https://dx.doi.org/10.20880/kler.2023.58.5.5

Structural Relations between Processing Variables and Text Quality while Reading and Writing from Multiple Documents

Chang, Sungmin Inha University
Department of Korean Language Education
Assistant Professor

* Multiple document literacy has been a rapidly expanding theme in South Korea over the past decade (e.g., Chang, 2017; Choi, 2013; Lee, 2013; Oh, 2020). This study aims to introduce theoretical and empirical research developed in South Korea to overseas researchers. It primarily cites international literature to enhance academic communication with them.

I. Introduction

Multiple-document processing is essential for learners to explore or make decisions on complex issues for which obtaining necessary information from a single source is not enough (Britt & Rouet, 2012). It is crucial to position, reinforce, evaluate, and organize ideas across multiple texts to comprehensively understand a topic (Anmarkrud et al., 2014; Goldman et al., 2012; Hagen et al., 2014; List & Alexander, 2017). Only reflective and adaptive learners with excellent self-regulating skills can achieve their intended learning outcomes (Stadtler, 2017). Therefore, it is necessary to systematically understand how the sub-variables of cognitive processing (e.g., planning, selecting documents, justifying own perspectives, and revision) generally influence each other while reading and writing from multiple documents.

Several models have been applied to systematically analyze the variables affecting multiple-document processing (Barzilai & Eshet-Alkalai, 2015; Bråten, Anmarkrud et al., 2014; Bråten, Ferguson et al., 2014; Kobayashi, 2009; List, 2021; Mateos et al., 2020; Stang Lund et al., 2019). However, despite the consensus on the importance of component skills underpinning cognitive processing while reading and writing from multiple documents, less attention has been paid

to the structural relationship between the subprocess variables and their outcomes. Most of the variables that these studies focused on belonged to the individual differences category (e.g., Bråten, Anmarkrud et al., 2014). Yet, predicting the results of multiple-document processing from individual difference variables (personal and intrinsic features such as working memory, epistemic beliefs, and gender) that are not observed during reading and writing from multiple documents has limitations in finding educational and practical implications.

This study aims to comprehensively and empirically analyze the structural relationship between the subprocess variables and their outcomes (i.e., written text product) while reading and writing from multiple documents based on two pillars: the Multiple-Document Task-based Relevance Assessment and Content Extraction (MD-TRACE) model of Britt and Rouet (2012) and Activity Tendency of Interdisciplinary Learning Literacy Questionnaire (ATILL-Q) of Chang (2021). The MD-TRACE model offers a useful framework for illustrating the subprocess variables involved in reading and writing from multiple documents, as problem-solving during purposeful reading overlaps with the strategic process of task-based writing (see RESOLV model, Rouet et al., 2017). ATILL-Q was developed based on the MD-TRACE model, and 16 items across five factors were ultimately validated for South Korean graduates.

II. Theoretical Background

1. MD-TRACE model

Britt and Rouet's (2012) MD-TRACE model is a representative model that systematizes a list of component skills that support cognitive processing while reading and writing from multiple documents. It has two sub-models: the task and documents models.

According to the task model, learners create task goals, action plans, and criteria while reading and writing from multiple documents. Learners also monitor their own progress, and if the product fails to meet the task goals, they may decide to continue reading and writing or change their goals. The effect of the task model on text quality has been a major issue in a series of cognitive processing theories (Bereiter & Scardamalia, 1987; Flower, 1987; List & Alexander, 2017; Rouet et al., 2001). The task model explains behavior in comprehending and using sources, such as text selection, duration of text access, document information access, and performance outcomes (List & Alexander, 2017).

The documents model consists of an intertext model and an integrated mental model. According to the intertext model, learners focus on document nodes (i.e., source information) and links (i.e., intertextual predicates), whereas according to the integrated mental model, they focus on the content information of each document and its coherent organization. The intertext model relates to the sources among external documents that learners attend to and identify and the evaluation of the sources' contributions. If the intertext model is not adequately performed, information overload or misinformation may occur (Braasch et al., 2013).

The integrated mental model is concerned with constructing coherent meaning by reconciling contradictions or discrepancies between documents. While the intertext model is a bottom-up process centered on comprehending external documents, the mental model is an internal strategic process centered on its use. Strategic processing during reading is ultimately externalized through coherent representation (e.g., argumentative writing; Anmarkrud et al., 2014; Barzilai et al., 2015; De La Paz & Felton, 2010; List et al., 2019). If the mental model is not strategically implemented, the learned content from multiple sources may remain a fragmented mixture (Goldman, 2004).

Britt and Rouet (2012) suggested three important decision steps supporting the construction of task and document models: (a) as-

sessing information needs, (b) selecting relevant material, and (c) assessing the product. Here, (a) and (c) mainly correspond to the task model and (b) to the documents model. In particular, the reader of the intertext model can also be the author of the integrated mental model. Creating the documents model is situated in the writing tasks, whereas comprehending the sources is a function of the rhetorical task as it is set.

2. ATILL-Q

Chang's (2021) ATILL-Q scale was developed based on the theoretical framework of the MD-TRACE model. Constructs were mainly established from the task model (i.e., planning and revision) and the documents model (i.e., intertext and mental models). This scale was validated through exploratory and confirmatory factor analyses.

The most significant difference, as validated, between ATILL-Q and the MD-TRACE model is that "considering counterarguments" appears as an independent factor. A new factor was created by linking the assessment of information needs in view of the audience (among the task model) and writing with consideration for counterarguments (among the integrated mental model). Although the element "considering counterarguments" is not a part of the MD-TRACE model, it can be included in this model (especially if the task is to argue) as multiple-document processing fundamentally presupposes a plural and interactive nature in which numerous facts, knowledge, claims, and opinions are represented and communicated with each other. Learning from counterarguments in disciplinary classrooms has been termed "multiperspectivism," for instance, in history learning (e.g., Hansen, 2011; Nygren et al., 2017).

Adding a new factor has led to the interpretation of the existing constructs derived from the MD-TRACE model to differ from the original intent of Britt and Rouet (2012). The meaning of some factors is partially limited; for instance, considering the audience in the task

model is not considering the contested perspective but rather taking into account their general understanding and interests. Furthermore, argumentation in the integrated mental model is to logically justify the writer's unique perspectives (i.e., persuasive argument), which is distinct from explicitly considering counterarguments (i.e., reflective or deliberative argument; Felton et al., 2009; Klein & Ehrhardt, 2015; Mateos et al., 2020; Nussbaum & Edwards, 2011). This separation contributes to a detailed capture of the difference between one-sided and two-sided communication that can be observed in multiple-document processing.

Table 1 describes the factors of multiple-document processing based on the MD-TRACE model and ATILL-Q.

Table 1. Factors of multiple-document processing

Factor		Description
Task model	Planning	Creating task goals and assessing information needs by considering attitudes toward the topic, expected cognitive products, and audience
	Revision	Assessing the product by comparing the written and intended text and taking appropriate actions after completing the first draft
Documents model	Intertext model	Selecting documents according to source features and evaluating their relevance to the task model
	(Integrated) mental model	Coherently articulating one's own perspectives by reconciling conflicts or inconsistencies between documents
Considering counterarguments		Considering alternative interpretations in balance and integrating them through appropriate responses

3. Present study

Based on the above theoretical background, this study sets subprocess variables while reading and writing from multiple documents, establishes a hypothesized model, and empirically verifies it by applying a structural equation model. In previous studies, path analysis

was applied to analyze the structural relationships between individual differences, processing variables, and multiple-document comprehension (e.g., Bråten, Anmarkrud et al., 2014). However, attention has rarely been paid to the relationships among subprocess variables, such as the task model, intertext model, and integrated mental model affecting the text quality as a result of reading and writing from multiple documents. Recently, there have been cases in which multiple-document processing frameworks were comprehensively proposed according to subprocess phases (e.g., List & Alexander, 2017, 2019; Richter & Maier, 2017; Rouet et al., 2017); however, how each phase structurally affects or is structurally affected has not been statistically identified.

The dependent variable is the text quality of essays produced by learners, which serves as external evidence of their multiple-document processing. Data were collected from high school students in South Korea who wrote argumentative essays on a given topic using multiple documents. Synthesis writing, such as essays, is considered to best explain deep learning (Cerdán & Vidal-Abarca, 2008; De La Paz & Felton, 2010; Du & List, 2021; Mateos et al., 2020; Wiley & Voss, 1999). Argumentative essays, in particular, are a writing product that externalizes the cognitive and affective results of multiple-document processing, with more selective cognitive properties and less personalized affective properties (List & Alexander, 2017).

The predictor variables were measured by assessing the degree of agreement with the modified ATILL-Q on a five-point scale. These reflect participants' habits and behavioral dispositions that indicate the deep-level strategies of multiple-document processing (Anmarkrud et al., 2014; Hagen et al., 2014; List et al., 2019).

The study aimed to resolve the following research questions (RQs) through verification of the hypothesized model:

RQ 1. Which variables have a direct effect on text quality?

RQ 2. How does the task model and documents model influence each

other?

RQ 3. How do the two variables of the task model, planning and revision, influence each other? Through which path do they have an indirect effect on text quality?

RQ 4. How do the two variables of the documents model, the intertext model and the mental model, influence each other? Through which path do they have an indirect effect on text quality?

RQ 5. How does the ability to consider counterarguments interact with other variables? Through which path do they have an indirect effect on text quality?

Regarding RQ1, it is assumed that the task model, documents model, and consideration of counterarguments will have a direct effect on text quality. This assumption is also the basic premise of the MD-TRACE model and ATILL-Q.

Regarding RQ 2, it is hypothesized that the documents model, externally identified through reading and writing behavior, will be influenced by the task model performed in the readers'/writers' minds. This assumption is supported by cognitivist theories (e.g., Anderson, 1985; Bereiter & Scardamalia, 1987; Flower & Hayes, 1980) that view reading and writing as meaning-construction processes. It is also related to a series of studies (Baaijen & Galbraith, 2018; De Milliano et al., 2012; Kieft et al., 2006; Limpo et al., 2014) that empirically analyzed the effects of planning, monitoring, evaluating, and revision on text quality. Furthermore, the task model is expected to have an indirect effect on text quality through the mediation of the documents model.

Regarding RQ 3, it is assumed that the two variables of the task model, planning and revision, will have a direct effect on each other. This assumption is supported by the fact that planning and revision are two axes of self-regulatory processing that are closely related and interact with each other (Galbraith & Torrance, 2004; Hayes, 2012; Rijlaarsdam et al., 2004; Zimmerman & Kitsantas, 2007). It is expected

that, in some cases, the level or pathway of impact between planning and revision that affects text quality may be different.

Regarding RQ 4, it is assumed that the intertext model will influence the integrated mental model within the documents model. This assumption is supported by the fact that creating the documents model is situated in the writing task; that is, comprehending the sources is a function of the rhetorical task as it is set (Lee & List, 2023; Mateos et al., 2020). Furthermore, since the dependent variable is the written text with coherent representations, such as essays, and it is common to first read sources and establish one's perspective regardless of writing skills in multiple-document processing (Gebril & Plakans, 2009), it is assumed that the task model, intertext model, and considering counterargument together influence text quality through the mediation of the mental model.

Finally, regarding RQ 5, it is assumed that considering counterarguments affects the mental model (i.e., text-generating behavior) and has direct and indirect effects on text quality. The results of Du & List (2021) support these hypotheses, suggesting that learners engage in conflict and achieve resolution more often than only recognizing the conflict and doing nothing. In contrast, the intertext model is relatively more deliberate, and its purpose is unclear (List, 2021); therefore, it is not expected to directly affect or be affected by considering counterarguments (even if these two can have a common effect on text quality). Moreover, since the argument is the core schema of self-regulated learning that decides the relevance between documents and between documents and task goals (Bråten & Strømsø, 2009; Felton et al., 2009; Klein & Ehrhardt, 2015; Mateos et al., 2020; Nussbaum & Edwards, 2011; Wiley et al., 2009), we can additionally expect considering counterarguments to be affected by the task model.

Figure 1 shows a model that depicts the paths of structural relations and a model that describes general procedures of actual performance. The sub-procedures in the latter model are reciprocally connected through the mediation of the monitoring process, including

the consideration of counterarguments.

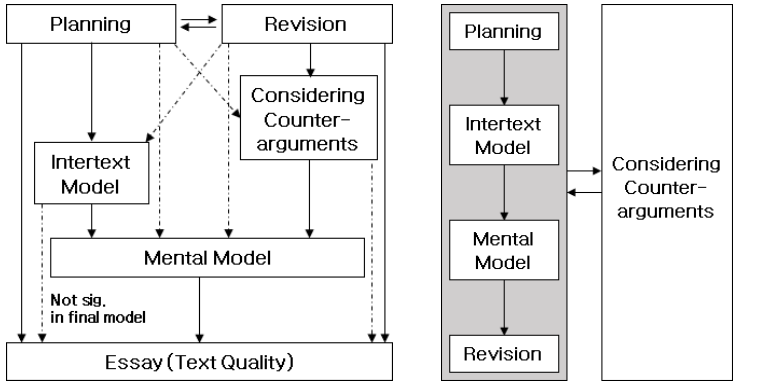

Figure 1. Hypothesized model

As can be seen from Figure 1, the paths of structural relations among the subprocess variables are not identical to the actual performance procedures. For example, in practice, revision is generally observed after the intertext and integrated mental model (i.e., drafting). However, in examining the structural relationships, the task model bearing non-visible properties influences writing products through mediation by the visually implemented documents model. This is associated with multiple-document processing bearing cyclical properties as a form of self-regulated learning (Burin et al., 2020).

III. Method

1. Participants

The participants consisted of 421 third-year students (62.2% women) from eight high schools in South Korea. It can be assumed that

they generally had sufficient prior knowledge related to topics in the regular curriculum of high school first- and second-graders. This design is based on the implications of previous research that identified the need for prior teaching to enable a minimum level of knowledge necessary to overcome the challenges of multiple-document processing (Gil et al., 2010).

General high schools in districts with relatively homogeneous socioeconomic status (i.e., middle class) were selected as partner schools for participant recruitment. All participants were native Korean speakers. We contacted the partner schools and recruited students with the approval of principals and teachers. The entire process of application and obtaining consent of the participants was conducted in the researcher's presence to avoid any vertical pressure from the teachers at the partner school; teachers at the partner school were not involved in anything other than the initial announcement regarding the recruitment of participants. Participants provided written informed consent. Following data collection, they were given an online gift certificate as a reward. Participants' anonymity was guaranteed throughout data collection.

2. Materials

1) Essay task

The dependent variable, "text quality," is the external output produced by learners as evidence of multiple-document processing. Essay writing was aimed at evaluating the ability to analyze and compare the given documents and to logically state claims by using them, focusing on topics that become social issues while dealing with various subjects in the high school curriculum. Participants were asked to write an argumentative essay on how to accommodate cultural diversity using 16 documents embodied in a digital environment (i.e., blog posts).

Participants were instructed to write 1,200 Korean characters

within 80 minutes. They were also instructed to use word processing software to allow for new representations to be created during the writing and immediately reflect without burdening working memory.

Participants were instructed not to submit the title of the essay. This decision was made considering that the title may not necessarily summarize the content of the text, and metaphorical expressions may be used depending on the author's personal writing style. Additionally, setting a title in advance may obstruct the development of the essay to be written later. Participants were asked to write a complete essay with an introduction, body, and conclusion so that the content, organization, and expression criteria of the text could be evaluated. Copying the text of the documents verbatim was forbidden to highlight the attributes of writing as distinct from reading.

2) Documents

The documents were divided into four categories based on their content relevance and author expertise (i.e., more relevant/higher expertise, more relevant/lower expertise, less relevant/higher expertise, and less relevant/lower expertise). A total of 16 documents were selected, with four texts in each category. Regarding content relevance, the texts were divided into those dealing with cultural diversity and cultural conflicts related to the essay task and those on topics that were not directly related to the essay task, such as social change, globalization, (anti-)terrorism, people with disabilities, and gender conflict. Regarding author expertise, following the experimental designs of McCrudden et al. (2016) and Bråten, McCrudden et al. (2018), texts by experts (professors) and journalists were presented separately, and the perspectives and developments between documents classified according to author expertise were balanced. Participants were allowed to conduct internet searches to explore the authors' backgrounds further.

Instead of providing only two documents with conflicting relationships that show opposite perspectives on the same issue (i.e.,

a single, central conflict) as in previous studies (e.g., Barzilai et al., 2018; Primor & Katzir, 2018), several documents linked in diverse relationships were provided together. This decision was based on empirical insights that each participant may have different points of interest in the documents and that the task may not be accepted as designed by the researchers (e.g., Du & List, 2021). To avoid being influenced by the arrangement order of the documents, which can be a superficial cue that can distort the experimental results (Rouet et al., 2011), the order was divided into four modes, and participants were randomly assigned to one of them.

Documents were extracted from among the original texts and partially revised to balance the number of characters and text features (i.e., text complexity, level of difficulty, and readability). To maintain the flow of traditional studies on text features (Bormuth, 1966; Dale & Chall, 1948; Flesch, 1948), some criteria were established, focusing on the quantitative factors working at the word or sentence level of the text surface layer. The number of characters in the 16 documents ranged from 968 to 2,261, the average sentence length ranged from 56.70 to 77.64 characters, and the highest grade word ratio ranged from .09 to .14, which do not deviate significantly from the range of middle school information texts (number of characters 586-4,354, average sentence length 40.11-76.69 characters, and the highest grade word ratio .05-.14, Choi et al., 2021).

3) Text quality measure

Text quality was measured through the analytical scoring of essays written by the participants. The quality of the writing output is equated with the outcomes of multiple-document processing because the text is essentially an artifact constructed to convey meaning (Braasch & Bråten, 2017). When writing an essay, it is important to consider not only content but also organization and expression and to effectively convey one's knowledge to the audience. This corresponds to Klein et al.'s (2016, p. 243) definition of knowledge-trans-

forming through discourse as "a dialogue between rhetorical and content problem-solving." Achieving deep learning about topics through multiple-document processing is not limited to accurately recalling factual information but also involves rhetorically connecting them, translating them into one's own language, and communicating them to others. Thus, the organization and expression, as well as the content, of the writing output are evidence of the learners' efforts to systematically discuss the topic.

Accordingly, we used overall writing skills, including organization and expression, as the evaluation criteria instead of solely using items related to multiple-document processing (such as argumentative reasoning), as done in other studies (e.g., Anmarkrud et al., 2014). We measured text quality by referring to a previously developed scoring rubric for analytically scoring participants' academic writing (Jacobs et al., 1981; Leki & Carson, 1997; Weigle, 2002; Weir, 1990). We measured constructs using the following criteria: reasonableness and clarity of claims, relevance and concreteness of reasons and evidence, novelty of approach (corresponding to content), structure of the entire text and paragraphs (corresponding to organization), adequacy of vocabulary and sentences, and mechanical accuracy (corresponding to expression). This evaluation framework targets writings by authors in their mother tongue and uses multiple documents as the context of the task.

Six experts scored each essay twice on a three-point scale, and the average scores were calculated. The six experts consisted of middle and high school teachers and researchers majoring in Korean L1 education in graduate school. To reduce differences in the interpretation of the criteria, an evaluation table with specific black marks for each item was employed (see Chang, 2022), and the understanding of the criteria was shared in a scorers' workshop prior to the actual scoring. The reflection ratios of content, organization, and expression to the total score were established at 50%, 30%, and 20%, respectively; however, they were not disclosed to the scorers to avoid distorted

results. After the individual scoring by three experts in the first round, three other experts re-scored the essays in the second round with a score deviation of two points for each criterion. The reliability between scorers (intraclass correlation coefficient) by item was .699-.920 and .750-.943 for the first and second rounds of scoring, respectively. The internal reliability (Cronbach's α) for each sub-criterion (consisting of two to three items) corresponding to content, organization, and expression was .884, .899, and .785, respectively, and the internal reliability (Cronbach's α) of the total seven items was .916. Since the scores of sub-constructs are formative indicators that affect the total score in different proportions, they were ultimately replaced by observed variables on a five-point scale.

4) Processing variables measure

A modified version of ATILL-Q was used to measure the processing variables. It is assumed that the higher the score, the stronger the multiple-document literacy tendency. The participants responded to ATILL-Q after writing the essay to prevent the self-explanation effect on them during reading and writing (Muñoz et al., 2006; VanLehn et al., 1992).

This study used a modified version of ATILL-Q—a measure of processing variables—that included three additional items, resulting in a total of 19 items: four items for planning, four for revision, four for the intertextual model, three for the mental model, and four for considering counterarguments. The original version, developed by Chang (2021), consisted of 16 items across five factors. To adhere to recommended measurement theory practices (DeVellis, 2016), the modified version added two items for revision and one for the integrated mental model, which originally had only two items each. The order of the items was arranged to aid the participants in recalling their performance in reading and writing from multiple documents, starting with planning (items 1-4), followed by intertextual model (items 5-8), mental model (items 9-11), considering counterarguments (items 12-15), and ending with revision (items 16-19). Each item was

measured on a five-point scale, with one indicating complete disagreement and five indicating complete agreement. All items for the modified ATILL-Q can be found in Appendix.

The internal consistency reliability of the observed variables for each latent variable (Cronbach's α) was .855 for planning, .888 for the intertext model, .856 for the mental model, .916 for considering counterarguments, and .900 for revision. The overall internal consistency reliability (Cronbach's α) was .919.

3. Procedure

Participants read multiple documents and wrote essays for 80 minutes in their own digital space while connected to the Zoom meeting. The participants were divided into 16 groups, with the number of members in each group ranging from 7 to 39. The essays were written in Korean using word processing software. The writing format of the essay task and the online link address of the blog were shared with the members via a group chat on Zoom. Participants submitted their essays to the researcher through individual chatting. Prior instruction was given to the participants on how to instantly check the number of written characters in real time.

After writing the essay for 80 minutes, participants freely responded to the 19 items of ATILL-Q, without time constraints, by accessing the website created through a Google survey. There was no missing or outlier data.

IV. Results

Table 2 presents the descriptive statistics (mean, standard deviation, skewness, and kurtosis) of the observed variables. The scores for all items are normally distributed.

Table 2. Descriptive statistics

Latent variables (Cronbach's α)	Observed variables	M	SD	Skewness (SE)	Kurtosis (SE)
Planning (α = .855)	Q1	3.66	.55	.034(.119)	-.755(.237)
	Q2	3.53	.62	.702(.119)	-.472(.237)
	Q3	3.75	.82	-.878(.119)	.332(.237)
	Q4	3.77	.51	-.297(.119)	-.164(.237)
Intertext model (α = .888)	Q5	3.84	.82	-.673(.119)	.201(.237)
	Q6	3.62	.80	.304(.119)	-.554(.237)
	Q7	3.95	.66	-.555(.119)	1.008(.237)
	Q8	3.87	.70	-.319(.119)	.135(.237)
Mental model (α = .856)	Q9	3.79	.62	.157(.119)	-.526(.237)
	Q10	3.88	.52	-.143(.119)	.425(.237)
	Q11	3.68	.62	-.627(.119)	.484(.237)
Considering counterarguments (α = .916)	Q12	3.13	.68	-.030(.119)	1.096(.237)
	Q13	3.20	.74	-.105(.119)	.021(.237)
	Q14	3.16	.84	-.512(.119)	.671(.237)
	Q15	3.37	.85	-1.006(.119)	.931(.237)
Revision (α = .900)	Q16	3.59	.77	-.685(.119)	-.072(.237)
	Q17	3.48	.89	-.338(.119)	-.690(.237)
	Q18	3.33	.75	-.092(.119)	-.308(.237)
	Q19	3.70	.64	-.237(.119)	.407(.237)
Text quality		3.32	.78	-.186(.119)	-.883(.237)

Table 3 displays the results of the exploratory factor analysis conducted to determine whether the intended factor structure is present in the modified ATILL-Q. The maximum likelihood method with Promax rotation was used. The factor loadings, intra-factor reliability (Cronbach's α), and squared multiple correlations (SMC) were examined. The loadings for classified factors exceeded .6, while that for

unclassified factors were less than .3, and the SMC exceeded .6. The reliability never increased after a specific item was removed.

Table 3. Exploratory factor analysis for the modified ATILL-Q

Item number	Factor loadings					SMC
	1	2	3	4	5	
Q14	.932	-.042	-.156	.090	.003	.809
Q13	.928	-.012	-.059	-.053	.074	.819
Q15	.817	.046	.096	.083	-.100	.844
Q12	.741	.026	.091	-.012	-.026	.651
Q7	-.038	.946	.026	.004	.164	.893
Q8	-.032	.884	.067	.030	.063	.842
Q6	-.074	.720	-.169	.183	-.113	.669
Q5	.172	.677	.098	-.086	-.111	.745
Q1	-.183	-.015	.909	.125	-.081	.759
Q3	.095	.213	.840	-.185	.016	.714
Q4	.062	-.004	.815	-.012	.003	.776
Q2	.005	-.202	.628	.221	.207	.766
Q16	.043	.017	.095	.992	-.183	.876
Q19	-.029	.058	.083	.781	.117	.875
Q18	.036	.052	-.125	.705	.124	.695
Q17	.104	.040	-.005	.680	.091	.794
Q10	.127	-.088	.123	-.144	.843	.764
Q11	-.089	-.035	-.005	.108	.821	.706
Q9	-.038	.139	-.081	.042	.799	.657
Eigen value	6.838	.959	2.286	1.822	.923	
Expl. Var. (%)	35.991	10.311	12.033	9.590	4.858	
Cum. Expl. Var. (%)	35.991	46.302	58.335	67.926	72.784	

KMO = .814, Bartlett's χ^2 = 7428.347 (df = 171, $p < .001$)

Table 4 presents the correlations between the factors (i.e., latent variables).

Table 4. Correlations between factors

	1	2	3	4	5
1. Planning	1.000	.587	.168	.322	.415
2. Intertext model	.587	1.000	.215	.442	.477
3. Mental model	.168	.215	1.000	.386	.412
4. Considering counterarguments	.322	.442	.386	1.000	.557
5. Revision	.415	.477	.412	.557	1.000

Verification of the measurement and structural model was conducted using IBM's SPSS Amos 22 software package. Parameter estimation followed the maximum likelihood method. Model fit was determined by the chi-square test, which should be non-significant, and the Tucker-Lewis index (TLI), comparative fit index (CFI), and root mean square error of approximation (RMSEA).

Estimation of the parameters of the measurement model for the five predictors by confirmatory factor analysis (in Figure 2) demonstrated favorable values at $\chi^2 = 357.414$ (df = 142, $p < .001$), TLI = .965, CFI = .984, and RMSEA = .060 (90% CI = .050-.071). When the arrow between planning and mental model (which is non-significant) is removed, the model fit is $\chi^2 = 358.93$ (df = 143, $p < .001$), TLI = .965, CFI = .983, and RMSEA = .060 (90% CI = .050-.070). The composite reliability and average variance extracted exceeded .7 (.913-.956) and .5 (.740-.843), respectively, demonstrating good convergent validity. The coefficients of determination between latent variables were all lower than the average variance extracted from disparate latent variables, demonstrating favorable discriminant validity.

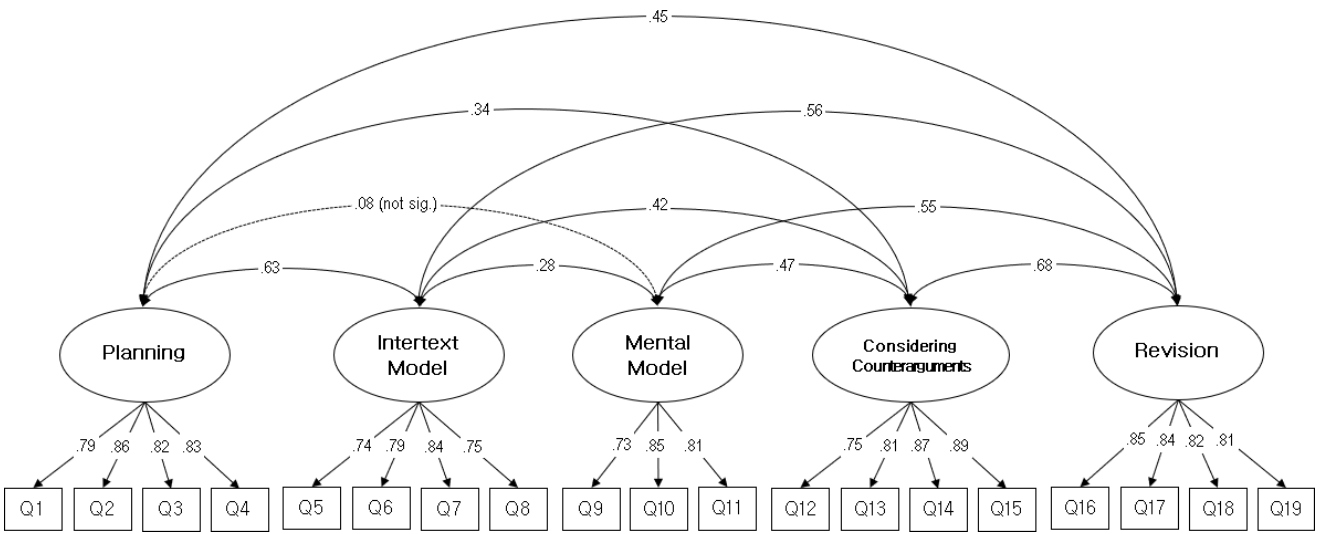

Figure 2. Standardized coefficients for the measurement model ($p < .001$)

When the data were applied to the hypothesized model, seven out of 13 causal paths were statistically significant, while the covariance between planning and revision was also significant. However, because six hypothesized paths (Planning → Mental model, Planning → Considering counterarguments, Intertext model → Text quality, Considering counterarguments → Text quality, Revision → Mental model, Revision → Text quality) were not significant, the model was re-examined, excluding those six paths. The final model (shown in Figure 3) was favorable with a $\chi^2 = 505.926$ (df = 162, $p < .001$), TLI = .949, CFI = .976, and RMSEA = .071 (90% CI = .062-.081).

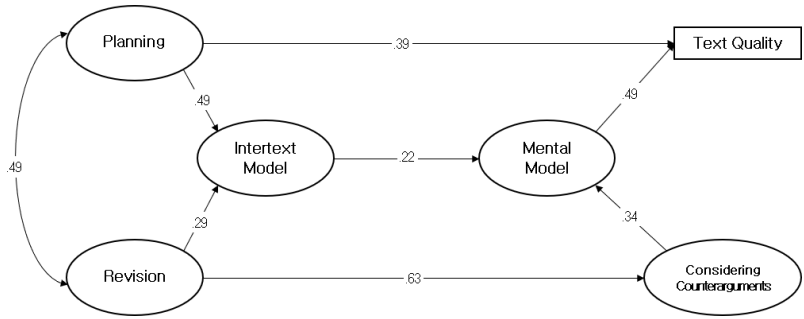

Figure 3. Standardized path coefficients for the final model ($p < .001$)

The results are summarized according to the five RQs as follows.

First, what variables have a direct effect on text quality? As can be seen from Figure 3, planning and mental model had a direct effect on text quality. The intertext model and considering counterarguments only had an indirect effect on text quality through the mediation of the mental model; revision also had an indirect effect on text quality through the mediation of the intertext model, considering counterarguments and the mental model.

Second, how do the task and documents models influence each other? The two variables of the task model, planning and revision, had direct and indirect effects on the two variables of the documents

model: the intertext model and the mental model. Planning had only an indirect effect on the mental model through the mediation of the intertext model. Similarly, revision only had an indirect effect on the mental model through the mediation of the intertext model or considering counterarguments in two paths.

Third, how do the two variables of the task model, planning and revision, influence each other, and through which path do they have an indirect effect on text quality? The correlation between planning and revision appeared to be significant, consistent with the traditional discussions of self-regulated learning theories. Planning had not only a direct effect on text quality but also an indirect effect through the mediation of the intertext and mental models. In contrast, revision only had an indirect effect on text quality in two ways through the mediation of the intertext model, considering counterarguments, and the mental model.

Fourth, how do the two variables of the documents model, the intertext model and the mental model, influence each other, and through which path do they have an indirect effect on text quality? The intertext model had a direct effect on the mental model but only an indirect effect on text quality through the mediation of the mental model.

Fifth, how does the ability to consider counterarguments interact with other variables, and through which path do they have an indirect effect on text quality? Considering counterarguments had a direct effect on revision and only an indirect effect on text quality through the mediation of the mental model. Considering counterarguments did not show a significant causal relationship with planning or the intertext model.

Overall, among the variance of the intertext model, considering counterarguments, the mental model, and text quality, 21.3%, 39.1%, 46.1%, and 48.4%, respectively, were explained by the predictors.

V. Discussion

This study comprehensively and empirically verifies the structural relationship between cognitive processing and writing products while reading and writing from multiple documents based on Britt and Rouet's (2012) MD-TRACE model and Chang's (2021) ATILL-Q.

As expected, a major finding of this study is that planning and revision, which are two variables of the task model, predict text quality. The documents model, which consists of the intertext and mental models, identifiable through visible verbal performance and external outputs, is also affected by planning and revision. This implies that non-visible and intrinsic self-regulatory processing, in addition to visible behavior, is essential for achieving intended outcomes in multiple-document processing. These results are in line with the long-standing discourse that emphasizes the importance of self-regulatory processing in reading and writing (e.g., Anderson, 1985; Bereiter & Scardamalia, 1987; Flower & Hayes, 1980). This suggests that improvements are needed in the existing writing activities of content subjects that require learners to unilaterally acquire information from reading materials and explain it independently of their rhetorical perspective without reflecting the development of cognitivist theories.

The results of the interaction between planning and revision correspond to a series of paradigms in which idea generation and review are closely interrelated and interactive (e.g., Galbraith & Torrance, 2004; Hayes, 2012; Rijlaarsdam et al., 2004; Zimmerman & Kitsantas, 2007). The learners' task representation is never one-off and changes continuously through the process of responding to additional representations or managing the constraints of verbal performance. Planning and revision have different paths to affect text quality. Planning had not only a direct effect on text quality but also an indirect effect

through the mediation of the intertext and mental models. In contrast, revision only had an indirect effect on text quality in two ways through the mediation of the intertext model, considering counterarguments, and the mental model.

The results of considering counterarguments only affected by revision provide an educational implication, especially for teaching multiple-document processing. To develop the two-sided communication ability required for multiple-document processing, planning alone is insufficient in considering task goals and the general understanding and interest of the audience. Both document reading and two-sided communication ability affect text quality through the mediation of text-generating behavior (i.e., the mental model) based on an argumentative schema. However, the intended outcome can be obtained only when high achievements are satisfied in both these two pathways of processing (see Kieft et al., 2006; Mateos et al., 2008; Rijlaarsdam et al., 2008).

Revision provides the implementation of various metacognitive strategies, including comprehension monitoring, epistemic monitoring, and cognitive product formation monitoring (List & Alexander, 2019); it is more than just error correction (Myhill & Jones, 2007). Therefore, various activities such as writing workshops for peer guidance, word processor software-utilized writing, steady recordings of process logs, and promoting epistemic circulation should be implemented to help learners continue to generate additional representations before the development of the final product (Huisman et al., 2019; Klein, 2015; Mateos et al., 2020; Nelson & Schunn, 2009).

Furthermore, the intertext model indirectly affected essay quality through the medium of the mental model. This demonstrates the importance of reconciling conflicts or inconsistencies between the documents one reads, establishing one's own perspective, and rhetorically conveying coherent texts to the audience. These processes go beyond merely attending to, identifying, and evaluating the contribution of sources in mind. While the literature on reading and writing from

multiple documents has mainly focused on effectively distinguishing relevant from irrelevant information (e.g., Anmarkrud et al., 2014; Bråten, Brante et al., 2018; Cerdán & Vidal-Abarca, 2008), empirical attention has recently been paid to which information is ultimately selected and incorporated into one's writing (e.g., Du & List, 2021; List, 2021; Mateos et al., 2020).

Argumentation with consideration of counterarguments (i.e., reflective or deliberative argument) is a representative example of a schema that maximizes the potential of the mental model and effectiveness of disciplinary learning (Anmarkrud et al., 2014; Gil et al., 2010; List & Alexander, 2019; Wiley & Voss, 1999). It is important to comprehensively integrate alternative interpretations and visualize arguments supporting learners' perspectives, especially in the educational context where the detailed processes of reading and writing from multiple documents should be publicly observed and continuously checked. Therefore, learners should be supported to not only personally acquire what they have understood through comprehending sources but also apply argumentative schemas to reconstruct what they have learned and communicate it socially to their audience.

While these results can contribute to the literature on multiple-document processing in various ways, some limitations should be noted. First, this study was conducted on participants with sufficient prior knowledge of the topic. However, although one has prior knowledge of the topic by having enrolled in the regular curriculum in high school, the qualitative level of prior knowledge established may vary depending on the teaching and learning methods. For example, if discourse activities such as presentations, discussions, essay writing, and project learning are applied to classrooms or if content knowledge is linked to everyday life or social issues, the qualitative level of prior knowledge held by participants who have learned in such ways may vary (Atlay et al., 2019; Muhonen et al., 2018; Sedova et al., 2019). A re-analysis may be conducted targeting participants

without sufficient prior knowledge or considering differences in the quality of prior knowledge.

Second, the measurement of sub-processing variables of reading and writing from multiple documents was conducted in the form of a self-report questionnaire given as a closed-ended question. To directly capture multiple-document processing, synchronous research methods such as concurrent think-aloud protocols and prompted pauses and asynchronous research methods such as retrospective protocols can be used (Leijten & Van Waes, 2013). However, the synchronous method may cause a self-explanation effect (Muñoz et al., 2006; VanLehn et al., 1992), which may interfere with multiple-document processing. Asynchronous methods are also limited in terms of recollection processes, reconstructive recall, and social desirability. Therefore, it is necessary to reduce the gap between self-reported responses to closed-ended questions and actual performance while minimizing these problems.

Third, although this study revealed structural relationships between subprocess variables while reading and writing from multiple documents, it did not explain which individual difference variables cause such paths to be significant. Personal difference variables affecting multiple-document processing include a variety of affective, motivational, and behavioral engagement factors in addition to prior knowledge, interest, and epistemic beliefs, which existing studies have primarily focused on (Barzilai & Ka'adan, 2017; Bråten, Brante et al., 2018; List & Alexander, 2019). Identifying a list of variables that can operate as moderating variables and provide optimal mediation while considering them may maximize the significance of this study. In the future, a comprehensive model that includes individual differences might be established, or a multi-group structural equation model with certain variables as moderating variables could be applied.

Fourth, although this study reveals structural relationships among subprocess variables, it is difficult to identify the detailed perfor-

mance of reading and writing from multiple documents by this means alone. In particular, owing to the cyclical properties of self-regulated learning that are intrinsic in multiple-document processing (Burin et al., 2020), the general path derived by applying a structural equation model and the procedures performed by participants are not identical. Therefore, further efforts are needed to reduce the gap and enhance the explanatory power of the verified model using research methods that allow simultaneous observation without directly interfering with multiple-document processing, such as keystroke logging, video observation, and eye-tracking; and by applying text analysis or versioning methods (Leijten & van Waes, 2013). One option may be to classify a cluster of participants according to the relationships between the variables comprising the model, select participants who are representative of each group, and qualitatively track and analyze the process.

Despite these limitations, this study contributes to the literature on multiple-document processing in several ways. First, it comprehensively verified the structural relationships between the subprocess variables that are directly related to the performance of reading and writing from multiple documents (rather than out-of-performance variables that are not directly relevant). Second, in addition to the visible verbal acts implemented through reading and writing, such as the documents model (i.e., intertext and mental models), non-visible and internal self-regulatory processing, such as planning and revision, were included as predictors. Third, this study comprehensively verified the components of the MD-TRACE model by considering the direct effects between different predictors and indirect effects affecting the writing products through different subprocess variables as mediators. Fourth, by accepting the theoretical background of reflective or deliberative arguments, we analyzed how considering counterarguments structurally relates to multiple-document processing and affects text quality. The findings provide empirical evidence to propose interdisciplinary teaching methods applicable to both lan-

guage arts and content subject lectures. This study can contribute to increasing the efficiency of learning by providing balanced guidance on the subprocess phases of reading and writing from multiple documents.

* Submitted 2023.10.30.
First revision recieved 2023.11.21.
Accepted 2023.12.5.

REFERENCES

Anderson, R. C. (1985). Role of the reader's schema in comprehension, learning and memory. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 372-384). International Reading Association.

Anmarkrud, Ø., Bråten, I., & Strømsø, H. I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. *Learning and Individual Differences*, 30, 64-76.

Atlay, C., Tieben, N., Hillmert, S., & Fauth, B. (2019). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap? *Learning and Instruction, 63*, 101211.

Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction, 36*(3), 199-223.

Barzilai, S., & Eshet-Alkalai, Y. (2015). The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36*, 86-103.

Barzilai, S., & Ka'adan, I. (2017). Learning to integrate divergent information sources: The interplay of epistemic cognition and epistemic metacognition. *Metacognition and Learning, 12*(2), 193-232.

Barzilai, S., Tzadok, E., & Eshet-Alkalai, Y. (2015). Sourcing while reading divergent expert accounts: Pathways from views of knowing to written argumentation. *Instructional Science, 43*(6), 737-766.

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review, 30*(3), 973-999.

Bereiter, C., & Scardimalia, M. (1987). *The psychology of written composition*. Erlbaum.

Bormuth, J. R. (1966). Readability: A new approach. *Reading Research Quarterly, 1*(3), 79-132.

Braasch, J. L. G., & Bråten, I. (2017). The discrepancy-induced source comprehension (D-ISC) model: Basic assumptions and preliminary evidence. *Educational Psychologist, 52*(3), 167-181.

Braasch, J. L. G., Bråten, I., Strømsø, H. I., Anmarkrud, Ø., & Ferguson, L. E. (2013). Promoting secondary school students' evaluation of source features of multiple documents. *Contemporary Educational Psychology, 38*(3), 180-195.

Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction, 30*,

9-24.

Bråten, I., Brante, E. W., & Strømsø, H. I. (2018). What really matters: The role of behavioural engagement in multiple document literacy tasks. *Journal of Research in Reading, 41*(4), 680-699.

Bråten, I., Ferguson, L. E., Anmarkrud, Ø., Strømsø, H. I., & Brandmo, C. (2014). Modeling relations between students' justification for knowing beliefs in science, motivation for understanding what they read in science, and science achievement. *International Journal of Educational Research, 66*, 1-12.

Bråten, I., McCrudden, M. T., Stang Lund, E., Brante, E. W., & Strømsø, H. I. (2018). Task–oriented learning with multiple documents: Effects of topic familiarity, author expertise, and content relevance on document selection, processing, and use. *Reading Research Quarterly, 53*(3), 345-365.

Bråten, I., & Strømsø, H. I. (2009). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse processes, 47*(1), 1-31.

Britt, M. A., & Rouet, J. F. (2012). Learning with multiple documents: Component skills and their acquisition. In J. R. Kirby & M. J. Lawson (Eds.), *Enhancing the quality of learning: Dispositions, instruction, and learning processes* (pp. 276-314). Cambridge University Press.

Burin, D. I., Gonzalez, F. M., Barreyro, J. P., & Injoque-Ricle, I. (2020). *Metacognitive regulation contributes to digital text comprehension in e-learning. Metacognition and Learning, 15*(3), 391-410.

Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology, 100*(1), 209-222.

Chang, S. (2017). Investigation on the predictive model of the writer's writing-from-sources as a learning acitivity. *Korean Language Education Research, 52*(2), 371-423.

Chang, S. (2021). Developing and validating a standardized scale to measure the activity tendency in interdisciplinary learning literacy. *Korean Language Education, 174*, 271-314.

Chang, S. (2022). Reading and writing in a digital multiple-document environment: The relationship between individual difference variables, multiple-document selection and use, and writing outcomes. *Journal of Reading Research, 65*, 41-80.

Choi, G. (2013). *(The) effects of group discussion about material texts on content generation* [Unpublished doctoral dissertation]. Korea University.

Choi, S., Park, T. Ryu, S. Jang, J., Kil, H., & Lee, I. (2021). *A study on the development of evaluation criteria for the readability of educational texts*. Korean Institute of

Curriculum and Evaluation.

Dale, E., & Chall, J. S. (1948). A formula for predicting readability: Instructions. *Educational Research Bulletin, 27*, 37-54.

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology, 35*(3), 174-192.

De Milliano, I., van Gelderen, A., & Sleegers, P. (2012). Patterns of cognitive self-regulation of adolescent struggling writers. *Written Communication, 29*(3), 303-325.

DeVellis, R. F. (2016). *Scale development: Theory and applications*. Sage Publications.

Du, H., & List, A. (2021). Evidence use in argument writing based on multiple texts. *Reading Research Quarterly, 56*(4), 715-735.

Felton, M., Garcia-Mila, M., & Gilabert, S. (2009). Deliberation versus dispute: The impact of argumentative discourse goals on learning and reasoning in the science classroom. *Informal Logic, 29*(4), 417-446.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221-233.

Flower, L. (1987). *The role of task representation in reading to write* (Technical Report, 6). University of California, National Center for the Study of Writing and Literacy.

Flower, L., & Hayes, J. (1980). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31-50). Erlbaum.

Galbraith, D., & Torrance, M. (2004). Revision in the context of different drafting strategies. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes* (pp. 63-85). Kluwer Academic Publishers.

Gebril, A., & Plakans, L. (2009). Investigating source use, discourse features, and process in integrated writing tests. *Spaan Fellow Working Papers in Second or Foreign Language Assessment,* 7(1), 47-84.

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Understanding and integrating multiple science texts: Summary tasks are sometimes better than argument tasks. *Reading Psychology, 31*(1), 30-68.

Goldman, S. R. (2004). Cognitive aspects of constructing meaning through and across multiple texts. In N. Shuart-Faris & D. Bloome (Eds.), *Uses of intertextuality in classroom and educational research* (pp. 317-352). Information Age Publishing.

Goldman, S. R., Braasch, J. L. G., Wiley, J., Graesser, A. C., & Brodowinska, K. (2012). Comprehending and learning from Internet sources: Processing patterns of better and poorer learners. *Reading Research Quarterly, 47*, 356-381.

Hagen, Å. M. M., Braasch, J. L. G., & Bråten, I. (2014). Relationships between spontane-

ous note-taking, self-reported strategies and comprehension when reading multiple texts in different task conditions. *Journal of Research in Reading, 37*(S1), S141-S157.

Hansen, H. L. (2011). Multiperspectivism in the novels of the Spanish Civil War. *Orbis Litterarum, 66*(2), 148-166.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369-388.

Huisman, B., Saab, N., van den Broek, P., & van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education, 44*(6), 863-880.

Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Newbury House.

Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2006). Writing as a learning tool: Testing the role of students' writing strategies. *European Journal of Psychology of Education, 21*(1), 17-34.

Klein, P. D. (2015). Mediators and moderators in individual and collaborative writing to learn. *Journal of Writing Research, 7*(1), 201-214.

Klein, P. D., Arcon, N., & Baker, S. (2016). Writing to learn. In C. A. MacArthur, S. Graham, & J. Fizgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 243-256). Guilford Press.

Klein, P. D., & Ehrhardt, J. S. (2015). The effects of discussion and persuasion writing goals on reasoning, cognitive load, and learning. *Alberta Journal of Educational Research, 61*(1), 40-64.

Kobayashi, K. (2009). The influence of topic knowledge, external strategy use, and college experience on students' comprehension of controversial texts. *Learning and Individual Differences, 19*(1), 130-134.

Lee, H. Y., & List, A. (2023). The role of relevance determinations in multiple text reading and writing: an investigation of the MD-TRACE. *Discourse Processes, 60*(1), 1-31.

Lee, Y. (2013). *A study on the aspects of text construction through discourse synthesis: Focusing on the relationship between task representation in writing and text construction* [Unpublished doctoral dissertation]. Yonsei University.

Leki, I., & Carson, J. (1997). "Completely different worlds": EAP and the writing experiences of ESL students in university courses. *TSOL Quarterly, 31*(1), 39-69.

Leijten, M., & van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication, 30*(3), 358-392.

Limpo, T., Alves, R. A., & Fidalgo, R. (2014). Children's high-level writing skills: Development of planning and revising and their contribution to writing quality. *British Journal of Educational Psychology, 84*(2), 177-193.

List, A. (2021). Investigating the cognitive affective engagement model of learning from multiple texts: A structural equation modeling approach. *Reading Research Quarterly, 56*(4), 781-817.

List, A., & Alexander, P. A. (2017). Cognitive affective engagement model of multiple source use. *Educational Psychologist, 52*(3), 182-199.

List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist, 54*(1), 20-39.

List, A., Du, H., & Wang, Y. (2019). Understanding students' conceptions of task assignments. *Contemporary Educational Psychology*, 59, 101801.

Mateos, M., Martín, E., Villalón, R., & Luna, M. (2008). Reading and writing to learn in secondary education: Online processing activity and written products in summarizing and synthesizing tasks. *Reading and Writing, 21*(7), 675-697.

Mateos, M., Rijlaarsdam, G., Martín, E., Cuevas, I., Van den Bergh, H., & Solari, M. (2020). Learning paths in synthesis writing: Which learning path contributes most to which learning outcome? *Instructional Science, 48*(2), 137-157.

McCrudden, M. T., Stenseth, T., Bråten, I., & Strømsø, H. I. (2016). The effects of topic familiarity, author expertise, and content relevance on Norwegian students' document selection: A mixed methods study. *Journal of Educational Psychology, 108*(2), 147-162.

Muhonen, H., Pakarinen, E., Poikkeus, A. M., Lerkkanen, M. K., & Rasku-Puttonen, H. (2018). Quality of educational dialogue and association with students' academic performance. *Learning and Instruction, 55*, 67-79.

Muñoz, B., Magliano, J. P., Sheridan, R., & McNamara, D. S. (2006). Typing versus thinking aloud when reading: Implications for computer-based assessment and training tools. *Behavior Research Methods, 38*(2), 211-217.

Myhill, D., & Jones, S. (2007). More than just error correction: Students' perspectives on their revision processes during writing. *Written Communication, 24*(4), 323-343.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science, 37*(4), 375-401.

Nussbaum, E. M., & Edwards, O. V. (2011). Critical questions and argument stratagems: A framework for enhancing and analyzing students' reasoning practices. *Journal of the Learning Sciences, 20*(3), 443-488.

Nygren, T., Vinterek, M., Thorp, R., & Taylor, M. (2017). Promoting a historiographic gaze through multiperspectivity in history teaching. In H. A. Elmersjo, A. Clark, &

M. Vinterek (Eds.), *International perspectives on teaching rival histories: Pedagogical responses to contested narratives and the history wars* (pp. 207-228). Palgrave Macmillan.

Oh, E. (2020). *A study on the educational design and development of multiple-document (MD) comprehension: Focusing on processes and strategies of exploration, integration, and critical evaluation* [Unpublished doctoral dissertation]. Ewha Womans University.

Primor, L., & Katzir, T. (2018). Measuring multiple text integration: A review. *Frontiers in Psychology, 9*, 2294.

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist, 52*(3), 148-166.

Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssen, T., Raedts, M., Van Steendam, E., Toorenaar, A., & Van den Bergh, H. (2008). Observation of peers in learning to write: Practise and research. *Journal of Writing Research, 1*(1), 53-83.

Rijlaarsdam, G., Couzijn, M., & Bergh, H. V. D. (2004). The study of revision as a writing process and as a learning-to-write process. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes* (pp. 189-207). Springer.

Rouet, J. F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist, 52*(3), 200-215.

Rouet, J. F., Ros, C., Goumi, A., Macedo-Rouet, M., & Dinet, J. (2011). The influence of surface and deep cues on primary and secondary school students' assessment of relevance in Web menus. *Learning and Instruction, 21*(2), 205-219.

Rouet, J. F., Vidal-Abarca, E., Erboul, A. B., & Millogo, V. (2001). Effects of information search tasks on the comprehension of instructional text. *Discourse Processes, 31*(2), 163-186.

Sedova, K., Sedlacek, M., Svaricek, R., Majcik, M., Navratilova, J., Drexlerova, A., Kychler, J., & Salamounova, Z. (2019). Do those who talk more learn more? The relationship between student classroom talk and student achievement. *Learning and Instruction, 63*, 101217.

Stadtler, M. (2017). The art of reading in a knowledge society: Commentary on the special issue on models of multiple text comprehension. *Educational Psychologist, 52*(3), 225-231.

Stang Lund, E., Bråten, I., Brandmo, C., Brante, E. W., & Strømsø, H. I. (2019). Direct and indirect effects of textual and individual factors on source-content integration when reading about a socio-scientific issue. *Reading and Writing, 32*(2), 335-356.

VanLehn, K., Jones, R. M., & Chi, M. T. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences, 2*(1), 1-59.

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Weir, C. J. (1990). *Communicative language testing*. Prentice Hall Regents.

Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal, 46*(4), 1060-1106.

Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology, 91*(2), 301-311.

Zimmerman, B. J., & Kitsantas, A. (2007). A writer's discipline: The development of self-regulatory skills. In S. Hidi & P. Boscolo (Eds.), *Writing and motivation* (pp. 51-69). Elsevier.

ABSTRACT

Structural Relations between Processing Variables and Text Quality while Reading and Writing from Multiple Documents

Chang, Sungmin

This study comprehensively and empirically examined a hypothesized model that incorporates the task model (i.e., planning and revision), documents model (i.e., intertext and integrated mental models), consideration of counterarguments as predictive variables, and text quality of essays as the dependent variable. Data were collected from 421 South Korean high school students who wrote argumentative essays using multiple documents and responded to a series of questionnaires measuring processing variables. Structural equation models were applied to analyze the data. The results indicated that the documents model predicted the participants' text quality, while the effect of the task model was mediated by the documents model and consideration of counterargument. Notably, the paths in which planning and revision affect text quality are different. Furthermore, the intertext model and consideration of counterarguments, as opposed to the mental model, did not directly affect text quality. These findings not only contribute to the literature on multiple-document processing but also provide empirical evidence to suggest an interdisciplinary approach applicable in both language arts and content subject domains.

KEYWORDS Multiple-document processing, MD-TRACE model, Task model, Documents model, Counterargument

Appendix. Items of Modified ATILL-Q

Answer the questions using the following scale: 1 = completely disagree, 5 = completely agree

1. I reference a wealth of material and create my own perspective that encompasses it.
2. I write essays to reveal my original perspective that is different from existing materials.
3. I read the material considering how it will be used in my writing.
4. I write essays while considering the level of understanding and interest of the intended audience.
5. I search for and read the material while judging the attributes (e.g., reliability, fairness) of source information.
6. I select the content information to use for my own writing, considering the attributes of the source information.
7. I search for and read the material to find objective examples or evidence to support my perspective.
8. I directly or indirectly cite the material to explicitly show that I have referenced it.
9. When writing an essay, I clearly present my perspectives or claims.
10. When writing an essay, I carefully consider the structure of the argument (i.e., claim, reason, evidence).
11. I write coherently by resolving the conflicts or inconsistencies between the materials I have read.
12. I expect people who have different attitudes toward topics as an audience.
13. I search for and use materials with different attitudes toward topics.
14. I give balanced consideration to materials that have different attitudes toward the topic.
15. I consider the expected counterargument.
16. I reflect on my writing more after completing the first draft.
17. After completing the first draft, I revise it, including changes in perspective or content.

18. After completing the first draft, I continuously revise the essay by looking for additional materials.
19. I adjust the structure or format of the essay to meet the requirements, such as word count or formatting guidelines.

To obtain the result, calculate the average of the following items:

- planning: items 1, 2, 3, 4
- intertext model: items 5, 6, 7, 8
- mental model: items 9, 10, 11
- considering counterarguments: items 12, 13, 14, 15
- revision: items 16, 17, 18, 19