between note-taking and memory. Considering that note-taking aims to help the understanding and recall of the lecture, we must learn more about these cognitive demands and efficient ways to reduce them.

Additionally, determining the best way to provide IP slides closely relates to the encoding effect because one of the reasons for examining the types or accessibility of IP slides is to enhance the effectiveness of students' note-taking behaviors (Katayama & Crooks, 2003; Katayama & Robinson, 2000; Kiewra & Benton, 1995). For example, partial slides are assumed to promote students' note-taking by making them curious about intentionally omitted keywords in such slides and letting them fill in those blanks. Therefore, the encoding effect may relate to the cuing effect (mentioned in 2.1) since Scerbo et al. (1992) and Titsworth and Kiewra (2004), who regarded partial slides as having the cuing effect, also argued that partial slides as lecture cues would encourage students to take notes. However, since there are only a few studies on this topic, there is a lack of experimental attempts to make a research design that links the IP slides that lecturers make and the notes that students take. A few of note-taking researchers have shown a special interest in how to produce and provide IP slides, for example, by using matrices or graphic organizers to promote students' note-taking (Kiewra & Benton, 1995). However, how students organized the notes on their own was rarely investigated. Nevertheless, many students often take supplementary or detailed notes in addition to the essential points that instructors want them to write down using the frames that they have prepared. Exceptionally, Robinson et al. (2006) presented a link between note-taking styles and partial slides. They conducted several experiments to discern the impact of different types of IP slides on students' conceptual learning and thereby verified students' better performance in partial graphic organizers compared to the other types. However, due to a lack of relevant research, more studies are needed to investigate the relationship between IP slides and students' note-taking.

2.3. Learning outcome measures

Most of the aforementioned studies on the impact of IP slides and/or note-taking during lectures have presented the results by conducting lecture content recall tests (Austin et al., 2004; Austin et al., 2002; Bretzing & Kulhavy, 1979; Buchko et al., 2012; Di Vesta & Gray, 1972, 1973; Fisher & Harris, 1973; Kiewra & Benton, 1995; Kiewra, 1985a, 1985b; Kirby & Pedwell, 1991; Peverly et al, 2007, 2014; Savoy et al., 2009). However, some researchers have argued that different types of tests would yield different results (Kiewra, 1985a; Kiewra et al., 1991; Oakhill & Davies, 1991). Thus, some recent studies have tended to refer to the total scores aggregated from more than two test types (Fisher & Harris, 1973; Kiewra, 1985a; Kobayashi, 2005) or final grades calculated from several types of test (Cornelius & Owen-DeSchryver, 2008; Szabo & Hastings, 2000; Worthington & Levasseur, 2015), rather than using only one type of test.

Apart from the types, the timing of test has also been configured as variables in research. Some studies have distinguished between immediate and delayed tests to compare their results (Coluccia, Gamboz, & Brandimonte, 2011; Di Vesta & Gray, 1973; Fisher & Harris, 1973, 1974; Katayama & Crooks, 2003; Kiewra, 1985a; Loaiza, McCabe, Youngblood, Rose, & Myerson, 2011; Williams et al., 2012), whereas other studies have conducted either immediate tests (Kiewra et al., 1991; Oakhill & Davies, 1991; Peverly et al, 2007, 2014) or delayed tests (Cornelius & Owen-DeSchryver, 2008; Kiewra & Benton, 1995; Savoy et al., 2009; Szabo & Hastings, 2000). It is probable that conducting and comparing immediate and delayed tests would yield more elaborate results for note-taking during slide-based lectures by explaining differences and/or commonalities between them (see Katayama & Crooks, 2003); meaning that if both the timing of the test and note-taking were adopted as variables, it will be likely to confirm how the encoding effect is affected by time.

2.4. Purpose

The present study aims to reveal the relationship between electronic IP slide accessibility during lectures and undergraduate students' learning outcomes by considering the effects of their individual differences with regard to note-taking and memory term. Although research has seldom included all of these variables together, they can be regarded as meaningful factors because they are supposed to reflect students' actual performance in lectures when using IP slides, and contribute to discerning the progress of the encoding effect in terms of memory. Regarding the measurement of note-taking, the present study configures the quantity of notes taken (i.e., counts the elements in students' notes) as relevant variables to avoid controversial presumptions about notes' quality. The memory term is differentiated into short- and long-term by conducting both immediate and delayed tests after each lecture session. The following three research questions suggest the present study's variables and design.

RQ1. Does accessibility to IP slide copies predict undergraduate students' learning outcomes?

RQ2. Does the quantity of undergraduate students' notes predict their learning outcomes?

RQ3. Does the memory term affect the impact of access to IP slide copies on undergraduate students' learning outcomes?

3. Methods

3.1. Participants

The participants were 76 undergraduate students (76% female) from a university in South Korea in their third (64%) and fourth (36%) grade from the Department of Education. A within-subjects design was employed, which means that they participated in 12 lecture units (15 min per unit) and tests divided into two sessions (see Table 1). Any student who missed any session was not included in the analysis; thus, the target of analysis eventually became 71 students and 852 cases. They were recruited from the same regular course; however, the experiment—including two sessions of lectures and tests—was conducted regardless of their GPA. Students could voluntarily decide to participate in the experiment and were not informed of the experimental conditions in detail (e.g. purpose, variables, sampling, etc.). The research design was approved in advance by the Institutional Review Board (SNUIRB no. 1411/002–005) and written consent was collected from all participants regarding its regulation.

The proper sample size was calculated by "longpower," which is one of the R packages with which power and sample size analysis can be conducted for multilevel and longitudinal data (Diggle, Heagerty, Liang, & Zeger, 2013; Donohue & Edland, 2016; Liu & Liang, 1997). As there are no correct tools with which to calculate the exact sample size for mixed-effects modeling, the current tool was only to calculate an approximate value that employed just one random structure in the original model. Since a minimum of 10 participants—the confidence interval (CI) indicated 7–16—were needed for the mixed-effects model structure (significant level = 0.05; expected power = 0.8), the total number of participants in the present study (n = 71) may be used for the analysis.

Table 1
Lecture and test information (n = 71).

Lecture information		Access to slides	Test information	
			Memory term	Max score
Session 1				
Unit 1	Argument and claim	None	Short-term (Immediate test)	7
Unit 2	Causes of a problem	Full		8
Unit 3	Thoughtful claims	Partial		8
Unit 4	Types of claim	Partial		7
Unit 5	Principles to find solutions	Full		7
Unit 6	Concept of values claims	None		8
Session 2				
Unit 7	Values in expression	None	Long-term(Delayed test)	7
Unit 8	Value-laden words	Full		8
Unit 9	Abstract concepts	Partial		8
Unit 10	Traps in using metaphor	Partial		7
Unit 11	Polarizing language	Full		7
Unit 12	Metaphor in language use	None		8

Note. Each session was 90 min long without breaks between units. Immediate tests were conducted 15 min after session 1; delayed tests were conducted one week after session 2. Test items were randomly assigned in an order, irrespective of the unit order.

3.2. Research design

A 3 × 2 within-subjects design (accessibility × memory term) was applied to the present study. The experiment was designed to confirm the effects of accessibility to IP slide copies on students' learning outcomes (RQ1) with consideration of the quantity of note-taking (RQ2) and the memory term of learning outcomes (RQ3). As Worthington and Levasseur (2015) proposed, three different scales were introduced for the "accessibility to the copies" predictor (RQ1): full slide copies (Full), no slide copy (None), and partial slide copies (Partial). In the "full slide copies" condition, copies of the full slides were distributed to students immediately prior to each lecture; in the "no copy" condition, students could not get any access to the slides and copies before and during each lecture; in the "partial slide copies" condition, students were offered copies of partial slides, meaning copies that intentionally missed several important parts of the full slides. The concept of accessibility to IP slides might be differently applied from the previous studies, in that some of them introduced the variable of accessibility to IP slides in terms of checking online accessibility or whether students downloaded the slides prior to each lecture (Babb & Ross, 2009; Chen & Lin, 2008; Grabe, 2005). However, uploading or downloading copies cannot cover the variability for some possible differences; for example, not all the students who downloaded copies may have previewed them before the lecture. The predictor was simplified and clarified, by distributing all copies immediately prior to each lecture to ensure the same condition for all students and remove their opportunity to read them in advance. Even though the treatment could contribute to enhancing the reliability and clarity, it might have low ecological validity in this respect.

Internal validity issues caused by the within-subjects design were considered (Levasseur & Sawyer, 2006) by applying two additional solutions—apart from maintaining the lecture content as similar as possible—counterbalancing by lecture units and including the random effects in the analysis model to control any possible effects from the design. Thus, orders of accessibility condition were counterbalanced across each lecture to minimize the order effect (see Table 1), and a mixed-effects model was applied for statistical treatment. Regarding the detailed ordering, ABBA counterbalancing—or ABCCBA for three conditions—(Shaughnessy, Zechmeister, & Zechmeister, 2012, pp. 233–235) was employed to avoid practice effects. In addition, six lecture units were conducted in succession in each session and items were randomly ordered during each test to avoid carryover and anticipation effects. The orders of the memory term were not counterbalanced, as the memory term was not the main predictor but rather the subsidiary predictor for examining the effects from accessibility conditions in the present study.

3.3. Procedures and materials

All lecture units in the same session were conducted consecutively without breaks on the same day and Session 2 was conducted one month after Session 1. Each session consisted of six 15-min lecture units on different topics (see Table 1) and the lectures were designed to maintain the same structure across unit and format, and were based on the same content source (Williams & Colomb, 2007). Lectures were conducted by the researcher using PowerPoint® slides. Slides were made in terms of some recommendations for an effective delivery (Bartsch & Cobern, 2003; Kosslyn, Kievit, Russell, & Shephard, 2012).

During each lecture, students were asked to do as they usually would, particularly with regard to note-taking. Their notes were written in the distributed copies (only in the full and partial copies conditions), or in supplementary blank paper (in either no copy conditions or some in full and partial conditions if desired) that was collected immediately after each session. An immediate test was conducted 15 min after Session 1 and a delayed test was conducted one week after Session 2. Participants did not have any time to review their notes or slide copies before the tests, because reviewing the notes should not act as a factor in the present study. Therefore, the materials collected from the experiment were IP slides, copies of students' private notes (both with and without IP slide copies), and test results.

Table 2
Item reliability by test.

	Accessibility condition		
	Full	None	Partial
Immediate test	0.53	0.69	0.66
Delayed test	0.66	0.71	0.59

3.4. Measures

3.4.1. Quantity of note-taking

The quantity of note-taking, which was one of the main variables in the present study, was calculated by counting the words and markers that the students had written in their notes. Since the Korean language is based on the unique word count system, we counted word spacing (called as “Eojeol” in Korean) as a unit for the word count variable. Regarding markers, any identifiable signs for structuring (e.g., numbering, bullets), emphasis or indexing (e.g., asterisks, underlining), connecting (e.g., drawing connecting lines), and summarizing or schematization (e.g., boxes and matrixes) were counted. Two independently working coders counted the number of words and markers respectively for each unit; the overall inter-coder agreement as measured by Cohen's Kappa was 0.89 for word count and 0.81 for marker count.

3.4.2. Scoring

Students' learning outcomes were measured in consideration of the various test types included, as proposed by previous research (reviewed in 2.3), in each set of test items for each unit: items per unit were composed of cued recall, recognition, and higher-order performance items respectively. In contrast to the factors of the types of items that did not act as variables but functioned by making the learning outcome measurements more equivalent and reliable, test timing was considered one of the predictors in terms of the research purpose (RQ3); both immediate and delayed tests were conducted and compared with each other. Thus, the tests were not conducted with each unit or each condition, but with each session (immediate/delayed) as a unit. The tests in each session were controlled across both experimental conditions (full/none/partial) and sessions (immediate/delayed), by taking measures that included the same types of item for each unit and assigning items to all six units in each session in an order that was irrespective of the unit order. The total scores for each condition (full/none/partial) were 15 points for both immediate and delayed tests, respectively (see Table 1). Table 2 indicates the item reliability (Cronbach's alpha) for each condition. Two raters scored test outputs and the inter-rater agreement—measured by the intraclass correlation coefficient—was 0.98.

3.5. Analysis

Since the research design of the present study includes multiple observations for the same participants and for the same lecture unit, the mixed-effects model was applied to the analysis to specify both the random and fixed effects. Researchers in areas such as psychology, linguistics, sociology, medicine, and biology have broadly used mixed-effects designs to model different types of random variation in a single analysis (Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013; Quené & van den Bergh, 2008). As Barr et al. (2013) proposed, the linear mixed-effects (LME) model was applied to the present study (that can handle both continuous and categorical data) to detect several fixed effects as core variables and thus investigate the main research questions, in consideration of random effects for controlling participants' idiosyncrasies that need to be taken into account. Furthermore, besides random effects from participants’ differences, the internal validity issue can be solved by adding a random effects structure as caused by differences between lecture units and controlling them during analysis (Linck & Cunnings, 2015). The LME models for the present study were configured using the statistical software package R 3.3.2 (R Core Team, 2016).

The LME model for the present study was built by lme4 (Bates, Maechler, Bolker, & Walker, 2015), an R software package, to calculate the regression of test scores on accessibility conditions, memory terms, and quantity of notes, with consideration for random effects due to participants' and lecture units’ idiosyncrasies. This model and its result was tested by lmerTest (Kuznetsova, Brockhoff, & Christensen, 2016), another R package specially developed to verify LME models made by lme4 in terms of backward elimination. While the model results made by lme4 produced estimates, standard errors, and t values, the model test results made by lmerTest yielded chi square (for random effects) or mean square values (for fixed effects), F values, elimination necessity determination, and the p values for the significance of each fixed or random effects. The following R codes were also used in the detailed analyses: LMERConvenienceFunctions (Tremblay & Ransijn, 2015), longpower (Diggle et al., 2013; Donohue & Edland, 2016; Liu & Liang, 1997), car (Fox & Weisberg, 2011), psych (Revelle, 2016), and irr (Gamer, Lemon, Fellows, & Singh, 2012).

4. Results

Among the four variables, accessibility to IP slide copies (Access) and the memory term (Term) were test variables in terms of the research design; the number of words (Noteword) and markers (Notemarker) in students' notes were covariates; and the score of each test to verify students’ learning outcomes (Score) was a response variable.

Table 3
Descriptive statistics of continuous variables (n = 71).

	Short-term memory effect			Long-term memory effect		
	Full	None	Partial	Full	None	Partial
Noteword						
mean	30.85	26.79	57.06	33.41	13.08	28.25
SD	29.88	18.96	42.71	40.18	14.78	33.50
skew	1.42	0.64	0.80	2.32	2.55	1.46
Notemarker						
mean	7.58	8.45	18.82	11.56	4.52	9.97
SD	7.24	5.23	13.48	9.15	4.00	9.57
skew	1.11	0.57	1.00	1.08	0.62	0.75
Score						
mean	9.04	10.54	8.68	5.81	9.58	7.47
SD	2.95	2.65	3.45	3.59	4.38	3.53
skew	−0.22	−0.31	−0.56	0.59	−0.45	−0.49

Table 3 shows descriptive information of three continuous variables in the present study: Noteword, Notemarker, and Score. Although two covariates of Noteword and Notemarker were supposed to be non-normally distributed considering large SDs and skewness (Table 3), they did not have to be transformed or processed since LME models do not need any assumption on the homogeneity of variance (Pinheiro & Bates, 2000, pp. 174–177). However, because the residuals should be approximately normally distributed according to the assumptions of the LME approach, the information of the model fit was verified: the density of model residuals, the Q–Q plot for the residuals, and the fitted values versus the standardized residuals provided positive grounds for the normal distribution of the model's residuals.

The main results were presented in Tables 4 and 5. As Barr et al. (2013) suggested, the maximal random effects structures were adopted in the final model by repeatedly conducting likelihood ratio tests using the ANOVA function (Baayen et al., 2008). Outliers were not excluded from the model as mixed-effects modeling "allows for mild *a priori* screening for outliers" (Baayen & Milin, 2010, p. 16). The model formula (Table 4) was built to verify both random and fixed effects. While random parts consisted of participants' (ID) and lecture units' (Unit) variance to control potential compounds by simply adding predictors, fixed parts were constructed to prove the main effects of the two test variables (Access and Test) and two covariates (Noteword and Notemarker). As random parts 'by Unit' corresponded to the six conditions differed by Access × Test, the random formula (Noteword + Notemarker | Unit) was aimed to control the idiosyncrasies of note-taking according to each condition. The model test was performed via backward elimination of the non-significant effects of LME models (Kuznetsova et al., 2016). The *F* and *p* values of test results were reported in Table 5. The *p* values for random effects were calculated based on the likelihood ratio test, whereas those for fixed effects were based on Satterthwaite's approximation. These *F* and *p* values were referred to as deciding the final significance for each variable rather than *t* values from the original model (Table 4) due to some possible errors caused by the quasi-linearity of LME models (Pinheiro & Bates, 2000, p. 70).

The mixed-effects analysis yielded both random and fixed effects; random effects accounted for the random variation caused by participants' and lecture units' variety, and fixed effects accounted for the main effects from targeted research questions (Tables 4 and 5). According to the results of random effects (Table 5), participants' random effects (by ID) were significant in terms of both IP slides' accessibility ($p < 0.05$) and the memory term ($p < 0.001$). However, regarding the random effects caused by unit difference, the by-

Table 4
Results from the final mixed-effects model (observation = 852).

Parameters	Fixed effects			Random effects			
				By ID		By Unit	
	Estimate	*SE*	*t*	Var.	SD	Var.	SD
Intercept	9.10	0.50	18.14	10.28	3.21	0.05	0.23
Access (Full)	−4.40	0.51	−8.55	4.54	2.13	–	–
Access (Partial)	−2.50	0.47	−5.34	1.88	1.37	–	–
Term (short)	0.61	0.49	1.25	3.71	1.93	–	–
Noteword	0.01	0.01	0.72	–	–	0.001	0.02
Notemarker	0.07	0.04	1.67	–	–	0.003	0.05
Access (Full): Term (short)	3.04	0.61	4.95	–	–	–	–
Access (Partial): Term (short)	−0.47	0.60	−0.79	–	–	–	–

Note. ID = identifiable number for each participant. Unit = lecture unit. Response variable = Score. The following factors were coded as abbreviated forms to get automatically dummy coded in R software: Access = Accessibility to IP slide copies (Full = full slide copies, None = no slide copy, Partial = partial slide copies), Term = Memory term (short = short-term effect, long = long-term effect). Model formula: Score ∼ Access ∗ Term + Noteword + Notemarker + (1 + Access + Term | ID) + (Noteword + Notemarker | Unit).

Table 5
Model test results by backward elimination (observation = 852).

	Chi. sq./ Mean sq.	*F*	Elim. No.	*p*
Random effects				
Notemarker: Unit	1.86	–	1	0.602
Noteword: Unit	3.53	–	2	0.172
Lecture: ID	17.32	–	kept	0.016*
Term: ID	22.00	–	kept	0.000***
Fixed effects				
Noteword	10.27	1.72	1	0.191
Access	221.73	37.21	kept	< 1e-07***
Term	127.29	21.36	kept	0.000***
Notemarker	148.05	24.84	kept	0.000***
Access: Term	100.03	16.79	kept	0.000***

Note. Model test condition: DDF approximation = Satterthwaite, alpha index of random effects = 0.1, alpha index of fixed effects = 0.05. Chi. sq. (Chi square) values for random effects and Mean sq. (mean square) values for fixed effects were reported. Final model formula: Score ~ Access + Term + Notemarker + (1 + Access + Term | ID) + Access: Term. *** < 0.001, ** < 0.01, * < 0.05.

unit variety in the quantity of words and markers in notes were not significant ($p > 0.1$). Thus, since the two note-taking variables did not indicate random effects between the six conditions, which differed by accessibility (none/partial/full) and memory term (immediate/delayed), the quantity of notes does not show any significant interaction caused by differences in accessibility and memory term.

The results of fixed effects (Table 4) indicated that accessibility to slide copies was significantly correlated with students' learning outcomes in cases of full copies (estimate = −4.40, $SE = 0.51$, and $t = -8.55$) and partial copies (estimate = −2.50, $SE = 0.47$, and $t = -5.34$) over the no copy condition. The model test results (Table 5) also showed that the fixed effect of accessibility to slide copies should be noted ($F = 37.21$, $p < 0.001$). Thus, accessibility to slide copies was a significant predictor for students' learning. The memory term was also a significant variable for students' performance ($F = 21.36$, $p < 0.001$) in terms of the model test results (Table 5); however, it was an obvious result, in that students should outperform in immediate tests because there was little decay over time (Katayama & Crooks, 2003). Rather, the results from the memory term (Table 5) were more meaningful when associated with another predictor; the main effect of accessibility to slide copies was qualified by interactions with the memory term ($F = 16.79$, $p < 0.001$). That is, as there were significant interactions between access to copies (full/none/partial) and test timing (immediate/delayed), the results from the memory term should be interpreted in terms of access to copies.

Regarding the two note-taking covariates, the number of words did not indicate any significant effect ($F = 1.72$, $p > 0.1$), while the number of markers was significantly correlated with students' learning outcomes ($F = 24.84$, and $p < 0.001$). That is, the number of words written in students' own notes taken during lectures could not predict their performance, whereas the number of markers in their notes could.

The present study interpreted the results of differences in least square (LS) means according to associated confidence intervals (CIs) and p values as one way of reporting effect sizes for results from the mixed-effects model (Larson-Hall & Plonsky, 2015). As mentioned above, access to slide copies, the memory term, and the interactions between them significantly correlated with students' learning outcomes. Table 6 demonstrates in more detail that access to slide copies was only significantly effective in the comparison between none and full copies (lower CI = 2.15, upper CI = 3.69, $p < 0.001$), and none and partial copies (lower CI = 1.99, upper

Table 6
Effect sizes: Differences of least squares means and their CIs.

	Differences of LS means			CIs		
	Estimate	*SE*	*t*	lower	upper	*p*
Access						
None – Full	2.9	0.39	7.57	2.15	3.69	< 2e-16***
None – Partial	2.7	0.37	7.42	1.99	3.45	< 2e-16***
Full – Partial	−0.2	0.37	−0.54	−0.94	0.54	0.593
Term						
L – S	−1.5	0.33	−4.62	−2.18	−0.87	< 2e-16***
Interaction: Access∗Term						
None/L – Full/L	4.4	0.50	8.91	3.45	5.41	< 2e-16***
None/L – Partial/L	2.6	0.46	5.75	1.72	3.51	< 2e-16***
Full/L – Partial/L	−1.8	0.46	−3.90	−2.73	−0.90	1e-04***
None/S – Full/S	1.4	0.48	2.95	0.47	2.36	0.004**
None/S – Partial/S	2.8	0.48	5.83	1.87	3.78	< 2e-16***
Full/S – Partial/S	1.4	0.51	2.77	0.41	2.41	0.006**

Note. L = Long-term effect, S = Short-term effect.

CI = 3.45, $p < 0.001$), whereas the difference between full and partial copies was unclear (lower CI = −0.94, upper CI = 0.54, $p > 0.1$). This result was better supported by explaining the interaction between access to slide copies and the memory term (Table 6). The differences in LS means between none and full copies, and between none and partial copies were consistent in both short- and long-term effect conditions, but the difference between full and partial copies assumed a contradictory result between the short- and long-term conditions; however, each was significant. While positive values appeared in the short-term condition (estimate = 1.4, lower CI = 0.41, upper CI = 2.41, and $p < 0.01$), negative values appeared in the long-term condition (estimate = −1.8, lower CI = −2.73, upper CI = −0.90, and $p < 0.001$). Thus, students in the full copies condition outperformed those in the partial copies condition during the immediate test and vice versa during the delayed test. Additionally, it was possible to compare their precision based on the width of the CI and its p value; for example, the difference between full and partial copies was more precise and thus had a stronger effect in the long-term condition (lower CI = −2.73, upper CI = −0.90, $p < 0.001$) than in the short-term condition (lower CI = 0.41, upper CI = 2.41, $p < 0.01$). That is, the differences between full/no/partial copies were more obvious in the long term than in the short term.

5. Discussion

Discussing the effects of IP slides has been studied for a long time (Levasseur & Sawyer, 2006); preceding research has shown skeptical or even negative results rather than positive ones, although surveys from students or lecturers have shown positive responses for their use or ease of access much more often. In this contradictory circumstance, it is necessary to reveal why access to electronic IP slides can be less effective and what factors contribute to these effects, rather than to conclude whether access to them is desirable. Therefore, this paper seeks appropriate factors other than accessibility to IP slides by investigating the impact of the quantity of notes that students take and the memory term on students' performance. The analysis' findings successfully explained the three research questions that were developed to verify their effects on students’ learning outcomes.

The first research question (RQ 1) dealt with the effects of providing IP slide copies, and the main results supported the assumption that access to slide copies predicted learning outcomes. Nevertheless, while the other two relationships (no copy vs. full copies/no copy vs. partial copies) were verified, the difference in effects between full and partial slide copies was unclear. After all, we might conclude that providing no slide copy would become the most effective condition for lectures in terms of both short- and long-term memory effects. This result is consistent with other research that has yielded skeptical or negative conclusions regarding access to IP slides (Babb & Ross, 2009; Grabe & Christopherson, 2005; Grabe, 2005; Le et al., 2010; Worthington & Levasseur, 2015). However, it is quite a complex situation, particularly when we refer to the survey results from the previous studies that reported students' positive responses to IP slides and easy access (Susskind, 2005, 2008; Szabo & Hastings, 2000). That is, it is possible that students want IP slide copies as a learning aid, but their effects are unclear or even detrimental. This contradiction can be explained to some degree; students in the “no slide copy” condition might become more anxious about missing important points and thus pay more attention during the lecture. However, it should be noted that the effects of the “no slide copy” condition might be temporary, as the present study was designed to investigate the effects during a relatively short period (15 min per unit). Thus, the results should be interpreted within the present study's design.

In addition, when combined with the third research question (RQ 3), some results from the first research question (RQ 1) suggested another possibility with which the aforementioned studies had not dealt: the long-term encoding effect from partial slide copies. Access to partial slides has been recognized as enabling the encoding effect (Annis, 1981; Cornelius & Owen-DeSchryver, 2008). However, the present study revealed its effect in more detail given that the effect of access to partial slides was significantly higher than that of full slides only in the delayed tests. On the contrary, access to full slides reported significantly better results in the immediate tests compared to partial slides. In other words, it is assumed that the encoding effect from access to partial slide copies would appear gradually and persist, whereas that from full slide copies would be better initially and then decrease. Moreover, when comparing the effect sizes to verify which is more effective, the effect of the partial copies in the long-term condition was stronger than the full copies in the short-term. In other words, we can expect that partial copies are more effective in the long-term condition than full copies in the short-term condition. Thus, it is possible that partial slide copies are preferable to full slide copies, not only for the stronger result, but also because the short-term effects cannot be more important than the long-term effects in an educational setting; the long-term encoding effect includes longer knowledge retention than the short-term encoding effect (Cornelius & Owen-DeSchryver, 2008). This interpretation was also advocated by Williams et al. (2012) and Katayama and Crooks (2003), in that they found positive effects for partial copies condition in delayed tests.

The findings mentioned above may be additionally interpreted in line with the relations between accessibility conditions and the quantity of notes (RQ 1 and RQ 2). In this study, the quantity of notes in the partial condition was not always greater than that of the other two conditions. However, this result is inconsistent with preceding research on the cuing effect. If partial slide copies act as organizational lecture cues as argued by Scerbo et al. (1992) and Titsworth and Kiewra (2004), students should take significantly more notes than in the other copy types. One possible answer may be the difference in the experimental condition regarding the way partial slides were constructed. There are many ways to make partial notes or slides in terms of percentage of blanks, types of frames, and other such factors. Partial slides in this study included approximately 10–15% blanks, which could be regarded as quite different from Kiewra (1985b) and Kiewra and Benton (1995). In addition, participants in the present study were instructed to take notes as they usually do in the lecture room when PowerPoint® slides were used; thus, they could sometimes take additional notes in addition to what the lecturer wanted them to fill in. For this reason, some participants who were offered full copies took more notes than those who were offered partial copies in terms of personal habits or lecture contents. Another interesting point found while interpreting the differences in quantity of notes was that participants with no copy always took the least notes both in the short- and long-term conditions. Since they had repeatedly experienced three accessibility conditions, they may have received relatively more stimulus when note-taking in the slide copy conditions than in the without copy condition. That is, they probably supposed that the contents without the slide copy would be less important than those with copies because they may believe that lecturers would only choose important points to make IP slides for. Furthermore, it may be that visible factors like electronic slides or their copies can stimulate students' note-taking behaviors. However, the difference in quantity of notes according to accessibility conditions was not a significant factor that influenced interpretation of the main results since the random effects from the quantity of words and markers did not show a significant interaction between accessibility conditions and memory term, as can be seen in Table 5 ($p > 0.1$).

The second research question (RQ 2) that investigated the effect of the quantity of note-taking, which is another differentiated point from preceding studies, also seemed significant according to the main results. However, only the number of markers was significant, unlike the number of words. This result is contradictory with some preceding studies (Aiken et al., 1975; Bui et al., 2013): they suggested that the number of words related to students' better performance. However, due to the difference in measuring response variables, it is not valid to directly compare the results with each other. While the preceding studies only measured the maintenance of the contents in memory, the measurement of the present study was closer to overall learning outcome: including recognition and higher-order performance as well as simple recall. In addition, the ways of counting words in preceding research may possibly include some markers since the researchers have not distinguished between words and markers. Considering this discrepancy, some interpretations may be inferred from the result. First, while the number of words almost purely related to the quantity of notes, the number of markers related more to how much students organized notes and how they considered the relationships between various content. The advantages from the explicit organization of IP slides or notes that lecturers offered were investigated by some studies (Apperson, Laws, & Scepansky, 2008, 2006; Farkas, 2005; Katayama & Crooks, 2003; Katayama & Robinson, 2000; Robinson et al., 2006). However, the merits of students' own attempts to organize their notes during slide-based lectures were seldom researched. Though not in the slide-based lectures, some researchers interested in the cognitive features during note-taking demonstrated that students' endeavor to make their notes more organized related to their better performance (Bretzing & Kulhavy, 1979; Piolat et al., 2005). However, they were focused on written organization such as summary or paraphrasing; markers have not been used as a predictor in the earlier studies. Nevertheless, unlike summary or paraphrasing, markers enable to quantitatively measure the organizational feature in notes. Furthermore, markers closely relate to the slide-based circumstances since students should use the contents already given in the lecturer's slide copies when they take notes: they can easily connect their own notes with the given notes using markers. Considering that markers include signs of structuring, emphasizing, connecting, and summarizing (see 3.4.1), the result that students' use of markers significantly related to their performance has implications for note-taking instructions and methods of constructing IP slides. As in one suggestion for note-taking instruction, students' additional endeavors to make their notes structured and connected between content must have greater emphasis, aside from what should be written down. This proposal can also be supported by Hidi and Klaiman (1983), who explored the differences in note-taking between experts and novices; experts' notes were organized and contained metacognitive choices, unlike novices' notes. Considering that note-taking is also regarded as a way of writing-to-learn in practice (Newell, 1984), note-taking instruction that emphasize structure and organization can be affiliated with other writing instruction, such as writing summaries or reports. The other educational implication is linked to how partial slides should be constructed. If students' own behavior of writing markers was effective, we can construct partial slides such that intentional misses are used for some markers or organizers as well as some keywords or concepts. We can also attempt more active methods of constructing IP slides to promote students' voluntary use of markers, such as quasi-cloze-test slides.

Another interpretation with regard to the results from the quantity of notes (RQ 2) involves the ineffectiveness of the quantity of words in students' notes. Interestingly, Norton and Hartley (1986), and Knight and McKelvie (1986) demonstrated contradictory results for the effects of students' own notes on their learning: significantly important source for an exam vs. less effective than not taking notes at all. This suggests the possibility that note-taking does not always guarantee positive effects. Instead, students' note-taking must become more efficient since note-taking is regarded as a complex multitasking activity that makes significant demands on the working memory (Bui & Myerson, 2014; Piolat et al., 2005). Thus, there is the possibility that too many words in notes may interfere with the encoding effect, unlike markers, as note-taking processes require additional cognitive effort aside from understanding lecture content and will thus need to become more precise and compact (Piolat et al., 2005).

6. Limitations and further research

Some limitations should be taken into consideration when interpreting the results. First, the present study lacks ecological validity due to its experimental design and how variables were controlled. Although the design was derived from an attempt to maximize reliability and validity, it was not ecologically appropriate that students participated in only two 90-min sessions, which seem a relatively short period, and that they were not allowed to read or study their notes. Thus, more ecologically approved research should be conducted to investigate students' actual learning circumstances.

Second, the present study was conducted with a within-subjects experimental design. Accordingly, additional issues should be considered with respect to the internal validity caused by the within-subjects design (Levasseur & Sawyer, 2006). For this reason, the present study attempted to control these issues as much as possible by adopting mixed-effects modeling and counterbalancing lecture content. For example, ABBA counterbalancing and the lecture/test schedule were helpful for avoiding practice, carryover, and anticipation effects, which repeated measurement can cause. As the mixed-effects model that the current study employed included random structures to control unit differences, the other order effects that were not controlled by the design could partly explain the results. Nevertheless, it should be noted that the current study could not apply the ideal design for repeated measurement, since the study was based on educational settings that have limited ability to repeat similar conditions as much as might be needed. Thus, treatments might not guarantee the avoidance of all the possible additional effects of ordering. Thus, the study's results should be interpreted in consideration of this sort of limitation.

Finally, regarding the effects of note-taking, the present study aims to focus on the encoding effect rather than the external storage effect. Research that advocates the external storage function has emphasized the effect of reviewing notes (Hartley, 1983; Kiewra, 1985b; Knight & McKelvie, 1986; Makany, Kemp, & Dror, 2009) and has insisted that it has stronger effects than encoding (Fisher & Harris, 1974; Kobayashi, 2005). Considering the research design that combines note-taking with the slide copies condition, the encoding effect was investigated as a priority in the present study. Nevertheless, further research about the external storage function of note-taking during slide-based lectures is also required to obtain a complete picture with which to enhance students' learning efficiency.

7. Conclusion

The provision of IP slides and students' note-taking are assumed closely related. However, previous research contains insufficient attempts to deal with both of these factors. The present study provides some insights into the effects of access to slide copies when students are allowed to take notes in whatever manner they prefer, corresponding to the actual lecture's circumstances. Some results partly supported previous studies. Students without any copy outperformed those with full or partial copies, which is in common with several preceding research results. However, other results suggested conflicting and/or original points. The number of markers in students' notes predicted students' learning outcomes, unlike the number of words. This result yielded educational implications regarding note-taking during slide-based lectures. Another significant point was made by the memory term conditions according to the distinction between short- and long-term effects. In particular, the long-term encoding effect from access to partial slides was verified: partial slide copies were more effective than full slide copies in the delayed tests but not in the immediate tests. However, further research is required to determine whether there are any significant influences from external storage effects and thus obtain a multidimensional understanding on the effects of IP slides and students' note-taking.

References

Aiken, E. G., Thomas, G. S., & Shennum, W. A. (1975). Memory for a lecture: Effects of notes, lecture rate, and informational density. *Journal of Educational Psychology, 67*(3), 439–444.

Annis, L. F. (1981). Effect of preference for assigned lecture notes on student achievement. *Journal of Educational Research, 74*(3), 179–182.

Apperson, J. M., Laws, E. L., & Scepansky, J. A. (2006). The impact of presentation graphics on students' experience in the classroom. *Computers & Education, 47*(1), 116–126. http://dx.doi.org/10.1016/j.compedu.2004.09.003.

Apperson, J. M., Laws, E. L., & Scepansky, J. A. (2008). An assessment of student preferences for PowerPoint presentation structure in undergraduate courses. *Computers & Education, 50*(1), 148–153. http://dx.doi.org/10.1016/j.compedu.2006.04.003.

Austin, J. L., Lee, M., & Carr, J. P. (2004). The effects of guided notes on undergraduate students' recording of lecture content. *Journal of Instructional Psychology, 31*(4), 314–320.

Austin, J. L., Lee, M. G., Thibeault, M. D., Carr, J. E., & Bailey, J. S. (2002). Effects of guided notes on university students' responding and recall of information. *Journal of Behavioral Education, 11*(4), 243–254.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28.

Babb, K. A., & Ross, C. (2009). The timing of online lecture slide availability and its effect on attendance, participation, and exam performance. *Computers & Education, 52*(4), 868–881.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29.

Baker, L., & Lombardi, B. R. (1985). Students' lecture notes and their relation to test performance. *Teaching of Psychology, 12*(1), 28–32.

Barbetta, P. M., & Skaruppa, C. L. (1995). Looking for a way to improve your behavior analysis lectures? Try guided notes. *The Behavior Analyst, 18*(1), 155–160.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bartsch, R. A., & Cobern, K. M. (2003). Effectiveness of PowerPoint presentations in lectures. *Computers & Education, 41*(1), 77–86.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Bretzing, B. H., & Kulhavy, R. W. (1979). Notetaking and depth of processing. *Contemporary Educational Psychology, 4*(2), 145–153.

Brobst, K. E. (1996). The process of integrating information from two sources, lecture, and text. (Doctoral dissertation, Teachers College, Columbia University, 1996). *Dissertation Abstracts International, 57*, 217.

Buchko, A. A., Buchko, K. J., & Meyer, J. M. (2012). Is there power in PowerPoint? A field test of the efficacy of PowerPoint on memory and recall of religious sermons. *Computers in Human Behavior, 28*(2), 688–695.

Bui, D. C., & Myerson, J. (2014). The role of working memory abilities in lecture note-taking. *Learning and Individual Differences, 33*, 12–22.

Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with Computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology, 105*(2), 299–309.

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

Castelló, M., & Monereo, C. (2005). Students' note-taking as a knowledge-construction tool. *L1: Educational Studies in Language and Literature, 5*(3), 265–285.

Chen, J., & Lin, T.-F. (2008). Does downloading PowerPoint slides before the lecture lead to better student achievement? *International Review of Economics Education, 7*(2), 9–18.

Coluccia, E., Gamboz, N., & Brandimonte, M. A. (2011). Normative data for a battery of free recall, cued recall and recognition tests in the elderly Italian population. *Neurological Sciences, 32*(6), 1103–1114.

Cornelius, T. L., & Owen-DeSchryver, J. (2008). Differential effects of full and partial notes on learning outcomes and attendance. *Teaching of Psychology, 35*(1), 6–12.

Daniels, L. (1999). Introducing technology in the classroom: PowerPoint as a first step. *Journal of Computing in Higher Education, 10*(2), 42–56.

Debevec, K., Shih, M.-Y., & Kashyap, V. (2006). Learning strategies and performance in a technology integrated classroom. *Journal of Research on Technology in Education, 38*(3), 293–307.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63*(1), 8–14.

Di Vesta, F. J., & Gray, G. S. (1973). Listening and note taking: II. Immediate and delayed recall as functions of variations in thematic continuity, note taking, and length of listening-review intervals. *Journal of Educational Psychology, 64*(3), 278–287.

Diggle, P. J., Heagerty, P. J., Liang, K.-Y., & Zeger, S. L. (2013). *Analysis of longitudinal data* (2nd ed.). Oxford, UK: Oxford University Press.

Donohue, M. C., & Edland, S. D. (2016). *Longpower: Power and sample size calculators for longitudinal data. R package version 1.0-16*.
Doumont, J.-L. (2005). The Cognitive style of PowerPoint: Slides are not all evil. *Technical Communication, 52*(1), 64–70.
Farkas, D. K. (2005). Explicit structure in print and on-screen documents. *Technical Communication Quarterly, 14*(1), 9–30.
Fisher, J. L., & Harris, M. B. (1973). Effect of note taking and review on recall. *Journal of Educational Psychology, 65*(3), 321–325.
Fisher, J. L., & Harris, M. B. (1974). Note taking and recall. *Journal of Educational Research, 67*(7), 291–292.
Fox, J., & Weisberg, S. (2011). *An {R} companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage. URL http://socserv.socsci.mcmaster.ca/jfox/Books/Companion/, Accessed date: 17 February 2007.
Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). *irr: Various coefficients of interrater reliability and agreement. R package version 0.84*. https://CRAN.R-project.org/package=irr/, Accessed date: 16 November 2001.
Grabe, M. (2005). Voluntary use of online lecture notes: Correlates of note use and note use as an alternative to class attendance. *Computers & Education, 44*(4), 409–421.
Grabe, M., & Christopherson, K. (2005). Evaluating the advantages and disadvantages of providing lecture notes: The role of internet technology as a delivery system and research tool. *The Internet and Higher Education, 8*(4), 291–298.
Hartley, J. (1983). Note-taking research: Resetting the scoreboard. *Bulletin of the British Psychological Society, 36*, 13–14.
Hidi, S., & Klaiman, R. (1983). Notetaking by experts and novices: An attempt to identify teachable strategies. *Curriculum Inquiry, 13*(4), 377–395.
James, K. E., Burke, L. A., & Hutchins, H. M. (2006). Powerful or pointless? Faculty versus student perceptions of PowerPoint use in business education. *Business Communication Quarterly, 69*(4), 374–396.
Jansen, R. S., Lakens, D., & IJsselsteijn, W. A. (2017). An integrative review of the cognitive costs and benefits of note-taking. *Educational Research Review, 22*, 223–233.
Katayama, A. D., & Crooks, S. M. (2003). Online notes: Differential effects of studying complete or partial graphically organized notes. *The Journal of Experimental Education, 71*(4), 293–312.
Katayama, A. D., & Robinson, D. H. (2000). Getting students "partially" involved in note-taking using graphic organizers. *The Journal of Experimental Education, 68*(2), 119–133.
Kiewra, K. A. (1985a). Students' note-taking behaviors and the efficacy of providing the instructor's notes for review. *Contemporary Educational Psychology, 10*(4), 378–386.
Kiewra, K. A. (1985b). Providing the instructor's notes: An effective addition to student notetaking. *Educational Psychologist, 20*(1), 33–39.
Kiewra, K. A. (1989). A review of note-taking: The encoding-storage paradigm and beyond. *Educational Psychology Review, 1*(2), 147–172.
Kiewra, K. A., & Benton, S. L. (1995). Effects of note-taking format and study technique on recall and relational performance. *Contemporary Educational Psychology, 20*(2), 172–187.
Kiewra, K. A., Dubois, N. F., & Christensen, M. (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*(2), 240–245.
Kirby, J. R., & Pedwell, D. (1991). Students' approaches to summarization. *Educational Psychology, 11*, 297–307.
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 Behavioural Science, 18*(1), 52–61.
Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30*(2), 242–262.
Kobayashi, K. (2006). Combined effects of note-taking/-reviewing on learning and the enhancement through interventions: A meta-analytic review. *Educational Psychology, 26*(3), 459–477.
Konrad, M., Joseph, L. M., & Eveleigh, E. (2009). A meta-analytic review of guided notes. *Education & Treatment of Children, 32*(3), 421–444.
Kosslyn, S. M., Kievit, R. A., Russell, A. G., & Shephard, J. M. (2012). PowerPoint(®) presentation flaws and failures: A psychological analysis. *Frontiers in Psychology, 3*, 230.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models. R package version 2.0-32*. https://CRAN.R-project.org/package=lmerTest/ Accessed 17.01.24.
Lahtinen, V., Lonka, K., & Lindblom-Ylänne, S. (1997). Spontaneous study strategies and the quality of knowledge construction. *British Journal of Educational Psychology, 67*(1), 13–24.
Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning, 65*(S1), 127–159. http://dx.doi.org/10.1111/lang.12115.
Larwin, K. H., & Larwin, D. A. (2013). The impact of guided notes on post-secondary student achievement: A meta- analysis. *International Journal of Teaching and Learning in Higher Education, 25*(1), 47–58.
Le, A., Joordens, S., Chrysostomou, S., & Grinnell, R. (2010). Online lecture accessibility and its influence on performance in skills-based courses. *Computers & Education, 55*(1), 313–319.
Levasseur, D. G., & Sawyer, J. K. (2006). Pedagogy meets PowerPoint: A research review of the effects of computer-generated slides in the classroom. *Review of Communication, 6*(1–2), 101–123.
Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning, 65*(suppl. 1), 185–207.
Liu, G., & Liang, K.-Y. (1997). Sample size calculations for studies with correlated observations. *Biometrics, 53*(3), 937–947.
Loaiza, V. M., McCabe, D. P., Youngblood, J. L., Rose, N. S., & Myerson, J. (2011). The influence of levels of processing on recall from working memory and delayed recall tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1258–1263.
Lowry, R. B. (1999). Electronic presentation of lectures: Effect upon student performance. *University Chemistry Education, 3*, 18–21.
Makany, T., Kemp, J., & Dror, I. E. (2009). Optimising the use of note-taking as an external cognitive aid for increasing learning. *British Journal of Educational Technology, 40*(4), 619–635.
Mantei, E. J. (2000). Using Internet class notes and PowerPoint in the physical geology lecture. *Journal of College Science Teaching, 29*(5), 301–305.
Murphy, T. M., & Cross, V. (2002). Should students get the instructor's lecture notes? *Journal of Biological Education, 36*(2), 72–75.
Newell, G. E. (1984). Learning from writing in two content areas: A case study/protocol analysis. *Research in the Teaching of English, 18*(3), 265–287.
Norton, L. S., & Hartley, J. (1986). What factors contribute to good examination mark?: the role of notetaking in subsequent examination performance. *Higher Education, 15*, 355–371.
Oakhill, J., & Davies, A.-M. (1991). The effects of test expectancy on quality of note taking and recall of text at different times of. *British Journal of Psychology, 82*(2), 179–189.
Peper, R. J., & Mayer, R. E. (1978). Note taking as a generative activity. *Journal of Educational Psychology, 70*(4), 514–522.
Peper, R. J., & Mayer, R. E. (1986). Generative effects of note taking during science lectures. *Journal of Educational Psychology, 78*, 34–38.
Peverly, S. T., Garner, J., & Vekaria, P. (2014). Both handwriting speed and selective attention are important to lecture note-taking. *Reading and Writing, 27*(1), 1–30.
Peverly, S. T., Ramaswamy, V., Brown, C., Sumowski, J., Alidoost, M., & Garner, J. (2007). What predicts skill in lecture note taking? *Journal of Educational Psychology, 99*(1), 167–180.
Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS.* New York, NY: Springer.
Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*(3), 291–312.
Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59*(4), 413–425.
R Core Team (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. URL https://www.R-project.org/.
Revelle, W. (2016). *psych: Procedures for personality and psychological research.* Evanston, Illinois, USA: Northwestern University. https://CRAN.R-project.org/

package=psych Version =1.6.12.
Robinson, D. H., Katayama, A. D., Beth, A., Odom, S., Ya-Ping, H., & Vanderveen, A. (2006). Increasing text comprehension and graphic note taking using a partial graphic organizer. *Journal of Educational Research, 100*(2), 103–111.
Savoy, A., Proctor, R. W., & Salvendy, G. (2009). Information retention from PowerPoint™ and traditional lectures. *Computers & Education, 52*(4), 858–867.
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*(4), 312–328.
Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2012). *Research methods in psychology* (9th. ed.). New York, NY: McGraw-Hill.
Stephenson, J. E., Brown, C., & Griffin, D. K. (2008). Electronic delivery of lectures in the university environment: An empirical comparison of three delivery styles. *Computers & Education, 50*(3), 640–651.
Susskind, J. E. (2005). PowerPoint's power in the classroom: Enhancing students' self-efficacy and attitudes. *Computers & Education, 45*(2), 203–215. http://dx.doi.org/10.1016/j.compedu.2004.07.005.
Susskind, J. E. (2008). Limits of PowerPoint's power: Enhancing students' self-efficacy and attitudes but not their behavior. *Computers & Education, 50*(4), 1228–1239. http://dx.doi.org/10.1016/j.compedu.2006.12.001.
Szabo, A., & Hastings, N. (2000). Using IT in the undergraduate classroom: Should we replace the blackboard with PowerPoint? *Computers & Education, 35*(3), 175–187.
Titsworth, B. S., & Kiewra, K. A. (2004). Spoken organizational lecture cues and student notetaking as facilitators of student learning. *Contemporary Educational Psychology, 29*(4), 447–461.
Tremblay, A., & Ransijn, J. (2015). *LMERConvenienceFunctions: Model selection and post-hoc analysis for (G)LMER models. R package version 2.10.* https://CRAN.R-project.org/package=LMERConvenienceFunctions/ Accessed 17.01.24.
Tufte, E. R. (2003). *The cognitive style of PowerPoint.* Cheshire, CT: Graphics 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*(3), 323–338.
Williams, J. M., & Colomb, G. G. (2007). *The craft of argument* (3rd ed.). New York, NY: Pearson.
Williams, R. L., & Eggert, A. C. (2002). Notetaking in college classes: Student patterns and instructional strategies. *The Journal of General Education: The Journal of General Education, 51*(3), 173–199.
Williams, W. L., Weil, T. M., & Porter, J. C. K. (2012). The relative effects of traditional lectures and guided notes lectures on university student test scores. *Behavior Analyst Today, 13*(1), 12–16.
Worthington, D. L., & Levasseur, D. G. (2015). To provide or not to provide course PowerPoint slides? The impact of instructor-provided slides upon student attendance and performance. *Computers & Education, 85*, 14–22.