

Written notes and listening comprehension: A correlation study

Joseph Siegel
Stockholm University

Abstract

Notetaking is a crucial aspect of learning in academic contexts, but as a relatively casual form of academic writing, it seldom receives pedagogic or research attention in the literature. Therefore, as more students study academic content through English as a second language (L2), research on student notetaking as a form of academic writing deserves attention. What students write in their notes and how they do so can play important roles in comprehension and learning. To address this gap, the present study examines 102 sets of notes and corresponding listening comprehension test scores to determine the relationships between four factors of quantity and quality in students' hand-written notes; namely, notations, words, information units, and efficiency ratio. Results indicate that total notations and total words written in notes do not impact overall test scores, while information units and higher efficiency ratios positively correlate to test scores. The paper closes with pedagogic advice for teachers and students operating in L2 academic contexts with a focus on how best to conceptualise and write notes.

Introduction

Taking notes is a traditional part of academic life, especially in secondary and tertiary education. Notetaking when listening to academic input not only helps concentration and learning but generates a source of information that can be used for tasks such as writing reviews, discussions, and essays as well as exam preparation more generally. The relationship between academic performance and notetaking has long been of interest to researchers in educational psychology, memory, study skills, notetaking strategies, and language use. Almost a century ago, Crawford (1925) provided empirical evidence via positive correlations between written notes and quiz scores that notes support learning. Later research (e.g., Clerehan, 1995; Dunkel, 1988) demonstrated differences between groups listening to and taking notes in their first (L1) and second language (L2).

More recently, the number of English medium instruction (EMI) courses has risen worldwide (e.g., Lasagabaster & Doiz, 2021; Macaro, et al., 2018; Murata, 2018). In response to the needs of L2 students for academic abilities in L2 English, the number of English for academic purposes (EAP) courses has also seen an increase. Since listening and taking notes in an L2 can be arduous, such EAP courses often include specific emphasis on listening comprehension and notetaking. Moreover, EAP textbooks target these crucial skills (e.g., Siegel, 2021a), particularly because much content in tertiary courses, including EMI, is delivered via the traditional lecture, which remains a dominant form of content delivery (e.g., Crawford Camiciottoli & Querol-Julian, 2016). In the European context, notetaking in EMI courses can be particularly important given the ample mobility of students to study in English in a number of countries, facilitated by student exchange programs like Erasmus.

Despite correlation studies in L1 contexts establishing a generally clear relationship between the quantity and quality of notes in relation to test performance, similar studies with L2 English learners are less available, with Dunkel (1988) as the one exception. The present study views the generation of notes on paper as a purposeful act of written composition. It takes a “reception-oriented” view of these writing processes through the semiotic meaning-making signs (e.g., letters, words, symbols, underlines, etc.) that students choose to include in their notes (Antia & Mafofo, 2021, p. 86). To determine the relationships between the quantity and quality of notes taken by L2 English users and their achievement on post-listening tests, the study examines notes at three levels: total notations, total words, total ‘information units’ (see below for a more comprehensive explanation), and efficiency rate. Using 102 sets of notes generated by students in three EAP classes, it aims to investigate, through correlation calculations, the contributions of each of these three levels to test achievement. The paper begins by reviewing key concepts related to notetaking followed by an overview of research on notetaking drawing from both L1 and L2 contexts, paying special attention to those studies that have included correlation calculations. The methodology for the study is then outlined, including background on the participants and procedures for note analysis. Results from the correlations on notations, words, information units, and efficiency ratio in relation to post-listening test scores are then displayed and discussed, leading to pedagogic implications for both L2 students and teachers.

Background

On Notes and Notetaking as Academic Writing

Notetaking is a broad strategy typically employed to support learning and memory. It is an informal type of academic writing that is often overlooked or taken for granted but is a form of writing that contributes significantly to learning (e.g., Kobayashi, 2005; 2006). As noted by John-Steiner (1997), individuals create meaning from different perspectives, including visual and verbal thinking. These processes can result in various semiotic representations of meaning in notes. Students also have a variety of notetaking options available to them. While a traditional view of notetaking is handwritten notes on paper, technology offers a wide array of alternatives, including typing on laptop computers or tablets, writing on a tablet surface with a digital pen, and taking photos, which might be integrated into a set of digital notes. Furthermore, teachers can hand out copies of PowerPoint slides to students to help support notetaking, and teaching sessions can be recorded by individual students or by the teacher, meaning they can be revisited. In addition, notes can be taken while listening, as in the present study, or during the reading process (e.g., Antia & Mafofo, 2021). All of these options and factors, as well as content, motivation, and language of delivery, may affect the approaches and strategies that students employ.

Since notetaking does not have the same type of formalised conventions as other types of academic writing (e.g., the 5-paragraph essay or a standard research report), it can be challenging to determine the quality of a listener’s notes in comparison to a pre-determined rubric or explicit set of criteria. Instead, individual notetakers typically express various conceptualizations of and purposes for notetaking (Badger et al., 2001) as well as various notetaking habits (e.g., Morehead et al., 2019). In addition, explicit pedagogic attention for notetaking is also lacking, as evidenced by reports from L1 contexts (e.g., Korn, 2018) as well as L2 students and teachers (e.g., Siegel, 2019a; 2021b). Furthermore, discrepancies related to which notetaking strategies and techniques are most effective are evident in the literature (e.g., Mueller & Oppenheimer, 2014; Luo et al., 2018; Morehead et al., 2019). As such, precisely the type of writing that is valuable in notes is open to debate. Questions related to how detailed notes should be, how to establish connections between discrete parts of notes, and at what level of expression notes should be written in order to be valuable to the notetaker have yet to be answered definitively. While this paper cannot claim to provide comprehensive answers, it probes such questions.

Regardless of whether notetaking while listening occurs in L1 or L2 contexts, the act involves multiple discrete language skills; specifically, listening, writing, and later, reading. Each student will have varying proficiency levels for each of these skills. More general academic skills such

as concentration, memory, handwriting speed, and multitasking are also involved in the process. The way one mentally approaches the acts of comprehending and thinking about input also play a role, as notetaking involves individuality and creativity in reconstructing and elaborating on messages delivered by a speaker. John-Steiner (1997) makes a theoretical distinction between visual and verbal thinking but emphasizes a common goal for the mind to order, reshape and remember experiences and information. At the same time, language itself can struggle to verbalise all conceivable ideas. Even visual thinkers are likely normalized by convention to “translate the visual shapes of their thoughts into words” (John-Steiner, 1997, p. 88). Verbal thinking, on the other hand, involves expressing thoughts in writing, often in short bursts of previously unrehearsed composition (John-Steiner, 1997, pp. 111-112). Notetaking has the potential to include concrete representations of both visual (e.g., pictures, symbols, highlighting, etc.) and verbal (e.g., words, abbreviations, etc.) thinking.

Another theoretical perspective relevant to notetaking is cognitive load (Craik & Lockhart, 1972). In an attempt to relieve potential cognitive overload from listening and memory, listeners can reduce their cognitive load (Craik & Lockhart, 1972) by writing down information for later use. This act not only creates a repository of information for later use but also frees up cognitive resources to continue paying attention to incoming speech and taking notes when necessary. As the listener writes down their notes, they are encoding information by transferring it from the aural form to written form (DiVesta & Gray, 1972). Encoding occurs in real-time as the notetaker listens and records information, in contrast to the storage effect of notes, which involves later review and use of the notes. In other words, the storage effect is separated in time from the actual listening input whereas encoding happens continuously within the listening event.

Within the encoding process, listeners must make a number of choices related to content, style, and expression as they simultaneously listen and write. For content-related choices, the listener must prioritize the various pieces of information they hear. They distinguish the more relevant from the less relevant and may decide to make a note of the former while paying less attention to the latter. Style decisions relate to overall structure of the notes, which might include the traditional outline format, the Cornell method (a system of notetaking that involves two columns on the page with specific space for main ideas, details, questions, and a summary; see Pauk, 2001, for a comprehensive description), a mind-map, or other structural (or random) designs (e.g., Kiewra, et al., 1985; Song, 2012), for example.

At the level of written expression in notes, students make conscious choices about the actual markings that appear on the paper. Generally speaking, notes are conceptualized as consisting mainly of words written on paper. The encoding of information in notes, at the most minute, itemised level, consists of letters, punctuation, symbols, etc. that combine in a bottom-up direction to form words, and then ideas. At this level during notetaking while listening, the notetaker has the option of writing notes verbatim or paraphrasing what the speaker says. Notetakers also choose whether and how to use symbols, abbreviations, and/or punctuation to encode meaning, create structure, and preserve the relationships between different ideas; for example, in relation to notes taken while reading (e.g., annotations), Antia and Mafofo (2021) demonstrate multiple uses of underlining, highlighting and use of question marks in their database.

Notetakers can also exercise their multi-lingual repertoire by engaging in translanguaging to take notes (i.e., using both L2 English and their L1(s)) in notes (e.g., Antia & Mafofo, 2021; Siegel, 2023). To express relevant content in notes, listeners must select from their mental bank of resources to generate representations of pieces of information they want to record and possibly the relationships between those pieces of information. This cache of resources includes the words they know (and related lexical aspects such as spelling and semantic meaning), the words the speaker has said (which can be approximated phonetically even if the listener does not know the correct spelling or semantic meaning), punctuation, symbols, abbreviations, and illustrative abilities for including images in notes.

Notetakers employ such recourses discretely, but their combination can lead to individual propositions, or ‘information units’ (IUs), a concept stemming from the field of psychology. According Anderson (2014), an IU represents the smallest amount of information that on its

own can be judged as true or false. The term IU is adopted for this paper, as it has been used and defined in other studies on notetaking (e.g., Dunkel, 1988). The concept is similar to the relatively casual terms 'points' or 'ideas' that have been used in other notetaking studies and discussions (e.g., Crawford, 1925; DiVesta & Gray, 1972; Einstein et al., 1985; Kiewra & Fletcher, 1984; Suritsky & Hughes, 1991), but the qualification that an IU is a proposition that can be judged as true or false renders it preferable to looser terms such as 'point' or 'idea' (see Siegel, 2018 for a discussion of benefits and drawbacks to various methods for determining note quality). From a linguistic perspective, Halliday (2014) points out that the closest grammatical unit to the IU is a clause, which contains multiple words. In speech, IUs can be marked by a change in tone or pitch and can contain new or previously known information (Halliday, 2014). The definition of an IU suggests that it must then contain multiple components (spoken or written), which could contain words (including abbreviations), symbols (e.g., punctuation, arrows, etc.), and/or illustrations. As such, both "visual thinking" and "verbal thinking" (John-Steiner, 1997), as well as a combination thereof, can manifest themselves in any single IU.

The concept of the IU has also been used to generate an efficiency index (Howe, 1970) or efficiency ratio (Dunkel, 1988). This measurement is calculated by dividing the number of IUs that appear in notes with the total number of words found in notes (e.g., Dunkel, 1988, p. 265). The higher the efficiency ratio, presumably, the more effective the notetaker was at combining individual words to produce meaningful ideas while limiting extraneous production. By focusing solely at the word-level, however, this calculation overlooks the potential contributions that semiotic symbols such as punctuation, arrows, stars, underlines, highlights, and illustrations (e.g., Antia & Mafofo, 2021; Kress, 2010) might make to generating IUs.

The relative value of these different components at the level of expression is open to question. For example, when taking notes, should students focus on writing down as many words as possible? How does each word or symbol contribute to forming the larger concept of the IU? How are the relationships between parts of the notes represented via symbols such as arrows, stars, and bullet points? Do such symbols contribute to the overall usefulness of notes on post-listening tests? Which are more useful to test performance: Total notations, total words, or total IUs? These are the questions this study aims to address.

Previous Correlation Research from L1 Contexts

A number of studies conducted in L1 contexts have provided empirical evidence that notetaking and test performance are positively correlated. Crawford (1925) demonstrated strong significant positive correlations between the number of 'points' in notes and the number of correct answers on post-listening quizzes. Participants in this study, however, ranged in academic level from university freshman to graduate students from two separate institutions, suggesting large gaps in skill-levels. DiVesta and Gray (1972) found that test performance was related to the number words students wrote down in a post-listening notetaking period. These researchers also emphasized that complete ideas in notes affected performance as opposed to individual words with no connection to a larger proposition or idea.

Likewise, Kiewra and Fletcher (1984) found strong positive correlations between test-related ideas in notes, and immediate test scores ($R = .72$) and delayed recall scores ($R = .78$). Einstein et al. (1985) found that students were able to recall 44% of the information they wrote down in their notes compared to only 6% of the information that they listened to but did not record in notes. More recently and in relation to revision of notes, Luo et al. (2016) demonstrated the strong relationship between additional notes added at lecture pauses and achievement on test scores. As these researchers observe, their findings reinforce those of previous studies conducted in L1 contexts, which find that notetaking is positively related to achievement. Peverly et al. (2003) focused their study on the relationship between college students' self-regulation, note taking (both at the macropropositional and additional information levels), study time and background knowledge. They found no significant correlations between the variables in connection to an essay task or a multiple choice listening test (Peverly et al., 2003, p. 343). However, these studies did not go beyond the conceptual level of idea, point, proposition, or IU to investigate how individual words and total notations written in notes might influence test performance. Furthermore, while they provide a valuable foundation for comparative purposes,

these studies all took place in contexts where English is the L1. As such, they provide strong evidence that notetaking is positively correlated with achievement in L1 contexts; this general finding, however, has yet to be confirmed in learning environments in which English is an L2.

Notetaking Research in L2 Contexts

Research on notetaking in L2 English contexts has increased recently, a trend likely due to the increase in EMI- and EAP-themed research. Previous studies have addressed topics such as notetaking habits of L2 English users (e.g., İpek, 2018; Siegel, 2021b), translanguaging practices in notes (e.g., Wang, 2021), and the effects of various forms of notetaking training (e.g., Tsai & Wu, 2010). Some studies have focused attention on close examination of note content, while others have investigated relations between notes and post-listening activities. Crawford (2015) analysed student notes for total notations, words and symbols and demonstrated increased quantity of each category following an intervention period. Sakurai (2018) also monitored the effects of a pedagogic notetaking intervention by measuring 'key points' (presumably similar to IUs but not defined), L1 words, L2 words, symbols and abbreviations. Similarly, Kusumoto (2019) analysed student notes through IUs and test scores following an intervention period. Her analysis involved t-tests to compare pre- and post-intervention results but did not include correlations of one measure to the other. None of these studies nor any others conducted in L2 settings appear to have measured correlations between the discrete items encoded in the notes and listening test performance.

Two studies focused attention on correlations between note content and performance and are therefore of particular relevance to the present paper. Carrell (2007) investigated the relationship between notes taken by students taking a computer-based version of TOEFL test-like items and reported that multiple-choice listening comprehension test scores were significantly correlated with number of content words in notes, the number of abbreviations and symbols, and the number of test answers in notes (p. 36). Dunkel (1988) analyzed notes taken by 66 native L1 English users and 63 L2 English users based on four measures: total number of words, IUs, test answerability, and efficiency ratio. Results of multicollinearity and stepwise multiple regression analysis among these different measures showed positive correlations of total words ($R = .53$) and IUs ($R = .43$) to correct answers on concept questions for the L2 English group. The correlation between IUs and detail questions was slightly lower at $R = .37$. These data suggest that, at least for broader concept questions, a higher number of total words leads to more correct answers when compared with number of IUs. However, her findings for the efficiency ratio in notes did not show a significant correlation between efficiency ratio and test performance, and the author points out that these findings stand in contrast to earlier results by Howe (1970), who found a significant correlation between the two measures (Dunkel, 1988, p. 272). Results generated by these two studies provide a useful comparative opportunity in relation to the data set collected for the present study.

Research Questions

In order to determine the relative value of notes scrutinized at various levels of the expression phase of notetaking, this study aims to address the following research questions in relation to notes taken by EAP students:

- (1) What is the relationship between the total notations written in notes and test performance?
- (2) What is the relationship between the total words written in notes and test performance?
- (3) What is the relationship between the total IUs written in notes and test performance?
- (4) What is the relationship between the efficiency ratio in notes and test performance?

Methods

Participants

The 34 participants in the study were taking the first of two required EAP classes at the upper secondary school level in Sweden. Students at this level are generally at or near the CEFR B1 level per Ministry of Education policy documents and must take two L2 English courses for

graduation. Students in Sweden typically begin formal English study in primary school, which continues up to and including secondary school. The students studied herein were from three intact classes taught by two different teachers. These teachers were participating in a professional development project focused on notetaking instruction for EAP students and volunteered to be involved in the data collection. Notably, notetaking was of interest as the Swedish Ministry of Education (Skolverket) refers to notetaking on its curriculum steering documents both implicitly (e.g., in relation to understanding and making sense of longer stretches of spoken and written text) and explicitly (e.g., that notetaking is one strategy to aid receptive skills) (Skolverket, 2021). The students who participated signed consent forms stating that their notes and test scores could be analyzed for research purposes. No personal information was collected. Each student was given an anonymous project code so that their notetaking performances could be compared from one data collection point to the others. These codes were assigned by the class teachers, who retained a project key. This key was not shared with the researcher in order to protect participant identities.

Data Collection and Analysis

At three separate points during the academic year, the 34 students took notes while listening to short TED Talks in English (see Table 1 for details about the videos). The three data collection points were 'progress check points' within the professional development project on notetaking mentioned above. Data was collected during the first week of the term, after eight weeks, and then again 16 weeks later. This generated a total of 102 note sets and 102 comprehension test scores (students missing any of these data points were removed from the analysis). They were instructed to take as many notes as possible, and that they would use their notes for post-listening tasks; namely, answering quiz questions and discussions in pairs. Students did not preview the quiz questions before listening. Ten minutes after the listening exercise was finished, students used their notes to complete online comprehension quizzes. Each quiz consisted of 16 items, eight of which were multiple choice and eight of which were fill-in-the-blank. When the quizzes were finished, the classroom teachers collected student notes, scanned them electronically and sent them via email to the researcher. The scanned notes were then transferred to a password protected external hard drive to facilitate manual analysis. Quiz scores were collected automatically in the online system. Fill-in-the-blank answers were then checked manually for spelling errors, paraphrasing, etc. and additional points were awarded as necessary.

Table 1. TED Talks used in data collection

Title	Speaker	Total IUs	Length
Clues to prehistoric time, found in blind cavefish	Male, North American accent	47	4:50
How we'll find life on other planets	Female, North American accent	46	5:25
Tracking our online trackers	Male, North American accent	47	6:33

Notes were analysed by the researcher in consultation with the teachers for total notations, total words and total IUs. The tallies for total words and total IUs were then used to calculate the efficiency ratio. Total notations included everything (i.e., all semiotic meaning-making signs (Antia & Mafofo, 2021)) the student wrote on the page; for example, each word, punctuation mark, symbol, underline, and numeral was counted as a single notation. For punctuation that typically comes in pairs, such as quotation marks or brackets, each set was counted as one notation (i.e., “ ” equated to one notation). The total words measurement excluded all punctuation and symbols (e.g., bullet points, arrows, stars, etc.) and instead focused on series of letters with blank space on either side. Abbreviations and numerals were counted as words since their intended use is typically to represent a complete word. For example, the expression “8h” (representing the phrase “eight hours” in one of the TED Talks) was counted as two words (8 = eight, h = hours). Figures 1 and 2 below provide additional examples of how note scoring was done.

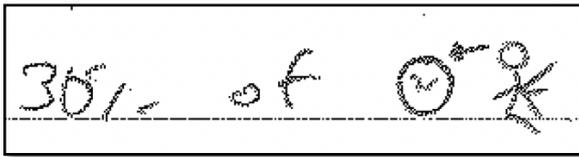


Figure 1. Example of note analysis A
Notations: six: 30; %; of; [circle image]; [arrow]; [stick figure of person]
Words: three: 30; %; of
IUs: one: representing this line from the TED Talk “Tracking our online trackers”: “that at any point in time more than 30 percent of the world’s population can go online”.

Table 2. Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
IUs	102	25,00	2,00	27,00	10,4608	5,25144
Total_Notations	102	154,00	5,00	159,00	72,4412	32,19390
Total_Words	102	128,00	4,00	132,00	58,0294	26,00398
Test_Scores	102	12,00	2,00	14,00	6,8725	2,51991
Efficiency_ratio	102	,45	,05	,50	,1893	,06723

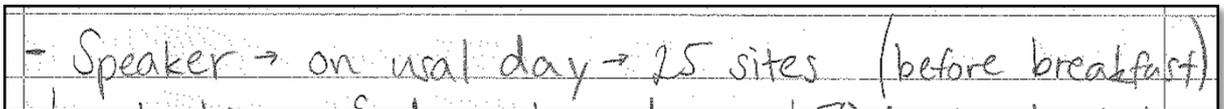


Figure 2. Example of note analysis B
Notations: 12: -; speaker; [arrow]; on; usual; day; [arrow]; 25; sites; (); before; breakfast
Words: eight: speaker; on; usual; day; 25; sites; before; breakfast
IUs: one: representing this line from the TED Talk “Tracking our online trackers”: “We are not even two bites into breakfast and there are already nearly 25 sites that are tracking me.”

In some note sets, students had crossed out certain words (e.g., ~~word~~). In these cases, the strikethrough was counted as a single notation and the word counted as a single word, equaling two separate notations. IU identification was based on the notion of whether an item of information or a proposition can be judged as true or false. Numbers generated by this quantitative analysis were then used with SPSS Version 27 to conduct bivariate Pearson correlation calculations. Data are available at 10.17045/sthlmuni.19244193.

Results

This section reports results of the correlation calculations related to the various notetaking quantity and quality measures described above in order to address the stated research questions. Table 2 displays descriptive statistics for each of the five categories of data. The total notations measure showed the largest range and standard deviation of the five measurements.

These observations are to be expected, since the notetaker could, in theory, write a nearly infinite number of items in their notes. In contrast, IUs and test scores have more modest ranges and standard deviations due to their delineated nature (i.e., a maximum total of 46 or 47 IUS, depending on the talk, and a maximum test score of 16).

Table 3 shows the overall correlation results. While several strong correlations between individual items in notes are evident, specifically between total words and IUs ($R = .805$), for instance, these were also expected to a large extent since each is a component of the others in a way. That is, in order to compose IUs, words are needed; thus, the strong correlation between total words and IUs is expected. The more words present, the more IUs are possible, up to a point (i.e., the 46 or 47 IUs identified in each talk). Another strong correlation is evident between total words and total notations ($R = .97$). Since words were counted as one type of notation, this was expected.

Of interest to the present research are the correlations of these various measurements with test scores. Weak positive correlations were found between total notations ($R = .069$) and total words ($R = .078$) and test scores. A stronger positive correlation is evident between IUs and test

Table 3. Correlations

		IUs	Total Words	Test Scores	Efficiency rate	
IUs	Pearson Correlation	1	,800**	,805**	,290**	,211*
	Sig. (2-tailed)		,000	,000	,003	,033
	N	102	102	102	102	102
Total_Notations	Pearson Correlation	,800**	1	,970**	,069	-,282**
	Sig. (2-tailed)	,000		,000	,491	,004
	N	102	102	102	102	102
Total_Words	Pearson Correlation	,805**	,970**	1	,078	-,304**
	Sig. (2-tailed)	,000	,000		,435	,002
	N	102	102	102	102	102
Test_Scores	Pearson Correlation	,290**	,069	,078	1	,253*
	Sig. (2-tailed)	,003	,491	,435		,010
	N	102	102	102	102	102
Efficiency_ratio	Pearson Correlation	,211*	-,282**	-,304**	,253*	1
	Sig. (2-tailed)	,033	,004	,002	,010	
	N	102	102	102	102	102

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

scores (.29). This correlation is weak but approaching a moderate linear relationship and is significant at $p < .01$. All correlations related to efficiency rate were significant at $p < .05$, with IUs ($R = .211$) with test scores ($R = .253$) being positively correlated. Total notations and total words negatively correlated with efficiency rate ($R = -.282$ and $R = -.304$, respectively), both significant at $p < .01$; in other words, as notations and words go up, efficiency rate goes down, indicating that the efficiency rate can become deluded by markings and words in notes that do not explicitly connect to the larger notion of propositional meaning represented by IUs. As such, these items in notes may be extraneous and represent effort expended with little to no benefits in terms of note quality or content recall.

Figures 3-6 are scatterplots visually displaying correlations between test scores and measures of note quantity and quality. Figures 3 and 4 show the relationships between test scores and total notations and total words, respectively. As evidenced by the nearly horizontal regression lines and low R^2 scores (.005 and .006), neither total notations nor total words seems to have much impact on test scores.

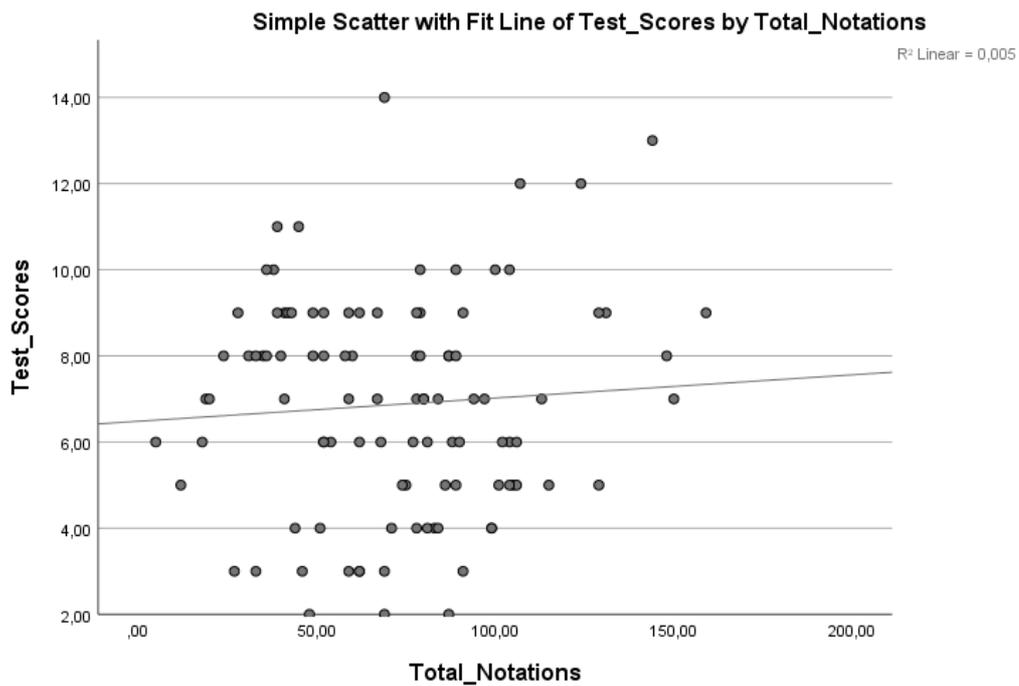


Figure 3. Scatterplot of Test scores and total notations

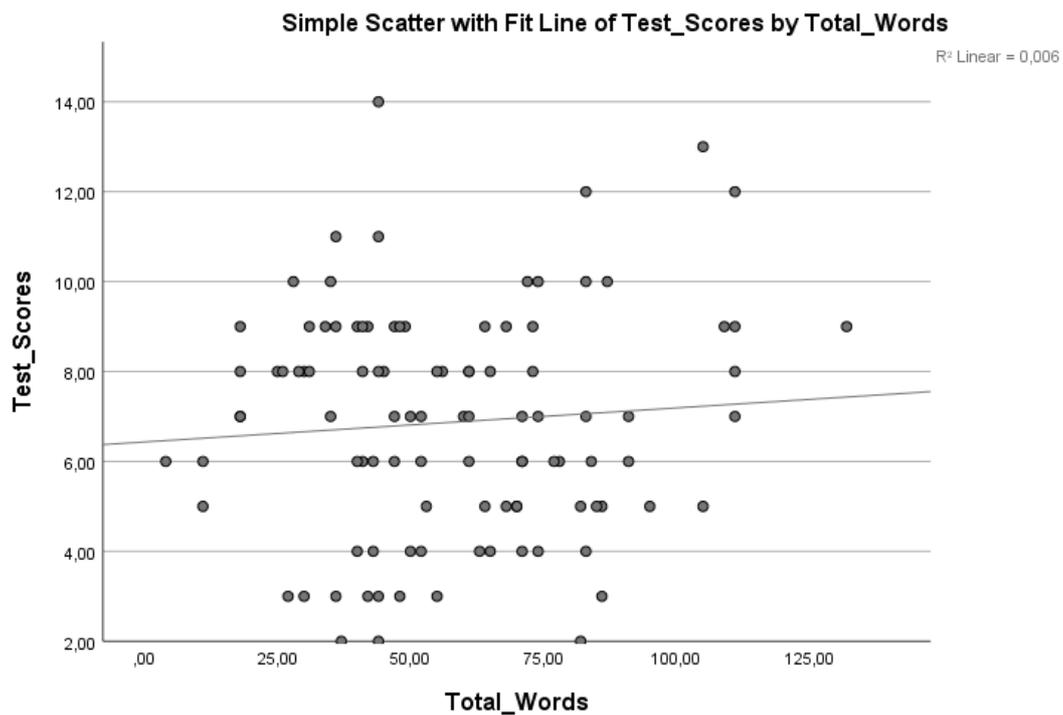


Figure 4. Scatterplot of Test scores and total words

Positive correlations between test scores and IUs and efficiency rate respectively, are evident in Figures 5 and 6. The regression lines are steeper than in Figures 3 and 4, representing stronger positive correlations. While the R^2 results for test scores and IUs and efficiency rates are relatively low, they provide some explanation of the factors that account for test performance and are much higher than for total words and total notations.

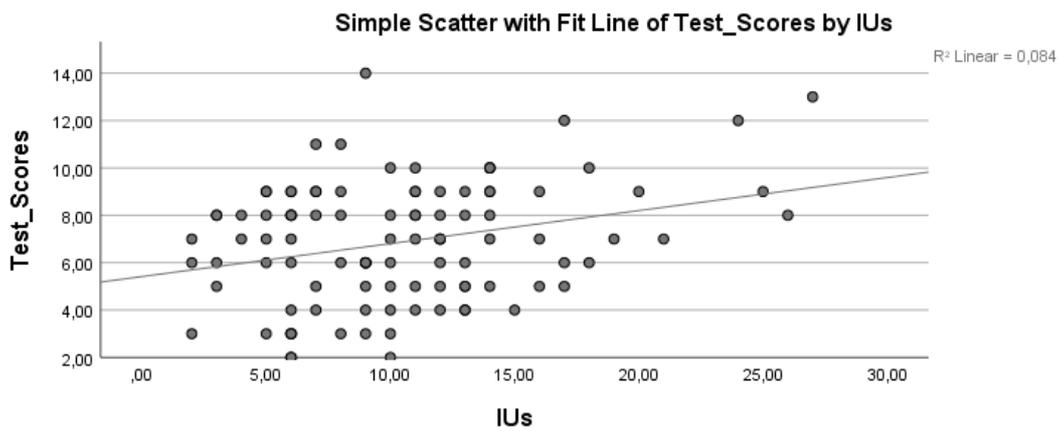


Figure 5. Scatterplot of Test scores and IUs

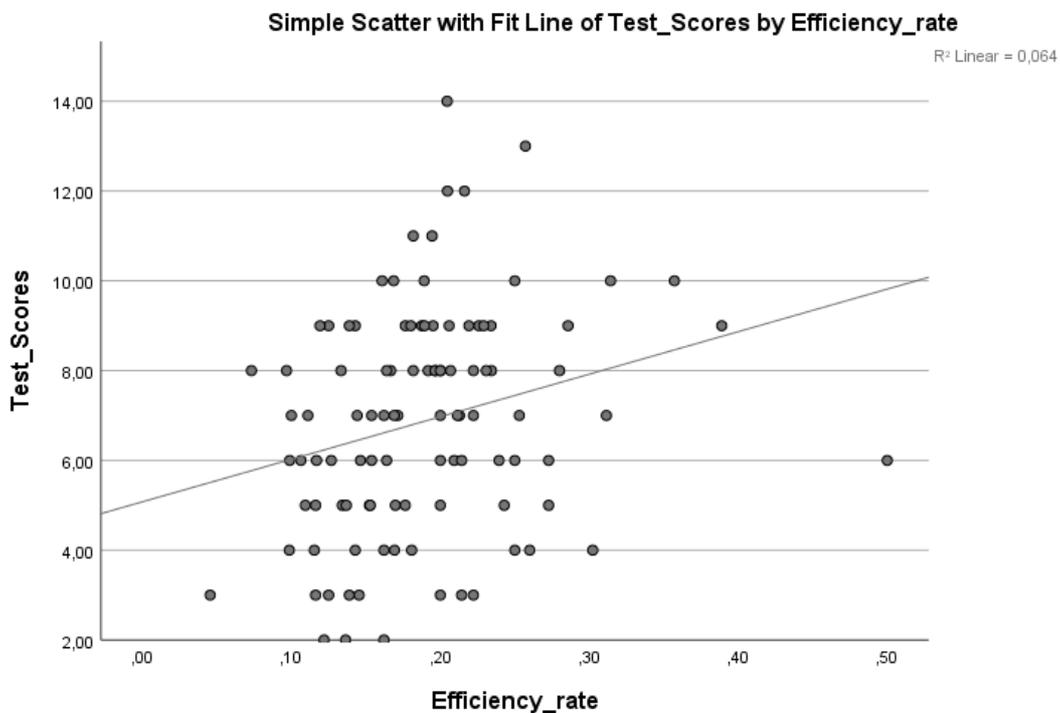


Figure 6. Scatterplot of Test scores and efficiency ratio

To summarise these findings, correlations of various measures of note quantity and quality suggest that total notations and total words in notes have little impact on test achievement. Stronger positive correlations were found between IUs and efficiency rate and test scores, indicating that notations and words on their own do not necessarily lead to test achievement but that these more discrete items need to be combined into fuller, more complete ideas in order to be useful for comprehension tests, particularly those with multiple choice and/or short gap fill questions.

Discussion

As evident in the data presented in the previous section, the relationships between total notations and total words with test scores are generally very low. These findings suggest that there is little purpose in simply writing as much as possible in notes. More does not always generate actionable material to stimulate recall for the purposes of answering test questions.

In the present study, the relationships between total IUs and efficiency rate has stronger positive correlations with test scores, indicating that items in notes that go beyond single words or symbols have more value when answering multiple choice and fill-in-the-blank questions. These findings contrast those in Carrell (2007), which involved TOEFL-like multiple choice items. In that study, correlations for total notations and content words were positively correlated with test items for both intervention and non-intervention groups. In terms of IUs, comparisons between the present study and Carrell's (2007) are not possible because that study did not include any propositional measurement, such as IUs, lecture points, or ideas. Dunkel (1988), however, found a moderate positive correlation ($R = .37$) between IUs and detail questions, a slightly stronger correlation than the $R = .29$ found here.

Therefore, what should students be aiming for in terms of the notes they take? The more IUs that appear in notes, the higher the efficiency rate will be. IUs themselves are comprised of notations that have some relation to each other. Focusing notetaking attention on maximising each notation and striving to have each one be part of a larger idea beyond its individual, discrete meaning would seem valuable advice. Of course, this is not possible with every notation. That would be unrealistic under normal notetaking conditions. It can, however, be a useful perspective to take during notetaking instruction and practice. Such a view aligns with Dunkel's (1988) statement that "terseness of note taking (involving the recording of lecture propositions) rather than mere quantity seems to be an essential ingredient" (p. 269). Notetakers should strive to compact the most propositions possible with the fewest notations. This can be achieved, for example, by focusing on content words as opposed to grammatical words (e.g., Crawford, 2015) and ensuring that any symbols and punctuation used contribute as much as possible to meaning (although they may also be used structurally to distinguish one point from another). By drawing on both visual and verbal representations, notetakers can maximise the semiotic meaning-making tools they have available. While the present study has examined the value of notes in relation to listening comprehension test items, measurements of note quantity and quality can also arguably be insightful when used to correlate with writing task performance, course grades, and content learning more generally.

Pedagogic implications

In terms of pedagogic implications, teachers may wish to model notetaking that demonstrates the capturing of complete ideas rather than potentially random isolated words or symbols that in and of themselves may struggle to stimulate recall (Dunkel, 1988). Teachers can also engage in reflective 'think-aloud' procedures after notetaking in which they explain to students their reasons for writing notes on particular information in certain ways. These discussions can also include questions to students so that they can examine their own notes, consider current strategies, and determine if they can make any changes for more efficient and effective notetaking in the future (Sakurai, 2018; Siegel, 2021b). Students themselves also might focus their attention on writing complete units of meaning that are composed of multiple parts (e.g., more than one word and/or symbol), since these likely represent larger and more meaningful propositions beyond single notations.

Depending on student L2 listening and writing proficiency as well as more general academic skills (e.g., multitasking, handwriting speed, selective attention), teachers may wish to incorporate different developmental activities for encoding (i.e., while simultaneously listening and taking notes). As Badger, et al. (2001) point out, students also have different conceptualisations of what notetaking entails and why they do it. The various techniques, habits, and purposes students apply when taking notes should be factored in as well. Students may express certain preferences, such as digital or longhand notetaking, use of abbreviation, paraphrasing, and/or various overall systems for taking notes (e.g., Badger et al., 2001; Morehead et al., 2019; Song, 2012). Others report certain challenges depending on the language of input and the language they use to write notes (e.g., Airey & Linder, 2006).

Teachers can incorporate these views and preferences, encouraging students to refine strategies they are comfortable with and introducing new techniques with the intent of helping students capture more IUs and increasing the efficiency rate of their notes. For lower proficiency students, teachers may wish to consider notetaking with smaller stretches of input. Siegel (2019b) presents a teaching sequence that focuses on short bursts of listening and notetaking

followed by reflection, discussion with others, and revision of notes in order to recognize opportunities to take notes more efficiently during future opportunities. As students become more comfortable with efficient and terse notetaking, longer stretches of texts can be used to increase endurance and more closely replicate authentic notetaking conditions such as those in EMI. Skeleton notes that include some but not all information are another option for scaffolding notetaking (e.g., Sakurai, 2018). Once the notes have been taken, activities focusing on the storage benefit of notes (DiVesta & Gray, 1972) can be engaged; for example strategies such as comparing notes with classmates, copying, reorganizing, and/or adding to notes can prove beneficial (e.g., Badger et al., 2001; Chen, 2021).

Limitations

These results were probably influenced by the types of questions included on the comprehension tests. Only multiple choice and fill-in-the-blank questions were included. These closed questions limited the ways in which students could respond. Other types of test questions or prompts, such as those that require short, written answers or more extended production via summary or compare/contrast paragraphs or essays might mean that students would take and utilise notes differently. Nonetheless, multiple choice and gap fill questions are widely used to test listening comprehension on international proficiency tests (e.g., TOEFL, IELTS; Carrell, 2007) and likely in more localized contexts as well. In addition, students took the comprehension tests approximately 10 minutes after the TED talks finished. As such, working memory, instead of and/or in addition to the notes themselves, likely affected test scores, at least for some participants. Future research should extend the time between the notetaking act and the recall stimulus in order to increase reliance on notes and reduce the impact of working memory capacity, similar to the immediate and delayed test battery used by Kiewra and Fletcher (1984). Results were also likely impacted by variations in students' topic background knowledge, motivation for taking notes, and concentration levels. While no strict controls for these variables were used in the study, all participants had similar educational experiences up to this point and were studying at the same grade level at the same school.

Conclusion

This study aimed to examine the relationship between content written in notes and its usability on post-listening comprehension tests. To investigate this informal type of academic writing, three sets of notes and corresponding test scores were collected from 34 EAP students, generating a total of 102 note samples and test results. Notes were analysed for total notations, total words, and total IUs. The notation and IU tallies were then used to calculate an efficiency ratio. These various measures of note quantity and quality were compared with test results via correlation calculations. Results showed little relationship between notations and words and test scores. IUs and efficiency rates correlated more strongly with test scores, suggesting that increasing these measures in notes, as opposed to simply writing isolated notations and/or words, leads to higher comprehension test scores, at least for the types of test items included in the study. Recommendations based on these findings include a focus on terseness in notes (Dunkel, 1988) during notetaking instruction and practice, making each notation count by contributing to a larger propositional unit of meaning.

References

- Airey, J., & Linder, C. (2006). Language and the experience of learning physics in Sweden. *European Journal of Physics*, 27(3), 553-560. <https://doi.org/10.1088/0143-0807/27/3/009>
- Anderson, J. (2014). *Cognitive Psychology and its implications* (8th ed.). Worth.
- Anita, B., & Mafofo, L. (2021). Text annotations: Examining evidence for a multisemiotic instinct and the intertextuality of the sign in a database of pristine self-directed communication. In S. Makoni, D. Verity & A. Kaiper-Marquez (Eds.), *Integrational linguistics and philosophy of language in the Global South* (pp. 84-103). Routledge.
- Badger, R., White, G., Sutherland, P., & Haggis, T. (2001). Note perfect: an investigation of how students view taking notes in lectures. *System*, 29, 405-417. [https://doi.org/10.1016/S0346-251X\(01\)00028-8](https://doi.org/10.1016/S0346-251X(01)00028-8)
- Carrell, P. L. (2007). Notetaking strategies and their relationship to performance on listening and communicative assessment tasks. (TOEFL Monograph Series No. MS-35). Princeton. <http://www.ets.org/Media/Research/pdf/RR-07-01.pdf>
- Chen, P. H. (2021). In-class and after-class notetaking strategies. *Active Learning in Higher Education*, 22(3), 245-260. <https://doi.org/10.1177/1469787419893490>
- Clerehan, R. (1995). Taking it down: Notetaking practices of L1 and L2 students. *English for Specific Purposes*, 14(2), 137-155. [https://doi.org/10.1016/0889-4906\(95\)00003-A](https://doi.org/10.1016/0889-4906(95)00003-A)
- Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671-684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Crawford, C.C. (1925). The correlation between college lecture notes and quiz papers. *The Journal of Educational Research*, 12(4), 282-291. <https://doi.org/10.1080/00220671.1925.10879600>
- Crawford, M. (2015). A study on note taking in EFL listening instruction. In P. Clements, A. Krause, & H. Brown (Eds.), *JALT2014 Conference Proceedings*, 416-424.
- Crawford Camiciottoli, B. & Querol-Julián, M. (2016). Lectures. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 309-322). Routledge.
- DiVesta, F. & Gray, S. (1972). Listening and note taking. *Journal of Educational Psychology*, 63(1), 8-14. <https://doi.org/10.1037/h0032243>
- Dunkel, P. (1988). The content of L1 and L2 students' lecture notes and its relation to test performance. *TESOL Quarterly*, 22(2): 259-278. <https://doi.org/10.2307/3586936>
- Einstein, G., Morris, J., & Smith, S. (1985). Note-taking, individual differences, and memory for lecture information. *Journal of Educational Psychology*, 77, 522-532. <https://doi.org/10.1037/0022-0663.77.5.522>
- Halliday, M.A.K. (2014). *Introduction to Functional Grammar* (4th. ed). Routledge.
- Howe, M. (1970). Note-taking strategy, review, and long-term retention of verbal information. *Journal of Educational Research*, 63(6), 285.
- İpek, H. (2018) Perceptions of ELT students on their listening and note taking skills. *International Online Journal of Education and Teaching (IOJET)*, 5(1), 206-217.

- John-Steiner, V. (1997). *Notebooks of the mind: Explorations of thinking*. Oxford University Press.
- Kiewra, K. A., & Fletcher, H. J. (1984). The relationship between levels of note-taking and achievement. *Human Learning: Journal of Practical Research & Applications*, 3(4), 273–280.
- Kiewra, K. A., Benton, S. L., Kim, S. I., Risch, N., & Christensen, M. (1995). Effects of note-taking format and study technique on recall and relational performance. *Contemporary Educational Psychology*, 20(2), 172-187. <https://doi.org/10.1006/ceps.1995.1011>
- Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination, *Contemporary Educational Psychology*, 30, 242-262. <https://doi.org/10.1016/j.cedpsych.2004.10.001>
- Kobayashi, K. (2006). Conditional effects of interventions in notetaking procedures on learning: A meta-analysis. *Japanese Psychological Research*, 48(2), 109-114. <https://doi.org/10.1111/j.1468-5884.2006.00311.x>
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kusumoto, Y. (2019). EFL students' perception of note-taking and the effect of note-taking instruction. *The Kyushu Academic Society of English Language Education Bulletin*, 47, 47-56.
- Lasagabaster, D., & Doiz, A. (Eds.). (2021). *Language Use in English-medium Instruction at University: International perspectives on teacher practice*. Routledge.
- Luo, L., Kiewra, K., & Samuelson, L. (2016). Revising lecture notes: How revision, pauses, and partners affect note taking and achievement. *Instructional Sciences*, 44, 45-67. <https://doi.org/10.1007/s11251-016-9370-4>
- Luo, L., Kiewra, K. A., Flanigan, A. E., & Peteranetz, M. S. (2018). Laptop versus longhand note taking: effects on lecture notes and achievement. *Instructional Science*, 46(6), 947-971.
- Macaro, E., Curle, S., Pun, J., An, J. & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, 51(1), 36-76. <https://doi.org/10.1017/S0261444817000350>
- Morehead, K., Dunlosky, J., & Rawson, K. (2019). How much mightier is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer (2014). *Educational Psychology Review*, 31(3), 753–780.
- Morehead, K., Dunlosky, J., Rawson, K., Blasiman, R. & Hollis, R. (2019). Note-taking habits of 21st century college students: Implications for student learning, memory, and achievement. *Memory*, 27(6), 807-819. <https://doi.org/10.1080/09658211.2019.1569694>
- Murata, K. (2018). *English-medium instruction from an English as a lingua franca perspective*. Routledge.
- Pauk, W. (2001). *How to Study in College* (7th ed). Houghton Mifflin.
- Peeverly, S., Brobst, K., 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(2), 335-346. <https://doi.org/10.1037/0022-0663.95.2.335>

- Sakurai, S. (2018). Promoting skills and strategies of lecture listening and note-taking in L2. *Fukuoka University Journal of Humanities*, 49(4), 1019–1046.
- Siegel, J. (2018). Did you take "good" notes? On methods for evaluating student notetaking performance. *Journal of English for Academic Purposes*, 35, 85-92.
- Siegel, J. (2019a). Collaborative action research on notetaking: Simultaneous cycles. *The European Journal of Applied Linguistics and TESOL*, 8(2), 77-100.
- Siegel, J. (2019b). Notetaking in EFL: A focus on simplification. *The Language Teacher*, 43(3), 20-24.
- Siegel, J. (2021a). Evaluating EAP notetaking textbooks: Eight key questions. *Journal of English for Academic Purposes*, 50, 100952.
- Siegel, J. (2021b). *Developing Notetaking Skills in a Second Language: Insights from classroom research*. Routledge.
- Siegel, J. (2023). Translanguaging options for notetaking in EAP and implications in EMI. *ELT Journal*, 77(1), 42-51. <https://doi.org/10.1093/elt/ccac027>
- Skolverket. (2021). *Ämnesplan i engelska* [Course plan for the English subject].
- Song, M. Y. (2012). Note-taking quality and performance on an L2 academic listening test. *Language Testing*, 29(1), 67-89. <https://doi.org/10.1177/0265532211415379>
- Suritsky, S., & Hughes, C. (1991). Benefits of notetaking: Implications for secondary and postsecondary students with learning disabilities. *Learning Disability Quarterly*, 14(1), 7-18. <https://doi.org/10.2307/1510370>
- Wang, X. (2021). An analysis of note-taking strategies: The effect of translanguaging on content comprehension and knowledge retention. *Journal of Language Teaching*, 1(3), 1-20. <https://doi.org/10.54475/jlt.2021.020>