



OPEN ACCESS

EDITED BY

Andrea Paula Goldin,
National Scientific and Technical Research
Council (CONICET), Argentina

REVIEWED BY

Jennifer Knellisen,
University of Wuppertal, Germany
Abdelhakim Boubekri,
Hassan Premier University, Morocco

*CORRESPONDENCE

Zahia Marzouk
✉ zahia.marzouk@kpu.ca

RECEIVED 31 October 2024

ACCEPTED 29 April 2025

PUBLISHED 06 June 2025

CITATION

Marzouk Z and Winne PH (2025) Recall or
transfer? How assessment types drive
text-marking behavior.
Front. Educ. 10:1510007.
doi: 10.3389/feduc.2025.1510007

COPYRIGHT

© 2025 Marzouk and Winne. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Recall or transfer? How assessment types drive text-marking behavior

Zahia Marzouk^{1*} and Philip H. Winne²

¹Educational Studies, Faculty of Arts, Kwantlen Polytechnic University, Surrey, BC, Canada, ²Faculty of Education, Simon Fraser University, Burnaby, BC, Canada

Introduction: Text marking is a widely used study technique, valued for its simplicity, and perceived benefits in enhancing recall and comprehension. This exploratory study investigates its role as an encoding mechanism, focusing on how marking impacts recall and transfer when learners are oriented toward different posttest items (recall or transfer).

Method: We gathered detailed data describing what learners were studying and how much they marked during studying. Participants were randomly assigned to one of four groups in a 2 × 2 factorial design. One independent variable, examples, determined whether participants were trained using examples of the types of information required to answer posttest items. The other independent variable, orientation, determined whether participants were instructed to prepare for a recall test or for an application (transfer) test.

Results: Statistical analysis revealed a detectable effect of study orientation (transfer vs. recall), $F = 2.076$, $p = 0.043$, partial $\eta^2 = 0.114$. Compared to learners oriented to study for recall, learners oriented to study for transfer marked information identified as examples ($F = 3.881$, $p = 0.051$, partial $\eta^2 = 0.028$), main ideas ($F = 7.348$, $p = 0.008$, partial $\eta^2 = 0.051$), and reasons ($F = 5.440$, $p = 0.021$, partial $\eta^2 = 0.038$). Moreover, a statistically detectable proportional relationship was found between total marking and transfer performance ($F = 5.885$, $p = 0.017$, partial $\eta^2 = 0.042$). Learners who marked more scored higher on transfer questions. Prior knowledge mediated approximately 52% of the effect, indicating that as prior knowledge increased, so did the frequency of marking.

Discussion: Orienting to study for a particular type of posttest item affected studying processes, specifically, how much learners marked and the categories of information they marked. While the frequency of marking was proportional to achievement, orienting to study for recall versus transfer posttest items had no effect on recall or transfer. Prior knowledge powerfully predicted how much learners marked text.

KEYWORDS

highlighting, underlining, recall, transfer, prior knowledge

1 Introduction

Highlighting and underlining are both popular and easy-to-use study strategies university students frequently use to manage the overwhelming amount of reading by isolating key concepts for focused study (Miyatsu et al., 2018; Morehead et al., 2016; Bell and Limber, 2009). Fowler and Barker (1974) suggested that underlining and highlighting are functionally identical and have similar effects on learning. For brevity, we refer to both as text marking.

Text marking remains one of the most popular study techniques among college students due to its simplicity and minimal effort to use (Miyatsu et al., 2018; Dunlosky et al., 2013), with many believing it aids focus and enhances comprehension (Nist and Kirby, 1986).

In the literature, text marking is viewed as (a) an encoding mechanism that encourages more extensive engagement with the material and (b) an external storage that isolates and highlights crucial information, making it easier to recall and access later. This study specifically examines text marking as an encoding mechanism.

2 A literature review

2.1 Text marking as an encoding mechanism

As an encoding mechanism, text marking facilitates information processing by prompting learners to actively search for and identify essential content that might otherwise be overlooked or processed less effectively (Weinstein and Mayer, 1983). This active engagement in selecting and marking information functions as a catalyst for other cognitive processes, supporting the storage of information in a recallable format (Leutner et al., 2007). Blanchard and Mikkelsen (1987) emphasized that marking helps learners focus attention on important information, reducing cognitive overload, and supporting retention. From this perspective, selecting information through marking should enhance recall by preventing learners from exceeding their processing capacity. This supports the idea that marking prompts more extensive cognitive involvement and enhances memory retention (Winne et al., 2015). In contrast, when learners use marking merely as a concentration technique, marking large portions of text without engaging in search and selection processes, it fails to facilitate extended cognitive processing, limiting its benefits (Nist and Hogrebe, 1987).

Research on text marking has primarily examined recall and comprehension, though comprehension tasks in some studies include inferential reasoning (Peterson, 1992). Recall has been measured using cloze tests (Ponce et al., 2018), fill-in-the-blank tasks (Yue et al., 2015), and sentence completion tasks (Jonassen, 1984). Multiple-choice recall tests were also used to assess short- and long-term retention (Peterson, 1992). Comprehension has typically been assessed using multiple-choice tests (Ben-Yehudah and Eshet-Alkalai, 2018; Hayati and Shariatifar, 2009; Leutner et al., 2007) and open-ended tasks, including summary writing (Ponce et al., 2018) and comprehension questions (Peterson, 1992).

Although comprehension has been widely studied, the nature and extent of processing varies across studies. Some research differentiates between factual recall and inferential comprehension (Ben-Yehudah and Eshet-Alkalai, 2018), while others do not specify whether measures assess simple retrieval or integration (Hayati and Shariatifar, 2009; Leutner et al., 2007). Open-ended assessments, such as summary writing (Ponce et al., 2018) and probed recall (Blanchard and Mikkelsen, 1987), require greater cognitive elaboration but are less frequently used.

While studies generally focus on recall or comprehension, none have explicitly examined text marking in relation to knowledge transfer. In some cases, comprehension tasks may require learners to infer unstated relationships, but no study has directly assessed transfer

as a distinct outcome. Given that transfer involves applying knowledge in new contexts, its inclusion in this study is warranted to examine whether text marking supports learning beyond simple recall and comprehension.

Despite its widespread use, the effectiveness of text marking as a learning strategy has been debated (Bisra et al., 2014). A meta-analysis by Ponce et al. (2022) found text marking positively impacted retention ($g = 0.36$, $p = 0.013$), suggesting learners who engage in marking remember key ideas better than those who only read the text. However, text marking had a limited effect on comprehension ($g = 0.20$, $p = 0.155$), indicating while it supports recall of specific information, it does not necessarily foster understanding. These findings align with Miyatsu et al. (2018), who suggested that the effectiveness of text marking is context-dependent, rather than universally beneficial.

While recent evidence highlights potential benefits, earlier research raised serious concerns about the overall effectiveness of text marking. Dunlosky et al. (2013) questioned text marking as a productive learning strategy, and Biwer et al. (2020) described it as a passive technique with minimal impact on long-term learning. One major concern is that learners may engage in text marking superficially, leading to a fluency illusion, which, in turn, results in an illusion of learning. The fluency illusion occurs when learners mistake the ease of processing marked text for actual comprehension and retention (Karpicke et al., 2009). This is particularly problematic in text marking because marking is relatively effortless, which can create a false sense of mastery even when learners have not processed the material sufficiently (Kornell et al., 2011).

This misinterpretation of ease of usage leads to metacognitive illusions where students misinterpret the immediate ease of processing as an indicator of long-term retention and comprehension of the material (Soderstrom and Bjork, 2015; Winne and Jamieson-Noel, 2002). As a result, learners may believe that simply marking text enhances learning even when they are not actively processing the content. However, we hypothesize text marking is not inherently ineffective. Its impact on learning depends on how it is used and whether learners engage actively in selecting, organizing, and integrating marked information versus passively marking.

Anderson and Pearson (1984) emphasized that the effects of text marking depend on how learners engage with information; marked text is better remembered when actively processed, rather than passively marked without further cognitive engagement. The Select–Organize–Integrate model (Fiorella and Mayer, 2015) provides a broader framework for understanding how learners process information to achieve meaningful learning. According to this model, meaningful learning requires selecting relevant information, organizing it into a structured representation, and integrating it with prior knowledge. If learners focus excessively on selecting, they may deplete cognitive resources needed for organizing and integrating, thereby limiting comprehension (Ponce et al., 2022). And if learners select irrelevant information, organizing and integrating it could hinder learning. This explains why text marking may not always enhance comprehension. When used passively, it primarily engages selection without promoting the organization and integration required for understanding. However, when learners select key information, then structure and integrate it with prior knowledge, text marking can function as more than just a selection tool to enhance learning. This distinction between surface and extended engagement may explain

inconsistencies in the literature. While Ponce et al. (2022) found text marking benefits recall but has a limited impact on comprehension, inconsistencies remain in prior research. Some studies found benefits (e.g., Hayati and Shariatifar, 2009), while others reported no effect (e.g., Johnson, 1988) or even detrimental outcomes (e.g., Peterson, 1992). These discrepancies may stem not only from differences in study design and comprehension measures but also from variations in how learners engage with the information they mark—whether they use marking superficially or as a gateway to more extensive processing. Thus, the effectiveness of marking depends not on the strategy itself but on how learners process marked content.

Given the fluency illusion's impact on metacognitive accuracy, researchers have proposed mechanisms to distinguish between perceived and actual learning. Delayed testing shifts reliance from encoding fluency to retrieval fluency, improving judgments of learning and better predicting long-term retention (Koriat, 2006). Similarly, using conceptual versus factual questions can help assess whether marked content is more extensively encoded, as fluency illusions tend to be strongest for direct recall tasks but weaker for tasks requiring inference-making and integration (Soderstrom and Bjork, 2015).

While prior research has extensively examined the role of text marking in recall and comprehension (Ayer and Milson, 1993; Ben-Yehudah and Eshet-Alkalai, 2018; Hayati and Shariatifar, 2009; Leutner et al., 2007), its relationship to knowledge transfer remains unclear. This is partly due to inconsistencies in how comprehension has been measured, as some studies distinguish between factual and inferential comprehension (e.g., Ben-Yehudah and Eshet-Alkalai, 2018; Peterson, 1992), while others do not specify cognitive processes underlying measures of comprehension (e.g., Hayati and Shariatifar, 2009; Leutner et al., 2007). However, no study has explicitly examined whether text marking supports knowledge transfer, leaving a critical gap in the literature. To address this, the present study investigates how text marking influences both recall and knowledge transfer.

2.2 Test expectancy

Research on how test expectancy influences learning strategies reveals learners adapt their study methods based on the type of test they anticipate. However, the role of text marking within this adaptation process has received less attention, despite its widespread use as a study strategy aimed at enhancing both recall and comprehension. Klauer (1984) conducted a meta-analysis that found specific instructional objectives improved learning of targeted information (intentional learning) but hindered incidental learning, whereas general study objectives enhanced incidental learning. Klauer suggested instructional objectives influence learners' intentions, which we interpret as establishing standards for metacognitively monitoring content that direct learners' attention to and operations on information, that is, how learners metacognitively monitor information and control study tactics.

The effects of test expectancy have been examined across various learning tasks, with mixed outcomes. While the test expectancy effect has been found in list-learning studies (Rivers and Dunlosky, 2021; Minnaert, 2003), it is inconsistent when the material studied is a text (Naveh-Benjamin et al., 2000). Some studies found no significant effects on recall or recognition of text information (e.g., Kulhavy et al., 1975), while others detected effects of test expectancy

on learners' performance (e.g., Minnaert, 2003). These mixed findings suggest the cognitive demands of different test types influence study behaviors in complex ways, which may also affect how learners engage with text marking as a strategy for encoding information.

McDaniel et al. (1994) highlighted students often inquire about test formats to adjust their study methods, indicating an intention to modify their study strategies based on anticipated tests. The encoding-strategy adaptation hypothesis proposed by Finley and Benjamin (2012) provided a framework for understanding these behaviors. According to this hypothesis, learners adapt their encoding strategies based on the expected test format, employing distinct approaches for cued-recall versus free-recall tests. In subsequent research, Tullis et al. (2013) conducted experiments that demonstrated how learners adaptively accommodate their encoding strategies to meet upcoming test demands. Their findings revealed participants shifted their strategies in alignment with test expectations, showing sophisticated metacognitive control. For example, learners who expected a cued-recall test focused more on linking cues with targets, while those expecting a free-recall test concentrated on connections between target words and used narrative strategies. These findings reveal learners exhibit sophisticated metacognitive control, shifting their encoding strategies to align with anticipated test demands (Finley and Benjamin, 2012; Cho and Neely, 2017). However, as far as we know, these findings have not been examined in the context of text marking, raising the question of whether learners adjust how they mark and process text based on test expectations.

Research suggests test expectancy influences how learners engage with text marking, shaping whether it functions as a surface-level strategy or a tool for comprehension. For example, Feldt and Ray (1989) found that learners expecting multiple-choice questions typically marked, and they reread text but refrained from taking notes, a behavior shared by many learners expecting free recall. This indicates learners may adjust their marking behavior based on the type of assessment they expect. However, these study tactics did not result in significant differences in test scores between groups, suggesting the chosen strategies did not directly impact performance outcomes.

Similarly, Abd-El-Fattah (2011) investigated how undergraduates adapt their study tactics based on the cognitive demands of expected test questions. Learners expecting deep-level questions stressed deep-level study tactics more than learners, who expected surface-level questions, while learners expecting surface-level questions primarily used surface-level tactics. Abdel Fatah also reported these strategies mediated the relationship between test expectancy and performance.

Surber (1992) explored the effects of test expectancy, text length, and topic on learners' study and text marking behaviors. He discovered learners studying shorter texts engaged in more text marking than those working with longer texts, suggesting task instructions and expectations, such as preparing for recall or transfer posttest questions, play a significant role in how learners interact with the material.

The effectiveness of text marking depends on learners' cognitive engagement and test expectancy. When preparing for recall-based tests, learners may mark extensively without deep engagement. In contrast, comprehension-based or transfer tasks may engage in more selective, organized and integrative marking to support more extensive processing of information. As the Select–Organize–Integrate model (Fiorella and Mayer, 2015) suggests, selection alone is insufficient for meaningful learning without organization and integration.

Task instructions, such as emphasizing recall versus comprehension, may play a role in shaping how text marking is used. Learners anticipating comprehension-based tests may engage in more meaningful selection, organizing the information and integrating it, while those expecting factual recall tests may mark passively, reinforcing the fluency illusion (Karpicke et al., 2009). In line with this, one of the objectives of our research is to investigate how different task instructions impact learners' text-marking practices and how this, in turn, influences their learning processes and outcomes. This focus will help us understand the relationship between orienting study instructions and the effectiveness of learners' engagement with the material.

2.3 Text marking as a trace of learning processes

While text marking has traditionally been studied as an encoding strategy that influences learning outcomes (as discussed in Section 2.1), it can also serve as a measurable trace of cognitive processes during learning. This perspective shifts the focus from its direct impact on learning outcomes to its role as an indicator of cognitive engagement, revealing how learners process information and respond to task instructions.

Most text marking research has treated text marking as an independent variable and investigated how text marking affected learning outcomes; text marking can also be seen as a trace of otherwise unobservable learning processes. This shift in perspective allows researchers to investigate not only whether text marking improves learning but also how the act of marking relates to learners' cognitive processing and metacognitive regulation. Research adopting this perspective may treat text marking as an outcome, a correlate of learning outcomes, or an intervening variable. In the last case, the amount and quality of text marking are influenced by instructional conditions and text marking, in turn, affects learning (e.g., Bell and Limber, 2009). By analyzing the frequency, patterns, and types of markings made by learners, researchers can enhance insights into how learners process information and regulate their study behaviors. Our research focused on text marking as a trace of learning processes that may be influenced by task instructions and then influence posttest scores.

Several studies analyzed the types of information learners mark. Dimensions, such as high-level and low-level sentences (e.g., Rickards and August, 1975) and superordinate and subordinate sentences (e.g., Johnson, 1988), have described marked information. However, using only these categories may be insufficient to explore relationships between learners' text marking and learning outcomes. Our research recognized eight categories of information in the text learners studied. We examined the frequency with which learners marked each type and whether the type of marked text predicted achievement.

The relationship between text marking and learning outcomes also depends on the skillfulness of learners' marking strategies. Most studies examining text marking asked participants simply to read and mark text without providing training about how and what to mark (e.g., Fowler and Barker, 1974). Studies that did train how to mark varied in the amount of training they provided. For instance, Hayati and Shariatifar (2009) provided a 60-min training session that included suggestions on when, how and what to mark. Amer (1994) trained participants for 90 min one time a week for 5 weeks to mark

text by following four steps adapted from Smith (1985). Cited in Amer (1994). Ponce et al. (2022) further supported this, showing students who received training in marking significantly outperformed those who only read or studied the text ($g = 1.02$, $p = 0.002$).

Leutner et al. (2007) varied training as an independent variable. They compared a group trained to use a text marking tactic to a group trained in that same tactic and self-regulation to a no training group. Trained participants outperformed those with no training, and participants in the learning tactic plus self-regulation group performed much better than learners trained only in the marking tactic. These findings suggest different modes of training can shape how learners execute marking and impact learning. Given mixed findings on the effectiveness of text marking, it is essential to examine how learners engage with marking as a process, rather than focusing solely on its outcomes. This study explores not only the impact of marking on learning but also how different task instructions shape marking behavior. Specifically, we investigate whether training influences the way learners mark text and whether these differences in marking behavior predict performance. By framing text marking as both an encoding strategy (Section 2.1) and cognitive trace (Section 2.3), we provide a more comprehensive understanding of its role in learning.

2.4 Research questions

Prior research on text marking and learning outcomes has produced mixed findings. While some studies suggest marking enhances recall (Ponce et al., 2022), others argue that without extended cognitive engagement, marking remains a passive strategy with limited benefits of comprehension and transfer (Dunlosky et al., 2013; Biwer et al., 2020). Studies on test expectancy indicate that learners adjust their study strategies based on anticipated test formats, which may influence how they engage with text marking (Finley and Benjamin, 2012; Abd-El-Fattah, 2011). Given these findings, our study investigates text marking as both an outcome of orienting instructions and a predictor of recall and transfer performance.

To examine text marking behavior, we used nStudy (Winne et al., 2019), a Chrome browser extension that tracks learners' marking activity. When a learner marks text, nStudy records the selected text, timestamp, and whether the marked content falls within predefined categories. For this study, we defined eight target categories: term, main idea, description, explanation, fact, consequence, reason, and example. This allowed us to analyze how learners marked different types of information under varying test expectations.

Because previous text-marking research has not examined transfer, and findings on recall were inconsistent, we examined the association between learners' text marking and performance on both types of questions. Rather than proposing explicit hypotheses, we draw on prior research to frame each research question, highlighting relevant findings that inform our investigation.

Prior research suggests test expectancy affects study behaviors, with learners adjusting strategies based on anticipated test formats (Finley and Benjamin, 2012; Abd-El-Fattah, 2011). Feldt and Ray (1989) found learners expecting a multiple-choice test marked and reread text but did not take notes, whereas those anticipating a free recall test engaged in marking, rereading, and taking notes. Similarly, Abd-El-Fattah (2011) found learners expecting deep-level questions used more effective study strategies, while those anticipating surface-level

questions relied on surface-level tactics. Despite these findings, prior research has not directly examined whether test expectancy influences the amount of marking of the specific types of information learners to choose to mark, leaving a gap in understanding how marking behaviors are adjusted based on expected assessment formats. To address this gap, we ask: RQ1: Does text marking vary depending on whether learners are oriented to expect a recall test or a transfer posttest?

Text marking can enhance recall but has limited effects on comprehension (Ponce et al., 2022). While prior research has examined comprehension, few studies identified cognitive processes underlying their comprehension measures. Some differentiate between literal and inferential comprehension (e.g., Ben-Yehudah and Eshet-Alkalai, 2018) but directly investigated whether text marking supports transfer. Given this gap, it is useful to examine whether marking contributes to learning beyond recall and comprehension. This leads to RQ2: Is there a relation between learners' text marking and their performance on both recall and transfer questions?

Research shows training can help learners mark strategically, rather than indiscriminately (Glover et al., 1980). While some studies suggest training enhances marking accuracy and selectivity (Glover et al., 1980), most examined its effects on performance, rather than how it shapes marking behavior (Amer, 1994; Hayati and Shariatifar, 2009; Leutner et al., 2007). These studies provided training in marking and included researcher feedback to guide learners' marking choices but did not systematically analyze what or how learners marked after training. Glover et al. (1980), one of the few studies to do so, suggested the need for further investigation into how training influences marking behavior. Given this gap, we ask: RQ3: Does training to identify categories of information affect learners' text marking?

Training in text marking has been shown to improve recall and comprehension, with students who received training on how to mark performing significantly better than those who only read the text (Ponce et al., 2022). However, the transfer has not been directly examined, leaving it unclear whether training enhances only recall and comprehension or also supports the application of knowledge in new contexts. Given that fluency illusions can lead learners to overestimate their learning when using text marking (Kornell et al., 2011), training may help mitigate these effects by encouraging more analytical engagement with marked text. Whether such training translates to improved transfer performance remains an open question. Therefore, we ask: RQ4: Does such training affect learners' posttest recall and transfer?

3 Methods

3.1 Participants

The participants were 140 undergraduates at a medium-sized Canadian university. In our sample, 65% were female, and a variety of disciplinary majors were represented (see Appendix A). The participants were recruited using flyers posted around the university.

3.2 Design

The participants were randomly assigned to one of four groups in a 2×2 factorial design. One independent variable, examples,

determined whether participants were trained using examples of the types of information required to answer posttest items. The other independent variable, orientation, determined whether participants were instructed to prepare for a recall test or for an application (transfer) test.

Orienting instructions were delivered using nStudy's bookmark feature. The group oriented to study and mark recall-related information was instructed: "Study the following text and highlight parts that would help you answer recall questions about gravity and weightlessness and their effects on human cardiovascular system." The group oriented to study and mark for transfer-related information was instructed as follows: "Study the following text and highlight parts that would help you answer application questions about gravity and weightlessness and their effects on human cardiovascular system."

3.3 Materials

A 1005-word text was created based on Fundamentals of Space Medicine (Clément, 2005) and resources on the Internet. The text described gravity, weightlessness, and their effects on the human cardiovascular system (see Appendix B).

The first author coded each sentence using one of eight categories of information: term (10), main idea (10), description (4), explanation (3), consequence (5), reason (6), example (9), or fact (10). Categories were defined as follows:

- A term: A sentence that included both term and definition
- A main idea: A sentence that represented the main idea of a paragraph
- A description: A sentence that provided a representation or account of something
- An explanation: A sentence that did not just provide an account but rather a clarification
- A consequence: A sentence that conveyed an effect, a result
- A reason: A sentence that presented a justification
- An example: A sentence that presented an illustration
- A fact: A sentence that simply stated a reality and was not categorized as any of the earlier categories.

The first author's coding of targets in the text was corroborated by two graduate students who individually read the text without codes and coded each sentence as representing one of the eight categories. Discrepant codes were discussed until a complete consensus was reached.

When learners select a text segment in nStudy, they can use a key command to open a menu and choose an option to mark the selected text. Once marked, the text is highlighted and copied into the learner's workspace as a quote. nStudy automatically records the selected text and timestamp, allowing precise tracking of marking activity. nStudy also enables researchers to predefine specific text targets within the text. When a learner marks any text falling within the boundaries of a predefined target category, the system logs a hit for that category. In this study, we defined eight target categories: term, main idea, description, explanation, fact, consequence, reason, and example. This functionality allowed us to analyze learners' selection patterns and compare the

information they marked to predefined information categories, providing insight into how they engaged with the material under different test expectations.

3.4 Measures

The participants completed a demographic questionnaire describing their sex, major, year of study, and how many years they had spoken English. Participants then completed several questionnaires that were not used in this study. A prior knowledge recall task assessed participants' knowledge of gravity, weightlessness, and their effects on the human cardiovascular system. Participants wrote bullet-point responses, which were later classified into two categories: (a) Related prior knowledge (content directly related to the reading text) and (b) general prior knowledge (broader knowledge about the topic).

To assess learning outcomes, participants completed two short-answer supply format posttests: (a) Recall posttest (eight items) required participants to retrieve declarative information from the text, and (b) transfer posttest (eight items) asked participants to use information in the text to solve a problem or make a prediction.

Transfer questions were presented first to avoid presenting information on which transfer was based. A scoring rubric was developed to determine which sentences in the text provided the information necessary to answer each question (see [Appendix C](#)).

3.5 Procedure

The participants were welcomed to the lab and given a letter of consent. All consented and then chose, without replacement, one card from a set of four to randomly assign themselves to groups. All participants were trained to use features in nStudy needed to complete the study by watching the first author highlighting using the nStudy software. Then, a PowerPoint presentation was used to train participants in the examples training group.

Participants in the example training groups were shown a 172-word text about Food Deserts and three questions. Two questions required recall, one about the main idea and another about a term. The third question required applying and transferring information found in two segments of the text: an explanation and a consequence. Participants in the examples group read about Food Deserts and then read the first recall question. After reading the question, the segment in the text providing the answer to that question was shaded red in the PowerPoint presentation. The same procedure was followed for the second recall question and then for the transfer question.

All participants were told recall questions require remembering information in the text, and application (transfer) questions require using information they studied to solve a problem or make a prediction.

After training, the participants were instructed to sit at any computer where they accessed a bookmark matching the card chosen previously. Each bookmark accessed one of four "surveys" created using the online questionnaire service FluidSurveys. Each survey corresponded to one cell of the 2×2 design. A "Next" button at the bottom of each webpage allowed participants to progress to the following page. Once it was selected, a participant could not

return to any previous page, thus preventing review. All participants studied the same text and responded to both transfer and recall posttests.

Participants studied for as long as they desired. A target was considered marked if learners marked all or any part of a sentence. When learners finished studying, they took the achievement tests. The study session took approximately an hour. After participants finished both posttests, they were compensated \$10 for their time.

3.6 Data analysis

3.6.1 Scoring and data preparation

Responses for prior knowledge were coded into two categories: related prior knowledge and general prior knowledge. Each valid response was assigned a score in only one of these categories, and total prior knowledge scores were computed by summing related and general scores. Posttest responses for recall and transfer were scored using a predefined rubric ([Appendix C](#)) developed based on key sentences in the reading text that contained relevant information needed to answer each question.

3.6.2 Analysis of text marking behavior

Learners' text marking behaviors were automatically recorded in nStudy, capturing marked sentences, timing, and whether markings corresponded to predefined content categories (e.g., terms, explanations, reasons, examples). The distribution and frequency of markings were analyzed to identify patterns across conditions. Text marking was examined both as an encoding strategy (how marking predicted posttest performance) and as an indicator of cognitive processing (how behaviors varied based on test expectancy and training).

3.6.3 Statistical analyses

To examine differences across experimental conditions, quantitative analyses were conducted to compare marking behaviors, prior knowledge, and learning outcomes. Multivariate and univariate statistical tests were used to assess the effects of test expectancy (recall vs. transfer posttest) and training (examples vs. no examples) on marking behaviors and learning performance. Regression analyses were conducted to explore whether specific marking behaviors predicted recall and transfer scores, and whether prior knowledge influenced these relationships.

3.7 Limitations

This study has several limitations. First, two issues related to the tasks used warrant consideration. Recall question 4 was as follows: In your house, why does a sock fall more slowly than a shoe? This question could be argued to assess generalization beyond recall. The other issue concerns transfer question 3: Two asteroids are travelling in space, where there is no air resistance, from the same starting point at a speed of 18.6 km/s. One is 10,000 k. the other is 15,000 k. which one would reach the Earth's atmosphere first and why? Two participants mentioned they could not tell whether "k" represented kilometers or kilograms. There is a possibility other participants shared this confusion, which may have influenced their responses.

Second, as in many studies, multiple statistical tests were applied to the same sample. We acknowledge this increases the probability of type I error, and results should be interpreted with appropriate caution.

Another limitation is the absence of a control group (i.e., students who did not engage in text marking). Including such a group could have strengthened the study's findings by providing a clearer comparison of the effects of marking on recall and transfer. Future research should consider incorporating a control group to isolate the impact of text marking more effectively.

Although the study discusses fluency illusion as a possible explanation for ineffective marking, it does not directly measure learners' perceptions of their understanding versus actual performance. Incorporating metacognitive judgments, such as confidence ratings, may provide a clearer assessment of the extent to which fluency illusions impact marking effectiveness.

While prior knowledge was accounted for, other individual differences, such as reading proficiency and other cognitive abilities, may have influenced marking behaviors and learning outcomes. Future research should consider controlling for these factors where theory points toward specific individual differences that moderate marking behavior, encoding and retrieval.

The training provided was brief and did not include active feedback or iterative practice. Prior research suggests more extended training with guided feedback could lead to more effective marking strategies. Future research should explore how different levels of training influence marking selectivity and learning outcomes.

It is possible some learners experienced difficulty using the nStudy marking tool, which may have affected their engagement with text marking. Since ease of use was not directly assessed, we cannot determine whether usability issues influenced marking behavior or learning outcomes. Future research should collect learner feedback on their experience with the tool to better understand its impact on study behavior.

Finally, it may be argued that our research questions should posit directional effects on manipulated variables. In light of mixed empirical findings for text marking and somewhat variable theoretical accounts, we treated our study as exploratory and, therefore, did not hypothesize the directionality of effects.

4 Results

Data were examined for normality of distributions and outliers. None of the variables was non-normally distributed (all skewness and kurtosis values ≤ 1.5). Four outliers were identified for the variable total marking. We decided to retain these scores to maximize sample size because other data for these cases were not atypical.

In Table 1, we present for each group the means and standard deviations of the frequency of learners' marking eight categories of information in the text (terms, examples, main ideas, explanations, facts, reasons, consequences, and descriptions) and total marks.

4.1 Effects of orientation to recall or transfer posttests on text marking

Given text marking serves as both an encoding strategy and a trace of cognitive processing, this analysis examines how test expectancy and training influence learners' marking behavior. Research suggests test expectancy shapes study strategies (Finley and Benjamin, 2012; Abd-El-Fattah, 2011), yet its impact on text marking remains unclear. Training in marking strategies has been shown to improve the selection of relevant content (Ponce et al., 2022), but its interaction with test expectancy in guiding marking behavior is less understood. Since prior knowledge enhances recall performance (Bransford and Johnson, 1972) and influences how learners engage with study strategies, we investigated it as a potential moderator of the relationship between test expectancy and marking behavior.

A 2×2 MANOVA was computed with exposure to training examples (yes, no) and orientation to the type of posttest (recall, transfer) as independent variables and the frequency of marking each category of information as dependent variables. Using Pillai's trace as the criterion, there was no statistically detectable effect due to exposure to training examples, $F(8, 136) = 0.846$, $p = 0.564$, partial $\eta^2 = 0.050$. The effect of orienting to study for transfer or recall was statistically detectable, $F(8, 136) = 2.076$, $p = 0.043$, partial $\eta^2 = 0.114$. Compared to learners oriented to study for recall, learners oriented to study for transfer marked more examples $\{F(1, 136) = 3.881$, $p = 0.051$, partial $\eta^2 = 0.028$, 95% CI $[-0.01, 0.90]\}$, main ideas $\{F(1, 136) = 7.348$, $p = 0.008$, partial $\eta^2 = 0.051$, 95% CI $[0.28, 1.76]\}$,

TABLE 1 Number of sentences marked for each sentence category and treatment group.

Sentence categories	Recall-examples		Transfer-examples		Recall-no examples		Transfer-no examples	
	$n = 33$		$n = 35$		$n = 36$		$n = 36$	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Terms	5.87	2.64	6.65	2.71	5.47	2.50	6.25	2.80
Examples	1.60	1.19	2.11	1.60	1.25	1.15	1.63	1.37
Main ideas	3.54	1.85	4.28	2.25	3.00	2.01	4.30	2.68
Explanations	1.06	0.96	1.42	0.94	1.22	1.01	1.08	0.90
Descriptions	0.90	0.97	1.14	1.14	1.00	1.01	0.97	1.15
Facts	2.42	2.35	3.17	2.59	2.47	2.24	3.41	2.98
Reasons	2.30	1.46	2.80	1.47	1.86	1.24	2.52	1.68
Consequences	1.54	1.14	1.20	1.25	1.13	1.04	1.30	1.23
Total Marks	19.27	9.13	22.80	9.94	17.41	9.33	21.50	10.91

and reasons [$F(1, 136) = 5.440, p = 0.021$, partial $\eta^2 = 0.038$, 95% CI [0.09, 1.07]].

An ANOVA was computed to compare total marking across groups oriented to study for recall versus transfer. There was a statistically detectable difference, $F(1, 136) = 5.327$ with $p = 0.022$, $\eta^2 = 0.037$ favoring the group oriented to study for transfer posttest items ($M = 22.14$, $SD = 10.39$, 95% CI [19.67, 24.61]) over the group oriented to study for recall items ($M = 18.30$, $SD = 9.21$, 95% CI [16.08, 20.52]). Providing examples of kinds of information appropriate to answer recall or transfer posttest items did not affect learners' marking. Orienting learners to study for recall or for transfer posttest items did affect total marking as well as some categories of information learners choose to mark.

The correlation between related prior knowledge (knowledge related to ideas in the text) and general prior knowledge (general knowledge about the theme of the text) was $r = 0.884, p < 0.01$. Therefore, related prior knowledge was used in all statistical tests involving prior knowledge, as it more directly reflects information presented in the text. To investigate whether treatments' effects on learners' text marking were moderated by prior knowledge, an ANCOVA was computed with orientation to the type of posttest as the independent variable, related prior knowledge as the covariate, and total marking as the outcome variable. There was a statistically detectable relationship due to related prior knowledge, $F(1, 137) = 9.074, p = 0.003$. Partial out related prior knowledge, the effect of orientation on learners' text marking was statistically detectable, $F(1, 137) = 4.649, p = 0.033$, partial $\eta^2 = 0.034$, 95% CI for adjusted mean difference [0.56, 7.12].

4.2 Effects of treatments on recall and transfer posttests

Although test expectancy has been shown to shape study strategies (Finley and Benjamin, 2012; Abd-El-Fattah, 2011), its impact on text marking and subsequent learning outcomes, particularly transfer, remains unexplored. Similarly, while training in marking strategies has been found to improve recall and comprehension (Ponce et al., 2022), its interaction with test expectancy and its effect on transfer performance have received little attention.

Given that marking supported retention in one study (Ponce et al., 2022) but has not been widely studied in the context of transfer, this analysis explores whether marking behavior predicts additional kinds of learning outcomes. Since prior knowledge enhances recall (Bransford and Johnson, 1972), we also examined its moderating role in the relationship between test expectancy, training, and learning performance.

Table 2 shows performance on the recall and transfer posttests for each treatment group. Scores for recall correlated with transfer scores, $r = 0.477; p < 0.001$ (two-tailed). Therefore, a MANOVA was calculated

with exposure to examples and orientation to study as independent variables and recall and transfer posttest scores as outcome variables. Using Pillai's trace as the criterion, no statistically detectable differences were observed ($p = 0.167$ for exposure to examples, $p = 0.759$ for orientation, $p = 0.656$ for the interaction).

Training learners to identify information categories relevant to answering posttest questions and orienting learners to the type of posttest had no statistically detectable effect on recall or transfer measures of achievement.

A MANCOVA was calculated with exposure to examples and orienting instructions as independent variables, total marking as a covariate, and recall and transfer posttest scores as outcome variables. Using Pillai's trace as the criterion, a statistically detectable difference between groups was observed, $F(2, 135) = 3.09, p = 0.048$, partial $\eta^2 = 0.044$, 95% CI [0.000, 0.101]. As reported in Table 3, a statistically detectable proportional relationship was found between total marking and transfer performance [$F(1, 135) = 5.885, p = 0.017$, partial $\eta^2 = 0.042$, 95% CI [-1.77, 1.95]]. Learners who marked more scored higher on transfer questions.

Another MANCOVA computed with exposure to examples and orienting instructions as independent variables, related prior knowledge as a covariate, and total recall and total transfer as outcome variables. Using Pillai's trace as the criterion, a statistically detectable relationship was found for related prior knowledge: $F(2, 135) = 14.029, p < 0.001$, partial $\eta^2 = 0.173$, 95% CI [0.075, 0.263]. No effect was detected for orienting instructions ($F(2, 135) = 0.478, p = 0.621$) or exposure to examples [$F(2, 135) = 0.598, p = 0.551$]. No interaction was found [$F(2, 135) = 0.307, p = 0.736$].

4.3 Relationship of text marking to recall and transfer posttests

Prior research on text marking suggests that, while it can aid memory retention, its effectiveness depends on how learners engage with the material (Weinstein and Mayer, 1983; Ponce et al., 2022). Studies indicate marking may support recall when key information is selected (Blanchard and Mikkelsen, 1987), but its role in transfer remains unexplored. To better understand this relationship, we examined text marking in terms of incidental marking (marking of non-essential content) and central marking (marking of recall- or transfer-relevant content), allowing us to differentiate between superficial selection and strategic engagement. Given prior knowledge enhances recall and comprehension (Bransford and Johnson, 1972) and learners' existing cognitive structures influence how they interact with and process (Rickards and August, 1975), we controlled for it to determine whether the effects of marking on performance were independent of existing knowledge. Regression analyses were used to investigate whether marking behavior directly predicted recall and transfer performance, addressing gaps in the

TABLE 2 Mean and standard deviation for all groups on total recall and transfer items.

Assessment type	Recall-examples, $n = 33$		Transfer-examples, $n = 35$		Recall-no examples, $n = 36$		Transfer-no examples, $n = 36$	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Total recall	5.42	1.78	5.48	1.78	5.11	2.21	4.55	2.07
Total transfer	10.48	3.77	10.57	3.76	9.94	3.77	9.56	3.19

literature on how marking contributes to higher-order learning outcomes beyond rote memorization (Fiorella and Mayer, 2015). Regression models were computed to predict each of the recall and transfer posttest scores using three predictors: (1) total incidental marks (marked information not relevant to answer recall or transfer questions), (2) recall central information (content needed to answer recall questions), and (3) transfer central information (content needed to answer transfer questions). Marking incidental information predicted scores on transfer $\{F(1, 136) = 4.571, p = 0.034, \beta = 0.179, 95\% \text{ CI } [0.014, 0.344]\}$ but not recall $\{F(1, 136) = 0.998, p = 0.320, \beta = 0.085, 95\% \text{ CI } [-0.083, 0.253]\}$. Marking information central to recall and central to transfer questions predicted learners' recall $\{F(1, 136) = 7.389, p = 0.007, \beta = 0.225, 95\% \text{ CI } [0.063, 0.387]\}$ and transfer $\{F(1, 136) = 6.658, p = 0.011, \beta = 0.215, 95\% \text{ CI } [0.050, 0.380]\}$, respectively. Also, marking information central to recall predicted transfer performance $\{F(1, 136) = 5.536, p = 0.020, \beta = 0.196, 95\%$

CI $[0.031, 0.361]\}$. Table 4 shows correlations among related prior knowledge, general prior knowledge, total marking, and marking of recall central content and transfer central content.

Three regression analyses were computed to investigate the direct effect of total marking on overall achievement, computed as the sum of recall scores plus transfer scores, and the mediated effect of total marking on achievement through related prior knowledge. Since the correlation between related prior knowledge and general prior knowledge was high and differences among their correlations with other variables were trivial, related prior knowledge was only used in the analysis. In Table 5, b and β coefficients for total marking controlling for prior knowledge are 0.048 (SE = 0.039) and 0.098, reduced from 0.100 and 0.041, respectively, when analyzing its direct effect (Meyers et al., 2013). Prior knowledge had a b -value of 0.910 (SE = 0.179) and a β -value of 0.404 when controlling for total marking.

TABLE 3 Result summary of MANCOVA for the interaction of examples/no examples and recall/ transfer and total recall and total transfer and the covariate total marking.

Source	Dependent variable	Type III sum of squares	df	MS	F	p	Partial eta squared	Observed power
Total marking	Total recall	10.324	1	10.324	2.665	0.105	0.019	0.367
	Total transfer	74.981	1	74.981	5.885	0.017	0.042	0.673
Transfer	Total recall	4.201	1	4.201	1.085	0.300	0.008	1.085
	Total transfer	6.437	1	6.437	0.506	0.478	0.004	0.506
Examples	Total recall	11.586	1	11.586	2.991	0.086	0.022	0.404
	Total transfer	15.104	1	15.104	1.186	0.278	0.009	0.191
Transfer \times Examples	Total recall	3.496	1	3.496	0.903	0.344	0.007	0.156
	Total transfer	2.338	1	2.338	0.184	0.669	0.001	0.071
Error	Total recall	522.924	135	3.874				

TABLE 4 Correlations matrix of prior knowledge and text marking.

Variables	Prior related	Prior general	Total marking	Recall central	Transfer central
Prior related	1.000	0.884**	0.263**	0.201*	0.310**
Prior general		1.000	0.263**	0.204*	0.284**
Total marking			1.000	0.848**	0.921**
Recall central				1.000	0.853**
Transfer central					1.000

** $p < 0.01$; * $p < 0.05$; prior related is a composite of prior general; recall central and transfer central sum to total marking.

TABLE 5 Regression analysis.

Model	R ²	Variables	b	SE-b	Beta (β)	Pearson r
1	0.035	Constant	13.239	0.921		
		Total Marking	0.100	0.041	0.204	0.204
2	0.062	Constant	2.098	0.403		
		Total Marking	0.057	0.018	0.263	0.263
3	0.182	Constant	11.330	0.927		
		Total Marking	0.048	0.039	0.098	0.204
		Prior knowledge	0.910	0.179	0.404	0.263

Model 1: Total Marking predicting achievement; Model 2: Total Marking predicting prior knowledge; Model 3: Total Marking and prior knowledge predicting achievement.

Aroian's test (Aroian, 1944) indicated the mediation effect was statistically detectable, $z = 4.617$, $p < 0.05$. A Freedman-Schatzkin test (Freedman and Schatzkin, 1992; $t = 4.816$, $p < 0.05$) indicated the effect of total marking on achievement was statistically reduced when prior knowledge was included as a mediator. The ratio of the indirect to the total direct effect indexes the relative strength of the mediated effect. The indirect effect, computed as the product of path coefficients in the mediated model (0.263×0.404), was 0.106. The total direct effect was 0.204. The ratio of the two is 0.52. Approximately 52% of the effect of total marking is mediated through prior knowledge. As prior knowledge increases, so does marking.

5 Discussion

Our study was designed to answer four main questions: RQ1: Does text marking vary if learners are oriented to expect a recall posttest or a transfer posttest? RQ2: Is there a relation between learners' text marking and their performance on recall and transfer questions? RQ3: Does training to identify information categories relevant to answer recall and transfer posttest items affect learners' text marking? RQ4: Does training affect learners' performance on posttest recall and transfer?

We also examined how prior knowledge shaped the relationship between text marking and learning outcomes, given the critical role of prior knowledge in guiding learners' interactions with text, influencing cognitive processing, and ultimately affecting performance.

5.1 Does text marking vary if learners are oriented to expect a recall or a transfer posttest?

Our findings show the effects of text marking varied depending on test expectancy. Learners who anticipated a recall test marked less text overall and selected fewer targets representing examples, main ideas, and reasons than learners expecting a transfer test. In contrast, learners oriented toward transfer engaged in more marking, possibly because transfer requires integrating information across different contexts, rather than simply retrieving isolated facts. These findings align with prior research on test expectancy effects (Finley and Benjamin, 2012; Abd-El-Fattah, 2011), which suggests learners adjust their study behaviors based on the type of test they expect. Our results extend this work by demonstrating test expectancy not only influences general study strategies but also affects how learners engage with text marking as a cognitive process.

This suggests educators should consider explicitly informing learners about the type of test they will take (e.g., recall vs. transfer) as test expectancy shaped both the quantity and type of information learners choose to mark in this study.

Future research should investigate these effects further and explore whether different types of prompts affect learners' text marking.

5.2 Is there a relation between learners' text marking and their achievement on recall and transfer questions?

Marking text central to recall and transfer items was associated with higher posttest scores. Our findings align with previous research,

which demonstrated marked information is better recalled than non-marked information (e.g., Winne et al., 2015; Yue et al., 2015). Ponce et al. (2022) metaanalysis revealed text marking positively impacts memory and comprehension, with an effect size of 0.36 for memory retention. These results, along with ours, challenge earlier studies that reported no effect of text marking on learners' recall (e.g., Johnson, 1988; Ponce et al., 2018; Peterson, 1992).

An interesting finding of our study is marking information incidental to transfer posttest items predicted better transfer. For instance, consider transfer question 5: A man is in a lift holding a ball. The lift suddenly breaks free, falls from the 30th floor and the man LETS GO OF the ball. Choose the option that best describes what happens after the man LETS GO OF the ball, then justify your choice.

- The ball will float in the lift
- The ball will fall on the floor
- The ball will continue to fall and will never reach the floor of the lift
- The man and the ball will be floating inside the lift

To answer this question, learners need to understand the underlined targets in this extract from the text:

"Weightlessness isn't actually a condition where an object has no weight, it's a condition when there is no stress or strain on an object due to gravity. For example, when you lie on the floor, you "feel" gravity – weight – because the floor reduces your acceleration due to gravity from a standard gravity of 1g to 0g. You aren't falling at all. An object affected by the force of gravity only is said to be in a state of free fall. In free fall, you would feel weightless because you're falling at a rate of 1g, the rate of acceleration due to gravity. Suppose you fall off a diving board. You fall at 1g, and you'd feel weightless".

Although the four unmarked sentences are not directly related to answering this transfer question, marking this kind of information predicted transfer. Transfer tasks, by definition, require learners to apply knowledge to new situations. We speculate content incidental to information central to transfer helps learners elaborate central information that benefits conceptual understanding underlying transfer posttest items. Marking incidental content may support the organization and integration of information and processes that enhance learning, as described in the Select–Organize–Integrate model (Fiorella and Mayer, 2015). These findings suggest instructors might encourage learners to mark not only directly relevant information when preparing for transfer tasks but also related contextual information. Marking such information could help facilitate the integration and organization of knowledge, making it easier to apply to transfer questions and new contexts.

Another possible interpretation is marking text, whether related or incidental, merely correlates with transfer performance and has no causal influence. Nevertheless, incidental marking may still play a role in transfer by helping learners elaborate on and connect key concepts, which could support understanding. We recommend more research to investigate how learners' marking of central and incidental information relates to transfer.

Not unexpectedly, our findings show that recall and transfer performance are correlated. This finding aligns with Bloom's taxonomy

(Kratwohl, 2002), where recall is a foundational cognitive process that supports higher-order processes like transfer. Theoretically, transfer requires retrieving, integrating, and adapting knowledge for new contexts, making recall a necessary but not sufficient condition. However, the differential relationship between marking behavior and recall versus transfer suggests that, while recall underpins transfer, the cognitive processes involved in each appear to be distinct.

Our results also show that orienting learners to study for a specific goal affects the quantity of marking, recall, and transfer achievement. Earlier research examining the effects of test expectancy on learners' choice of study tactics also reported an effect of orienting learners to the type of test on how learners' study. For instance, Feldt and Ray (1989) found learners expecting a multiple-choice test underlined and reread text but did not take notes, while learners expecting a free recall test underlined, reread and took notes. Building on the idea that test expectancy shapes learners' marking behaviors, these findings further suggest instructors should not only inform students about the type of test they will take but also provide guidance on effective marking strategies tailored to different assessment formats. Encouraging learners to strategically mark and engage with material based on their study goals may enhance both recall and transfer.

Considering our findings and those of Feldt and Ray, future research could examine how providing clear goals for learners influences marking behaviors and how changed marking affects recall and transfer tasks.

5.3 Does training to identify information categories relevant to answer recall and transfer posttest items affect learners' text marking?

Providing learners with examples of kinds of information to mark was not enough to influence what they marked. Glover et al. (1980) found extended training, which involved practicing marking with immediate feedback, increased accuracy in marking key information and reduced extraneous marking. Their study examined learners' marking behavior after training, analyzing both what they marked and the accuracy of their markings. Thus, as far as we know, it is the only study reporting findings that can be compared with ours. In contrast, other studies (e.g., Amer, 1994; Hayati and Shariatifar, 2009; Leutner et al., 2007) provided training in marking and included researcher feedback to guide learners' marking choices during training. However, those studies did not systematically examine changes in marking behavior. Instead, they reported the effects of marking training on outcomes. Our finding suggests simply providing examples of key information to mark may not be sufficient. Instead, educators should consider incorporating structured and extended training in marking, including explicit instruction coupled with practice and feedback, to help learners mark more effectively for recall and transfer tasks.

Future research should investigate how training to search for and mark particular kinds of information affects learners' marking and achievement.

5.4 Does training affect learners' performance on posttest recall and transfer?

Briefly showing learners examples of kinds of information relevant to answering recall and transfer posttest items did not affect

recall or transfer. Other studies found benefits for training in comprehension and essay performance. This aligns with research suggesting minimal guidance in marking strategies does not necessarily improve learning outcomes. In contrast, studies that reported benefits of marking training for comprehension and essay performance typically involved (a) a lengthy duration of the training, which ranged between 60 min for one session (Hayati and Shariatifar, 2009) to 90 min per session for five weeks (Amer, 1994); (b) clear steps and explicit suggestions about when to mark and when not to, and how and what to mark (Hayati and Shariatifar, 2009; Amer, 1994; Leutner et al., 2007); and (c) coupling training a text marking tactic with self-regulation (Leutner et al., 2007). None of these features characterized our study, which involved only a brief exposure to examples without direct instruction, practice, or feedback. This may explain the lack of effects for briefly showing examples. We infer learners may need explicit and longer training to mark text effectively, which, in turn, could enhance recall and transfer performance. Building on the earlier suggestion that structured and extended training may be necessary for effective marking, these findings further highlight the importance of integrating explicit instruction, guided practice, and feedback into marking training. Rather than relying on brief exposure to examples, instructors should consider sustained training approaches that incorporate various instructional supports, such as self-regulation strategies, clear marking guidelines, and extended practice, as seen in prior studies. Future research should explore how variations in training duration, instructional scaffolding, and learner support impact the effectiveness of marking as a tool for enhancing recall and transfer performance.

5.5 The role of prior knowledge in text marking and achievement

Prior knowledge was included as a covariate in our analyses, given its well-documented influence on learning outcomes (Simonsmeier et al., 2022). Learners with higher prior knowledge of the topic performed better on recall than learners with less prior knowledge. This accords with the general finding that more prior knowledge improves recall (e.g., Bransford and Johnson, 1972).

Our findings also indicate prior knowledge mediated the effect of marking on both recall and transfer achievement posttests. Specifically, learners with higher prior knowledge benefitted more from marking text than those with lower prior knowledge. This aligns with findings from Annis and Davis (1978), who reported text marking was more effective when learners were already familiar with the topic. These results reflect the Matthew effect, where individuals with greater initial advantages, in this case, more prior knowledge, tend to gain greater advantages than learners with lesser initial advantages (Merton, 1968). Since providing learners with examples of relevant information to mark did not affect their performance in our study, it is unlikely such minimal guidance can compensate for differences in prior knowledge.

Overall, our findings suggest instructors should ensure learners have sufficient prior knowledge of the content to be studied before engaging in text-marking activities. To maximize the benefits of marking, educators should consider ensuring that learners have sufficient knowledge of the topic before engaging in marking tasks. Providing opportunities to build prior knowledge before introducing

marking activities may help learners mark more effectively and improve recall and transfer.

Future research should explore how different instructional approaches can help learners with lower prior knowledge benefit from text marking.

6 Conclusion

Our findings highlight the role of text marking in recall and transfer, demonstrating that both test expectancy and prior knowledge shape marking behaviors and the impact of text marking on learning outcomes. Learners expecting a transfer test marked more, with marking central content enhancing performance. Notably, incidental marking also predicted transfer, suggesting engaging with related but non-essential information supports learning beyond precisely what is marked. Prior knowledge emerged as a key factor, playing an important role in the effects of marking and reinforcing the need for structured training, particularly for learners with lower prior knowledge.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Simon Fraser University Office of Research Ethics. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ZM: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Writing – original draft, Writing – review & editing. PW: Conceptualization, Formal analysis, Funding acquisition,

Methodology, Resources, Software, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Funding used in carrying out the research was received by Professor Philip Winne from the Social Sciences and Humanities Research Council of Canada. Funding for publishing the article is the knowledge mobilization grant provided by Kwantlen Polytechnic University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2025.1510007/full#supplementary-material>

References

- Abd-El-Fattah, S. M. (2011). The effect of test expectations on study strategies and test performance: a metacognitive perspective. *Educ. Psychol.* 31, 497–511. doi: 10.1080/01443410.2011.570250
- Amer, A. A. (1994). The effect of knowledge map and underlining training on the reading comprehension of scientific texts. *Engl. Specif. Purp.* 13, 35–45.
- Anderson, R. C., and Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. *Handbook Read. Res.* 1, 255–291.
- Annis, L. F., and Davis, J. K. (1978). Study techniques: comparing their effectiveness. *Am. Biol. Teach.* 40, 108–110.
- Aroian, L. A. (1944). The probability function of the product of two normally distributed variables. *Ann. Math. Stat.* 18, 265–271.
- Ayer, W. W., and Milson, J. L. (1993). The effect of notetaking and underlining on achievement in middle school life science. *J. Instr. Psychol.* 20, 91–95. (Accessed May 20, 2015).
- Bisra, K., Marzouk, Z., Guloy, S., and Winne, P. H. (2014). A meta-analysis of the effects of highlighting or underlining while studying. Philadelphia, USA: American Educational Research Association.
- Bell, K., and Limber, J. E. (2009). Reading skills textbook marking and course performance. *Lit. Res. Instr.* 49, 56–67. doi: 10.1080/19388070802695879
- Ben-Yehudah, G., and Eshet-Alkalai, Y. (2018). The contribution of text highlighting to comprehension: a comparison of print and digital reading. *J. Educ. Multimedia and Hypermedia* 27, 153–178. Available at: https://www.researchgate.net/publication/321052022_The_contribution_of_text-highlighting_to_comprehension_A_comparison_of_print_and_digital_reading. (Accessed September 15, 2024).
- Biwer, F., Oude Egbrink, M. G., Aalten, P., and de Bruin, A. B. (2020). Fostering effective learning strategies in higher education—a mixed-methods study. *J. Appl. Res. Mem. Cogn.* 9, 186–203. doi: 10.1016/j.jarmac.2020.03.004
- Blanchard, J., and Mikkelsen, V. (1987). Underlining performance outcomes in expository text. *J. Educ. Res.* 80, 197–201. doi: 10.1080/00220671.1987.10885751
- Bransford, J. D., and Johnson, M. K. (1972). Contextual prerequisites for understanding: some investigations of comprehension and recall. *J. Verbal Learn. Verbal Behav.* 11, 717–726. doi: 10.1016/S0022-5371(72)80006-9

- Cho, K. W., and Neely, J. H. (2017). The roles of encoding strategies and retrieval practice in test-expectancy effects. *Memory* 25, 626–635. doi: 10.1080/09658211.2016.1202983
- Clément, G. (2005). *Fundamentals of space medicine*. 2nd Edn. New York: Springer.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Public Interest* 14, 4–58. doi: 10.1177/1529100612453266
- Feldt, R. C., and Ray, M. (1989). Effect of test expectancy on preferred study strategy use and test performance. *Percept. Mot. Skills* 68, 1157–1158.
- Finley, J. R., and Benjamin, A. S. (2012). Adaptive and qualitative changes in encoding strategy with experience: evidence from the test-expectancy paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 632–652. doi: 10.1037/a0026215
- Fiorella, L., and Mayer, R. E. (2015). Eight ways to promote generative learning. *Educ. Psychol. Rev.* 28, 717–741. doi: 10.1007/s10648-015-9348-9
- Fowler, R. L., and Barker, A. S. (1974). Effectiveness of highlighting for retention. *J. Appl. Psychol.* 59, 358–364. doi: 10.1037/h0036750
- Freedman, L. S., and Schatzkin, A. (1992). Sample size for studying intermediate endpoints within intervention trials or observational studies. *Am. J. Epidemiol.* 136, 1148–1159.
- Glover, J. A., Zimmer, J. W., Filbeck, R. W., and Plake, B. S. (1980). Effects of training students to identify the semantic base of prose materials. *J. Appl. Behav. Anal.* 13, 655–667.
- Hayati, A. M., and Shariatifar, S. (2009). Mapping strategies. *J. Coll. Read. Learn.* 39, 53–67. doi: 10.1080/10790195.2009.10850318
- Johnson, L. L. (1988). The effects of underlining textbooks sentences on passages and sentence retention. *Read. Res. Instr.* 28, 18–32.
- Jonassen, D. H. (1984). Effects of generative text processing strategies on recall and retention. *Human Learn.: J. Practical Res. Applications* 84, 182–192.
- Karpicke, J. D., Butler, A. C., and Roediger, H. L. (2009). Metacognitive strategies in student learning: do students practise retrieval when they study on their own? *Memory* 17, 471–479. doi: 10.1080/09658210802647009
- Klauser, K. J. (1984). Intentional and incidental learning with instructional texts: a meta-analysis for 1970–1980. *Am. Educ. Res. J.* 21, 323–339.
- Koriat, A. (2006). *Metacognition and consciousness*. Haifa, Israel: Institute of Information Processing and Decision Making, University of Haifa, 289–325.
- Kornell, N., Rhodes, M. G., Castel, A. D., and Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: dissociating memory, memory beliefs, and memory judgments. *Psychol. Sci.* 22, 787–794. doi: 10.1177/0956797611407929
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: an overview. *Theory Pract.* 41, 212–218. doi: 10.1207/s15430421tp4104_2
- Kulhavy, R. W., Dyer, J. W., and Silver, L. (1975). The effects of notetaking and test expectancy on the learning of text material. *J. Educ. Res.* 68, 363–365. doi: 10.1080/00220671.1975.10884802
- Leutner, D., Leopold, C., and Den Elzen-Rump, V. (2007). Self-regulated learning with a text highlighting strategy. *Zeitschrift für Psychologie/J. Psychol.* 215, 174–182. doi: 10.1027/0044-3409.215.3.174
- McDaniel, M. A., Blischak, D. M., and Challis, B. (1994). The effects of test expectancy on processing and memory of prose. *Contemp. Educ. Psychol.* 19, 230–248.
- Merton, R. K. (1968). The Matthew effect in science: the reward and communication systems of science are considered. *Science* 159, 56–63.
- Meyers, L. S., Gamst, G. C., and Guarino, A. J. (2013). *Performing data analysis using IBM SPSS*. Hoboken, NJ: John Wiley & Sons.
- Minnaert, A. E. (2003). The moderator effect of test anxiety in the relationship between test expectancy and the retention of prose. *Psychol. Rep.* 93, 961–971. doi: 10.2466/pr0.2003.93.3.961
- Miyatsu, T., Khuyen, N., and McDaniel, M. (2018). Five popular study strategies: their pitfalls and optimal implementations. *Perspect. Psychol. Sci.* 13, 390–407. doi: 10.1177/1745691617710510
- Morehead, K., Rhodes, M. G., and DeLozier, S. (2016). Instructor and student knowledge of study strategies. *Memory* 24, 257–271. doi: 10.1080/09658211.2014.1001992
- Naveh-Benjamin, M., Craik, F. I., Perretta, J. G., and Tonev, S. T. (2000). The effects of divided attention on encoding and retrieval processes: The resiliency of retrieval processes. *Q. J. Exp. Psychol.* 53, 609–625. doi: 10.1080/027249800410454
- Nist, S. L., and Hogrebe, M. C. (1987). The role of underlining and annotating in remembering textual information. *Read. Res. Instr.* 27, 12–25. doi: 10.1080/19388078709557922
- Nist, S. L., and Kirby, K. (1986). Teaching comprehension and study strategies through modeling and thinking aloud. *Read. Res. Instr.* 25, 254–264.
- Peterson, S. E. (1992). The cognitive functions of underlining as a study technique. *Read. Res. Instr.* 31, 49–56.
- Ponce, H. R., Mayer, R. E., Loyola, M. S., López, M. J., and Méndez, E. E. (2018). When two computer-supported learning strategies are better than one: an eye-tracking study. *Comput. Educ.* 125, 376–388. doi: 10.1016/j.compedu.2018.06.024
- Ponce, H. R., Mayer, R. E., and Méndez, E. E. (2022). Effects of learner-generated highlighting and instructor-provided highlighting on learning from text: a meta-analysis. *Educ. Psychol. Rev.* 34, 989–1024. doi: 10.1007/s10648-021-09654-1
- Rickards, J. P., and August, G. J. (1975). Generative underlining strategies in prose recall. *J. Educ. Psychol.* 67, 860–865. doi: 10.1037//0022-0663.67.6.860
- Rivers, M. L., and Dunlosky, J. (2021). Are test-expectancy effects better explained by changes in encoding strategies or differential test experience? *J. Exp. Psychol. Learn. Mem. Cogn.* 47, 195–207. doi: 10.1037/xlm0000949
- Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., and Schneider, M. (2022). Domain-specific prior knowledge and learning: a meta-analysis. *Educ. Psychol.* 57, 31–54. doi: 10.1080/00461520.2021.1939700
- Smith, B. D. (1985). *Bridging the gap: College reading*. Scott, Foresman and Company.
- Soderstrom, N. C., and Bjork, R. A. (2015). Learning versus performance: an integrative review. *Perspect. Psychol. Sci.* 10, 176–199. doi: 10.1177/1745691615569000
- Surber, J. R. (1992). The effect of test expectation, subject matter, and passage length on study tactics and retention. *Lit. Res. Instr.* 31, 32–40.
- Tullis, J. G., Finley, J. R., and Benjamin, A. S. (2013). Metacognition of the testing effect: guiding learners to predict the benefits of retrieval. *Mem. Cogn.* 41, 429–442. doi: 10.3758/s13421-012-0274-5
- Weinstein, C. E., and Mayer, R. E. (1983). The teaching of learning strategies. *Innovation Abstracts* 5:2–4.
- Winne, P. H., and Jamieson-Noel, D. (2002). Exploring students' calibration of self-reports about study tactics and achievement. *Contemp. Educ. Psychol.* 27, 551–572. doi: 10.1016/S0361-476X(02)00006-1
- Winne, P. H., Marzouk, Z., Ram, I., Nesbit, J. C., and Truscott, D. (2015). Effects of orienting learners to study terms or processes on highlighting and achievement. Limassol, Cyprus: European Association for Learning and Instruction.
- Winne, P. H., Teng, K., Chang, D., Lin, M. P.-C., Marzouk, Z., Nesbit, J. C., et al. (2019). nStudy: software for learning analytics about processes for self-regulated learning. *J. Learn. Analytics* 6, 95–106. doi: 10.18608/JLA.2019.62.7
- Yue, C. L., Storm, B. C., Kornell, N., and Bjork, E. L. (2015). Highlighting and its relation to distributed study and students' metacognitive beliefs. *Educ. Psychol. Rev.* 27, 69–78. doi: 10.1007/s10648-014-9277-z