Reading & Writing - Journal of the Reading Association of South Africa
ISSN: (Online) 2308-1422, (Print) 2079-8245

The cause-effect relation of latent variables in scientific multi-text reading comprehension: Testing the sequential mediation model

Authors:
Hsiao-Hui Lin[1]
Yuh-Tsuen Tzeng[2]
Hsueh-Chih Chen[3]
Yao-Hsuan Huang[3]

Affiliations:
[1]College of Education, National Taiwan Normal University, Taipei, Taiwan

[2]Center for Teacher Education, National Chung Cheng University, Chia-Yi, Taiwan

[3]Department of Educational Psychology and Counseling, National Taiwan Normal University, Taipei, Taiwan

Research Project Registration:
Project Number:
MOST 107-2511-H-003-035-MY3

Corresponding author:
Yao-Hsuan Huang,
896010049@ntnu.edu.tw

Dates:
Received: 15 Jul. 2019
Accepted: 11 Dec. 2019
Published: 30 Mar. 2020

How to cite this article:
Lin, H.-H., Tzeng, Y.-T., Chen, H.-C. & Huang, Y.-H., 2020, 'The cause-effect relation of latent variables in scientific multi-text reading comprehension: Testing the sequential mediation model', *Reading & Writing* 11(1), a256. https://doi.org/10.4102/rw.v11i1.256

Background: The issue of science is seldom brought into focus because of the way developing assessments of students' multiple text reading comprehension.

Objectives: This study tested the sequential mediation model of scientific multi-text reading comprehension (SMTRC) by means of structural equation modelling (SEM), and aimed to advance the scientific multi-text reading comprehension assessment (SMTRCA), with a focus on discussing the causal relationship of potential variables in SMTRC.

Method: Test items included 10 closed-ended and 7 open-ended questions and were categorised into four subscales: information retrieval (IR), information generalisation (IG), information interpretation (IIP), and information integration (IIG).

Results: The confirmatory factor analysis results showed that there was an acceptable goodness-of-fit among the SMTRCA, indicating that the construct validity was good. Furthermore, the Cronbach's α of the test items was 0.88, indicating good internal consistency. In addition, using 1535 students, structural equation modelling was applied to analyse the relations of the latent variables. The findings showed that when readers are performing multi-text reading comprehension, IR will simultaneously have direct influences on IG, IIP and IIG. Moreover, through IG, IR had an indirect impact on IIP; through IIP, IG had an indirect impact on IIG; through the two intermediate mediators of IG and IIP, IR had an indirect impact on IIG.

Conclusion: In our data-driven model, multi-text reading comprehension is a hierarchical and complex cognitive process. That is to say, when an individual is engaging in multi-text reading comprehension, they will not just follow a single approach, but will deal with several cognitive processing routes simultaneously. Recommendations are made for future research to explore the cognitive model of scientific multi-text reading comprehension and to determine whether there are differences among multiple groups, as well as standard setting to define the cut-off scores of the criterion-referenced model, to develop an assessment reporting system of scientific multi-text reading comprehension, and strategies for scientific multi-text reading.

Keywords: Cause-effect relation of latent variables; Scientific multiple text reading comprehension; Sequential mediation model; Cognitive mechanism of reading comprehension; Psychology studies; Science.

Introduction

As scientific knowledge is accessed and internalised by students in various ways these days, the reading of multiple scientific texts has come to be an inevitable routine in everyday academic life. Therefore, the research reported in this study highlights the pivotal place of scientific multiple-text reading comprehension. However, research on the processes of multiple-text reading comprehension is fairly recent, having only been carried out in the past 20 years or so. A decade ago, Bråten, Strømsø and Britt (2009) pointed out that interpretation of the theory of multiple-text reading comprehension still tended to be individualistic, with different researchers propounding their own views, from Hartman's concept of intertextuality within literary theory, Spiro's cognitive flexibility theory, to the documents model which is now highly influential. According to current developments of reading comprehension assessments, the primary concern is to evaluate readers' ability to hold in memory and understand the content of a single text (Royer et al. 1996) and, further, to evaluate readers' ability to make intertextual inferences (Bråten & Strømsø 2010; Bråten et al. 2009), as measured by an intertextual inference verification task (InterVT). However, on the way to developing assessments of students' multiple-text reading comprehension, the issue of science is seldom brought into focus. In other words, what are the latent variables that can be used

as key concepts of multiple text reading comprehension? The assessment model can be rigorously tested by using empirical data to establish the construct validity of the Scientific Multi-Text Reading Comprehension Assessment (SMTRCA), which can help to objectively evaluate students' reading comprehension proficiency of multiple scientific texts. That is the first aim of the study.

Past research into reading comprehension has concentrated mostly on the subject of reading strategies, whereas the mechanisms of the processes of reading comprehension have received relatively less attention (Goldman 2015). Perfetti, Rouet and Britt (1999) advanced the notion that while an individual is engaging in multiple-text reading, they can firstly deal with the parsing of words, phrases, grammar and meaning of multiple texts to build an intertext model, which is then integrated with the individual's prior knowledge to set up a situation model. After analysing the literature on the subject of multi-text reading comprehension, Lin and Tzeng (2018) put forward the process of multi-text reading comprehension consisting of four stages: (1) the first stage is information retrieval (IR), meaning that a reader can interpret words and phrases and analyse the forms of sentences; (2) the second stage is information generalisation (IG), meaning that a reader can extract the gist of a text; (3) the third stage is information interpretation (IIP), meaning that a reader can establish a situation model surrounding the deep-level comprehension of a text; (4) the last stage is information integration (IIG), meaning that a reader can deduce and add the information of content to the comprehended situation to express their positive or negative judgment of a text.

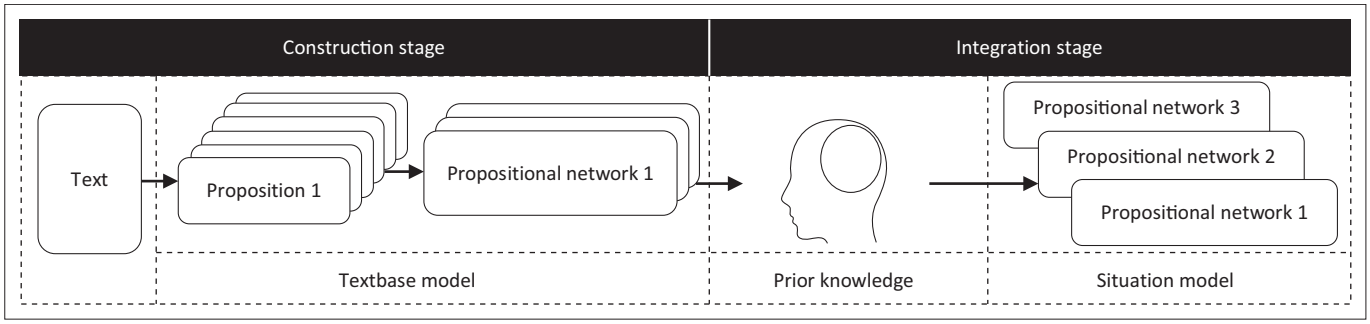
In addition, Chung's (2000) research findings showed that the meanings of a series of logical conjunctions are retrieved appropriately at the stage of IR, which will benefit the individual in constructing the macrostructure of the text (i.e. the situation model). Gil et al. (2010) further proposed that if an individual is capable of summarising the main points of a text, they will not only understand the text more profoundly, but also completely integrate the information from the different texts in the subsequent stage. Other research (Bråten & Strømsø 2009; Wiley & Voss 1999) has shown that if readers, when explaining or developing their argument, are capable of restating it in their own words, that will give them a deeper level of integration comprehension. It is within this context that our research in the field of the mechanisms of the processes of scientific multiple-text reading comprehension will test the mediation model and the sequential mediation model by means of structural equation modelling (SEM) and product of coefficients (Sobel 1982) to determine whether IR will indirectly influence IIG through these two intermediate mediators, viz. IG and IIP. This is the second aim of the study, which will examine various interactions between the subcomponents of reading comprehension. To summarise, the research questions of the current study are as follows:

1. Does the SMTRCA have good construct validity and internal consistency?
2. Concerning the various interactions between the subcomponents of reading comprehension:
    a. Will IR exert an indirect effect on IIP through IG (IR → IG → IIP)?
    b. Will IR exert an indirect effect on IIG through IG (IR → IG → IIG)?
    c. Will IR exert an indirect effect on IIG through IIP (IR → IIP → IIG)?
    d. Will IG exert an indirect effect on IIG through IIP (IG → IIP → IIG)?
    e. Will IR exert an indirect effect on IIG through these two intermediate mediators, IG and IIP (IR → IG → IIP→ IIG)?

Before proceeding to the research itself, we first provide a more detailed discussion of the theoretical framework underpinning multi-text reading comprehension.

Framework for the assessment of multi-text reading comprehension

The construction-integration (CI) model, proposed by Kintsch (1998), explains the cognitive processing patterns of single-text reading comprehension and consists of two cognitive stages: the construction stage and the integration stage, and two mental representations: the textbase model and the situation model (see Figure 1). Wolfe and Goldman (2005) state that readers construct multiple layers of mental representations, including the surface form (the specific words, sentences and layout of the text), the meaning of the text itself (textbase) and the interpretation or model of the world referred to by the text (the situation model).



Source: Lin, H.H. & Tzeng, Y.H., 2017, 'Developing and validating a scientific multi-text reading comprehension assessment: Evidence from texts describing relationships between climate changes and the Three Gorges Dam', *Bulletin of Educational Psychology* 49(2), 215–241

FIGURE 1: Process of single-text reading comprehension.

Numerous studies on multi-text reading comprehension have been published recently. The documents model proposed by Perfetti et al. (1999) has had a significant influence on current research. It is an extension of the CI model and includes the intertext model and the situation model. The intertext model primarily represents the meaning of each text, which includes the main information of the intratext, viz. the main information within a single text, the relationships and differences between the intertext information. In the situation model, readers can combine previously obtained knowledge with the intertext information, then come up with an explanation of the textbase model. They are then able to reorganise the new incoming information and continually turn it into the most up-to-date situation model (Johnson & Seifert 1999; Rouet 2006). In addition to a series of higher-order processes and skills of cognition, multi-text comprehension also includes the utilisation of prior knowledge, sources, deeper-level strategies, task awareness, documentary expertise and personal epistemology. That is, multi-text reading comprehension is the ability of an individual to locate, evaluate and apply the diverse sources of information when constructing and communicating an integrated and meaningful representation for a particular topic, plan or situation (Bråten & Strømsø 2010; Bråten et al. 2009). Besides dealing with intratext reading comprehension, multi-text reading further requires linking the intertext information in order to develop the situated cognition of the text. Lin and Tzeng (2017) indicate that the cognitive process of multi-text reading comprehension (see Figure 2) could be discussed from five aspects: prior knowledge, information retrieval, information generalisation, information interpretation and information integration.

Prior knowledge

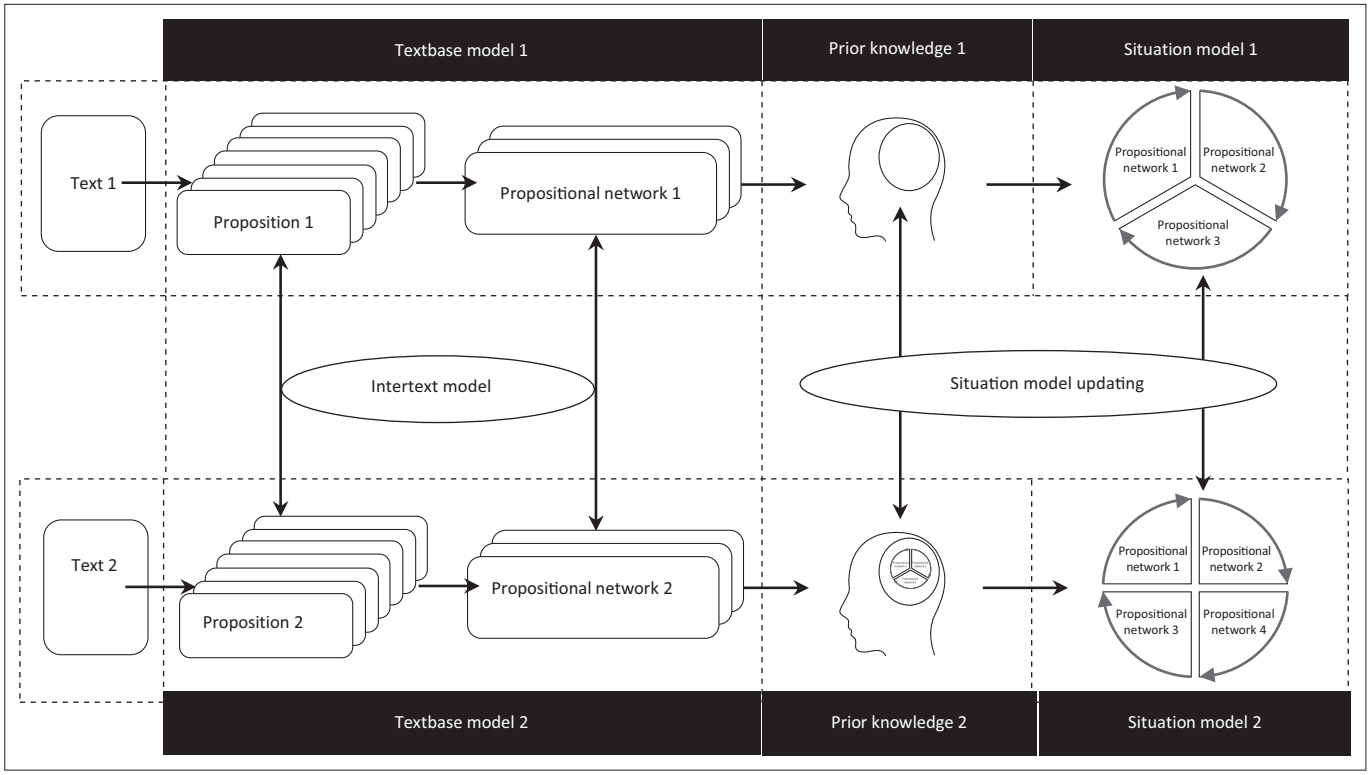
When readers are engaged in text reading, their prior knowledge relevant to the text can help them retrieve important information in the text and activate the basic model to assist their comprehension (Rouet et al. 1996). At the construction stage of the CI model, the text information can activate the relevant knowledge in readers' long-term memory to proceed to interpret what they are reading. However, at the integration stage, readers can establish a coherent situation model on the basis of their prior knowledge.

Information retrieval

Readers can retrieve prior knowledge from their long-term memory that is relevant to the text, and then attempt to interpret words, phrases and sentences, in order to comprehend the surface meaning of the text, that is, the microstructure. This kind of understanding is 'direct understanding' or literal comprehension, which is situated at the shallow level of information processing.

Information generalisation

In the next stage, readers can extract the main concepts of the microstructure built during the previous stage, and establish an overall comprehension of the content of the text. The scope of concepts that is being extracted at the information generalisation stage does not go beyond the content of the text. At this point readers should have the ability to consolidate the information and subsequently facilitate the establishment of a coherent situation of text.



Source: Lin, H.H. & Tzeng, Y.T., 2018, 'Developing and validating a scientific multi-text reading comprehension assessment: In the text case of the dispute of whether to continue the fourth nuclear power plant construction in Taiwan', *Journal of Applied Measurement* 19(3), 320–337

FIGURE 2: Process of multi-text reading comprehension.

Information interpretation

Readers can recombine the main ideas of an article without creating any other new information. At this stage, they have to activate their prior knowledge, construct deeper comprehension, infer the implicit meaning of the text, and restate the text information in their own words in order to establish a situation model that conforms to the gist of the full text.

Information integration

At this stage, readers can deduce and also add the information in the content to the comprehended situation to express their positive or negative judgment of a text (Wolfe & Goldman 2005). In the meantime, they can contrast and compare the intertext information to engage in the evaluative comprehension and critical comprehension of the content of the text. That is to say, a reader can deduce their own viewpoint to analyse the text content. This explains what the situation model is; as noted by Kintsch (1998), it includes the content of macropropositions and the context of causal relations.

According to the model of multi-text reading comprehension advanced by Lin and Tzeng (2017), in the first place, readers can retrieve information to interpret words and phrases and analyse the forms of sentences, and then understand the surface meaning of a text. Next, readers can extract the main concepts to integrate into propositional networks and build the superficial level of description of meaning, that is, the gist of the text, bearing in mind that the scope of the concepts is not beyond the content of the text. This is called 'information generalisation' (IG). Furthermore, readers can integrate the information of propositions that are extracted from the previous two stages with their prior knowledge. According to the prerequisite of not adding any new information, they can infer the implicit meaning of the text and establish a situation model which conforms to a deep-level comprehension of the full text. This is called 'information interpretation' (IIP). In the end, readers can deduce and add their own viewpoints to the comprehended situation to express their positive or negative judgments on the text and even engage in the evaluative comprehension and critical comprehension of the text content. This is called 'information integration' (IIG).

In accordance with the theory put forward by Linderholm, Kwon and Therriault (2014), to be successful in multiple-text comprehension (MTC), readers must keep track of more information and task goals compared to when reading a single text. Based on the complexity of metacognitive processes, multiple-text comprehension is likely a task that places additional burdens on available cognitive resources, that is to say, intertextual information integration is more complex than single-text information integration. Therefore, the assessment tool of the study is driven by the cognitive mechanisms of multi-text reading comprehension proposed by Lin and Tzeng (2017). In the cognitive mechanisms, three primary concepts, proposed by Bråten and Strømsø (2010), were utilised to design the SMTRCA, namely the sentence verification task (SVT), the intratextual inference verification task (IntraVT) and the intertextual inference verification task (InterVT). The full assessment, SMTRCA, is categorised into four sub-assessments, including IR, IG, IIP and IIG.

Interactions in the reading comprehension process

In the past, research on reading comprehension usually emphasised the strategies of reading, for instance what strategies will improve students' reading performance? Because of stressing the test outcomes of reading, the strategies of teaching mainly focused on words, sentences and even the classification of paragraphs besides summaries and contexts by which students can infer new vocabulary (Goldman 2015). In fact, the topic of the interactions involved in the reading comprehension process is also worth paying attention to, but research in this area is lacking. In a study involving 577 students in Hong Kong, Chung (2000) examined whether logical connectives and paragraph headings would influence reading comprehension. Results showed that although logical connectives do not make it possible to understand the microstructures of texts, they can make it easier to understand the macrostructures. In other words, if readers can conduct good retrieval of the lexical meanings of logical connectives at the IR stage, that will assist them in establishing macrostructures of the texts in subsequent stages.

In addition, Gil et al. (2010) invited 53 students to read seven different science texts about climate change. Besides neutral descriptions, these texts also presented conflicting information. After reading the seven science texts, the subjects were instructed to write summaries or argumentative essays according to different group conditions. They were requested to perform three tasks, namely a SVT, IntraVT and InterVT task. Results showed that students who wrote summaries obtained higher scores on the SVT task than those who wrote argumentative essays for assessing memory of the text, deeper understanding of each single text, and the ability to draw inferences across texts in the IntraVT and InterVT tasks. Likewise, in their compositions, students in the summary condition covered the text materials more completely and merged information from the different sources. To put it another way, if readers can reinforce the macrostructures of texts and summarise the main points of the texts, they will not only understand the texts more deeply, but will also be able to integrate information from different texts more completely at subsequent stages.

Wiley and Voss (1999) invited 64 students from the University of Pittsburgh to take part in an experiment, in which the researchers provided several historical texts about Ireland to the students. After reading the texts, the students were instructed to act as historians and construct narratives about the transformation of the population in Ireland. In the process of the experiment, the expressions used in narratives would change into expressions used in summaries, explanations and arguments according to different groups. Results showed that students given the task of constructing arguments gained better understanding than other students, and students

given the tasks of constructing explanations and arguments obtained much more casual understanding than students given the task of narratives since readers can construct their own forms of expression to help their deeper and more integrated understanding when providing explanations and developing arguments. In addition, the study presented by Bråten and Strømsø (2009) also had a similar result in which readers given the tasks of constructing arguments and summarising information were more able to build deep-level and integrated understanding from the texts than those given the task of producing a general overview.

In brief, these four processes of multi-text reading comprehension, IR, IG, IIP and IIG, not only show that there is a subtle time sequence, in which each process occurs step by step, but it also seems that they are likely to have direct or indirect influences on each other. This study specifically explores in what way these four processes of multi-text reading comprehension exert direct or indirect effects on one another, as reflected in research questions 2(a) – 2(e).

Method

Research framework

To begin with, the SMTRCA was refined further in this study. That is, the items of the SMTRCA were dealt with by item analysis in order to modify them. In its modified form, a formal reading comprehension test was administered, and then the answer sheets were collected and analysed, using construct validity and reliability tests. Confirmatory factor analysis (CFA) was applied to test the construct validity. By testing the goodness-of-fit between the theoretical model and experiment data, we could ensure that the constructs of the assessment would be supported by empirical data and would possess construct validity. To check the reliability, we tested the internal consistency of the assessment items, and then checked the reliability of the assessment. In the second stage of the study, there were three sequential mediation models to be tested and compared in order to select the best one and verify research questions 2(a) – 2(e) for research question 2. We also tested whether IR exerted an indirect effect on IIG through the two intermediate mediators of IG and IIP.

Research participants

In the present study, there were 165 learners from Grades 5–9 from Taiwan who completed the pretest scale, and the relevant data of the assessment were modified through item analysis. Thereafter, we invited 1535 learners from Taiwan in Grades 5–9, including 796 boys and 739 girls, whose average age was 13 years old and whose native language was Chinese, to take the formal test. To balance the impact of the text sequence, the test question books were divided into two versions, A and B. In other words, half of the students were given version A (*Climate change is related to the Three Gorges Dam vs Climate change is unrelated to the Three Gorges Dam*) and the other half were given version B (*Climate change is unrelated to the Three Gorges Dam vs Climate change is related to the Three Gorges Dam*).

Development of the scientific multi-text reading comprehension assessment

Test domain

The propositions of the assessment primarily discussed a scientific issue, namely the relationship between climate change and the Three Gorges Dam. The relevant fields of research included meteorology, physics, geology, Earth science, eco-environment, and so on.

Development of the test question book

The unit of the reading texts of the assessment was 'group questions'. Each set of group questions consisted of one text and a set of items which were related to the text. The SMTRCA was composed of two texts. The style of writing for each text was an exposition of 983 to 1000 words. On the basis of the mixture of complementary texts and conflicting texts proposed by Hartman and Allison (1996), the texts held two opposing viewpoints on the topic of 'the relationship between climate change and the Three Gorges Dam', that is to say, whether or not the Three Gorges Dam has contributed to climate change.

Items development

The SMTRCA consisted of two sets of items: 12 multiple-choice format items relating to IR and IG and nine constructed-response format items designed for IIP and IIG. Each constructed-response format item provided answer hints, and a clear focus statement, to activate students' prior knowledge and existing base model, which made them capable of practising dynamic mental activities, such as rhetoric, concepts and the interpretation and inferencing of semantic viewpoints, and then further organising their answer tasks. The framework of items is detailed as follows:

1. Provide instructions to students that they should answer the questions based on the text. Example: Why did the lakes such as Dongting Lake and Poyang Lake in the middle and lower reaches of the Yangtze River shrink and dry up?
2. Students were asked to explain the author's views and reasons based on the information in the text and to express their opinions on the author's point of view. Example: Do you think that the author's description that the Three Gorges Dam contributed to successive years of drought in the Yangtze River region is reasonable or not? Please explain your reasons.

Analysis and results of the pretest

In the study, item analysis was applied to test the pretest data, and the items of the pretest were selected and modified in accordance with the criteria of the elimination of items as follows: (1) test of homogeneity: to calculate the correlation between individual items and the full assessment, SMTRCA, and to delete those items with values less than 0.3 (Nunnally & Bernstein 1994); (2) test of consistency: to delete those items

which, if deleted, would increase the value of the Cronbach's α; (3) principal axis factoring (PAF) out of exploratory factor analysis (EFA): to delete those items with factor loading values less than 0.3; (4) estimating critical ratio: to delete those items that do not reach the significance level. Analytical results showed that four items should be deleted, specifically the first and second of the multiple-choice format items and the fourth and seventh of the constructed-response format items.

Analysis and results of the formal test question book

Constructs and items of the formal test question book

As mentioned above, following item analysis and EFA, four items were eliminated and 17 were retained. The retained items were: IR with three items, IG with seven items, IIP with three items, and IIG with four items. Among them, IIG was further divided into intratextual information with two items and intertextual information with two items (see Table 1).

Scoring criteria and regulations of execution

The SMTRCA is a group-administed reading comprehension test. Every subject has a formal test question book and an answer sheet. The test takes 90 min, including the answering time (85 min), and the time for explaining the instructions of the test question book (5 min). The dichotomous scoring model is adopted by IR and IG, where each correct response scores 1 mark and each incorrect response scores 0. The range of scores of IR is from 0 to 3, whereas that of IG is from 0 to 7. In addition, the polytomous scoring model is applied by IIP and IIG, where the rating is based on reader-response performance and the total score is 4. The rating levels are as follows: (1) subjects who are unable to correctly construct earn 1 score; (2) those who are able to roughly construct earn 2 marks; (3) those who are able to correctly construct earn 3 marks; (4) those who are able to correctly and completely construct earn 4 marks. The range of scores of IIP is from 1 to 12, and that of IIG is from 1 to 16.

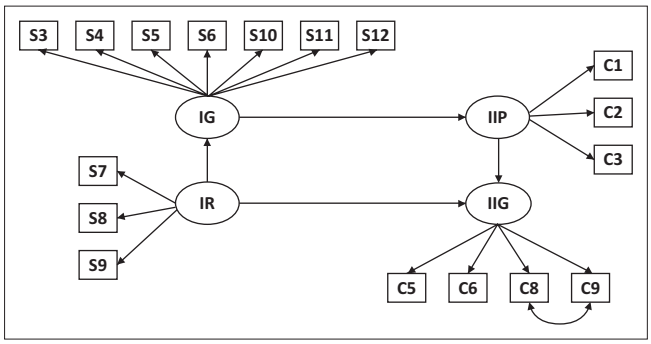
Data processing and statistical analysis

After collecting the data of the formal test, item analysis, test of internal consistency and test of construct validity were firstly carried out, and then three sequential mediation models were tested in the next phase to verify their goodness-of-fit, which are described as follows.

Test of model 1: In model 1, there were two latent serial mediators, IG and IIP, merged into a sequential mediation model by which we tested whether IR of multiple-text reading comprehension had an effect of sequential mediation on IIG (abc: IR → IG → IIP → IIG). First, this study postulated that the predictor variable (IR) would initiate the first serial mediator (IG) and subsequently the second serial mediator (IIP) with the criterion variable (IIG) as the final consequent.

Test of model 2 (Figure 3): In model 2, there were two latent serial mediators, IG and IIP, merged into a sequential mediation model by which we tested whether IR of multiple-text comprehension not only had a direct effect on IIG (f'1: IR → IIG) but also exerted an effect of sequential mediation on IIG (abc: IR → IG → IIP → IIG). First, this study postulated that the predictor variable (IR) would initiate the first serial mediator (IG) and subsequently the second serial mediator (IIP) with the criterion variable (IIG) as the final consequent.

Test of model 3 (Figure 4): In model 3, we tested whether IR not only had a direct effect on IIG (f'1: IR → IIG) but also exerted three indirect effects on IIG, namely (1) ae: IR → IG → IIG; (2) dc: IR → IIP → IIG; (3) abc: IR → IG → IIP → IIG). Therefore, the sum to which a direct effect plus three indirect effects of IR will be exerted on IIG equals f'1 + ae + dc + abc which represents the total effect that IR has on IIG.



IG, information generalisation; IIP, information interpretation; IR, information retrieval; IIG, information integration; C, item number; S, item number.

FIGURE 3: The sketch of Model 2: The sequential mediation model with information retrieval as the predictor variable.

TABLE 1: Summary of dimensions and primarily tested abilities in scientific multi-text reading comprehension assessment.

Assessment dimension	Examinee ability	Sub-dimension	Item numbers	Number of items
Information retrieval	• Retrieving knowledge and memory relevant to the aim of reading a text • Explaining the key words and phrases of a text • Analysing the key sentence patterns of a text • Understanding the key semantics and substance of a text	-	S7, S8, S9	3
Information generalisation	• Extracting the main idea of a text • Inducing the key information of a text • Summarising the main points of a text • Constructing the effect of a text and the macrostructure of relations among paragraphs	-	S3, S4, S5, S6, S10, S11, S12	7
Information interpretation	• Extracting a propositional network and integrating it into prior knowledge • Constructing a situation model conforming to the substance of a text • Inferring implicit meanings in text • Restating text information in one's own words	-	C1, C2, C3	3
Information integration	• Contrasting and comparing intertextual information and integrating it into prior knowledge • Constructing a coherent situation model involving macropropositions • Deducing one's own viewpoint in accordance with text information	Intratextual	C5, C6	2
	• Judging and analysing intratextual and intertextual information	Intertextual	C8, C9	2
Total	-	-		**17**

Comparison of the three sequential mediation models: Three competing sequential mediation models (model 1, 2, and 3) were put forward in the study and a chi square difference test was carried out to test whether there were differences between them. The indices, such as χ^2, Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), were then referred to in order to identify which model best represents the cognitive mechanisms of multiple text reading comprehension.

Ethical consideration

Center for Taiwan Academic Research Ethics Education Certificate of Completion, Certification No. P106031740. Ethical Clearance Number: 107WFA0310690

Results

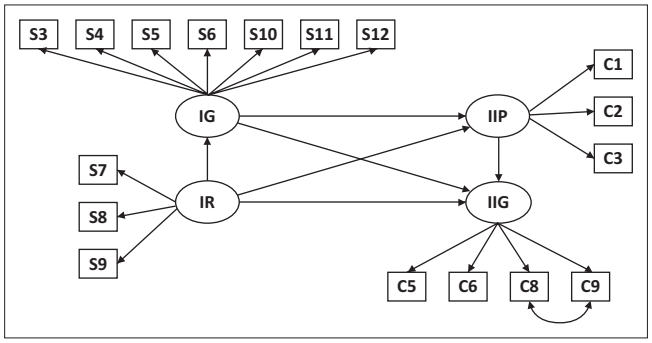
Tests of reliability and validity

Internal consistency reliability

In the analysis of internal consistency, the results showed that the Cronbach's α of the full assessment, SMTRCA, is 0.88, the Cronbach's α values of the sub-assessments range between 0.66 and 0.86, specifically 0.66 (IR), 0.68 (IG), 0.83 (IIP) and 0.86 (IIG) in sequence. Among these values, since the items for each sub-assessment are fewer than those of the full assessment, the Cronbach's α value of each construct is less than that of the full assessment, but the reliability of Cronbach's α of the full assessment is still within an acceptable range.

Construct validity

In the analysis of composite reliability, results showed that the values are as follows: 0.61 (IR), 0.68 (IG), 0.83 (IIP) and 0.88 (IIG), all of which are more than 0.6 (Fornell & Larcker 1981), indicating that the composite reliability of SMTRCA is good. In addition, the maximum likelihood method was applied to conduct CFA. According to the statistical indices of model fit suggested by Kline (2015) – including: (1) chi-squared significance (χ^2), (2) the 90% confidence level of RMSEA (root mean square error of approximation), (3) comparative fit index (CFI) and (4) standardised root mean square residual (SRMR) – evaluation of the model fit was performed to examine whether the theoretical construct model of the assessment would be verified and supported by the empirical data. In terms of the results of the analysis, except that the chi-squared distribution reached significance level ($\chi 2 = 269.18$, $p < 0.001$), where there is significant difference between the hypothetical model and observed values, the other statistical indices are all located within ideal fit ranges, as follows: SRMR is 0.032 (≤ 0.05), RMSEA is 0.038 (≤ 0.05), the value of the 90% confidence level is between 0.032 and 0.044 (close fit) and CFI is 0.975 (≥ 0.90). Therefore, the construct validity is good.



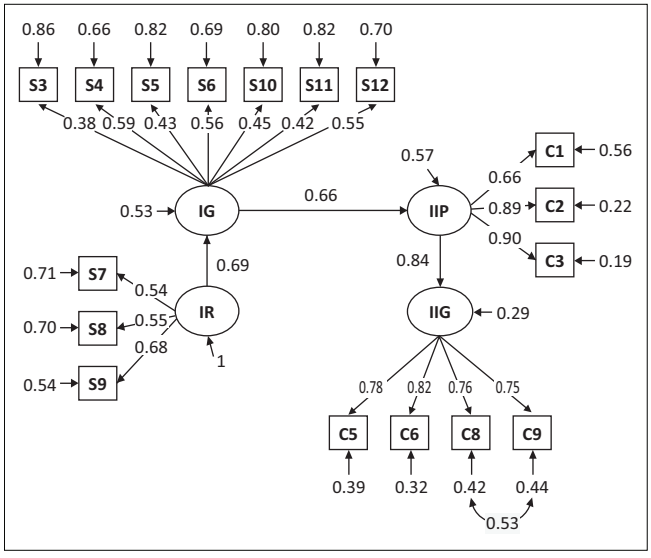
IG, information generalisation; IIP, information interpretation; IR, information retrieval; IIG, information integration; C, item number; S, item number.

FIGURE 4: The sketch of Model 3: The sequential mediation model with information retrieval as the predictor variable.

Tests of the mediation model

Model 1: Analytical resuflts of the sequential mediation model

In the study, IR was utilised to conduct the analysis of the effect of sequential mediation to test H1: IG and IIP will mediate IR in sequence to influence IIG. In the analytical results, except that the chi-squared distribution reached the significance level ($\chi 2 = 294.772$, $p < 0.001$), the fit indices of model 1 are all located within ideal fit ranges, where the AIC is 22847, the BIC is 23033, SRMR is 0.037 (≤ 0.05), RMSEA is 0.040 (≤ 0.05), the value of the 90% confidence level is between 0.034 and 0.045 (close fit) and CFI is 0.975 (≥ 0.90). In model 1, H1: IR exerts an indirect influence on IIP through IG (a*b: IR → IG → IIP, $p < 0.001$); H2: IG exerts an indirect influence on IIG through IIP (b*c: IG → IIP → IIG, $p < 0.001$); H3: IR exerts an indirect influence on IIG through these two latent mediators, IG and IIP, in sequence (a*b*c: IR → IG → IIP → IIG, $p < 0.001$). These values are all supported by the empirical data (see Figure 5, Table 2).



Note: Solid lines indicate statistically significant paths, and dashed lines indicate statistically nonsignificant paths.

IG, information generalisation; IIP, information interpretation; IR, information retrieval; IIG, information integration; C, item number; S, item number.

FIGURE 5: Model 1: The sequential mediation model with information retrieval as the predictor variable.

TABLE 2: The testing results of the mediation effect of model 1.

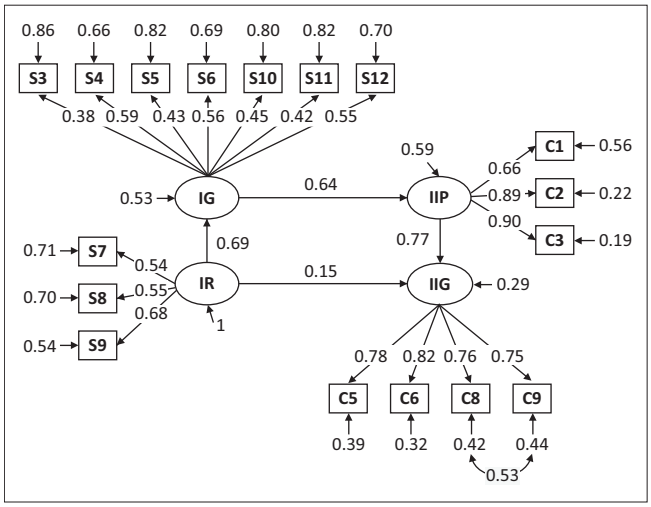
Mediation test	Estimate	Standard error	z	$P(>\|z\|)$	Standard all
a*b	1.796	0.187	9.611	0.000	0.452
b*c	2.160	0.233	9.253	0.000	0.554
a*b*c	1.414	0.146	9.663	0.000	0.381

Analytical results of Model 2

In the analytical results, except that the chi-squared distribution reached the significance level ($\chi 2$ = 276.084, $p < 0.001$), the fit indices of Model 2 are all located within ideal fit ranges, where AIC is 22830, BIC is 23021, SRMR is 0.033 (≤ 0.05), RMSEA is 0.038 (≤ 0.05), the value of the 90% confidence level is between 0.032 and 0.044 (close fit) and CFI is 0.974 (≥ 0.90). In Model 2, H1: IR exerts a direct influence on IIG (f': IR → IIG, $p < 0.001$); H2: IR exerts an indirect influence on IIP through IG (a*b: IR → IG → IIP, $p < 0.001$); H3: IG exerts an indirect influence on IIG through IIP (b*c: IG → IIP → IIG, $p < 0.001$); H4: IR exerts an indirect influence on IIG through these two latent mediators, IG and IIP, in sequence (a*b*c: IR → IG → IIP → IIG, $p < 0.001$). All of these values are supported by the empirical data (see Figure 6, Table 3).

Analytical results of Model 3

In the analytical results, except that the chi-squared distribution reaches the significance level ($\chi 2$ = 269.181, $p < 0.001$), the fit indices of Model 3 are all located in ideal fit ranges, where AIC is 22827, BIC is 23028, SRMR is 0.032 (≤ 0.05), RMSEA is 0.038 (≤ 0.05), the value of the 90% confidence level is between 0.032 and 0.044 (close fit) and CFI is 0.975 (≥ 0.90). In Model 3, except that H3 is not supported by the empirical data, that is, IR exerts an indirect influence on IIG through IG (a*e: IR → IG → IIG, $p > 0.05$), all the others are backed by the empirical data. H2: IR exerts an indirect influence on IIP through IG (a*b: IR → IG → IIP, $p < 0.001$); H4: IR exerts an indirect influence on IIG through IIP (d*c: IR → IIP → IIG, $p < 0.01$); H5: IG exerts an indirect influence on IIG through IIP (b*c: IG → IIP → IIG, $p < 0.001$); H6: IR exerts an indirect influence on IIG through these two latent mediators, IG IIP, in sequence (a*b*c: IR → IG → IIP → IIG, $p < 0.001$) (see Figure 7, Table 4).



IG, information generalisation; IIP, information interpretation; IR, information retrieval; IIG, information integration; C, item number; S, item number.

Note: Solid lines indicate statistically significant paths, and dashed lines indicate statistically nonsignificant paths.

FIGURE 6: Model 2: The sequential mediation model with information retrieval as the predictor variable.

TABLE 3: The testing results of the mediation effect of model 2.

Mediation effect	Estimate	Standard error	z	p (>|z|)	Standard all
a*b	1.796	0.183	9.510	0.000	0.439
b*c	1.915	0.215	8.913	0.000	0.492
a*b*c	1.248	0.133	9.410	0.000	0.336

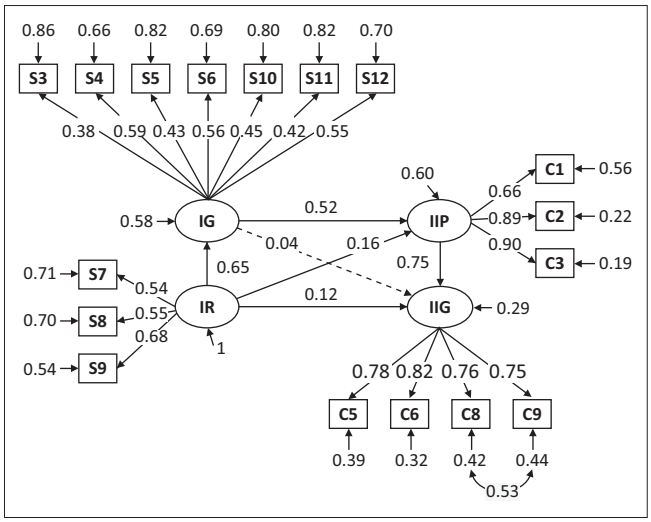
Comparison of the three sequential mediation models

In the study, the chi-square difference test was applied to test whether there are differences among these three sequential mediation models. Analytical results showed that there are statistically significant differences among them (Table 5). The chi-squared distribution (χ^2) of Model 3 is kept to a minimum, indicating that Model 3 is the best of the three.

Discussion

General discussion

The cognitive ability that is evaluated by the SMTRCA includes: (1) readers' lexical comprehension of scientific text information (IR), (2) readers' gist comprehension of



IG, information generalisation; IIP, information interpretation; IR, information retrieval; IIG, information integration; C, item number; S, item number.

Note: Solid lines indicate statistically significant paths, and dashed lines indicate statistically nonsignificant paths.

FIGURE 7: Model 3: The sequential mediation model with information retrieval as the predictor variable.

TABLE 4: The testing results of the mediation effect of model 3.

Mediation effect	Estimate	Standard error	z	p (>|z|)	Standard all
a*b	1.317	0.195	6.765	0.000	0.336
a*e	0.094	0.126	0.748	0.455	0.252
d*c	0.446	0.166	2.680	0.007	0.121
b*c	1.496	0.231	6.488	0.000	0.386
a*b*c	0.927	0.144	6.450	0.000	0.252

TABLE 5: The results of the chi-square difference test.

Model	df	AIC	BIC	χ^2	$\triangle\chi 2$	$\triangle df$	Pr (> Chisq)
Model 1	115	22830	23021	276.08	18.688	1	0.000***
Model 2	114	22847	23033	294.77	-	-	-
Model 1	115	22830	23021	276.08	25.591	3	0.000***
Model 3	112	22827	23028	269.18	-	-	-
Model 2	114	22847	23033	294.77	6.904	2	0.032*
Model 3	112	22827	23028	269.18	-	-	-

AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; df, degrees of freedom; Pr (> Chisq), p-value.

*, $p < 0.05$; **, $p < 0.01$;***, $p < 0.001$.

scientific text information (IG), (3) readers' construction of the situation model which conforms to the substance of the full text and keeps coherent (IIP) and (4) readers' evaluative and critical comprehension which is both intratextual and intertextual to scientific text information (IIG). Besides the results of the study and a general discussion, the research applications and research suggestions that we will provide for academic groups in the future can be seen below.

Scientific multi-text reading comprehension assessment, an assessment of reliability and validity

Internal consistency reliability

The analytical results of Cronbach's α showed that except for IR and IG, the statistics of IIP and IIG exceed 0.70 and the alpha value of the full assessment is 0.88, indicating that the SMTRCA possesses internal consistency reliability.

Construct validity

First, the items of the construct share a high proportion of variance in common, known as convergent validity. The values of composite reliability and the ratio of explained variance over total variance (Kline 2015) are all more than 0.6 for the four factors of the SMTRCA. According to the criteria proposed by Fornell and Larcker (1981), the SMTRCA's convergent validity is acceptable. Second, the results of CFA showed that except for the chi-square distribution reaching the significance level, which means that there is an obvious difference between the hypothetical model and the observed statistic, both SRMR and RMSEA out of the absolute fit indices and CFI out of the incremental fit indices fall within ideal fit ranges. In other words, the model, which is composed of these four constructs of the SMTRCA, IR, IG, IIP and IIG, is supported by the empirical data, denoting that the SMTRCA has good construct validity.

Model 3 is considered the best model for cognitive mechanisms of multi-text reading comprehension

All the chi square distributions of these three competitive models reached the significance level. However, the values of the chi-square distribution are the statistic on which a sample will depend, and the model will be rejected easily if the sample size becomes large enough. Moreover, for these three competing models, all of the SRMR and the 90% confidence interval of RMSEA are less than 0.05 (close fit), and their CFI values are all more than 0.9, indicating that these three competitive models all lie within ideal-fit ranges. Finally, following a chi-square difference test, the results of these three sequential mediation models showed that there are statistically significant differences among the three models, and that the chi-square statistic and the AIC index of Model 3 is the least of the three, indicating that Model 3 is the best.

Primary effects among variables

Results (see Table 6) showed that IR exerts direct effects on IG ($a = 0.65$, $p < 0.001$), IIP ($d = 0.16$, $p < 0.01$) and IIG ($f'1 = 0.12$, $p < 0.05$), indicating that when readers are performing multi-text reading comprehension, IR will simultaneously have direct influences on IG, IIP and IIG; in other words, the cognitive mechanisms of multi-text reading comprehension will not just have a single approach, but will have several processing routes, which are part of a complex cognitive process. Moreover, the primary effects of IR on IG, IIP and IIG are significant but exhibit a trend showing a steady decrease in regression coefficients, which is in agreement with the viewpoint on the cognitive mechanisms of multi-text reading comprehension (IR → IG → IIP → IIG) put forward by Lin and Tzeng (2017). That is to say, IR is basically essential for any other type of comprehension, including IG, IIP and IIG, and, without it, the others cannot really operate. Furthermore, IR exerts a greater effect on IG than IR on IIP; similarly, IR will exert a greater effect on IIP than IR on IIG. The situation will correspond with the field of teaching where simply increasing the amount of students' vocabulary and reading can only help improve their ability to retrieve information but will have a limited effect on development of the advanced level of reading comprehension, such as IIP and IIG. In terms of the best model of the study, Model 3, the regression coefficients of IR on IG, IR on IIP and IR on IIG are 0.65, 0.16, and 0.12, which shows a descending trend, demonstrating not simply that when teaching reading, training students' reading skills and teachers' teaching strategies of reading are considered to be more important, but also that the necessity of literal comprehension (viz. IR) must not be ignored. In addition, our study found that the direct effect of IG on IIG does not appear to be sufficiently significant ($e = 0.04$, $p > 0.05$), but the indirect effect of IG on IIG relies on the full mediation of IIP. That is to say, the strength of an individual's ability to extract the major

TABLE 6: Summary of model 3 results with information retrieval as the predictor variable (IR → IG → IIP → IIG).

Variable	M1 (IG)				M2 (IIP)				Y (IIG)				Effect	SE	z	p (>|z|)
	Path	Coeff.	SE	z	Path	Coeff.	SE	z	Path	Coeff.	SE	z				
X: IR	a	0.65	0.078	7.914***	d	0.16	0.237	2.669**	f'1	0.12	0.178	2.567*	-	-	-	-
M1: IG	-	-	-	-	b	0.52	0.312	6.804***	e	0.04	0.203	0.744	-	-	-	-
M2: IIP	-	-	-	-	-	-	-	-	c	0.75	0.046	15.278***	-	-	-	-
Total indirect effect:	-	-	-	-	-	-	-	-	-	-	-	-	0.373	-	-	-
Indirect 1 effect: IR → IG → IIP → IIG	-	-	-	-	-	-	-	-	-	-	-	-	0.252	0.144	6.450***	0.000
Indirect 2 effect: IR → IG → IIG	-	-	-	-	-	-	-	-	-	-	-	-	0.121	0.166	2.680***	0.007
Direct effect	-	-	-	-	-	-	-	-	-	-	-	-	0.120	0.178	2.567***	0.010
Total effect	-	-	-	-	-	-	-	-	-	-	-	-	**0.493**	-	-	-

IR, information retrieval; SE, standard error; IG, information generalisation; IIP, information interpretation; IIG, information integration; Coeff., path coefficient; M, mediator; Y, dependent variable.
*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

concepts of a text (IG) will not exert a direct effect on IIG of the text; instead, an individual should combine these extracted concepts with their prior knowledge and describe the situation of the text in their own words (IIP), which will then help them to move forward to the intratextual and intertextual IIG.

Indirect effects of mediators and serial mediators

Results (see Table 6) showed that IR exerts an indirect effect on IIP through the partial mediation of IG (IR → IG → IIP, indirect effect = 0.336, $p < 0.001$); IR exerts an indirect effect on IIG through IIP (IR → IIP → IIG, indirect effect = 0.121, $p < 0.01$); IR exerts an indirect effect on IIG through the partial mediation of these two serial mediators, IG and IIP (IR → IG → IIP → IIG, indirect effect = 0.252, $p < 0.001$). Furthermore, IG exerts an indirect effect on IIG through the full mediation of IIP (IG → IIP → IIG, indirect effect = 0.386, $p < 0.001$). Additionally, the hypothesis that IR will exert an indirect effect on IIG through IG, was not supported by the empirical data (IR → IG → IIG, a*e = 0.025, $p > 0.05$).

As previously mentioned, derived from empirical research, multi-text reading comprehension is a hierarchical and complex cognitive process which operates from the information input stage, involving the comprehension of words, phrases and the patterns of sentences (IR) and the induction of the extracted gist (IG), to the information output stage, including the inference of connotations and the expression of main information (IIP), as well as the deduction of opinions and the criticism and analysis of texts (IIG). That is to say, when an individual is engaging in multi-text reading comprehension, they will not simply follow a single approach, but will deal with several cognitive processing routes at the same time.

Conclusion

Limitations and future research suggestions

In terms of the present study, related future research, which will be derived from the framework of the SMTRCA, includes: exploring the cognitive model of scientific multi-text reading comprehension and whether there are differences among multiple groups, standard setting to define the cut-off scores of the criterion-referenced model and develop an assessment reporting system of scientific multi-text reading comprehension, as well as developing strategies and techniques for scientific multi-text reading.

Exploring the cognitive model of scientific multi-text reading comprehension

The reading comprehension of both single-text (Kintsch 1988) and multi-text (Perfetti et al. 1999) retrieving information in the textbase model and then interpreting information in the situation model are serial processes. However, in this study, the primary effects of IR on IG, IIP and IIG are significant, and roughly represent that the cognitive mechanisms of multi-text reading comprehension may have several approaches. Therefore, some recommendations are given according to these findings in this study. First, IR is basically essential for reading comprehension, because individuals who perform better in IR will have a more excellent performance in IG, IIP and IIG. Second, the cognitive process of multi-text reading comprehension needs to be extended, because it may be done by parallel processing after IR.

Standard setting to build the cut-off scores of the criterion-referenced model

In order to build standardised comprehension tests to help teachers understand whether students have 'scientific literacy' of science related to basic language and writing skills, and to teach scientific reading comprehension, we suggest that standard setting can be done to build the cut-off scores of the criterion-referenced test in the future. In much of the past research, we found that the response performance of evaluation is equalised to the same scale, and cut-off scores are decided through the relative position of a subject's evaluation performance and ranked by percentage. This testing type is based on the Norm-Referenced model with a relative standard. In other words, the performance level of a subject will be classified into the other level according to the difference of the comparison group in which the subject attends, or, more exactly, the performance of a subject is in a state of change pending the extent of the comparison group. If the cut-off scores selected by the Norm-Referenced model are of great significance, the premise should be to satisfy these hypotheses that the tested population is similar, and the two-way specification tables on which items are based should be the same (Cizek & Bunch 2007). To overcome the limitation of the Norm-Referenced model which does not provide a precise explanation for a subject's performance level, and to ensure that a subject's performance will not differ according to the difference of the comparison group, the criterion-referenced model which provides an absolute standard should be adopted in the future. That is to say, cut-off scores are built from the standard setting (Cizek 2006; Eckes 2009) of performance, and students' performance levels and the extent of their knowledge and ability are made in advance, in order to describe the content standard of what students 'should know' and 'should be able to do' and the levels of the content standard that students will reach.

Developing an assessment reporting system of scientific multi-text reading comprehension

To give information feedback, including teachers' teaching of reading and students' reading improvement, we suggest that an assessment reporting system should be developed in the future. The Many-Facet Rasch Measurement can analyse a rater's perspective on ability performance, for example assessment constructs, in detail and adjust a subject's performance response to complement the variability in rater severity. In line with this, we offer some suggestions. The Many-Facet Rasch Measurement should be utilised to estimate subject ability in the future, and a subject's statistical control of performance response will be adjusted by severity so that the evaluation of scientific multi-text reading comprehension will be independent and beyond rating variability. In addition, on the basis of performance level descriptions and cut-off scores formulated by standard

setting, an assessment reporting system should be developed to provide analytical feedback which will fulfil a diagnostic function so that subjects will understand their strengths and weaknesses in scientific multi-text reading comprehension to which both teachers' teaching of reading and students' reading improvement will refer.

Exploring whether there are differences among multiple groups in the cognitive model of scientific multi-text reading comprehension

Specifically, further research should be done to explore other possible models of reading comprehension with more investigated variables, such as examinees' gender, grade levels and type of text content. By means of testing and comparing the three sequential mediation models mentioned above, the results were as follows: (1) IR will have an indirect effect on IIP through IG; (2) IR will have an indirect effect on IIG through IIP; (3) IR will have an indirect effect on IIG through these two mediators occurring in time sequence, IG and IIP. However, although the cognitive process of multi-text reading comprehension could be roughly seen through these research results, an analysis of multi-group structural equation modelling should be applied to test the goodness-of-fit of invariance across gender and grades on the cognitive model of scientific multi-text reading comprehension, including factor invariance, item intercept invariance, factor variance and covariance invariance across gender, and error variation invariance. That is to say, conducting the test of invariance across groups would help to explore the impact of the categorical variables on scientific multi-text reading comprehension.

Exploring the strategies and skills of scientific multi-text reading

As Adams and Lowery (2007) pointed out, besides reading processes, for example general literacy and understanding, science reading relies more heavily on prior knowledge in science and reading skills. Therefore, following the definition of the cognitive model of scientific multi-text reading comprehension, the strategies and skills of science reading which are based on language reading will be developed to help support students' reading processes from textbooks, thereby enabling them to build effective learning strategies and so achieve the goal of lifelong learning.

Limitations

First, this study can only provide a glimpse of the cognitive process of multi-text reading comprehension. There are four intratextual items and two intertextual items in IIG of SMTRCA, and, in accordance with the opinions put forward by researchers (Lin & Tzeng 2017; Linderholm et al. 2014), the ability of intertextual integration is more difficult and complicated. Therefore, we think that there should be a latent construct, but we still have no idea what it is, and what its association with IR, IG and IIP is. Second, because we used multiple scientific texts that were limited to the topic of Earth science in this study, there is no discussion on analysing the moderating effects of type of text content on the dependent variables of the scientific multi-text reading comprehension causal model. The limitations of this study are possible directions for future research, and we hope to continue to explore this important issue further.

Acknowledgements

This work was financially supported by the Ministry of Science and Technology in Taiwan.

Competing interests

The authors have declared that no competing interests exist.

Authors' contributions

All authors contributed equally to this work.

Funding information

This work was financially supported by the Ministry of Science and Technology in Taiwan.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References

Adams, T.L. & Lowery, R.M., 2007, 'An analysis of children's strategies for reading mathematics', *Reading and Writing Quarterly* 23(2), 161–177. https://doi.org/10.1080/10573560601158479

Bråten, I. & Strømsø, H., 2010, 'When law students read multiple documents about global warming: Examining the role of topic-specific beliefs about the nature of knowledge and knowing', *Instructional Science* 38(6), 635–657. https://doi.org/10.1007/s11251-008-9091-4

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. https://doi.org/10.1080/01638530902959646

Bråten, I., Strømsø, H.I. & Britt, M.A., 2009, 'Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts', *Reading Research Quarterly* 44(1), 6–28. https://doi.org/10.1598/RRQ.44.1.1

Chung, J.S., 2000, 'Signals and reading comprehension – Theory and practice', *System* 28(2), 247–259. https://doi.org/10.1016/S0346-251X(00)00010-5

Cizek, G.J., 2006, 'Standard setting', in S.M. Downing & T.M. Haladyna (eds.), *Handbook of test development*, pp. 225–258, Lawrence Erlbaum Associates, Mahwah, NJ.

Cizek, G.J. & Bunch, M.B., 2007, *Standard setting: A guide to establishing and evaluating performance standards on tests*, Sage, Thousand Oaks, CA.

Eckes, T., 2009, 'Many-facet Rasch measurement', in S. Takala (ed.), *Reference supplement to the manual for relating language examinations to the Common European Framework of Reference for languages: Learning, teaching, assessment (Section H)*, pp. 1–52, Council of Europe/Language Policy Division, Strasbourg.

Fornell, C. & Larcker, D.F., 1981, 'Evaluating structural equation models with unobservable variables and measurement error', *Journal of Marketing Research* 18(1), 39–50. https://doi.org/10.2307/3151312

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. https://doi.org/10.1080/02702710902733600

Goldman, S.R., 2015, 'Reading and the Web: Broadening the need for complex comprehension', in R.J. Spiro, M. DeSchryver, M.S.P. Morsink & P.T. Hagerman (eds.), *Reading at a crossroads*, pp. 89–103 Routledge, New York.

Hartman, D.K. & Allison, J., 1996, 'Promoting inquiry-oriented discussions using multiple texts', in L.B. Gambrell & J.F. Almasi (eds.), *Lively discussions!: Fostering engaged reading*, pp. 106–133, International Reading Association, Newark, DE.

Johnson, H.M. & Seifert, C.M., 1999, 'Modifying mental representations: Comprehending corrections', in H. Van Oostendorp & S. Goldman (eds.), *The construction of mental representations during reading*, pp. 303–318, Lawrence Erlbaum Associates, Mahwah, NJ.

Kintsch, W., 1988, 'The role of knowledge in discourse comprehension: A construction-integration model', *Psychological Review* 95(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W., 1998, *Comprehension: A paradigm for cognition*, Cambridge University Press, Cambridge, MA.

Kline, R.B., 2015, *Principles and practice of structural equation modeling, fourth edition*, 4rd edn., Guilford Press, New York.

Lin, H.H. & Tzeng, Y.H., 2017, 'Developing and validating a scientific multi-text reading comprehension assessment: Evidence from texts describing relationships between climate changes and the Three Gorges Dam', *Bulletin of Educational Psychology* 49(2), 215–241. https://doi.org/10.6251/BEP.2017-49(2).0003

Lin, H.H. & Tzeng, Y.T., 2018, 'Developing and validating a scientific multi-text reading comprehension assessment: In the text case of the dispute of whether to continue the fourth nuclear power plant construction in Taiwan', *Journal of Applied Measurement* 19(3), 320–337.

Linderholm, T., Kwon, H. & Therriault, D.J., 2014, 'Instructions that enhance multiple-text comprehension for college readers', *Journal of College Reading and Learning* 45(1), 3–19. https://doi.org/10.1080/10790195.2014.906269

Nunnally, J.C. & Bernstein, I.H., 1994, *Psychometric theory*, McGraw-Hill, New York.

Perfetti, C.A., Rouet, J.F. & Britt, M.A., 1999, 'Toward a theory of documents representation', in H. Van Oostendorp & S.R. Goldman (eds.), *The construction of mental representations during reading* (pp. 99–122), Erlbaum, Mahwah, NJ.

Rouet, J.F., 2006, *The skills of document use: From text comprehension to web-based learning*, Lawrence Erlbaum Associates, Mahwah, NJ.

Rouet, J.F., Britt, M.A., Mason, R.A. & Perfetti, C.A., 1996, 'Using multiple sources of evidence to reason about history', *Journal of Educational Psychology* 88(3), 478–493. https://doi.org/10.1037/0022-0663.88.3.478

Royer, J.M., Carlo, M.S., Dufresne, R. & Mestre, J., 1996, 'The assessment of levels of domain expertise while reading', *Cognition and Instruction* 14(3), 373–408. https://doi.org/10.1207/s1532690xci1403_4

Sobel, M.E., 1982, 'Asymptotic confidence intervals for indirect effects in structural equation models', *Sociological Methodology* 13(1982), 290–312. https://doi.org/10.2307/270723

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. https://doi.org/10.1037//0022-0663.91.2.301

Wolfe, M.B. & Goldman, S.R., 2005, 'Relations between adolescents' text processing and reasoning', *Cognition and Instruction* 23(4), 467–502. https://doi.org/10.1207/s1532690xci2304_2