IEICE TRANS. INF. & SYST., VOL.E101–D, NO.4 APRIL 2018

PAPER Special Section on Intelligent Information and Communication Technology and its Applications to Creative Activity Support

Investigative Report Writing Support System for Effective Knowledge Construction from the Web

Hiroyuki MITSUHARA[†a)], ***Member***, **Masami SHISHIBORI**[†], ***Nonmember***, ***and*** **Akihiro KASHIHARA**[††], ***Member***

SUMMARY Investigative reports plagiarized from the web should be eliminated because such reports result in ineffective knowledge construction. In this study, we developed an investigative report writing support system for effective knowledge construction from the web. The proposed system attempts to prevent plagiarism by restricting copying and pasting information from web pages. With this system, students can verify information through web browsing, externalize their constructed knowledge as notes for report materials, write reports using these notes, and remove inadequacies in the report by reflection. A comparative experiment showed that the proposed system can potentially prevent web page plagiarism and make knowledge construction from the web more effective compared to a conventional report writing environment.

key words: *web page plagiarism, copy and paste, investigative report, knowledge construction, web*

1. Introduction

The web has become an indispensable source of information. In other words, knowledge construction from the web has been popularized. However, such knowledge construction occasionally fails because not all information on the web is credible (well-authenticated) [1]. Several studies (e.g., [2]) have pointed out that Wikipedia, an open web-based encyclopedia that appears well-organized, contains low-credibility information due to its openness (being collaboratively edited by the public). The low credibility of the web, potentially due to the widespread prevalence of social media, content farms, and fake new sites, has been recognized as a social problem [3].

In educational institutions (e.g., universities), students often use the web as their main information source when writing investigative reports. Investigative report writing, which attaches importance to collecting as many facts as possible rather than expressing opinions, can be considered knowledge construction that requires students to browse many web pages and externalize knowledge. Therefore, to make knowledge construction more effective, students should pay attention to the credibility of the information, i.e., they should evaluate the authenticity of information (e.g., text and images) in web pages before writing their reports. For example, information comparison is an effective way to evaluate credibility [4], [5]. However, students may simply copy and paste information without confirming its credibility and use it without a citation, i.e., plagiarism [6]. In addition to potentially using unauthenticated information in their report, web page plagiarism can violate copyrights. Thus, we consider that web page plagiarism should be eliminated for effective knowledge construction (report writing).

Plagiarized reports, which will reveal ineffective knowledge construction, have been recognized as a social problem; therefore, various countermeasures have been proposed [7]. For example, plagiarism detection tools (e.g., [8]) not only reduce teacher workload, they also act as a deterrent. However, teachers eventually must examine all potentially plagiarized reports to determine web page plagiarism for themselves.

In this study, rather than detecting plagiarized reports from many submitted reports, we aim to prevent web page plagiarism and make students evaluate information by restricting their ability to copy and paste text from web pages when writing investigative reports. Thus, we developed an investigative report writing support system.

The remainder of this paper is organized as follows. Section 2 outlines our investigative report writing model. The proposed system is described in Sect. 3. Section 4 discusses an experimental evaluation of the system. Conclusions and suggestions for future work are presented in Sect. 5. Note that this paper has been revised based on our previous work [9] and includes new perspectives relative to investigative report writing, experimental results, and considerations.

2. Investigative Report Writing Model

In this study, knowledge construction in investigative report writing is simply defined as "remembering facts (term explanations) that comprise the report's topic in their own words (commentary)." Such remembered facts are referred to as constructed knowledge. For example, if the assignment is to "Investigate deep learning," students are expected to remember facts regarding not only "deep learning" but also "neural networks" "pattern recognition," "AlphaGo," etc.

Knowledge construction, which is a complicated cognitive activity (process), has been represented by various models. Marchionini [10] proposed a search activity model that can be applied to knowledge construction from the web. This model consists of three activities, i.e., lookup, learn, and investigate. The knowledge construction process be-

Manuscript received March 15, 2017.
Manuscript revised July 18, 2017.
Manuscript publicized January 19, 2018.

[†]The authors are with Tokushima University, Tokushima-shi, 770–8506 Japan.
[††]The author is with The University of Electro-Communications, Chofugaoka, Chofu-shi, 182–8585 Japan.
a) E-mail: mituhara@is.tokushima-u.ac.jp

DOI: 10.1587/transinf.2016IIP0023

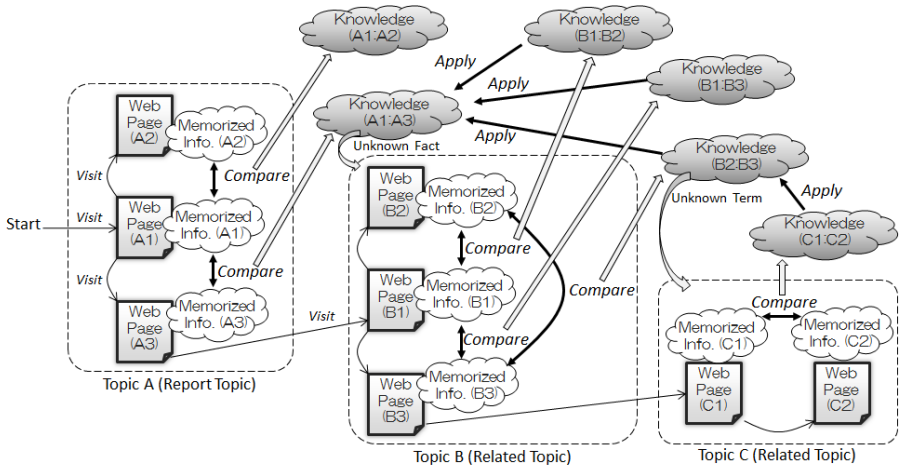

Fig. 2 Knowledge construction through web browsing.

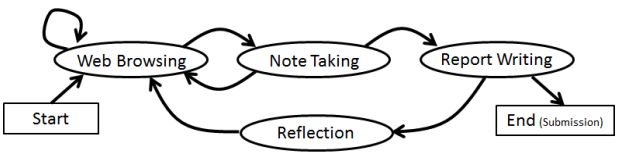

Fig. 1 Investigative report writing model.

gins with the search for knowledge in an indefinite state and draws to an end through continuously exploring the web. Kashihara et al. [11] proposed a knowledge construction process model that categorizes the movement between web pages into six activities, i.e., supplement, elaborate, compare, justify, rethink, and apply. Referring to these models, we previously proposed an investigative report writing model that consists of four activities, i.e., web browsing, note taking, report writing, and reflection [12]. Figure 1 shows an overview of that model.

2.1 Web Browsing (Knowledge Construction)

A student browses many web pages to collect the information (facts) needed for an investigative report, starting by using a search engine to browse the web. We revised our previous model by dividing web browsing into three activities. Figure 2 shows the revised model focusing on web browsing.

(1) Information memorization

For each web page, the student memorizes relevant text and images. We regard information memorization as the initial activity for investigative report writing. The memorized information may not be included in the constructed knowledge, rather it is considered a primary source for knowledge construction. If memorized information is used directly in a report, the report will be considered plagiarized.

(2) Information comparison

To verify facts from the web, the student browses multiple pages about the same topic and then compares their memorized information. In other words, they construct knowledge by recognizing similarities and discrepancies among the information. Constructed knowledge must differ from memorized information, and this is achieved through information comparison (interpretation). Thus, information comparison is an indispensable activity for evaluating the information and determining facts. Note that, in this model, information compared simultaneously is limited to two web pages.

(3) Knowledge application

When discovering an unknown fact (or term) in information comparison (or information memorization), the student shifts to secondary information comparison (i.e., browses web pages that include information about the unknown fact and compares the information) to verify the fact. Then, they apply the knowledge constructed from the secondary information comparison to the knowledge being constructed, i.e., the secondary information complements the knowledge using the verified fact (most credible fact). Thus, knowledge application is a complementary activity for information comparison.

2.2 Note Taking (Knowledge Externalization)

The student externalizes the knowledge constructed (or being constructed) through information comparison by writing notes as digital data in order to remember the constructed knowledge. The notes (i.e., externalized knowledge) become the materials (interim media) for report writing. Students can move from web browsing to note taking each time they compare information and apply the knowledge.

2.3 Report Writing (Knowledge Reconstruction and Re-externalization)

After sufficient notes have been taken, the student writes a report. During the writing process, the student compares and applies the notes (constructed knowledge) while considering the report configuration. Report writing can be considered knowledge reconstruction by scrutinizing constructed knowledge and knowledge re-externalization. According to

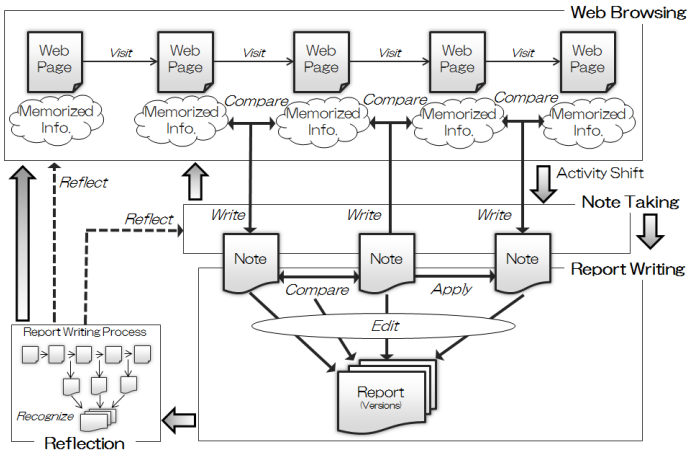

Fig. 3 Iterative process of the proposed investigative report writing model.

this model, the report can be considered a well-organized knowledge base (glossary) that comprises the report's topic. Through this activity, the student can discover inadequacies (e.g., discrepancies) in the report.

2.4 Reflection

To eliminate inadequacies, the student must return to web browsing. For this web browsing to be efficient, the student should determine web pages to be compared again by reflecting on (e.g., looking over) their browsing history and notes.

Reports are not completed through one round of these activities, i.e., multiple iterations are required, as shown in Fig. 3.

3. System

In this study, we consider that preventing web page plagiarism encourages students to evaluate the authenticity of information in web pages and write better investigative reports that include as many facts as possible. In other words, preventing plagiarism leads to more effective knowledge construction.

Based on the proposed model, we developed an investigative report writing support system, called "ReNote," that focuses on preventing web page plagiarism.

3.1 Restricting Copy and Paste

From the perspective of conventional investigative report writing, copy and paste can be performed during note taking (from web pages to notes) and report writing (notes to reports) in the proposed model.

In the proposed model, copy and paste is only restricted during note taking. If students are prevented from copying and pasting information from web pages to notes, the report cannot be plagiarized. Students must write their reports (i.e., construct, externalize, and re-externalize knowledge) according to the proposed model. However, copy and paste is allowed during report writing. This differs from web page plagiarism because the information source is the constructed knowledge externalized as notes. We do not restrict copying and pasting of images during note taking because it is difficult to represent images as notes. Therefore, copying and pasting an image is acceptable as long as the source is identified. Thus, we refer to restricting rather than prohibiting copy and paste. If students transcribe text from web pages verbatim into their notes, this equates to web page plagiarism. However, such typing will incur a heavy burden; therefore, copy-and-paste restrictions are expected to act as a deterrent to plagiarism.

Unless students can write reports efficiently following the proposed model, a large amount of time will be required. As reports typically have submission deadlines, the proposed model must make the various activities easier.

3.2 Configuration and User Interface

In most cases, students browse web pages using a browser, write reports using word processor software, and submit reports (files) via a learning management system or e-mail. Thus, multiple independent applications are used prior to submitting the report. Copy and paste between the browser and word processor is essential from the perspective of exchanging data in ordinary work environments. Therefore, we integrate the minimum required functions into one system rather than integrating such independent applications.

ReNote adopts a client/server configuration (Fig. 4).

3.2.1 Client System

The client system is implemented using Microsoft Visual C# as a Windows application. The main functions of the client system are described below.

(1) Web browsing

Two tabbed browser components (TBC) are arranged in one window (the main window, (Fig. 5 (A)). This function reduces the burden of information memorization and makes information comparison easier. Traditionally, students will

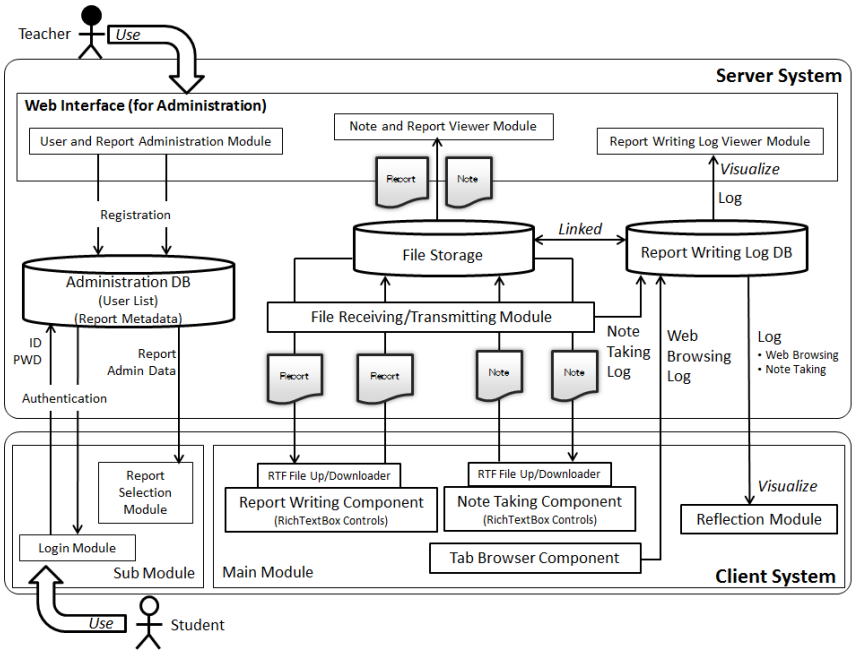

Fig. 4 System configuration.

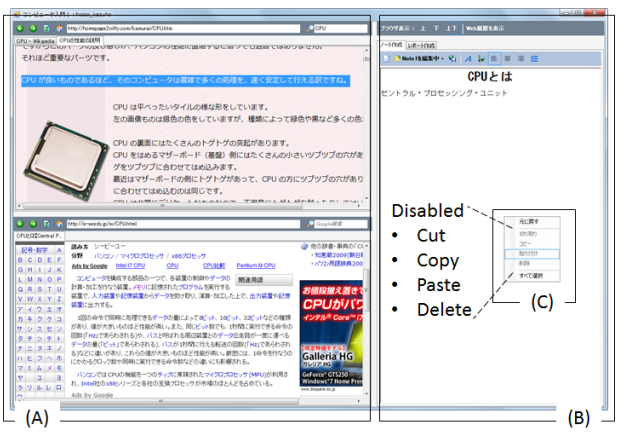

Fig. 5 User interface for web browsing and note taking.

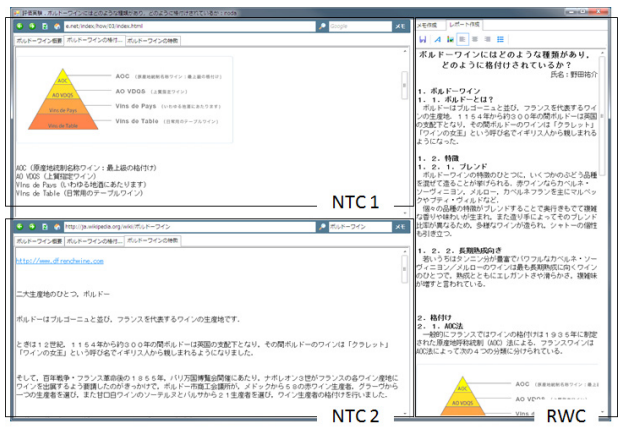

Fig. 6 User interface for report writing.

memorize information from two pages by switching tabs on a web browser that does not display multiple pages simultaneously. Through this function, students (users) can instantly switch between pages in each TBC and easily compare information in two web pages displayed in the TBCs. Each TBC has buttons to switch between pages and web searching, and provides basic web browsing functionality (e.g., forward and back).

(2) Note taking

The note taking component (NTC) is positioned to the right of the TBCs (Fig. 5 (B)). The NTC is implemented using RichTextBox controls and provides text formatting functions (e.g., alignment and font size) and basic functions for note taking (e.g., undo). Furthermore, with the NTC, notes can only be saved in Rich Text Format. With this function, students can instantly externalize knowledge constructed by information comparison and cannot edit the notes using other software.

Copy-and-paste restrictions are implemented in the NTC. Students must type the text and create the notes. The copy-and-paste restrictions are realized by setting the ShortcutsEnabled property in the RichTextBox control to False. The Clipboard.GetImage method determines whether the copied information (clipboard data) is text or other data. When it is an image, copying and pasting is permitted (ShortcutsEnabled=True). If the copied information is text, the “Paste” option in the NTC pop-up menu is disabled (Fig. 5 (C)).

(3) Report writing function

For report writing, the NTC switches to the report writing component (RWC), and the two TBCs switch to the NTC (Fig. 6). The RWC has the same functionality as the NTC. The created notes are stored on the tabs on the respective NTCs. Through this function, the students can instantly

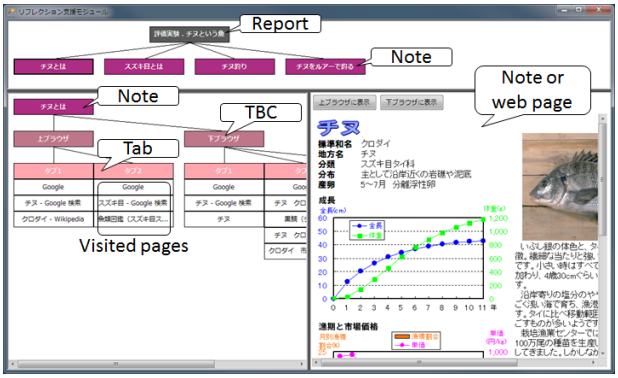

Fig. 7 User interface for reflection.

switch between notes and create the report while easily comparing and applying their externalized knowledge (the content of the two notes).

(4) Reflection function

Through reflection, the report writing process (web browsing process and related notes) are visualized in a separate window (the reflection window, (Fig. 7)). The notes (titles) comprising the report are displayed in creation order from the left, and the web pages (titles) are displayed by the TBC or tab from the top in the visited order. The web pages displayed when creating the note are emphasized and displayed at the lowest level. With this function, students can oversee the entire process and return to web browsing while recognizing the knowledge to be constructed.

3.2.2 Server System

The server system, implemented using a Linux, Apache, MySQL, and PHP configuration, sends and receives notes and reports to/from the client system and manages web browsing process after user authentication. In addition, the server system provides a web interface that enables teachers to create report assignments and confirm the status of student submissions.

3.3 System Usage Flow

Here, we focus on student operations and ReNote behavior, and show the system use flow.

(1) User authentication

When starting ReNote (the client system), a student is required to input their ID and password for user authentication.

(2) Report selection

If user authentication is successful, the report assignments (e.g., theme, deadline, and submission status) are displayed in a list. When the student selects the saved report, the previous state is reproduced. When the student creates a new report, the search engine pages (on the TBCs) and a blank note (on the NTC) are displayed in the main window.

(3) Web browsing

By inputting a search query or URL, a search result is displayed in the TBC. By clicking a link in the search result or displayed web pages, the linked page is displayed in the TBC. The student can freely increase the number of tabs in each TBC.

(4) Note taking

When two web pages displayed in the TBCs are combined for the first time, a blank note is displayed in the NTC to encourage knowledge externalization. Due to copy-and-paste restrictions, the student externalizes knowledge to the NTC by inputting text using the keyboard.

(5) Report writing

By clicking the “Write Report” tab at the top of the NTC, the student can transition to report writing. The student writes reports using the same operations as the NTC, and, as there are no restrictions on copy and paste in the RWC, text and images in the NTC (notes) can be copied and pasted to the RWC (reports).

(6) Reflection

By clicking the “Display browsing process” tab at the top right of the main window, the student can transition to reflection. The student reviews the visualized report writing process (the structure of notes and visited web pages) in the reflection window. In addition, by clicking the title box, the student can compare notes and web pages in the main window.

(7) Report submission

The student can save and submit a created report as many times as needed within the submission deadline. The report file is not saved on the client computer and cannot be submitted other than via ReNote. In other words, students cannot submit plagiarized reports that were written in any application other than ReNote or by other people.

3.4 Related Studies

Since the advent of the web, knowledge construction from the web has remained a central theme [13], [14]. Many studies have focused on cognition in knowledge construction. Gerjets et al. [15] stated the importance of information comparison (inter-category and horizontal category comparisons) through learning with hypermedia, such as the web, and they reported that a comparison tool improved problem-solving performance. AlAgha [16] developed a web navigational learning tool by defining a goal-based knowledge construction model. This tool encourages learners to plan navigation paths by visualizing their mental model (semantic structure of the content of their visited web pages) as a graph and then adapts hyperlinks and annotations on the visited pages to the navigation paths (information needs). Although sharing some attributes with these studies, ReNote focuses on information comparison as a central activity in knowledge construction from the web.

ReNote is positioned as an e-notebook system for externalizing constructed knowledge as digital data. For example, NetNotes [17] is also an e-notebook system that allows text and images in web pages to be copied and pasted to notes while preserving their visual characteristics (e.g.,

layout) and link information to make information collection more efficient. WebAnnot [18], which is a Firefox add-on annotation tool that can be used as an e-notebook system, enables users to annotate web pages semantically based on ontologies of annotation objectives. ReNote is not multifunctional as an e-notebook system but differs from existing systems in that it prevents web page plagiarism (i.e., copy-and-paste restrictions) in investigative report writing.

In terms of reflection support, there are systems that visualize knowledge construction from the web. For example, Saito et al. [19] developed a system that visualizes web browsing processes based on a search operation schema and enables students to reflect on their constructed knowledge by reviewing the browsing process. O'Rourke et al. [20] developed a reflection support system focusing on report writing. With this system, report content is analyzed, and the meaningful distance between paragraphs is calculated, and, by visualizing the report structure in map form, it is possible to generally reflect on the logical flow of the report. ReNote does not visualize web browsing processes in detail or analyze report content. In addition, ReNote differs from existing systems in that it visualizes the report writing process, including interim media (e.g., notes).

4. Experimental Evaluation

We conducted a comparative experiment to examine whether ReNote prevents web page plagiarism and facilitates effective knowledge construction.

4.1 Method

4.1.1 Participants

The subjects, 30 undergraduate and graduate students in information engineering, were divided into three even groups based on a pre-questionnaire about their report writing styles (e.g., dependence on copy and paste) and interests in assigned report topics.

(1) Group A (10 subjects)

These 10 subjects wrote reports with their usual report writing environment (computer and software) and submitted reports via e-mail.

(2) Group B (10 subjects)

These 10 subjects wrote and submitted reports with ReNote without using the reflection function.

(3) Group C (10 students)

These 10 subjects wrote and submitted reports with full ReNote functionality. Note that these participants were not forced to use the reflection function.

4.1.2 Report Writing Assignments

The subjects were given an assignment in each of three weeks and were required to submit their reports within one week. There were no rules regarding the report content (e.g., minimum number of words).

- First-week assignment (A1): "Investigate how to catch black seabream."
- Second-week assignment (A2): "Investigate the authenticity of beauty effect by collagen."
- Third-week assignment (A3): "Investigate the types and classifications of Bordeaux wine."

Topic A1, which was unknown to most subjects and included many possible facts, could encourage them to browse many pages to collect facts. Topic A2, which was not interesting to most subjects, could encourage information comparison to determine information authenticity. Topic A3, which was interesting but unfamiliar for most subjects, could encourage web browsing to satisfy interest.

4.1.3 Procedure

Prior to the experiment, the subjects in groups B and C were simply instructed about the methods to use ReNote, including the copy-and-paste restriction, precluding report writing, and submission with external software.

(1) Report writing (and submission)

Although given each assignment at the same time, the subjects could start report writing at any time within one week. The subjects timed their report writing manually and declared a rounded time when submitting their reports.

(2) Surprise test

For each assignment, the subjects were given a surprise test 11 days after submitting a report. The test consisted of 10 ten-choice-questions about facts that comprise the report topic. If the subject answered all questions correctly, they were given a score of 10 points. For example, a question for A2 was "What generates gelatin when added to collagen? Choose the correct one."

(3) Questionnaire

Immediately after the A3 test, the subjects were given a post-questionnaire (five-scale Likert).

4.2 Results

We analyzed the results with non-parametric statistics assuming an unpredictable population distribution and heteroscedasticity resulting from the variety of report writing styles and situations (e.g., other urgent assignments).

4.2.1 Reports

Table 1 shows the medians and mean ranks for report writing time (T_RepW), the number of written characters in the report ($N\_CharRep$), copy-and-paste rate in the report ($R\_CpPasRep$), and the efficiency of report writing, which excludes manifest copy and paste (E_RepW). $R\_CpPasRep$ was calculated for each report (i) for each subject (j) as follows.

$$R\_CpPasRep(i, j) = \frac{N\_CharEstCpPas(i, j)}{N\_CharRep(i, j)}$$

Here, $N\_CharEstCpPas$ is the number of estimated

Table 1 Medians and mean ranks regarding reports.

	A1			
	T_RepW (min.)	*N_CharRep*	*R_CpPasRep*	*E_RepW*
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
Group A	60 (10.90)	1,153 (18.50)	0.098 (19.65)	16.57 (20.70)
Group B	105 (18.85)	658.5 (11.40)	0.045 (11.75)	5.86 (10.10)
Group C	80 (16.75)	873 (16.60)	0.054 (15.10)	8.08 (15.70)
	A2			
	T_RepW (min.)	*N_CharRep*	*R_CpPasRep*	*E_RepW*
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
Group A	30 (7.50)	910.5 (20.50)	0.138 (17.05)	23.93 (23.10)
Group B	120 (22.0)	426 (10.25)	0.212 (18.15)	2.75 (7.60)
Group C	60 (17.0)	731.5 (15.75)	0.091 (11.30)	7.39 (15.80)
	A3			
	T_RepW (min.)	*N_CharRep*	*R_CpPasRep*	*E_RepW*
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
Group A	55 (9.20)	844 (17.20)	0.419 (19.50)	10.55 (18.40)
Group B	90 (19.65)	590.5 (12.20)	0.217 (14.10)	4.77 (11.60)
Group C	60 (17.67)	822 (17.10)	0.196 (12.90)	8.21 (16.50)

Table 2 Medians and mean ranks of test scores.

	A1	A2	A3
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
Group A	3 (14.00)	3.5 (13.55)	2.5 (11.40)
Group B	4 (16.55)	4 (16.40)	3 (15.80)
Group C	4.5 (15.95)	4 (16.55)	4.5 (19.30)

copied and pasted characters. The estimated characters were counted manually based on the result of a plagiarism detection tool.

E_RepW was defined as the number of characters that were not copied and pasted (i.e., other than the estimated copied-and-pasted characters) written per minute and calculated as follows.

$$E\_RepW(i,j) = \frac{N\_CharRep(i,j) \times (1 - R\_CpPasRep(i,j))}{T\_RepW(i,j)}$$

For *T_RepW*, the medians of groups B and C were greater than those of group A for all assignments. From multiple comparisons (Steel-Dwass test), significant differences were found between groups A and B ($p = 0.001$) and A and C ($p = 0.024$) for A2, and between groups A and B ($p = 0.020$) for A3. For *N_CharRep*, the medians of groups B and C were less than that of group A for all assignments. Multiple comparisons revealed a significant difference between groups A and B ($p = 0.041$) for A2. For *R_CpPasRep*, the medians of groups B and C were less that of group A, except for A2; however, variance analysis (Kruskal-Wallis test) revealed no significant differences. For *E_RepW*, the median of group A was greater than the medians of groups B and C, for all assignments. Multiple comparisons revealed significant differences between groups A and B ($p = 0.033$) for A1 and between groups A and B ($p = 0.001$) and B and C ($p = 0.033$) for A2.

4.2.2 Test Scores

Table 2 shows the medians and mean ranks of the test scores (*TScore*). For *TScore*, the medians of groups B and C were greater than that of group A for all assignments. However, variance analysis revealed no significant differences.

4.2.3 Notes

Table 3 shows the medians and mean ranks of the number of notes (*N_Note*), the number of written characters in notes (*N_CharNote*), the number of visited (browsed) pages during report writing (*N_VisitPage*), and the copy-and-paste rate in notes (*R_CpPasNote*). *R_CpPasNote* was calculated analogous to *R_CpPasRep*—the target simply changes from reports to notes. Note that *N_VisitPage* included revisited pages—*N_VisitPage* of group A could not be collected. For *N_Note*, only the median of group C was greater than that of group B for A1. For *N_CharNote*, the median of group C was greater than that of group B for all assignments. In contrast, for *N_VisitPage*, only the median of group C was greater than that of group B for A1. The Mann-Whitney U test revealed a significant difference for only *N_CharNote* for A1. For *R_CpPasNote*, the medians of groups B and C were relatively high (higher than expected). There were no significant differences relative to *R_CpPasNote* between the groups. Each report, except seven, included at least one copied-and-pasted sentence. The maximum rate was 1.0 by two subjects, i.e., these subjects copied and pasted text to notes despite using ReNote.

4.2.4 Reflection

Table 4 shows the usage situations of the reflection function, such as the number of times the function was activated (*N_ActRef*), time taken for reflection (*T_UseRef*), and the number of web pages or notes revisited directly from

Table 3 Medians and mean ranks regarding notes.

	A1			
	N_Note	*N_CharNote*	*N_VisitPage*	*R_CpPasNote*
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
Group B	3 (9.20)	345.5 (7.20)	63.5 (10.00)	0.206 (8.30)
Group C	5 (11.80)	940 (13.80)	88 (11.00)	0.406 (12.70)
	A2			
	N_Note	*N_CharNote*	*N_VisitPage*	*R_CpPasNote*
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
Group B	3 (11.20)	426 (9.00)	47.5 (11.50)	0.421 (9.90)
Group C	3 (9.90)	667.5 (12.00)	40 (9.50)	0.411 (11.10)
	A3			
	N_Note	*N_CharNote*	*N_VisitPage*	*R_CpPasNote*
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
Group B	2.5 (10.35)	590.5 (9.40)	41.5 (11.30)	0.440 (11.25)
Group C	2.5 (10.65)	972 (11.60)	34.5 (9.70)	0.248 (9.75)

Table 4 Medians and mean ranks regarding reflection function.

	N_ActRef	T_UseRef (min.)	*N_RevisitPageNote*
	Med. (M. rank)	Med. (M. rank)	Med. (M. rank)
A1	4 (22.50)	117 (22.45)	3.5 (19.20)
A2	1 (13.90)	29 (12.85)	1 (14.75)
A3	0 (10.10)	0 (11.20)	0 (12.55)

Table 5 Questionnaire result.

Question "*Do you agree that . . . ?*" Options = {1: Definitely no, 2: No, 3: Neutral, 4: Yes, 5: Definitely Yes}	Group A	Group B	Group C
Q1. You depended on copy and paste during report writing.	5	-	-
Q2. Enabling (for Group A)/Restricting (for Groups B and C) copy and paste did not give you burdens.	5 (23.55)	1 (10.95)	1.5 (12.00)
Q3. Enabling/Restricting copy and paste made your report writing efficient.	5 (27.23)	1 (11.59)	1 (12.18)
Q4. Enabling/Restricting copy and paste increased effect of knowledge construction through your report writing.	2 (8.18)	4 (24.18)	3 (18.64)
Q5. You compared information within web pages in note taking.	-	5 (10.50)	5 (10.50)
Q6. You compared the contents of notes in report writing.	-	4 (10.70)	4 (10.30)
Q7. Note taking made your report writing efficient.	-	4 (9.50)	4 (11.50)
Q8. Note taking made your report writing effective.	-	4 (9.65)	4 (11.35)
Q9. Visualized report writing process was useful for report writing.	-	-	4
Q10.The reflection function improved the quality of your report.	-	-	4
Q11.You used the reflection function easily.	-	-	3.5

the function (*N_RevisitPageNote*). T_UseRef was counted accumulatively while the function (reflection window) was displayed. The medians of all items decreased with each assignment—finally reaching 0 for A3.

4.2.5 Questionnaire

Table 5 shows the main questions in the questionnaire and the medians and mean ranks (for partial questions) of the subjects' replies. The median of Q1 for only group A indicated that many subjects depended on copy and paste in this experiment. For all groups, Q2, Q3, and Q4 asked about copy and paste. For Q2 and Q3, the medians of group A were significantly greater than those of groups B and C. In contrast, for Q4, the medians of groups B and C were greater than that of group A. Multiple comparisons revealed significant differences between groups A and B ($p = 0.005$) and groups A and C ($p = 0.002$) for Q2, between groups A and B ($p < 0.001$) and A and C ($p < 0.001$) for Q3, and between groups A and B ($p < 0.001$) and A and C ($p = 0.010$) for Q4. For groups B and C, Q5-Q8 asked about note taking and report writing. All medians were greater than or equal to 4. The Mann-Whitney U test did not reveal a significant difference between groups B and C. The medians for Q9-Q11 about reflection for only group C were 4, 4, and 3.5, respectively.

4.3 Considerations

The differences in *T_RepW* may result from the copy-and-paste restriction because the subjects in groups B and C had to type text as notes. The differences in *E_RepW*, which may also result from the restriction, can be regarded as a measure of the efficiency by which subjects construct knowledge. The medians of *T_RepW* and *E_RepW* for Q2 and Q3 indicate that the restriction was a burden for these two groups, which made report writing inefficient. On the other hand, the high median for Q1 indicates that subjects in group A could write their reports efficiently by copying and pasting. Thus, we think that copy and paste should be restricted but can contribute to efficient report writing.

The lower medians of *N_CharRep* for groups B and C may result from note taking and report writing. We read the reports from all subjects to understand how they copied and

pasted. The subjects in group A tended to copy and paste text, including long sentences, directly from web pages. Regardless of this tendency, the medians of *R_CpPasRep* for group A were low. This was because they tended to partly alter the copied-and-pasted text while keeping its meaning or adding their own opinion on the report topics. This alternation can be inferred from the high Q1 median since the subjects in group A must have tentatively copied text from web pages to their reports and then altered the pasted text. On the other hand, although some subjects in groups B and C tended to type text from web pages verbatim into their notes, the others tended to condense or itemize text. The condensed or itemized text can be regarded as their constructed knowledge; namely, they remembered facts, which comprised the report's topic, in their own words. These tendencies may have been reflected in the medians of *TScore*, which indicate that subjects in groups B and C constructed knowledge more effectively. In addition, the medians of Q4 indicate that many subjects realized that they should not copy and paste to achieve effective knowledge construction.

In general, longer report writing time is expected to entail more effective knowledge construction. The lower medians of groups B and C for *E_RepW* may indicate that the subjects spent more time condensing or itemizing text, and the *N_CharRep* was accordingly decreased. For effective knowledge construction, the copy-and-paste restriction does not simply increase *T_RepW* but rather increases the time of condensing or itemizing text.

The values of *N_Note* were far less than expected. Subjects in groups B and C did not take notes frequently. We read the notes from all subjects to understand what they typed. The medians of *R_CpPasNote* show that subjects tended to type text from web pages verbatim into their notes, although many condensed or itemized text components were found. This tendency may indicate that they considered such note taking (verbatim typing) to be a reasonable method in their report writing process. In groups B and C, the medians of *R_CpPasRep* were less than those of *R_CpPasNote*. This may indicate that the subjects did not copy and paste but scrutinized (i.e., compared and applied) text within notes consciously to write better reports, and such a process can considered successful report writing (knowledge reconstruction) that includes note taking. From the favorable medians of Q5 and Q6, we consider that ReNote promotes information comparison (for web pages and notes) for effective report writing. Furthermore, from the favorable medians of Q7 and Q8, we consider that, although time-consuming and burdensome, note taking is indispensable for efficient and effective report writing.

The copy-and-paste restriction did not work satisfactorily and web page plagiarism could not be prevented. However, compared to a typical report writing environment, we think that note taking and the copy-and-paste restriction contribute to effective report writing (i.e., knowledge construction) by encouraging students to condense or itemize text within web pages.

The medians of Q9 and Q10, which were favorable, may indicate that the subjects recognized the need to reflect for effective report writing. The median of Q11 was relatively favorable. On the other hand, the medians of *N_ActRef*, *T_UseRef*, and *N_RevisitPageNote* may indicate that the reflection function was not used frequently. The subjects in group C may have had the ability to write reports successfully without relying on the function, i.e., they may have become increasingly familiar with report writing by using ReNote for each assignment. Although *N_RevisitPageNote* for *N_VisitPage* and *N_Note* in group C is limited, the subjects were encouraged to revisit web pages or notes via the function. The medians of *N_CharNote* for group C may indicate that the reflection function encouraged the revision of notes. Thus, we consider that the reflection function can contribute to effective report writing but should be improved for more frequent use.

Overall, from the above results and considerations, we conclude that ReNote should be improved but can potentially prevent web page plagiarism using the copy-and-paste restriction and making investigative report writing (knowledge construction) from the web more effective. In other words, ReNote does not necessarily increase efficiency but does increase effectiveness in investigative report writing.

5. Conclusion

This paper has described an investigative report writing support system that can potentially prevent web page plagiarism to facilitate effective report writing, i.e., effective knowledge construction, based on our previously proposed investigative report writing model. The proposed system attempts to prevent plagiarism by restricting copying and pasting from web pages to notes (materials for report writing) and adopting an approach that differs from detecting plagiarism from many submitted reports. The results of a comparative experiment demonstrate that the proposed system has possible beneficial effects for effective knowledge construction.

There may be advantages and disadvantages relative to copy-and-paste restrictions, especially in terms of student load and knowledge construction efficiency. As long as there is no plagiarism in the final report, there may be no problem copying and pasting while note taking (accumulation of report materials). In other words, there is a concern that restricting copy and paste burdens students and thus reduces report writing efficiency. In the experiment, some subjects typed text from web pages verbatim into notes regardless of the copy-and-paste restriction. This may indicate that copy and paste is a reasonable activity for students and should not be restricted in report writing. If a report writing assignment attaches importance to opinion expression, copy and paste (with citation) should be allowed to facilitate efficient report writing. However, from the reports produced in the experiment, we found that the amount of copied and pasted text decreased in their reports, i.e., subjects scrutinized note content and revised text using their own words. As a result, submitted reports did not contain significant amounts of copied and pasted text. We interpret

that the proposed system is useful for preventing web page plagiarism in final reports.

Note that the reports written by the subjects were not considered for academic credits. In addition, our definition of knowledge construction from the web may be simplified compared to reality. Therefore, we must examine whether the system can prevent web page plagiarism and increase the effectiveness of knowledge construction relative to more realistic and complicated investigative report writing.

Acknowledgments

This study was supported in part by Grant-in-Aid for Scientific Research (B) No. 26282047 from the Japan Society for the Promotion of Science. We express our thanks to K. Hosoi and T. Komatsu for their efforts in this study.

References

[1] D.J. Grimes and C.H. Boening, "Worries with the Web: A Look at student use of web resources," College and Research Libraries, vol.62, no.1, pp.11–22, Jan. 2001.
[2] P. Denning, J. Horning, D. Parnas, and L. Weinstein, "Wikipedia risks," Communications of ACM, vol.48, no.12, p.152, Dec. 2005.
[3] M.J. Metzger and A.J. Flanagin, "Credibility and trust of information in online environments: The use of cognitive heuristics," J. Pragmat., vol.59, part B, pp.210–220, Dec. 2013.
[4] M. Meola, "Chucking the Checklist: A contextual approach to teaching undergraduates web-site evaluation," Libraries and the Academy, vol.4, no.3, pp.331–344, July 2004.
[5] S.Y. Rieh and B. Hilligoss, "College students' credibility judgments in the information-seeking process," Digital Media, Youth, and Credibility, Foundation Series on Digital Media and Learning, The MIT Press, Cambridge, MA, 2008.
[6] N. Dalal, "Responding to plagiarism using reflective means," International Journal for Educational Integrity (IJEI), vol.11, no.4, June 2015.
[7] H. Maurer, F. Kappe, and B., Zaka, "Plagiarism-A Survey," J Univers. Comput Sci., vol.12, no.8, pp.1050–1084, Aug. 2006.
[8] C.J. Neill and P. State, "A web-enabled plagiarism detection tool," IT Professional, vol.6, no.5, pp.19–23, Sept. 2004.
[9] H. Mitsuhara, M. Shishibori, and A. Kashihara, "Preventing web page plagiarism in investigative report writing," Proc. of 7th International Conference on the Applications of Digital Information and Web Technologies (ICADIWT2016), pp.162–171, April 2016.
[10] G. Marchionini, "Exploratory search: from finding to understanding," Communications of the ACM, vol.49, no.4, pp.41–46, April 2006.
[11] A. Kashihara and S. Hasegawa, "A model of meta-learning for web-based navigational learning," International Journal of Advanced Technology for Learning, vol.2, no.4, pp.198–206, 2005.
[12] K. Hosoi, H. Mitsuhara, and Y. Yano, "Fundamental support and reflection support for report writing from web," Workshop Proc. of the 18th Intl. Conf. on computers in education, pp.117–124, Nov. 2010.
[13] J.R. Mill and M.J. Hannafin, "Cognitive strategies and learning from the world wide web," Educational Technology Research and Development (ETR&D), vol.45, no.4, pp.37–64, Dec. 1997.
[14] J.-J. Lo, Y.-C. Chan, and S.-W. Yeh, "Designing an adaptive web-based learning system based on students' cognitive styles identified online," COMPUT EDUC., vol.58, no.1, pp.209–222, Jan. 2012.
[15] P. Gerjets, K. Scheiter, and J. Schuh, "Information comparisons in example-based hypermedia environments: supporting learners with processing prompts and an interactive comparison tool," Educ Technol Res Dev., vol.56, no.1, pp.73–92, Feb. 2008.
[16] I. AlAgha, "KnowledgePuzzle: "A browsing tool to adapt the web navigation process to the learner's mental model"," J Educ Techno Soc., vol.15, no.3, pp.275–287, July 2012.
[17] Y.J. Reimer and S.A. Douglas, "Implementation challenges associated with developing a web-based E-notebook," JODI., vol.4, no.3, Feb. 2003.
[18] F. Azouaou, H. Mokeddem, L. Berkani, A. Ouadah, and B. Mostefai, "WebAnnot: a learner's dedicated web-based annotation tool," Int. J. Technology Enhanced Learning, vol.5, no.1, pp.56–84, Jan. 2013.
[19] H. Saito and K. Miwa, "Construction of a learning environment supporting learners' reflection: A case of information seeking on the Web," COMPUT EDUC., vol.49, no.2, pp.214–229, Sept. 2007.
[20] S.T. O'Rourke, R.A. Calvo, and D.S. McNamara, "Visualizing topic flow in students' essays," Educ Techno Soc., vol.14, no.3, pp.4–15, July 2011.

Hiroyuki Mitsuhara received the B.E. and M.E. degrees from Kindai University in 1998 and 2000, and then he received the Ph.D. degree from Tokushima University in 2003. He is currently an associate professor at Tokushima University. His research interests include web-based learning, game-based learning, entertainment computing, and human-computer interaction, and disaster education.

Masami Shishibori was born in 1965. He received his BS Degree in 1991, his MS Degree in 1993 and PhD Degree in 1997, from Tokushima University, Japan. He is currently a Professor in the Department of Computer Science, Graduate School of Science and Technology, Tokushima University, Japan. His research interests include multimedia processing, information retrieval, and natural language processing.

Akihiro Kashihara received his BS and MS in Computer Science and System Engineering from Tokushima University in 1987 and 1989, and his PhD in Information and Computer Science from Osaka University in 1992. He was an Assistant/Associate Professor of the Institute of Scientific and Industrial Research in Osaka University from 1992 to 2004. He is currently a Professor of Graduate School of Informatics and Engineering, The University of Electro-Communications. His main research interests lie in modelling of learning process and designing intelligent learning environments.