The Internet and Higher Education 55 (2022) 100873

Contents lists available at ScienceDirect

The Internet and Higher Education

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

Individual versus collaborative note-taking: Results of a quasi-experimental study on student note completeness, test performance, and academic writing

Matthew Courtney [a], Jamie Costley [b], Matthew Baldwin [c], Kyungmee Lee [d], Mik Fanguy [c,d,*]

[a] Graduate School of Education, Nazarbayev University, Nur-Sultan, Kazakhstan
[b] National Research University Higher School of Economics, Institute of Education, Moscow, Russia
[c] School of Digital Humanities and Computational Social Sciences, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea
[d] Lancaster University, The Department of Educational Research, County South, Lancaster, Lancashire LA1 4YD, United Kingdom

A R T I C L E I N F O

Keywords:
Note-taking
Collaborative learning
Collaboration
Korea
Academic writing

A B S T R A C T

There is research showing benefits to both collaboration and note-taking, but a lack of research into how they may both work together in an online context. More specifically, there is a gap in the research looking at how collaborative note-taking and individual note-taking can be compared when considering the quality of the notes taken, and how note-quality can impact student performance. The present study looks at the online note-taking behavior and performance of 186 graduate students studying at a Korean university. The results indicate that students who collaborate perform better than individual note-takers on measures of recall of course content, but that individual note-takers perform better on tasks focused on academic writing. Furthermore, the findings suggest that note-quality has no effect on collaborative note-takers' recall of course content, and a slight negative impact on their writing, while individual note-takers benefit from higher quality notes for both recall and writing.

1. Introduction

The practice of note-taking is widespread in higher education, where it is done to improve student learning and performance (Chen, 2019; Wu, 2020). Note-taking can be used informally as well as formally, with students taking notes from their own initiative, being suggested or encouraged to do so by the instructor, or as a compulsory part of the class (Grabe & Christopherson, 2005). The reason for note-taking's effectiveness is that it helps students to recall and engage with the content of the lectures (Salame & Thompson, 2020). This is because note-taking appears to not only improve student learning (van de Sande, Abramson, & Judson-Garcia, 2017), but also performance (Luo, Kiewra, Flanigan, & Peteranetz, 2018). More specifically, a traditional and still widely thought of benefit of note-taking is its application as an aid to students' ability to remember what was being taught during a lesson (Aiken, Thomas, & Shennum, 1975; Tindale & Winget, 2017). This process can act in two ways, with note-taking aiding the students' ability to transfer the information to their long-term memory through the writing process, as well as giving them a written document that they can refer to at a later stage (Di Vesta & Gray, 1972; Kiewra, 1985). The two processes are linked, though with an important distinction: the distinction between the processes engaged in (collaboration while creating notes) and the subsequent products created (the notes) (Petko, Schmid, Müller, & Hielscher, 2019).

Note-taking is not only considered to have benefits for the retention of information but has also been shown to drive student focus on the class contents they are studying (Kane et al., 2017), students' in-class achievement (van de Sande et al., 2017), and their depth of understanding of the course materials (Kiewra, 1987). It may be the case that despite the benefits of note-taking, it places additional mental strain on learners during the class (Chen, 2019). This is because students are attempting to understand the contents of a lecture whilst also writing it down (Shi, Yang, Yang, Liu, & Yang, 2020). For this reason, some studies found and there is some conjecture that taking notes in a group might lead to more benefits than taking notes independently (Chen, 2019; Harbin, 2020). There are two reasons for this: 1) the learners may

* Corresponding author at: KAIST, School of Digital Humanities and Computational Social Sciences, 291, Daehak-ro, Guseong-dong, Yuseong-gu, Daejeon, South Korea.

E-mail address: mik@kaist.edu (M. Fanguy).

https://doi.org/10.1016/j.iheduc.2022.100873
Received 21 December 2021; Received in revised form 13 June 2022; Accepted 27 June 2022
Available online 1 July 2022

interact regarding the contents of the notes, allowing them to understand the contents more thoroughly, and 2) as the burden of note-taking can be shared among group members, the cognitive strain of attempting to write while listening is reduced for each constituent member, which may enable students to grasp the classes' contents more comprehensively (Kirschner, Sweller, Kirschner, Zambrano, & J., 2018; Orndorff III, 2015). However, it should be noted that the original research by Chen (2019) and Harbin (2020) purporting the relative benefits rests on qualitative interviews and conjecture, respectively.

As many aspects of higher education have moved online, so has the practice of note-taking (Nakayama, Mutsuura, & Yamamoto, 2017), and this trend has been accelerated by emergency online education due to the pandemic (Lee et al., 2022). Recently, shareable online tools and platforms have enabled learners to collaborate when taking notes in a manner that allows greater retention of as well as engagement with the contents of a lesson (Steimle, Brdiczka, & Mühlhäuser, 2009). Within this context and because of the affordances of technology, there is an increased interest in collaborative note-taking online in higher education. However, this interest brings with it questions as to what advantages (if any) collaborative note-taking brings learners, as well as how aspects of note-taking lead to learning gains and increases in student performance (Adeniran, Masthoff, & Beacham, 2019).

2. Literature review

2.1. The relationship between the volume and completeness of notes

There is evidence that individuals who write more notes will include more of the information that they are attempting to cover (Chen, 2019). As with individuals, groups that add more words to a set of notes are likely to capture more relevant course content. More specifically, higher volume in note-taking tends to lead to the addition of a greater range of contents and concepts from the class that is being taught (Adeniran et al., 2019; Doberstein, Hecking, & Hoppe, 2019). Prior work has shown that collaborative note-takers are likely to notice and rectify incorrect or incomplete information in the notes taken by group members, thereby improving the accuracy and completeness of the notes (Kam et al., 2005). Furthermore, previous research into collaborative note-taking has shown a consistent positive relationship between the volume of words and the note-taking completeness (Fanguy, Baldwin, Shmeleva, Lee, & Costley, 2021a). Finally, the process of revising notes with peers and adding more information has been shown to lead to the creation of higher-quality, more complete notes (Luo, Kiewra, & Samuelson, 2016). Groups will likely take greater amounts of notes than individuals in sum because of many members working together (Luo et al., 2016). However, this may lead to each constituent member of the group producing fewer notes and, compared with individual note-taking counterparts, revising less content (Fanguy et al., 2021a). On the other hand, research has also suggested that learners are more motivated and engaged by working together, which may lead to each constituent member within a group matching or even surpassing individual note-takers' productivity, word-for-word (Kam et al., 2005).

2.2. Completeness of notes and retention of knowledge

The degree to which students write more voluminous and complete notes has been shown to impact the degree to which students retain information from a lesson (Haynes, McCarley, & Williams, 2015; Kiewra, 1987). Individual students as well as groups who have access to more complete notes perform better on measures of retention and attention to detail (Raver & Maydosz, 2010; Volet, Summers, & Thurman, 2009). This is because a higher-quality learning artifact leads to greater retention of contents and better learning performance (Butson & Thomson, 2014). Furthermore, research suggests that more complete notes reduce the cognitive burden students feel when trying to recall information (Hadwin, Kirby, & Woodhouse, 1999). The reason for this is learners can divide up the workload of writing notes and more closely concentrate on the class materials (Tindale & Winget, 2017). This means that students have to spend less mental effort (cognitive load) on writing the notes. This reduction of mental effort to produce notes can be transferred to using their mental effort to focus on learning the materials that are taught in the class by referring to their notes (Kirschner et al., 2018).

However, some research has shown that simple transcription of contents (which would be akin to high completeness) may lead to lower levels of learning in some cases as students are overly focused on recording notes and not on learning (Mueller & Oppenheimer, 2014). However, this is disputed by other research suggesting that cases where learners have access to their notes, which are complete transcriptions, lead to better performance (Bui, Myerson, & Hale, 2013). In terms of comparing group (or collaborative) note-takers to individual note-takers, there is evidence that collaborative note-takers retain more information from their notes than individual note-takers (Johnson, Johnson, & Smith, 2014; Orndorff III, 2015).

2.3. Completeness of notes and academic writing

It has been shown that students who have both produced and have access to more thorough notes produce academic writing of better quality than those who do not (Wilson, 2014). In addition, there is some evidence that when students generate high-quality notes, these notes will lead to better writing performance (Ju & Kim, 2020). Furthermore, students who take larger amounts of high-quality notes demonstrate improved performance in writing essays that require high-order application, analysis, and synthesis of content (Waite, Lindberg, Ernst, Bowman, & Levine, 2018). Benton, Kiewra, Whitfill, and Dennison (1993) found that note-taking volume and completeness were correlated with longer and more coherent and cohesive compare-and-contrast essays, provided that the notes were able to be referenced while writing the essay. Similarly, Slotte and Lonka (2001) found that any form of note-taking (i.e., underlining only, verbatim, summarizing, and concept mapping) was superior to *not* taking notes in terms of the effects on students' ability to write a coherent essay.

It has been suggested that the degree to which students provide more information to one another during collaborative processes, the better their academic writing will be (Fanguy et al., 2021a). Furthermore, there is evidence that more comprehensive group note-taking processes before writing may lead to higher quality student essays (Pospelova, 2021). However, previous research looking at the differences between the performance of (1) constituent collaborative note-takers and (2) individual note-takers has shown that individual note-takers perform better in their academic writing (Fanguy et al., 2021a). There are several reasons suggested for this effect: it may be that learners prefer to work alone in some contexts (Retnowati, Ayres, & Sweller, 2017), that they are negatively affected by relative cognitive transactional costs associated with collaboration (Kirschner, Paas, & Kirschner, 2009), by the discomfort of having their contributions edited/changed/deleted within a collaborative document (Blau & Caspi, 2009; Lund & Smørdal, 2006), or by a lack of writing practice (Fanguy et al., 2021a).

2.4. Comparing the learning outcomes of individual to collaborative note-takers

The technology enabling learners to take collaborative notes online came about as a result of advances in cloud computing in the mid-2000s, with the Google Docs service becoming fully available to the public in 2009 (Irshad & Johar, 2015). Because the practice of taking collaborative notes online is rather new, research on the efficacy of this strategy is still in its infancy, and the results have been mixed. Kam et al. (2005) introduced Livenotes software to enable students to take notes on an interactive whiteboard, and the researchers compared the amounts of notes taken and the subsequent quiz scores of individual and collaborative note-takers. Interestingly, the study results showed that individual note-takers took fewer notes than collaborative note-takers, but the study found no significant differences in the quiz scores of the two treatment groups. Orndorff III (2015) conducted an experimental study comparing the learning outcomes of individual and collaborative note-takers enrolled in social science courses at a university. In the study, collaborative note-takers, on average, earned grades nearly a full letter grade higher than their individual note-taking counterparts, although the means of assessment (exams, essays, etc.) varied among the classes that were analyzed. Another experimental study (Baldwin, Fanguy, & Costley, 2019) comparing the learning outcomes of collaborative note-takers and a control group who were asked to take notes individually found no difference in quiz scores or writing scores; however, the researchers noted that in this study the collaborative note-takers took very few notes during the semester, and the notes of the control group were not analyzed. Recently, a study by (Fanguy et al., 2021a) found that individual note-takers did better on measures of writing performance, while collaborative note-takers did better on measures of recall (i.e., quizzes). Altogether, differences in the amount of notes produced from collaborative note-taking versus individual note-taking are mixed. One possible explanation for the mixed findings is that the ultimate and more meaningful benefit of collaborative note-taking may not be increases in the amount of notes, but instead, increases in the completeness of notes.

2.5. The present study

While prior research on note-taking has shown benefits for both individuals and groups, it remains an open question as to what drives these positive effects. It could be the case that the act of writing notes itself (the processes engaged in) allows learners to retain more of the information that they transcribed, or that the document the learners produce (subsequent product created) serves as a type of scaffolding that allows learners to perform better. In this respect, it is important to look at how both the volume and the quality of the student notes that are produced interact with learner performance. Furthermore, it is important to distinguish the types of performance that could benefit from note-taking. The first is that high-quality (complete) individual or collaborative notes might lead to learners being able to retain more of the information embedded in the notes themselves. The second is that note quality might also drive performance on the application of skills learned in the class, like academic writing. The present study adds further nuance to our understanding of this relationship by asking how document quality in the form of completeness interacts with these relationships and helps to explain the effects of note-taking from both individual and group perspectives. Specifically, we ask the following three research questions:

RQ1: How do student volume, completeness, test performance, and writing differ across individual and collaborative conditions for each of the ten weeks?

RQ2: For the individual note-taking condition, what is the effect of volume on completeness, and completeness on student test and writing performance?

RQ3: For the collaborative group, what are the within-group effects for volume on test performance and writing, and the between-group effects for volume on completeness, and completeness on test performance and writing?

3. Methodology

3.1. Participants and learning context

This study examined 186 students taking notes online in 10 sections of a graduate scientific writing course at a university located in South Korea. Each course section consisted of 8–25 students, and all students were majoring in science or engineering. Among the 10 sections, 4 were designated as individual note-taking sections, and 6 were designated as collaborative note-taking sections. Each section (or tranche) experienced the exact same course content, though delivered at different times during the week due to timetabling. The collaborative note-taking condition comprised 27 groups of 3–5 members each. The groups were made up of 4 (9 groups) or 5 students (17 groups), with one group of 3 students (it was a group of 4, but a student dropped out). Smaller groups have been shown to increase learner-to-learner interaction in online contexts (Caspi, Gorsky, & Chajut, 2003). Furthermore, in research into collaborative note-taking, students have expressed a preference for groups of around the size used in the present study (Orndorff III, 2015). Among all participants, 128 were enrolled in a master's degree program, and 58 were enrolled in a doctoral program. There were 138 males and 48 females, and participants had an average age of 25.5 ($SD = 2.5$), with the youngest participant being 22 and the oldest being 36.

All students that were a part of the present study had the same instructional circumstances, except for the type of note-taking they were engaged in. They watched the same online videos, did the same activities, and took the same assessment. Students were separated into two conditions on account of their personal availability to the classes that were run at different times in the week: individual note-taking ($n = 64$) and collaborative note-taking ($n = 123$) groups. The treatment groups were similar in their compositions in terms of gender and age, and participants in the two groups did not differ significantly in terms of gender ($\chi^2(1) = 0.639$, $p = .424$), age ($t = -0.907$, $p = .365$), or academic writing proficiency level measured via a pre-test quiz ($t = 0.998$, $p = .319$). The pre-test quiz consisted of 10 items and was administered at the beginning of the semester. The pre-test assessed students' prior knowledge of topics covered during the 10 weeks of course instruction. Having determined that the individual and collaborative groups are generally equivalent in terms of gender, age, and academic writing ability, an examination of the differences in online collaborative note-taking behavior between these groups was justified.

In the graduate scientific writing course that was examined in the present study, students learned how to compose a manuscript for publication in an academic journal in their field. Lectures for the course were uploaded on the course learning management system in streaming video format (Fanguy, Lee, & Churchill, 2021b). The course consisted of 10 instructional weeks, each consisting of 4–8 videos, totaling 56 videos for the semester. The duration of these videos averaged about 12 min, within a range of 4:56 and 24:50.

In each week of instruction, participants were asked to take notes on the contents of the instructional videos on the learning management system. Learners in the individual note-taking group were asked to take notes individually, and those in the collaborative group were asked to do so in small self-selected groups of 3 to 5 students. The two treatment groups each took notes in Google Docs that were created and monitored by the course instructor. Accordingly, participants in the individual note-taking condition each took notes in 10 Google Docs corresponding to 10 weeks of instruction. Likewise, participants in the collaborative note-taking condition took notes in small groups in 10 shared Google Docs corresponding to 10 weeks of course instruction. In order to motivate students to participate in note-taking, 10% of the course grade was given to students who added notes to the weekly note-taking documents in both the individual and collaborative note-taking conditions, with students receiving a weekly score of "0" or "1" based on whether they contributed any amount of notes for each instructional week. Near the end of a given week of instruction, students were required to take an online quiz testing their knowledge of the content conveyed in the lecture videos of that week. Quiz items covered a variety of concepts from the course lectures including ethical issues related to scientific research and communication, the submission and peer review processes when publishing a manuscript in an academic journal, and academic writing conventions including style and grammar. The Google Docs that students used when taking notes were later data-mined for the number of words that were written and were also scored using a rubric to measure their completeness, i.e., the extent to which the notes represented relevant concepts from the course lectures.

3.2. Measures

3.2.1. Volume

The word counts contributed by each participant to the final version of each of the 10 sets of notes taken during the course were used as the volume variable in the present study. In the individual note-taking condition, the total word count contributed in each note-taking document served as the volume variable. Likewise, in the collaborative condition, the total word count contributed in *each* note-taking document served as the volume variable. The total number of words contributed to each Google Doc was counted using a customized computer program written in Python language (Fanguy & Chang, 2021).

3.2.2. Completeness

Completeness is a variable that measures the extent to which the student notes accurately represent meaningful units of information presented in the weekly video lectures. Each set of notes produced by each group was tripled rated, with one rating from three different teaching assistants (TAs) of the course using a rubric created by the course instructor and tailored to the concepts covered in the lecture videos for each instructional week. The TAs used the rubric to assess whether informational units from each lecture video were included or not included in the students' notes. In order to improve the reliability of the completeness grading done by the TAs, the course instructor held ten weekly norming sessions on Zoom with a total of five course TAs. These norming meetings lasted from 60 to 90 min each. During the meetings, the instructor explained the rubrics to the TAs, and then rated one sample set of notes together with the TAs, before discussing the scores given for each item on the rubric. Next, TAs were required to rate another sample set of notes on their own and then reassembled with one another and the instructor to discuss discrepancies in their scores. After this was done, the TAs were given three sample sets of notes to score on their own, and once again gathered back together with one another and the instructor to discuss discrepancies in scoring. Then, each of the five TAs was asked to score a proportion of the total number of notes students produced for a given instructional week. The documents were assigned to allow for each document to be rated by three different TAs.

The maximum possible score for each week varied according to how many instructional units were included in the lecture videos of each instructional week. Note-taking completeness had generally high alpha coefficients at = 0.88 (100 items, *I*), 0.95 (84 *I*), 0.86 (63 *I*), 0.94 (77 *I*), 0.80 (35 *I*), 0.84 (53 *I*), 0.87 (41 *I*), 0.88 (51 *I*), 0.97 (168 *I*), and 0.98 (258 *I*), respectively, with item-total(rest) correlations which were positive for all informational units for all 10 assessments (see CTT R package's reliability function; Willse, 2018). All informational units used in all rubrics can be viewed in the file labeled "Completeness Rubric" at https://osf.io/5t8vw/?view_only=3514f73b64b1497a9948e1a544d565bc

3.2.3. Quiz scores

Ten quizzes were administered in order to test students' recall and comprehension of concepts from the online lecture videos during the ten instructional weeks of the course. Each of the quizzes comprised 8–30 multiple-choice items testing students on their knowledge of the topics covered in the lecture videos in the corresponding week of instruction. Students were allowed only one attempt per quiz, and quizzes were timed so that students had two minutes to answer each item. Students were allowed to take the quiz at any time during the instructional week, but a final deadline of 6 pm on the Friday of each instructional week was imposed. Students were permitted to select multiple correct answer options for each item and were given partial credit for instances in which fewer than the total number of correct answer options were chosen. In cases where an incorrect answer choice was selected, students received a score of 0 for the quiz item to disincentive indiscriminate guessing when students did not comprehend the concepts being assessed. The scores for each quiz were equally weighted so that each quiz (test) accounted for 3% of the total course grade. As there were a total of 10 quizzes during the course, quizzes accounted for 30% of the total grade points for the course. The weekly quiz scores had Cronbach's alpha coefficients (number of items in brackets) of α= 0.67 (13 I), 0.63 (10 I), 0.63 (15 I), 0.68 (8 I), 0.80 (20 I), 0.65 (12 I), 0.78 (18 I), 0.59 (11 I), 0.64 (13 I), and 0.86 (30 I), respectively, with all items exhibiting positive item-rest correlations. These coefficients and associated item properties indicate that the quizzes provided a moderately reliable measure of the concepts taught in each week of instruction. For more information about the quiz items and their relationship to the instructional content of the course, refer to the following URL under the tab labeled "Quiz Items and Video List": https://osf.io/5t8vw/?view_only=3514f73b64b1497a9948e1a544d565bc

3.2.4. Individual writing assignments

Students enrolled in the scientific writing course were required to submit five individual writing assignments that corresponded to the five major components of a journal manuscript: Introduction, Methodology, Results, Discussion & Conclusion, and Abstract. Every writing assignment was rated twice, once by each of two raters who scored each submission on a scale of 1–10 using rubrics that were adapted from those proposed by Clabough and Clabough (2016). In cases of moderate divergence between scores given by the two raters, a third rater was asked to review the assignment and provide a final score. To ensure reliability of scoring, a norming session was held among the raters, separately scoring 10 randomly-selected writing assignments. In the norming session, the raters discussed instances where scoring diverged, and scoring was calibrated acceptably. The two course instructors simultaneously evaluated 20% of all the writing assignments submitted in all six of the course sections of the course. Each writing assignment accounted for 10% of the course grade point total, and the five assignments collectively accounted for 50% of the course grade point total. There was an additional individual assignment worth 10% of the course point total that was not included in the study. The assignment was to construct a References section at the end of the manuscript, and this task was not deemed to involve actual research writing skills. Previous to this experiment, the majority of the TAs who graded and the instructors who regraded the assignments had been working together for multiple semesters. After the process of rater training was completed a subset of 130 documents were double rated, and inter-rater Kappa was found to be 0.862.

3.3. Statistical methods

To answer RQ1 concerning a comparison of volume, completeness, performance, and writing for control and experimental groups, a series of Levene's tests (Fox & Weisberg, 2019) for homogeneity of variance (center = median) was employed prior to undertaking paired two-sample Wilcoxon tests (R Core Team, 2021). Effect sizes were assessed by way of r with the following interpretation: $r < 0.10$ (negligible), $0.10 \leq r > 0.30$ (small), $0.30 \leq r > 0.50$ (medium), and $0.50 \leq r$ (large) (Kassambara, 2021). Box plots were used to illustrate the data (Kassambara, 2020). Finally, to compare test performance by group size, an examination of differences in means between students in groups of 4 and students in groups of 5 was made by way of 10 independent sample t-tests (for each week) with alpha set at 0.05.

To answer RQ2 concerning the effects each week in the individual note-taking condition, ten single-level models (Fig. 1) were run.

Prior to answering RQ3, an examination of the variance components of volume, performance, and writing attributable to between-group, between-person, and within-person (weekly) effects was undertaken.

This descriptive examination involved two steps. First, for this

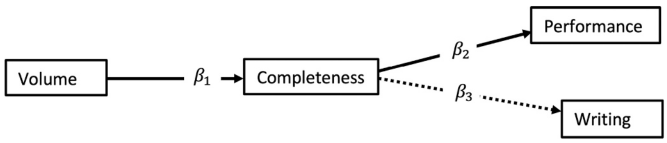

Fig. 1. Single level temporal model for volume, completeness, performance, and writing for the control group (individuals).
Note. Coefficient β_3, (represented by the dotted line) is also tested for weeks 2, 4, 6, 8, and 9. $\beta1$ and $\beta3$ were also tested for weeks 1, 3, 5, 7, and 10.

analysis, three main null models were specified. Specifically, intra-class correlations coefficients (ICCs) for volume and performance involved nested data structures with 10 data points for each week nested in students, and three to five students nested in 24 different student groups. ICCs for writing involved an examination of variance for the same data structure, though with *five* datapoints for weeks 2, 4, 6, 8, and 9, and three to five students nested in 24 different student groups. Second, the proportion of variance in volume, performance, and writing attributable to between-group effects for each week was also undertaken. Here 10 null models were specified for each week (experimental group) for all three outcomes. All null models were specified with the assistance of the R lme4 package (Bates, Maechler, Bolker, & Walker, 2015). The null models provide a broader understanding of the variance components of the variables of interest, and the results can be compared to studies with similar designs.

Thereafter, to answer RQ3 specifically, ten linear mixed-effects models were run (see Fig. 2). For the between-group model component, variables were aggregated up by group.

All temporal models (Figs. 1 and 2) were run with the assistance of the R lavaan package (Rosseel, 2012). Effect sizes for outcomes in all temporal models were interpreted in accordance with $f^2 = R^2/(1\text{-}R^2)$, with the following interpretation: $f^2 < 0.02$ (negligible), $0.02 \leq f^2 > 0.15$ (small), $0.15 \leq f^2 > 0.35$ (moderate), and $0.35 \leq f^2$ (large) (Cohen, 1992). For the temporal models, levels of statistical significance for effects were interpreted as follows: $^{\dagger}p < .10$ (of some interest), $^{*}p < .05$, $^{**}p < .01$, and $^{***}p < .001$.

All data preparation and analysis were undertaken with the assistance of the open-source R programming software (R Core Team, 2021). Prior to carrying out the single- and multi-level path temporal models (Goldstein, 2003), all variables were standardized. This was done to achieve scale comparability across variables.

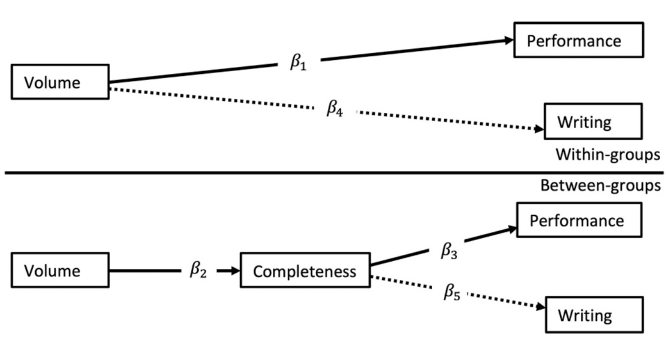

Fig. 2. Multilevel temporal model for volume, completeness, performance, and writing for the collaborative group (individuals in groups).
Note. Coefficients β_1, β_2, and β_3 (represented by the solid lines) are tested for weeks 1, 3, 5, 7, and 10; coefficients β_4, and β_5 (see dotted lines) are also tested for weeks 2, 4, 6, 8, and 9.

4. Results

4.1. RQ1: volume, completeness, performance, and writing for individual and collaborative conditions

Figs. 3, 4, 5, and 6 provide an illustration of the differences in volume, completeness, performance, and writing for the individual and group conditions for each of the 10 weeks.

Results suggest that, in terms of volume of words, on average, students in the individual group produced more words than constituent students in the collaborative group with effect sizes moderate (5 instances) to large (5 instances). However, in terms of note-taking completeness, students in the collaborative group outperformed students in the individual condition with effect sizes small (1), moderate (4), and large (5). In terms of test performance, the collaborative group outperformed the individual group for 7 of the 10 weeks, though these effects were small. Finally, in terms of writing performance, the individual group outperformed the collaborative group for four of the five weeks with effects small (2) and moderate (2).

4.2. RQ2: volume on completeness, and completeness on performance and writing (individual group)

Table 1 presents the results for RQ2 concerning the temporal model for effects in the individual group.

Results suggest that volume had a predominantly large and statistically significant effect on completeness for the instructional period. In addition, completeness tended to have a small to moderate statistically significant effect on performance for six of the ten weeks of instruction. Finally, completeness had a small to moderate statistically significant effect on writing performance for two of the five weeks for which the effect was examined.

4.3. RQ3: For the experimental group, what are within- and between-group effects each week in the proposed model?

As explained in the methodology, prior to examining RQ3, an assessment of the overall within-person, between-person, and between-group variation was explored.

Of the total variance in student volume of words, 66.4% can be attributed to within-person (weekly) effects, 33.6% to between-person effects, and 0.00% to between-group effects. Of the total variance in student test performance, 82.4% can be attributed to within-person (weekly) effects, 16.3% to between-person effects, and 1.4% to between-group effects. Of the total variance in writing performance, 69.3% can be attributable to within-person (weekly) effects, 25.2% to between-person effects, and 5.5% to between-group effects.

As explained, an examination of the within- and between-group variance was also explored for each of the 10 weeks. Table 2 presents the two-level ICCs for volume, performance, and writing across the duration of the course.

Results suggest that the proportion of variance in volume attributable to group effects ranged between 0 and 6.9%. In addition, the proportion of variance in performance due to group effects ranged between 0 and 20.1%, while the proportion of variance in writing performance ranged between 0 and 16.6% across the instructional period.

Table 3 presents the results of the temporal effects for the proposed model for the experimental group (see Fig. 2).

Results suggest that, within-groups, volume only had a statistically significant effect on performance for Week 3, and volume had no statistically significant effect on writing for any of the five weeks that this effect was tested. In terms of between-group effects, volume tended to have a moderate to large and statistically significant effect on completeness for all ten weeks. Further, although completeness tended to have a small effect on performance, these effects did not reach statistical significance. Finally, between-groups, completeness tended to

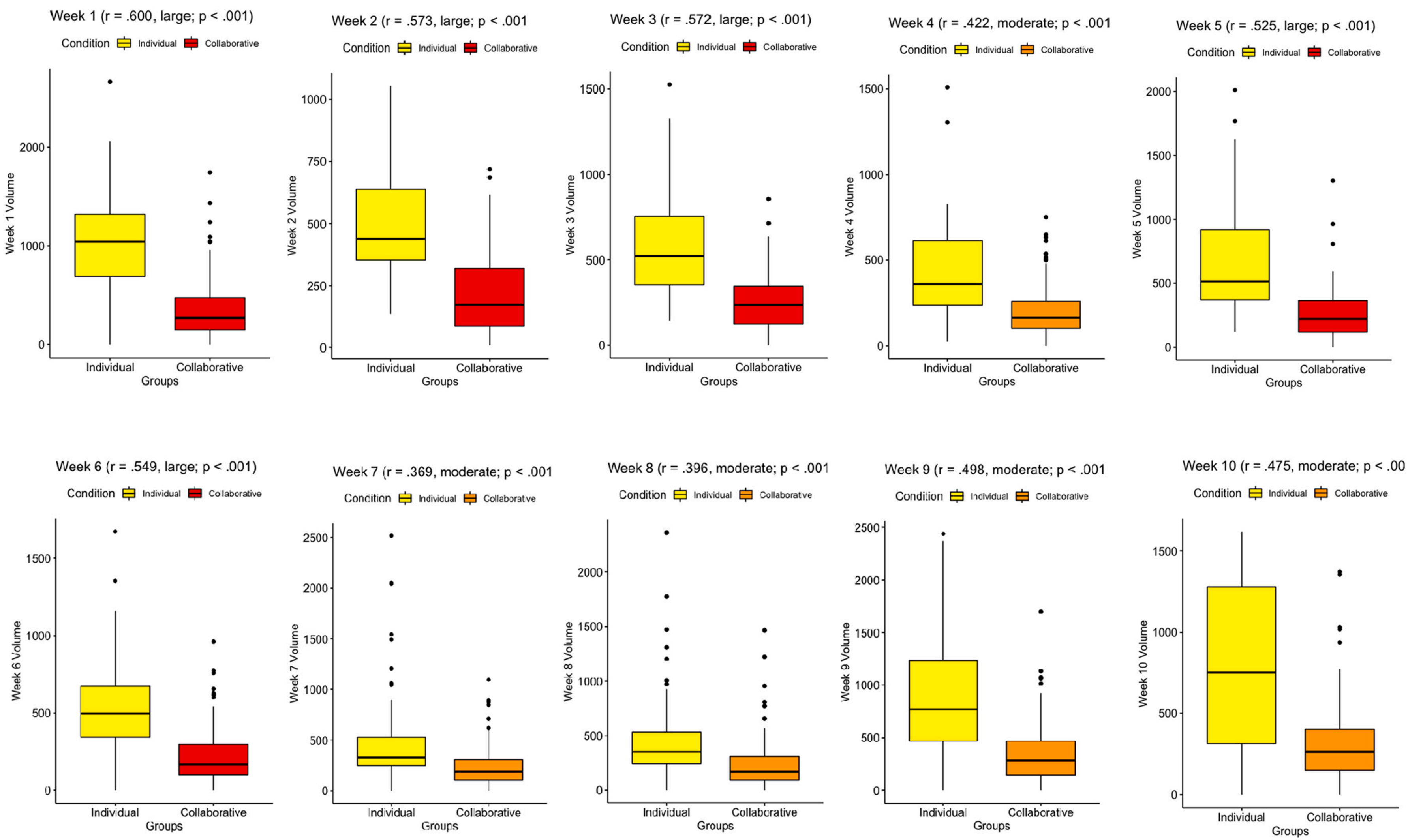

Fig. 3. Box plots of weekly volume by individual and collaborative conditions.

Note. For the interpretation of the r effect size, small ($0.10 \leq r < 0.30$) = yellow orange, moderate ($0.30 \leq r < 0.50$) = orange, large ($0.50 \leq r$) = red; Levene's tests indicated unequal variances: W1 $F = 18.56$ ($p < .001$); W2 $F = 7.65$ ($p < .01$); W3 $F = 30.09$ ($p < .001$); W4 $F = 15.25$ ($p < .001$); W5 $F = 29.85$ ($p < .001$); W6 $F = 8.85$ ($p < .01$); W7 $F = 9.88$ ($p < .01$); W8 $F = 6.20$ ($p < .05$); W9 $F = 36.84$ ($p < .001$); W10 $F = 78.38$($p < .001$).

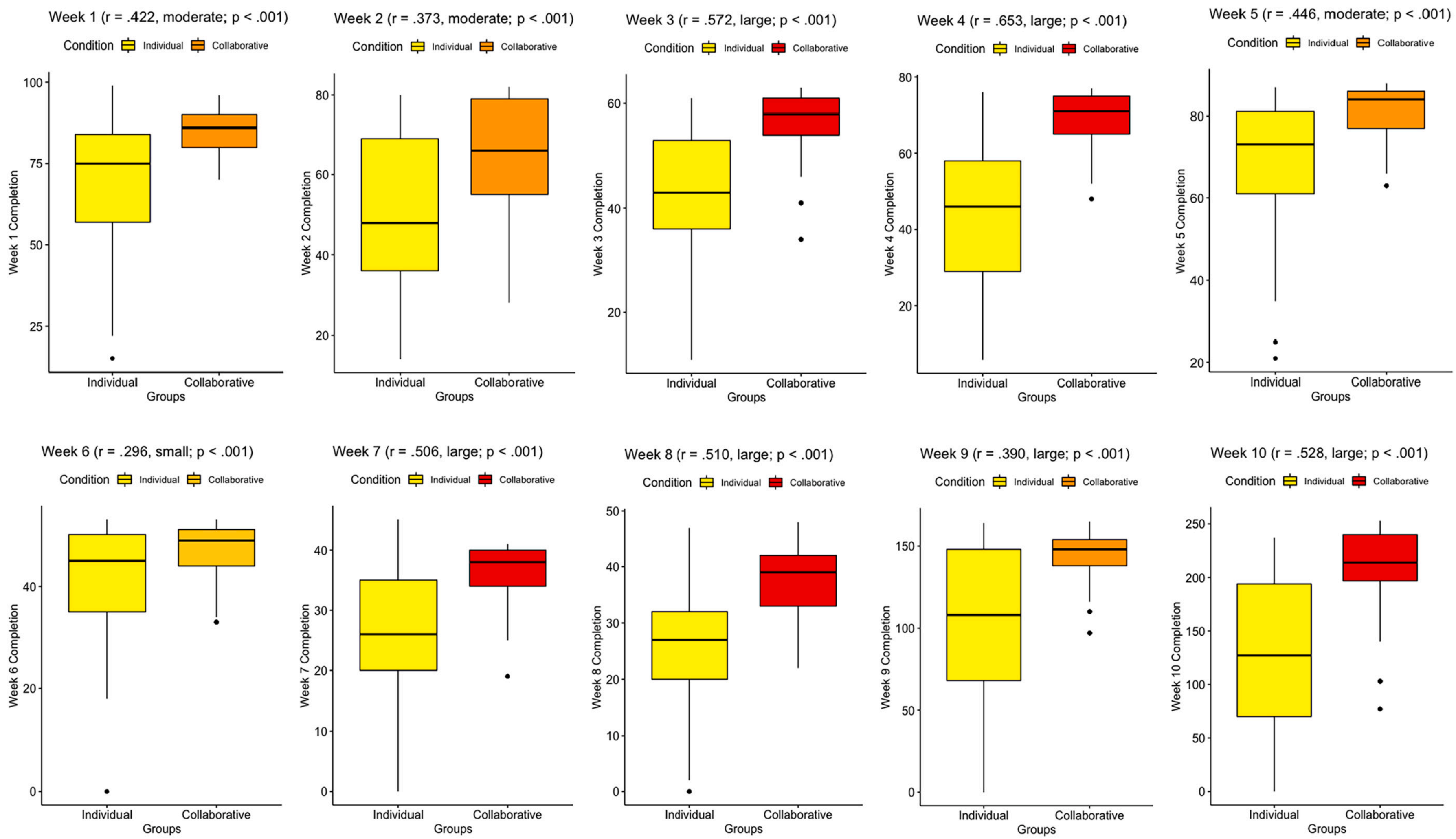

Fig. 4. Box plots of weekly completeness by individual and collaborative conditions.

Note. For the interpretation of the r effect size, small ($0.10 \leq r < 0.30$) = yellow orange, moderate ($0.30 \leq r < 0.50$) = orange, large ($0.50 \leq r$) = red; Levene's tests indicated unequal variances: W1 $F = 50.51$ ($p < .001$); W2 $F = 12.60$ ($p < .001$); W3 $F = 34.47$ ($p < .001$); W4 $F = 51.81$ ($p < .001$); W5 $F = 37.24$ ($p < .001$); W6 $F = 23.45$ ($p < .01$); W7 $F = 43.45$ ($p < .001$); W8 $F = 3.85$ ($p > .05$); W9 $F = 112.62$ ($p < .001$); W10 $F = 40.44$ ($p < .001$).

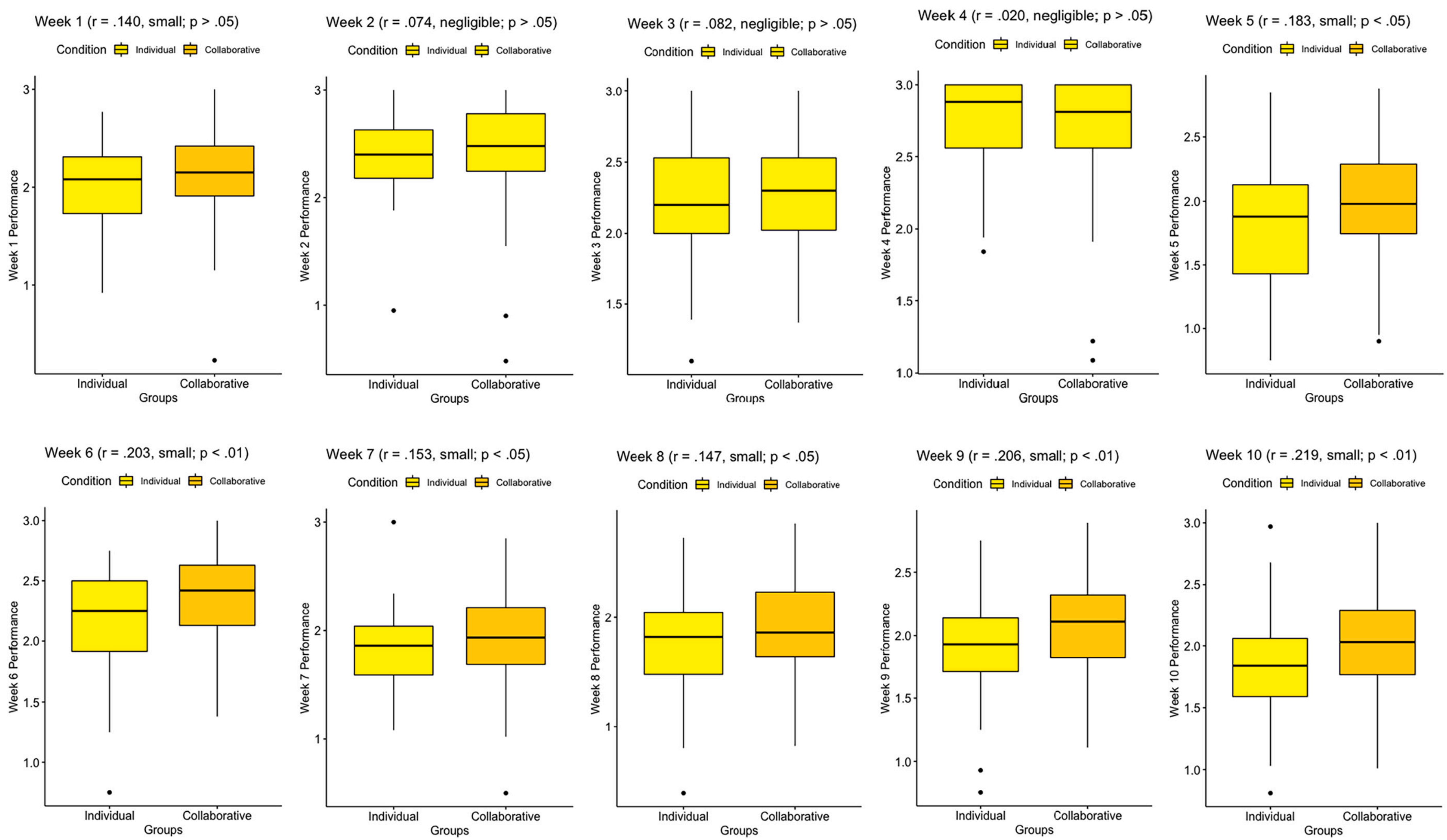

Fig. 5. Box plots of weekly test performance by individual and collaborative conditions.
Note. For the interpretation of the r effect size, small ($0.10 \leq r < 0.30$) = yellow orange, moderate ($0.30 \leq r < 0.50$) = orange, large ($0.50 \leq r$) = red; Levene's tests indicated unequal variances: W1 $F = 50.51$ ($p < .001$); W2 $F = 12.60$ ($p < .001$); W3 $F = 34.47$ ($p < .001$); W4 $F = 51.81$ ($p < .001$); W5 $F = 37.24$ ($p < .001$); W6 $F = 23.45$ ($p < .01$); W7 $F = 43.45$ ($p < .001$); W8 $F = 3.85$ ($p > .05$); W9 $F = 112.62$ ($p < .001$); W10 $F = 40.44$ ($p < .001$); note that the test performance of students in groups of 4 was equivalent to that of students in groups of 5 for 9 of the ten total weeks (for week 3, $t[68.17] = -2.59$, $p = .012$ (groups of 4 mean = 2.16, groups of 5 = 2.35) as evidenced by independent sample t-tests with alpha set at 0.05.

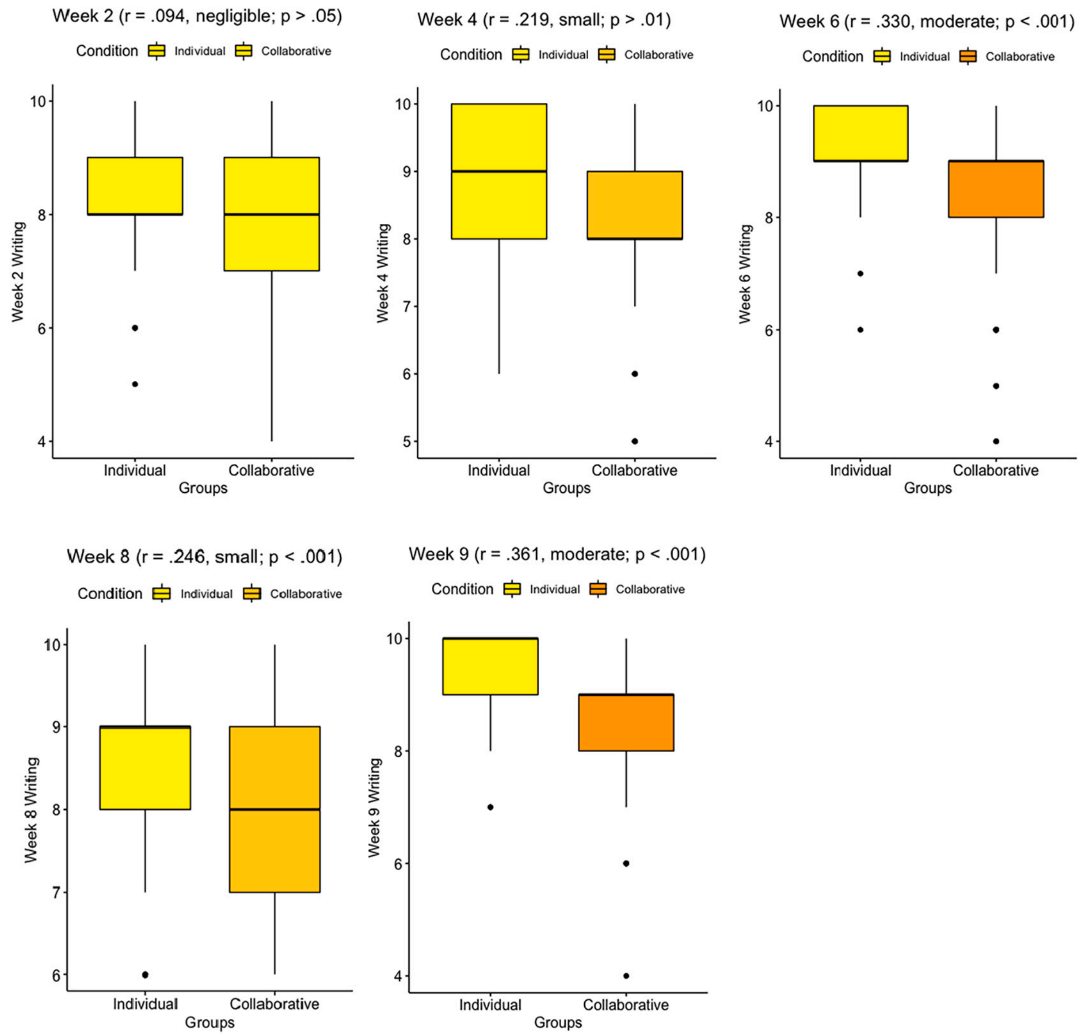

Fig. 6. Box plots of weekly writing by individual and collaborative conditions.
Note. For the interpretation of the *r* effect size, small ($0.10 \le r < 0.30$) = yellow orange, moderate ($0.30 \le r < 0.50$) = orange, large ($0.50 \le r$) = red; Levene's tests indicated unequal variances: W1 $F = 0.110$ ($p > .05$); W2 $F = 3.65$ ($p > .05$); W3 $F = 0.01$ ($p > .05$); W4 $F = 0.11$ ($p > .05$); W5 $F = 0.97$ ($p > .05$); W6 $F = 0.54$ ($p > .05$); W7 $F = 3.80$ ($p > .05$); W8 $F = 1.60$ ($p > .05$); W9 $F = 0.04$ ($p > .05$); W10 $F = 0.25$ ($p > .05$).

Table 1
Effects for temporal models for volume, completeness, test performance, and writing for individual group.

IVs	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
Effect on completeness [β_1]										
Volume of words	0.815***	0.868***	0.782***	0.737***	0.662***	0.655***	0.455**	0.696***	0.833***	0.887***
$R^2(f^2)$	0.664 (**1.976**)	0.753 (**3.049**)	0.611 (**1.571**)	0.543 (**1.188**)	0.438 (**0.779**)	0.429 (**0.751**)	0.207 (0.**261**)	0.485 (**0.942**)	0.694 (**2.268**)	0.787 (**3.695**)
Effect on test performance [β_2]										
Completeness	0.425***	0.353**	0.353***	0.034	0.232†	0.035	0.152	0.294*	0.299*	0.234*
$R^2(f^2)$	0.181 (0.**221**)	0.125 (0.143)	0.125 (0.143)	0.001 (0.000)	0.054 (0.057)	0.001 (0.000)	0.023 (0.024)	0.086 (0.094)	0.089 (0.098)	0.055 (0.058)
Effect on writing [β_3]										
Completeness	–	0.124	–	**0.405*****	–	0.160	–	0.241*	0.221†	–
$R^2(f^2)$	–	0.015 (0.015)	–	0.164 (**0.196**)	–	0.025 (0.026)	–	0.058 (0.062)	0.049 (0.052)	–

Note. *Control n = 65;* R^2 = total variance explained in outcome variables; $f^2 = R^2/(1-R^2)$ with small (0.02, underlined), medium (0.15, **bold**), and large (0.35, **bold and underlined**); $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$ in **bold**; $^{***}p < .001$ **bold and underlined**; all values represent standardized beta coefficients; writing tasks: Introduction (Week 2), Method (Week 4), Results (Week 6), Discussion (Week 8), Abstract (Week 9).

Table 2
ICCs for volume, performance, and writing for 10 weeks.

Metric	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10
Volume	0.021	0.025	0.069	0.000	0.000	0.000	0.026	0.016	0.000	0.000
Performance	0.029	0.000	0.201	0.014	0.000	0.127	0.000	0.098	0.054	0.089
Writing	–	0.106	–	0.166	–	0.030	–	0.000	0.127	–

Note. Experimental $n = 123$; all estimates represent two-level intra-class correlations.

Table 3
Effects for weekly temporal models for volume, completeness, test performance, and writing for collaborative group.

IVs	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
Effect on test performance [β_1,within-group effects]										
Volume of words	0.012	0.207†	0.237*	0.121	0.181†	0.062	0.105	0.000	0.105	0.102
$R^2(f^2)$	0.000 (0.000)	0.043 (0.045)	0.056 (0.059)	0.015 (0.015)	0.033 (0.034)	0.004 (0.004)	0.011 (0.011)	0.000 (0.000)	0.011 (0.011)	0.010 (0.010)
Effect on writing [β_4,within-group effects]										
Volume of words	–	0.021	–	−0.117	–	0.023	–	0.035	−0.089	–
$R^2(f^2)$	–	0.000 (0.000)	–	0.014 (0.014)	–	0.001 (0.001)	–	0.001 (0.001)	0.008 (0.008)	–
Effect on note completeness [β_2,between-group effects]										
Volume of words	**0.569***	**0.536****	**0.529****	**0.572****	**0.495****	**0.475****	**0.529****	**0.687***	**0.711***	**0.769***
$R^2(f^2)$	0.323 (**0.477**)	0.287 (**0.468**)	0.280 (**0.389**)	0.327 (**0.486**)	0.245 (**0.325**)	0.226 (**0.292**)	0.280 (**0.389**)	0.472 (**0.894**)	0.505 (**1.020**)	0.591 (**1.445**)
Effect on test performance [β_3,between-group effects]										
Completeness	0.065	0.177	0.317†	0.164	−0.097	0.218	0.085	−0.031	0.272	0.149
$R^2(f^2)$	0.004 (0.004)	0.031 (0.032)	0.100 (0.111)	0.027 (0.028)	0.009 (0.009)	0.048 (0.050)	0.007 (0.007)	0.001 (0.001)	0.074 (0.080)	0.022 (0.022)
Effect on writing [β_5, between-group effects]										
Completeness	–	0.227	–	−0.345*	–	−0.299*	–	−0.146	−0.094	–
$R^2(f^2)$	–	0.052 (0.055)	–	0.119 (0.135)	–	0.089 (0.098)	–	0.021 (0.021)	0.009 (0.009)	–

Note. *Experimental n = 110,* number of groups = 24*;* R^2 = total variance explained in outcome variables; $f^2 = R^2/(1-R^2)$ with small (0.02, underlined), moderate (0.15, **bold**), and large (0.35, **bold and underlined**); $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$ in **bold**; $^{***}p < .001$ **bold and underlined**; otherwise stated, all values represent standardized beta coefficients; writing tasks: Introduction (Week 2), Method (Week 4), Results (Week 6), Discussion (Week 8), Abstract (Week 9).

have a small negative and statistically significant effect on writing midway through the program.

5. Discussion

The present study investigates the effects of students' note-taking activities (supported by web-based writing applications, specifically, Google Docs) on their learning. To establish a comprehensive and in-depth understanding of the effects of different types of note-taking activities, this study compared individual and collaborative note-taking and how the completeness of students' notes is associated with student performances on quizzes and academic writing tasks. First, it was noted that individual note-takers wrote a higher number of words on average than constituent collaborative note-takers who were able to lean on the contribution of groupmates. Though, very different trends were observed when looking at completeness and student performance by condition. The students who were working collaboratively consistently generated higher levels of note completeness than those working individually as well as tending to perform better in weekly course quizzes.

Constituent note-takers writing with the support of multiple groupmates need not produce a higher volume of words. Conversely, as the individual note-takers had more ground to cover by themselves, they had higher volume when compared to each member of a collaborative group. In a similar vein, in the present study, collectively, collaborative note-takers exhibited a higher level of completeness than the individual note-takers. Again, as with volume, this is somewhat intuitive: the more hands that touch the document, the more course content the notes will likely cover. This effect has also been discussed in extant literature, which suggests that one of the primary benefits of collaborative note-taking is that it leads to a more complete higher-quality learning artifact (Luo et al., 2016). However, the present study builds on previous research comparing individual and collaborative note-takers' note-quality as well as the relationship between volume and completeness by looking directly at the relationships and comparing them. Previous studies have not looked at these relationships directly and have hinted at the possible effects. The present study shows that these relationships are more complex than previously suggested, but that collaborative note-takers will consistently have higher completeness and that this completeness is driven by many hands producing a higher level of volume than students who take notes individually.

However, overall, there were clear differences between individual and collaborative note-takers when it came to quiz performance. Across nine of the ten weeks, the collaborative note-takers had higher scores on quizzes, with seven of those weeks showing a statistically significant difference (see Fig. 5). This falls in line with extant literature that claims collaboration in general should help with the retention of information, as collaboration leads to more rigorous engagement with the contents and a higher quality learning artifact (Nokes-Malach, Richey, & Gadgil, 2015). Furthermore, with regard to note-taking specifically, previous research suggests that those who collaborate when taking notes will retain more information and perform better on quizzes (Johnson et al., 2014; Orndorff III, 2015). Another possible explanation for the advantages of collaborative note-taking is that the submitted notes are less likely to be written verbatim when students edit and revise the notes of other students and that the likelihood that one or more collaborators who write notes in their own words increases as more collaborators are added to the group. Rewriting information in students' own words has been shown to potentially lead to greater learning performance when note-taking (Aragón-Mendizábal et al., 2016).

Interestingly, while the collaborative note-takers performed better on quizzes than the individual note-takers, the relationship between completeness and quiz scores told a more complex story for the two experimental conditions. For the individual note-taking condition, there was a consistent positive association between completeness and quiz scores across all ten weeks, with six of the ten weeks having a positive correlation between completeness and quiz scores. This falls in line with current research into individual note-taking, which claims that the more of the class's contents the students write down, the more of that information they will retain (Haynes et al., 2015; Kiewra, 1987). However, while previous research suggests that collaboration might drive performance differences, we looked directly at the benefits of collaboration broadly in that we have shown a somewhat consistent retention of information advantage of group note-takers over individual note-takers. Furthermore, more specifically, completeness has a consistent positive effect on test performance for both individual and collaborative note-takers. We suggest that the advantages in test performance may be driven by the higher levels of completeness that are found among collaborative note-takers.

While collaborative note-takers exhibited higher levels of note completeness and quiz performance than individual note-takers, the relationship between completeness and quizzes for the collaborative note-takers is lacking. There were no statistically significant positive relationships between completeness and quiz scores for any of the ten weeks. This contrasts with previous research looking at data from the same project as the present study which suggests that completeness does affect performance (Costley, Courtney, & Fanguy, 2022), as well as previous studies suggesting that higher quality notes will allow learners to retain more information and perform better on quizzes (Tindale & Winget, 2017). A possible explanation for this finding is that there may be a saturation point of completeness at which further improvement yields diminishing returns for students' recall performance. As shown previously in Fig. 4, on average, collaborative notes were more complete than individual notes in every instructional week in the study. Moreover, variability of completeness among individual note-takers was relatively higher than that of collaborative note-taking groups. Therefore, it is possible that the recall performances of individual note-takers were distinguished by the completeness of the notes they created, as there was a substantial number of individual note-takers who took relatively incomplete notes; on the other hand, such a distinction did not exist among collaborative note-taking groups, as they tended to take highly complete notes throughout the study. Finally, the discrepancy may also be the result of the different modelling approach employed for each condition. Further research is needed to tease these effects out.

Individual note-takers had higher writing scores for all five of the compared measurements in the present study, with three of those differences being statistically significant. This difference has been noted in previous research looking at the whole sample used here (Fanguy et al., 2021a), with the suggested explanation being that individual note-takers get more practice writing (they write more total volume individually) than collaborative note-takers, which leads to higher writing scores for individual note-takers. This explanation seems plausible in the present study as the differences between the writing scores of individual and collaborative note-takers showed a tendency to increase from week to week during the semester, suggesting that the effect builds over time, as would be expected if it were due to practice. However, one problem with this explanation is that the content and style of writing differ considerably between lecture notes and academic writing assignments, so it is unclear that increased "practice" with writing notes would yield improvements in academic writing. Nevertheless, a number of studies have found that individual learners who take a higher volume of notes tend to produce higher quality academic writing on subsequent assignments (Benton et al., 1993; Ju & Kim, 2020; Waite et al., 2018; Wilson, 2014), whether by practice or some other means. In terms of the two research models, there were substantial differences between how completeness contributed with writing performance when comparing individuals to collaborative note-takers. Among individual note-takers, in all five cases there was a positive association between completeness and writing, with two of those cases being statistically significant. We speculate that the consistently positive yet inconsistently significant relationship between completeness and writing may be associated with systemic differences in the content and assessment. However, we cannot be sure and consider this as an avenue for further research. Nevertheless, this general positive relationship for individuals falls in line with research suggesting that higher quality note-taking will allow for better essay writing because of the connections students can draw between the notes and their own writing, and the higher order thinking skills they acquire (Waite et al., 2018). Furthermore, the present study develops these ideas in more detail by operationalizing the quality of individual note-taking by measuring completeness, and looking at the direct effect of completeness on how the students write.

In the most surprising finding of the present study, for collaborative note-takers, in four out of five weeks, there was a negative association between completeness and writing, with two of these associations being statistically significant. Although speculative, it may be the case that in the present study, collaborative note-taking in *a subset* of groups involved tasking individual constituent members with the responsibility of taking notes for specific lecture videos each week. This may have resulted in each member of such groups only paying careful attention to the contents of his/her specifically assigned video. Even if a group took highly complete notes, each member may have better understood the course concepts covered in *the notes he/she took* compared to the concepts covered in the notes taken by fellow group members. Therefore, instances of delegation in a subset of groups may have inflated group note completeness: while an individual constituent student may have deepened their focus on recording and understanding the notes that they contributed to, this may have come at the expense of overall group contemplation of the concepts both read and expressed in the notes themselves. On the other hand, in the individual note-taking condition, students who took highly complete notes would have paid closer attention to concepts from all of the course videos for a given week, enabling them to apply those concepts as skills when writing an essay, leading to higher writing scores. Such a result implies that the recording of notes may be more important to a learner's ability to *apply* what he/she has learned in a course. Here, we note that quizzes and writing assignments are fundamentally different forms of assessment, and that note-taking affects the outcomes of each form of assessment in different ways. In the course examined in the present study, quizzes served as a measure of comprehension and recall of contents from course lectures, whereas writing assignments represented the application of those contents to the real-world task of explaining, through writing, the results of scientific research. We suggest that access to high-quality notes are a driver of learning performance on quizzes, which explains the superior results of collaborative note-takers in this regard, whereas the process of recording notes may drive a learner's ability to apply these concepts in the form of academic writing, which explains the superior results of the individual note-takers in this regard.

There is a distinction in the literature that needs to be drawn when looking at the benefits of both collaboration and note-taking. That is the distinction between the knowledge gained and the skills gained from both processes. A measurement of knowledge gained looks at the facts, or the meaning of concepts or processes, and can be defined as "declarative knowledge" (Dochy & Alexander, 1995). On the other hand, a measurement of skills (knowledge application) looks at how students can use the knowledge that they have been taught (Glaser, 1990). The value of this distinction is that, in previous research, we have suggested that the differences in performance between individual and collaborative note-takers was driven by the interaction between the act of note-taking leading to better performance on the application of skills and that the increased gains collaborative learners saw were from their collaboration leading to better retention of information (Fanguy et al., 2021a). Thus, it may be best for instructional and pedagogical designers to make use of online group note-taking activities as a means to build content knowledge while utilizing individual tasks to enhance productive writing competencies. To sum, the inclusion of online group note-taking activities in programs of learning should be carefully considered alongside the goals and intended learning outcomes of the course of learning.

6. Conclusion

Following online lectures with note-taking assignments can be an excellent pedagogical aid to mitigate challenges students may face when studying online. From the students' perspective, taking notes online in a shared format with their instructor and peers while watching lecture videos may enable them to remain focused and more effectively engage with course content. The produced notes can also become useful learning artifacts students can utilize to recall the course content and review for exams (or prepare knowledge applications). From the instructors' perspective, online collaborative note-taking can also be a great relief to ensure students' engagement with their lecture videos and to monitor students' learning progress by visiting and checking their notes online at different points throughout the course. Collaborative note-taking activities, supported by online writing applications that allow multiple students to contribute to the notes both synchronously and asynchronously, can be particularly useful in creating a sense of social presence in distance learning contexts.

The effectiveness of collaborative note-taking (and collaborative writing in general) in online or blended learning contexts has been previously argued and demonstrated by previous researchers. The present study has built on such research findings of the positive outcomes of collaborative note-taking activities. It does this in two ways, by providing the first in-depth comparison of individual and collaborative note-taking that includes how these different learning contexts interact with the quality of the notes taken and student performance. Furthermore, the present study also provides a nuanced investigation of how different aspects of note-taking in the form of completeness interact with the retention of information as well as essay writing from the perspective of individual and collaborative note-taking.

However, the present study further demonstrates that while collaborative note-taking provides a number of benefits, it is not a panacea, and the benefits it can provide to learners are nuanced. In light of this, the present study presents two important implications for practitioners with regard to note-taking in online courses. The first implication acknowledges an important benefit of collaborative note-taking: when students take notes collaboratively, they tend to produce notes of sufficiently high quality to benefit their ability to recall concepts from the course material. This is meaningful because prior research has noted that students are often very poor note-takers, taking very sparse notes that often fail to include or expound on salient points from the lecture (Boyle, 2010). The present study suggests that collaborative note-taking provides an antidote, as students working together in small groups have the ability to combine their efforts to create sufficiently voluminous and complete notes to improve their recall. The second implication reveals a caveat to collaborative note-taking: despite the benefits to students' recall, collaborative note-taking may not be the most effective way for students to learn to apply knowledge as a skill. As groups divide up responsibilities in constructing collaborative notes, constituent members may focus heavily on their assigned parts at the expense of contemplating the contents of the notes written by other members. Applying the knowledge one has learned to a real-world problem, such as writing a scientific manuscript, requires more than just recall, as the student must synthesize concepts learned throughout the course and integrate them into their own knowledge schemas in order to apply them to the task at hand.

In the present study, individual note-takers were better able to apply their knowledge to the skill of writing because taking notes on all parts of the course instruction may have given them increased opportunities to reflect on the complete set of course lectures. On the other hand, collaborative note-takers may have had a harder time in synthesizing all the instructional contents into their existing knowledge schemas since these learners tended to strongly focus on certain parts of the instruction while neglecting others that were assigned to their group members. Such narrow focus on certain parts of the lesson seems to have impeded collaborative note-takers' ability to apply the knowledge learned in the course. Therefore, we recommend that instructors encourage all members of collaborative note-taking groups to contribute notes to each section or part of the instruction rather than assigning parts of the instruction to constituent group members. If collaborative note-takers interactively construct notes on each part of the instruction, i.e., each lecture video, each member will have the opportunity to deeply reflect on each concept conveyed throughout the course, leading to an increased ability to integrate this information into his/her existing knowledge schema.

To the best of the authors' knowledge, the present study represents the first attempt to measure the completeness of the notes students compose when working individually and in groups. Measuring the completeness of the notes enabled us to analyze the effects of completeness on subsequent measures of learning. Although these effects were somewhat conflicting, we provided some possible explanations. Nevertheless, the study results do not offer any definite explanations of the lack of correlation between the quality of collaborative notes and collaborative note-takers' learning outcomes. This comes down to the limitations of the present study, which has not examined the nature of collaborative note-taking activities in light of small group formation and interactions and individual contributions to the process of writing lecture notes online. For example, as speculated in the earlier section, each note-taker could have overly focused on their assigned sections of online lectures without fully engaging with lecture content. In addition, individual students' attempts to recall information recorded by other group members, which may be manifested as their "note-reading" activities, have not been investigated in this study. Furthermore, the present study does not look at the differences between how the information from a lesson may be presented in note form. For example, information may be transcribed verbatim, or paraphrased. Notes may contain information that does not directly cover information from the lecture; for example, students may reflect on the information in the lecture or add further information on the topic. Future research can look into this question and ask how different layers of information in note form affect subsequent learner performance. Thus, the extent to which students interacted with other students' contributions in constructing their notes (e.g., the deletion of specific students' notes) and how the forms of revisions and suggestions might result in improved quiz performance and writing should be examined in future research. In addition, this study was unable to analyze how frequently or how long students accessed their notes when reviewing for quizzes or assignments because the Google Docs platform used in this study does not provide such access-related data. This presents an opportunity for future research if such data could be gleaned from other note-taking platforms or if students self-reported their reviewing habits in interviews or surveys. Finally, similar studies in the future should also look to ensure equivalence between groups in terms of personality (i.e., introversion) as more introverted members may take more from learning individually considering the transactional costs associated with group work. Careful consideration for and measurement of the types of course content (e.g., implicit vs explicit, and mechanical vs. stylistic) could also constitute a new line of enquiry.

Declaration of Competing Interest

None.

Data availability

Data will be made available on request.

References

Adeniran, A., Masthoff, J., & Beacham, N. (2019, June). Model-based characterization of text discourse content to evaluate online group collaboration. In *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 3–8). Cham: Springer.

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. https://doi.org/10.1037/h0076613

Aragón-Mendizábal, E., Delgado-Casas, C., Navarro-Guzmán, J. I., Menacho-Jiménez, I., & Romero-Oliva, M. F. (2016). A comparative study of handwriting and computer typing in note-taking by university students. *Comunicar. Media Education Research Journal, 24*(2). Retrieved from https://www.scipedia.com/public/Aragon-Mendizabal_et_al_2016a.

Baldwin, M., Fanguy, M., & Costley, J. (2019). The effects of collaborative note-taking in flipped learning contexts. *Journal of Language and Education, 5*(4), 25–35. https://doi.org/10.17323/jle.2019.9726

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benton, S. L., Kiewra, K. A., Whitfill, J. M., & Dennison, R. (1993). Encoding and external-storage effects on writing processes. *Journal of Educational Psychology, 85* (2), 267–280. https://doi.org/10.1037/0022-0663.85.2.267

Blau, I., & Caspi, A. (2009). What type of collaboration helps? psychological ownership, perceived learning and outcome quality of collaboration using google docs. In *, vol. 12. Proceedings of the Chais Conference on Instructional Technologies Research* (pp. 48–55). Israel: Open University. no. 1.

Boyle, J. R. (2010). Note-taking skills of middle school students with and without learning disabilities. *Journal of Learning Disabilities, 43*(6), 530–540. https://doi.org/10.1177/0022219410371679

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. https://doi.org/10.1037/a0030367

Butson, R., & Thomson, C. (2014). Challenges of effective collaboration in a virtual learning environment among undergraduate students. *Creative Education, 5*(16), 1449–1459. https://doi.org/10.4236/ce.2014.516162

Caspi, A., Gorsky, P., & Chajut, E. (2003). The influence of group size on nonmandatory asynchronous instructional discussion groups. *The Internet and Higher Education, 6* (3), 227–240. https://doi.org/10.1016/S1096-7516(03)00043-5

Chen, P. H. (2019). In-class and after-class lecture note-taking strategies. *Active Learning in Higher Education, 22*(3), 245–260. https://doi.org/10.1177/1469787419893490

Clabough, E. B., & Clabough, S. W. (2016). Using rubrics as a scientific writing instructional method in early stage undergraduate neuroscience study. *Journal of Undergraduate Neuroscience Education, 15*(1), A85. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5105970/.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Costley, J., Courtney, M., & Fanguy, M. (2022). The interaction of collaboration, note-taking completeness, and performance over 10 weeks of an online course. *The Internet and Higher Education, 52*(1), Article 100831. https://doi.org/10.1016/j.iheduc.2021.100831

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63*(1), 8. https://doi.org/10.1037/h0032243

Doberstein, D., Hecking, T., & Hoppe, H. U. (2019, September). What can interaction sequences tell us about collaboration quality in small learning groups?. In *In international conference on web-based learning* (pp. 61–71). Cham: Springer. https://doi.org/10.1007/978-3-030-35758-0_6.

Dochy, F. J. R. C., & Alexander, P. A. (1995). Mapping prior knowledge: a framework for discussion among researchers. *European Journal for Psychology of Education, 10*(3), 225–242. https://doi.org/10.1007/BF03172918

Fanguy, M., Baldwin, M., Shmeleva, E., Lee, K., & Costley, J. (2021a). How collaboration influences the effect of note-taking on recall of contents and writing performance. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2021.1950772

Fanguy, M., & Chang, J. (2021). Operationally defining turn-taking in collaborative online documents. *Proceedings of the 14th International Conference on Computer-Supported Collaborative Learning-CSCL 2021*.

Fanguy, M., Lee, S. Y., & Churchill, D. G. (2021b). Adapting educational experiences for the chemists of tomorrow. *Nature Reviews Chemistry, 5*(3), 141–142. https://doi.org/10.1038/s41570-021-00258-5

Fox, J., & Weisberg, S. (2019). *An {R} companion to applied regression* (3rd ed.). Thousand Oaks CA: Sage. Retrieved from https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Glaser, R. (1990). Toward new models for assessment. *International Journal of Educational Research, 14*, 475–483. https://doi.org/10.1016/0883-0355(90)90028-7

Goldstein, H. (2003). Multilevel modelling of educational data. In *Methodology and epistemology of multilevel analysis* (pp. 25–42). Dordrecht: Springer.

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. https://doi.org/10.1016/j.iheduc.2005.09.002

Hadwin, A. F., Kirby, J. R., & Woodhouse, R. A. (1999). Individual differences in notetaking, summarization, and learning from lectures. *Alberta Journal of Educational Research, 45*(1). https://doi.org/10.11575/ajer.v45i1.54623

Harbin, M. B. (2020). Collaborative note-taking: A tool for creating a more inclusive college classroom. *College Teaching, 68*(4), 214–220. https://doi.org/10.1080/87567555.2020.1786664

Haynes, J. M., McCarley, N. G., & Williams, J. L. (2015). An analysis of notes taken during and after a lecture presentation. *North American Journal of Psychology, 17*(1), 175–186. Retrieved from https://www.researchgate.net/profile/Joshua_Williams4/publication/272417797_An_Analysis_of_Notes_Taken_During_and_After_a_Lecture_Presentation/links/54e3a2000cf2dbf60693a790.pdf.

Irshad, M. B. M., & Johar, M. G. M. (2015). A study of undergraduate use of cloud computing applications: special reference to Google Docs. *European Journal of Computer Science and Information Technology, 3*(4), 22–32. Retrieved from https://www.researchgate.net/profile/Irshad-M-B-M/publication/297757609_A_STUDY_OF_UNDERGRADUATE_USE_OF_CLOUD_COMPUTING_APPLICATIONS_SPECIAL_REFERENCE_TO_GOOGLE_DOCS/links/56e2d2a308aee84447bf332f/A-STUDY-OF-UNDERGRADUATE-USE-OF-CLOUD-COMPUTING-APPLICATIONS-SPECIAL-REFERENCE-TO-GOOGLE-DOCS.pdf.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2014). Cooperative learning: improving university instruction by basing practice on validated theory. *Journal on Excellence in University Teaching, 25*(4), 1–26. Retrieved from https://karlsmithmn.org/wp-content/uploads/2017/08/Johnson-Johnson-Smith-Cooperative_Learning-JECT-Small_Group_Learning-draft.pdf.

Ju, M., & Kim, J. (2020). A qualitative analysis of the effect of note-taking on writing summary sentences. *Korean language education, 168*, 47–77 (in Korean). Retrieved from https://kiss.kstudy.com/thesis/thesis-view.asp?key=3754431.

Kam, M., Wang, J., Iles, A., Tse, E., Chiu, J., Glaser, D., & Canny, J. (2005, April). Livenotes: A system for cooperative and augmented note-taking in lectures. In *Proceedings of the Sigchi Conference on Human Factors in Computing Systems* (pp. 531–540). CHI.

Kane, M. J., Smeekens, B. A., von Bastian, C. C., Lurquin, J. H., Carruth, N. P., & Miyake, A. (2017). A combined experimental and individual-differences investigation into mind wandering during a video lecture. *Journal of Experimental Psychology: General, 146*(11), 1649–1674. https://doi.org/10.1037/xge0000362

Kassambara, A. (2020). Ggpubr: "ggplot2" based publication ready plots. In *R package version 0.4.0*. Retrieved from https://CRAN.R-project.org/package=ggpubr.

Kassambara, A. (2021). Rstatix: pipe-friendly framework for basic statistical tests. In *R package version 0.7.0*. Retrieved from https://CRAN.R-project.org/package=rstatix.

Kiewra, K. A. (1985). Students' note-taking behaviors and the efficacy of providing the instructor's notes for review. *Contemporary Educational Psychology, 10*(4), 378–386. https://doi.org/10.1016/0361-476X(85)90034-7

Kiewra, K. A. (1987). Notetaking and review: the research and its implications. *Instructional Science, 16*(3), 233–249. Retrieved from https://link.springer.com/content/pdf/10.1007/BF00120252.pdf.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review, 21*(1), 31–42. https://doi.org/10.1007/s10648-008-9095-2

Kirschner, P. A., Sweller, J., Kirschner, F., Zambrano, R., & J.. (2018). From cognitive load theory to collaborative cognitive load theory. *International Journal of Computer-Supported Collaborative Learning, 13*(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y

Lee, K., Fanguy, M., Bligh, B., & Lu, X. S. (2022). Adoption of online teaching during the COVID-19 Pandemic: A systematic analysis of changes in university teaching activity. *Educational Review, 74*(3), 460–483. https://doi.org/10.1080/00131911.2021.1978401

Lund, A., & Smørdal, O. (2006). Is there a space for the teacher in a wiki?. In *Proceedings of the 2006 International Symposium on Wikis* (pp. 37–46). ACM.

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. https://doi.org/10.1007/s11251-018-9458-0

Luo, L., Kiewra, K. A., & Samuelson, L. (2016). Revising lecture notes: How revision, pauses, and partners affect note taking and achievement. *Instructional Science, 44*(1), 45–67. https://doi.org/10.1007/s11251-016-9370-4

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science, 25*(6), 1159–1168. https://doi.org/10.1177/0956797614524581

Nakayama, M., Mutsuura, K., & Yamamoto, H. (2017). Effectiveness of student's note-taking activities and characteristics of their learning performance in two types of online learning. *International Journal of Distance Education Technologies (IJDET), 15* (3), 47–64. https://doi.org/10.4018/IJDET.2017070104

Nokes-Malach, T. J., Richey, J. E., & Gadgil, S. (2015). When is it better to learn together? Insights from research on collaborative learning. *Educational Psychology Review, 27*(4), 645–656. https://doi.org/10.1007/s10648-015-9312-8

Orndorff, H. N., III (2015). Collaborative note-taking: the impact of cloud computing on classroom performance. *International Journal of Teaching and Learning in Higher Education, 27*(3), 340–351. Retrieved from https://files.eric.ed.gov/fulltext/EJ1093744.pdf.

Petko, D., Schmid, R., Müller, L., & Hielscher, M. (2019). Metapholio: A mobile app for supporting collaborative note taking and reflection in teacher education. *Technology, Knowledge and Learning, 24*(4), 699–710. https://doi.org/10.1016/j.compedu.2019.03.006

Pospelova, T. (2021). The collaborative discussion model: developing writing skills through prewriting discussion. *Journal of Language and Education, 7*(1 (25)), 156–170. https://doi.org/10.17323/jle.2021.10748

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: *R Foundation for* Statistical Computing. Retrieved from https://www.Rproject.org/.

Raver, S. A., & Maydosz, A. S. (2010). Impact of the provision and timing of instructor-provided notes on university students' learning. *Active Learning in Higher Education, 11*(3), 189–200. https://doi.org/10.1177/1469787410379682

Retnowati, E., Ayres, P., & Sweller, J. (2017). Can collaborative learning improve the effectiveness of worked examples in learning mathematics? *Journal of Educational Psychology, 109*(5), 666–679. https://doi.org/10.1037/edu0000167

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/.

Salame, I. I., & Thompson, A. (2020). Students' views on strategic note-taking and its impact on performance, achievement, and learning. *International Journal of Instruction, 13*(2), 1–16. https://doi.org/10.29333/iji.2020.1321a

van de Sande, C., Abramson, J., & Judson-Garcia, J. (2017). An exploration of note-taking in an online calculus course. *Journal of Computers in Mathematics and Science Teaching, 36*(1), 75–99. Retrieved from https://www.learntechlib.org/primary/p/174372/.

Shi, Y., Yang, H., Yang, Z., Liu, W., & Yang, H. H. (2020, August). The effects of a collaborative learning approach with digital note-taking on college students' learning achievement and cognitive load. In *Proceedings of the International Conference on Blended Learning* (pp. 187–198). Cham: Springer.

Slotte, V., & Lonka, K. (2001). Note taking and essay writing. In *In writing as a learning tool* (pp. 131–143). Dordrecht: Springer. https://doi.org/10.1007/978-94-010-0740-5_8.

Steimle, J., Brdiczka, O., & Mühlhäuser, M. (2009). Collaborative paper-based annotation of lecture slides. *Educational Technology & Society, 12*, 125–137. Retrieved from https://www.jstor.org/stable/10.2307/jeductechsoci.12.4.125.

Tindale, R. S., & Winget, J. R. (2017). Learning while deciding in groups. In *The oxford handbook of group and organizational learning*. Retrieved from https://psyarxiv.com/8ufgh/download?format=pdf.

Volet, S., Summers, M., & Thurman, J. (2009). High-level co-regulation in collaborative learning: How does it emerge and how is it sustained? *Learning and Instruction, 19*(2), 128–143. https://doi.org/10.1016/j.learninstruc.2008.03.001

Waite, B. M., Lindberg, R., Ernst, B., Bowman, L. L., & Levine, L. E. (2018). Off-task multitasking, note-taking and lower-and higher-order classroom learning. *Computers & Education, 120*, 98–111. https://doi.org/10.1016/j.compedu.2018.01.007

Willse, J. T. (2018). CTT: Classical test theory functions. In *R package version 2.3.3*. Retrieved from https://CRAN.R-project.org/package=CTT.

Wilson, K. (2014). Note-taking in the academic writing process of non-native speaker students: is it important as a process or a product? *Journal of College Reading and Learning, 29*(2), 166–179. https://doi.org/10.1080/10790195.1999.10850077

Wu, J. Y. (2020). The predictive validities of individual working-memory capacity profiles and note-taking strategies on online search performance. *Journal of Computer Assisted Learning, 36*(6), 876–889. https://doi.org/10.1111/jcal.12441

Further-reading

Tindale, R. S., Stawiski, S., & Jacobs, E. (2008). Shared cognition and group learning. *Work group learning: Understanding, improving and assessing how groups learn in organizations* (pp. 73–91). Psychology Press.