國立政治大學資訊管理學系

碩士學位論文

以眼動實證研究探討個人差異於教育輔助平台視覺分析上之影響

The impact of individual differences on visual analytics of an orchestration platform: An empirical study using eye-tracking

指導教授：林怡伶博士

研究生：李明緯 撰

中 華 民 國 108 年 7 月

DOI:10.6814/NCCU201901137

ACKNOWLEDGEMENT

To my family,
To participants, joining the user study to support my data collection,
To my advisor, Dr. Yi-Ling Lin, giving me guidance and helping me to complete my research,
To the thesis committee, Dr. Yen-Chun, Chou & Dr. I-Chin Wu,
To the community that is exploring learning analytics.

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201901137

摘要

本研究著重於探討學習目標導向、視覺化圖表格式(折線圖、柱狀圖、雷達圖)與學習類型(程序性學習及推論學習)對學生在線上複習平台中複習紙本程式考試表現的影響。我們透過使用者研究及眼動儀，探討自行開發的視覺化系統之可行性。此研究總共募集了 34 位曾經至少修習過一堂 Java 程式設計課的受測者，並收集了問卷資料、系統紀錄、眼動追蹤數據等相關資料進行後續分析。我們的實驗透過使用迴歸模型驗證學習目標導向、視覺化圖表格式以及學習類型對於使用者在視覺化分析上認知的影響，進而提出以實證研究分析視覺化學習的可行性。我們的實驗結果顯示具有較高學習目標導向的使用者在視覺化分析的輔助下，相對應會有較高的學習表現與學習認知。然而實驗結果也顯示，雷達圖因為組成較為複雜，會對使用者複習時的效率有負面影響。在學習類型方面，實驗結果顯示在視覺化分析的輔助下，使用者在資訊檢索類型的複習表現較推理發想類型更為優越。

關鍵詞：學習分析、圖表理解、資訊視覺化、學習目標導向、紙本考試、教育科技協作、眼動追蹤

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

ABSTRACT

We examined the impact of learning goal orientation, visualization format (line, bar and radar chart) and type of learning task (search fact vs. inference generation) upon a viewer's perception of reviewing paper-based exams in an online virtual assessment environment. A lab experiment was conducted with an eye-tracker. System log, eye-tracking data and questionnaires were collected from 34 students who have taken at least one Java programming course. Our experiments demonstrate the empirical research practicality by using a regression model to validate the effect of learning goal orientation, format and task on user perceptions of visualization analytics. Our results show that the viewers with a high degree of learning goal orientation would have better learning perception of visualization material. Radar graph, however, would have a negative influence on the review performance due to its complicated composition. We also found that with the assistance of visualization analytics, users perform more efficiently on search fact tasks rather than inference generation tasks when reviewing programming exams.

Keywords: learning analytics, graph comprehension, information visualization, learning goal orientation, paper-based assessment, classroom orchestration technology, eye tracking

國立政治大學 National Chengchi University

Table of Contents

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201901137

List of Tables

DOI:10.6814/NCCU201901137

List of Figures

DOI:10.6814/NCCU201901137

Chapter 1 INTRODUCTION

1-1 Background and Motivation

With the development of the Internet and other various technologies, learning styles and environments have changed since last decade. Nowadays, the learning behavior is not limited to the classroom. Instructors can manage classroom activities from any distance through the instructional design that connects different systems (namely web services). As a result, learners can use multiple ways to engage in the learning material. In the context of education, we use "orchestration" to refer to the integrated process, and "classroom orchestration" is defined as how a teacher manages multi-layered activities (i.e., offline and online). Classroom orchestration discusses how and what research-based technologies have been adopted and should be integrated within the physical classrooms (Dillenbourg, 2013; Dillenbourg & Jermann, 2010). For decades it has been used in different types of learning environments. Several innovative systems have been proposed in classroom orchestration to improve students' learning performance (Brusilovsky, Hsiao, & Folajimi, 2011; Denny, Luxton-Reilly, & Hamer, 2008; Hsiao, Bakalov, Brusilovsky, & König-Ries, 2013; Hsiao & Lin, 2017). Most of these studies present innovative Web-based tools based on the concepts of social navigation as well as open student models. User usage and implementation performance results in classroom were also provided to validate the effectiveness of these systems.

While the classroom orchestration provides students with abundance of materials corresponding to various aspects of their learning, the benefit may not be fully realized without proper guidance. Rather than "one-size-fits-all" solutions (such as ordering questions in a fixed sequence), an adaptive guidance should be provided given that students typically have different starting knowledge and learn at different paces (Hsiao, Sosnovsky, & Brusilovsky, 2008). To support adaptive guidance, most classroom

DOI:10.6814/NCCU201901137

orchestration systems come with analytic dashboards which help teachers monitor students' engagement and effectiveness toward a specific subject. Visualizations in dashboards not only summarize general performance indicators like scores, but also visualize advanced indicators like interactions between students and learning content, time spent and corresponding resource using in a virtual classroom (Govaerts, Verbert, Duval, & Pardo, 2012; Hsiao, Pandhalkudi Govindarajan, & Lin, 2016; Hsiao & Brusilovsky, 2012; Lu & Hsiao, 2016). These works show that a dashboard with visualizations guide students to the suitable learning material as well as significantly increases the quality of students' learning and motivation to work with non-mandatory learning content.

Although orchestration technologies have changed the education environment, it is commonly agreed that there is a need to gain insights into students' perceptions on assessments and discover how they behave while dealing with assessment tasks with different requirements (Papamitsiou & Economides, 2015). Some previous works have attempted to investigate adaptive navigation support for self-assessment questions in larger classes with a broader range of question difficulty. Specifically, a series of works were proposed to concentrate on the context of paper-based programming exams, particularly given the fact that paper-based exams are still one of the most practical assessments in large programming courses in school (Hsiao, 2016; Hsiao, Huang, & Murphy, 2017; Hsiao et al., 2008; Paredes, Huang, Murphy, & Hsiao, 2017). These works connected paper-based assessments to the online virtual assessment environment and showed students' performance in the exam in order to provide adaptive user interface for programming exams.

Just as in the previous works, the present study also focuses on the domain of Java programming language, which is now still the language of choice in most introductory programming classes. We implemented an online Java exam reviewing system called

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

Topic Combination Analysis Visualization (TCAV) to provide different kinds of visualization analytics of students' behavior in exams. To be more specific, we extracted topics from predefined rubrics of each question in a Java paper-based exam and applied adaptive visualizations to support the interpretation of students exam performance.

This is a study based on a Java programming orchestration platform with visualization analytics. We attempted to enhance learning awareness to programming learners by providing elaborated visualization results. Students can find patterns between their behavior and performance during the paper-based exam so that they can prepare for the next exam more efficiently. We are interested in whether visualization analytics can benefit the students and, if so, which visualization format is most effective. We are also interested in whether individual differences in psychology can influence the impact of the visualization analytics.

1-2 Research Questions

Learning analytics have been defined as the "*measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs*" (Siemens & Gasevic, 2012). Learning analytics have combined research, methods, and techniques from several fields such as orchestrated learning, information visualization, psychology and HCI. Illustrating student academic performance and providing dedicated feedback have been two of the most frequently adopted tasks associated with learning analytics. To investigate learning achievement, the role of some main psychological indicators are involved in previous studies, including self-efficacy (Wang, Shannon, & Ross, 2010) and locus of control (Joo, Lim, & Kim, 2013). Further, indicators peculiar to the context of learning have been widely applied in recent studies, such as learning engagement

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

(Hu & Hui, 2012) and learning goal orientation (Debicki, Kellermanns, Barnett, Pearson, & Pearson, 2016). These indicators have thoroughly investigated the individual differences and perceptions of various learning contexts.

Although establishing lead indicators of academic performance are essential steps for learning analytics, there has been a gap in empirical studies which have sought to evaluate the impact and transferability of this initial work across domains and contexts. Despite the fact that there are a large number of learning analytics tools developed with innovative approaches and accompanied by elaborate dashboards, most of them are generally not developed from theoretically established instructional strategies, especially those which utilize the trace data and feedback from students (Gašević, Dawson, & Siemens, 2015).

We argued that there is a disheartening lack of necessary empirical research in the field of learning analytics, especially for the validity of orchestration technology learning tools. Few studies have focused on the influence of individual differences and user perceptions in learning analytics tools empirically. Orchestration technology tools should be developed and investigated under the consideration of both psychological indicators and user perceptions. Specifically, we were interested in two factors, learning genres and visualization formats, and we investigated the effects of these two factors between individual differences and students' perceptions. As a result, the goals of this study were a) to depict students' performance in the paper-based programming exam with different visualizations b) to explore the factors that influence the effectiveness of students when they view different visualization formats, and c) to examine the effect of students' personality and their learning comprehension of the visualization analytics.

This study attempts to answer the following research questions:

RQ1. Does the visualization format, task type and individual indifferences influence students' comprehension of the visualization analytics on the

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

proposed orchestration technology platform?

RQ2. Does the visualization format and individual indifferences influence students' understanding of the visualization analytics on the proposed orchestration technology platform?

RQ3. Does the visualization format and individual indifferences influence students' perceived learning of the proposed orchestration technology platform?

1-3 Research Method

In this study, we introduced WPGA (Paredes et al., 2017) as an exam grading tool for the Java programming course taught at the National Chengchi University. During the semester, grading data (i.e., exam questions with answers and corresponding topic rubrics) were collected from WPGA for the development of TCAV, a learning analytics tool for paper-based programming exams. We adopted an exploratory data analysis to design the TCAV prototype. The process is summarized in figure 1.1.

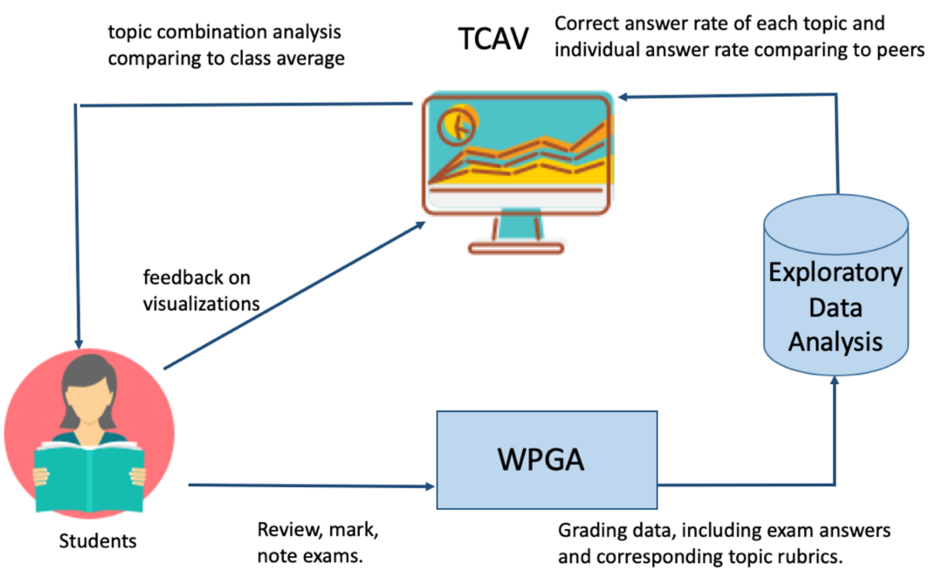


Figure 1.1 The exploratory data analysis framework.

TCAV provides visualizations of students' performance in a series of paper-based programming exams. A lab experiment was conducted to collect both objective and

DOI:10.6814/NCCU201901137

subjective data, including questionnaires and eye-tracking data. An eye-tracker was employed to collect real-time data with several settings from students. Individual differences such as personality were collected through questionnaires. This data was used to identify the impact and capability of adopting such orchestration technology. The participants were student volunteers from students who have taken at least one Java programming course prior to this study.

Adaptive information visualizations of paper-based exams is a main function of TCAV. Since the visualizations were presented as a guidance of reviewing exams, the understanding of how the visualizations were generated and composed is crucial. Corresponding to the increased prevalence of analytic dashboards and visualizations, graph comprehension has been widely used in order to evaluate the efficiency of the visual analytics interface (Ratwani & Boehm-davis, 2008; Ratwani & Trafton, 2008). For our purposes, we defined learning comprehension, which came from graph comprehension, to capture how well students learned from the proposed visualizations. To measure how the visualization provided knowledge for students, we used the concept of perceived learning, which was adapted from the concept of perceived usefulness in TAM (Davis, 1989).

In the research field of cognitive science, there are two broad genres of learning: procedural and inferential, which represent different processes of knowledge creation and use. Integrating this into our context, we divided the process of reviewing exam results into two kinds of task types: fact-retrieval tasks and inference-generation tasks. As the task of reviewing exams varies, students' perceptions and learning comprehension to the visualization analytics should be considered while designing the learning analytics tool.

Empirical studies that investigate the impact of individual differences and the visualization formats on the effectiveness of orchestration technology platform are

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

scarce. This study focuses on students' learning comprehension and perceptions of visualization learning analytics tools. Instead of using only subjective survey data, we also collected secondary data from eye-tracking analysis to conduct a leaning analysis of students' initiatives toward the programming exam.

The remainder of this paper is organized as follows. Chapter 2 presents the literature review of orchestration technology and focuses on the dashboard and the visualization analytics. The impact of the individual differences, perception and eye-tracking analysis are also included. Chapter 3 describes the concept of learning comprehension, understanding of visualizations and proposes hypotheses in regards to utilizing the proposed visualization analytics tool. Chapter 4 describes the research system and experiment procedure. Chapter 5 describes the collected data and operationalization of variables. Chapter 6 presents the estimation methodology and results of our empirical analysis. Chapter 7 discusses the practical and research implications, limitations, and potential directions for future research. Chapter 8 concludes the paper.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

Chapter 2 LITERATURE REVIEW

Learning analytics is an emerging approach toward the educational context. Its main idea is bridging the computer science and sociology/psychology of learning to ensure that interventions and organizational systems serve the needs of all stakeholders (G. Siemens & Baker, 2012). With increasing numbers of education technologies for programming language learning, there is an abundance of approaches related to learning analytics in recent studies. The present study proposes personalized visualization analytics of an orchestration technology platform. Other than the effect of the visualizations alone, we were interested in how individual differences reflect on the visualization analytics, thus, we tried to investigate the influence of learning goal orientation, one of the common used indicators that measures the learning performance. The literature review presents these approaches in four categories: orchestration in learning analytics, dashboard analysis & visualizations, individual differences and individual perception of learning analytics visualizations.

2-1 Orchestration in Learning Analytics

Classroom orchestration defines how a teacher manages multi-layered activities. It discusses how and what research-based technologies have been adopted and should be done in the physical classrooms (Dillenbourg, 2013). For several decades, classroom orchestration has been used in different types of learning environments. For example, PeerWise offers an innovative approach that enhances standard teaching and learning practice by requiring students to participate in the construction and evaluation of multiple choice questions (MCQs) (Denny et al., 2008). QuizMap also combines open student modelling and social-based adaptive navigation support, an approach that is based on the "collective wisdom" of a student community to guide students to the right questions as successfully as classic knowledge-based guidance (Brusilovsky et al.,

DOI:10.6814/NCCU201901137

2011). Progressor as well was an innovative Web-based tool based on the concepts of social navigation and open student modeling that helped students to find the most relevant resources in a large collection of parameterized self-assessment questions on Java programming　Hsiao, Bakalov, Brusilovsky, & König-Ries, 2013).

More specifically, series of previous work was done to deal with practical paper-based programming language learning exams. Some examples can be found through QuizJET (Java Evaluation Toolkit), a system for authoring, delivery, and evaluation of parameterized questions for Java (I. Hsiao et al., 2008), and EduAnalysis which tested an intelligent semantic indexing for paper-based programming problems by integrating physical classroom learning assessments to online visual learning analytics (Hsiao & Lin, 2017). Programming Grading Assistant (PGA) also supports an automatic semantic partial credit assignment approach through scanning the paper-based exam results into a mobile app and providing an interface for teachers to calibrate recognition results (Hsiao, 2016). Web Programming Grading Assistant (WPGA) is a Web-based system to facilitate grading traditional paper-based exams in today's majority classes. It connects paper-based assessments to the online virtual assessment environment and ensures teachers the flexibility to continue using paper exams without having to learn new content authoring tools (Hsiao, Huang, & Murphy, 2017; Paredes, Huang, Murphy, & Hsiao, 2017).

Within the reported literatures, it is clear that the integration of a new learning technology in the real classroom is important when implementing classroom orchestration technology. There is a need to design the learning environments as well as the system from the very beginning. It is also very important that the system itself can interact with any other web service or simply consider the learning environment as web services. Therefore, to provide a better learning environment for programming education, we concentrated on the effect of building an online learning analytics tool

DOI:10.6814/NCCU201901137

for paper-based programming exam.

2-2 Dashboards and Visualizations in Learning Analytics

To support orchestration technology, most of classroom learning analytic systems come with analytic dashboards. Dashboards help teacher interpret students' performance in classroom orchestration activities. Students can also evaluate and adjust their learning strategies via reviewing personalized information such as their own learning progress on dashboards. LOCO-Analyst was a learning analytics tool aimed at providing educators with feedback on the performance of the learning activities taking place in a web-based learning environment. The study showed that educators value the mix of textual and graphical representations of different kinds of feedback provided by the tool (L. Ali, Hatala, Gašević, & Jovanović, 2012). Visualizations in a dashboard not only just summarize general performance indicators like scores, but also visualize interactions between students and the learning content (Hsiao et al. 2016; Lu and Hsiao 2016). It strives to help both students and teachers to find patterns, and contribute to awareness and self-monitoring. The Student Activity Meter emphasized awareness of time spent and corresponding resource use in a virtual classroom (Govaerts et al., 2012). The Temporal Learning Analytics Visualization (TLAV) tool also aims to visualize time spent on activities, but instead focuses on the time aggregation according to the correctness of a submitted answer during an assessment procedure (Papamitsiou & Economides, 2015).

According to these previous works, a dashboard with visualization benefits students through suitable learning material as well as significantly increases the quality of students' learning and motivation to work with non-mandatory learning content. Some studies also reported that the selection of appropriate display formats for users could be based on data types (categorical or quantitative), tasks types (comparison or

DOI:10.6814/NCCU201901137

identification of trends or totals), and user backgrounds (experts or casual users of graphs). For different user personality types, dashboards are more effective when they present flwxbility, i.e. allowing users to switch between alternative presentation formats (Helfman & Goldberg, 2007; Yigitbasioglu & Velcu, 2012). In our work, we implemented different formats of visualization formats in the hope of helping students with different personality types review their Java programming exam. Different task types were also considered as a key factor when we designed the learning analytics tool.

2-3 Visual Analytics in Learning Environment

Data visualizations can support organized and easy-to-understanding depiction of primary data. In orchestration technology, a well-designed dashboard is often adopted to visualize mass data and to assist users according to their needs, abilities and preferences. Graphs are a collection of nodes and links. Each node represents a single data point and each link represents how two nodes are connected. This way of representing data is well suited for scenarios involving relationships and correlations of data. Graph visualization is the visual representation of the nodes and links of a graph and can be presented as an image, picture or interactive multimedia with different sizes and colors. Further, graph visualization provides useful and efficient ways to understand the data. A better depiction of quantitative information can be derived from a well-designed graph visualization (Freedman, Shah, & Vekiri, 2005). As such, graph visualization is used extensively in different fields as various applications(Cui, Zhou, Qu, Wong, & Li, 2008; Gansner & North, 1999).

With the amount of analytic dashboards with graph visualization increasing, the graph comprehension has also been widely discussed in order to evaluate the efficiency of the visualization analytics interface (Ratwani & Boehm-davis, 2008). When asked to extract information from a graph, users generally have some stored knowledge that

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

they use to comprehend the graph, despite the fact that different graph types represent information differently (Ratwani & Trafton, 2008). Green & Fisher (2010) explored the impact of individual differences in three personality psychometric factors (Locus of Control, Extraversion, and Neuroticism) on interface interaction and learning performance behaviors in both an interactive visualization and a menu-driven web table. The results demonstrated that all three measures predicted completion times and the number of insights participants reported while completing the tasks in each interface. In the study of Ziemkiewicz et al.(2011), they observed the correlation between Locus of Control and the layout style and conducted a user study with four visualizations that gradually shift from list-based to spatial-based. The results demonstrate that participants with an internal locus of control perform more poorly with spatial-based visualizations, while those with an external locus of control perform well with such visualizations.

The effect of different graph visualization types has also been investigated in recent studies. Ali & Peebles (2013) reported three experiments investigating the ability of undergraduate college students to comprehend 2 × 2 “interaction” graphs from two-way factorial research designs. The results of the three experiments demonstrated the effects (both positive and negative) of Gestalt principles of perceptual organization on graph comprehension. Toker, Conati, Carenini, & Haraty (2013) investigated the impact of four user characteristics (perceptual speed, verbal working memory, visual working memory, and user expertise) on the effectiveness of two common graph visualization formats: bar graphs and radar graphs. The results showed that different characteristics may influence different factors that contribute to the user’s overall experience and effectiveness with a bar and radar graphs. Research has shown that people differ substantially in their ability to understand graphically presented information. Individuals with high graph literacy usually make more elaborate

DOI:10.6814/NCCU201901137

inferences when viewing graphical displays (Yashmina Okan, Garcia-Retamero, Cokely, & Maldonado, 2011). Novice graph viewers often neglect the relevance of important elements of graphs and interpret graphs incorrectly (Mazur & Hickam, 1993). In our study, the concept of graph comprehension could help to capture how well students reviewed and learned the Java programming topics from the visualizations. Moreover, when students reviewed the visualization analytics, understanding of why certain visualizations were generated and what the embedded information was mainly presented were also very important.

The effect of personality trait on user behavior in a learning environment has also been widely applied in past research. In the research field of information management, self-efficacy and perceived ease of use have been extensively used to determine a person's behavior. Self-efficacy is a measurement of how individuals believe their ability to achieve specific goals. Bandura (1993) stated that perceived self-efficacy operated as an important contributor to academic development. Prior research on technology acceptance behavior had examined the effects of self-efficacy and enjoyment on ease of use (Venkatesh, 2000). Locus of control, which referred to the degree of individual's perception about how well they can control over the underlying causes of events in their lives, was also widely used in studies about online learning and distance learning. Previous studies have confirmed this psychological construct plays an important role in learning achievement, satisfaction and persistence in online learning context because learners' capabilities to apply proper time management, continuous monitoring and self-evaluation is more important than a teacher's role in such learning tasks (Cascio, Botta, & Anzaldi, 2013; Joo et al., 2013).

Much of the existing studies in the visualization field investigate the impact of individual differences through the influence of user's cognitive abilities. Conati & Maclaren (2010) find that a user's perceptual speed predicts whether a star graph or

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

heatmap will be most effective for a given user. Similarly, Toker, et al.(2012) investigate the impact of four user cognitive abilities (perceptual speed, verbal working memory, visual working memory, and user expertise) on the effectiveness of bar graphs and radar graphs and find that certain user characteristics have a significant effect on task efficiency, user preference, and ease of use. However, in the work of Ziemkiewicz & Kosara (2009), the results show that factors of Big Five personality (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) can lessen the significant effect of visual metaphor on accuracy for simple visualization tasks. The research by Green et al. find effects on interface performance from three psychometric measures: locus of control, neuroticism, and extraversion (2010). Ziemkiewicz et al. (2011) further suggests that locus of control can influence an individual's use of a complex visualization system.

Taking a more learning perspective, with regard to measurement of academic or learning performance, learning engagement and goal orientation were two variables which had received a great deal of attention in organizational research. Learning engagement is a concept extended from work engagement (Xanthopoulou, Bakker, Demerouti, & Schaufeli, 2009) and could be seen as a behavioral factor that can influence the learning outcome. Chen (2017) extended the job demands-resources (JD-R) model which was proposed by Crawford et al. (2010) to evaluate the positive relationship between learning engagement and learning performance. The results found that learning engagement is positively associated with learning performance. Furthermore, the results also strengthened the solid finding that work engagement improves job performance. Goal orientation is the primary aim of individual toward developing or validating one's ability in an achievement settings (VandeWalle, 1997) and it has been applied in many studies in IS and HCI domains toward Web-based distance learning contexts (Chang, 2005; Klein, Noe, & Wang, 2006; Payne,

DOI:10.6814/NCCU201901137

Youngcourt, & Beaubien, 2007). Yi & Hwang (2003) extended the technology acceptance model (TAM) by adding learning goal orientation into the domain of technology acceptance in order to predict the use of Web-based information systems. The results demonstrated that learning goal orientation was positively related to the self-efficacy on a particular formation. Previous studies have showed that learning goal orientation exerts a significant effect on system use over behavioral intention, therefore, we adopt learning goal orientation as a factor of learning performance in this study.

2-4 Eye-tracking analysis

With the increase in computing power, as well as new data processing methods, data is accumulated more quickly than ever before. The variety of data makes it possible to conduct more comprehensive and user-adaptive analysis. To date, learning analytics has been focused on the investigation of the effects of operations performed by users. The analysis based on tracking data from the interactions between users with educational content has been considered to be a promising approach for advancing our understanding of the learning process in learning analytics (Gašević et al., 2015). Human eye movement is a sequence of activity in which the viewer focuses on specific information to support the mental or physical activities (Malcolm & Henderson, 2009), and is one of the most common and profitable types of tracking data which is widely used in the field of HCI. Li, Pelz, Shi, & Haake (2012) used an eye- tracker to model eye movement behavior in a medical examination context. Expert medical practitioners' examination processes were recorded in order to provide guidance to the novices. Besides behavior modeling, eye tracking analysis is also used to evaluate the effectiveness of users' characteristics on graph comprehension. Steichen, Carenini, & Conati (2013) explored users' eye gaze patterns while interacting with bar graphs and line graphs to predict the users' visualization tasks, as well as user cognitive abilities

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

including perceptual speed, visual working memory, and verbal working memory. The results showed that predictions based on eye gaze data are significantly better than a baseline classifier.

Drawing on the reviewed literatures, visualization is necessary and would be beneficial to the students. It is proved through previous research to be better at reflecting user's demand while developing an user-adaptive system and support more timely and effective feedback through monitoring information about learning (Gibson & de Freitas, 2016). However, the visualization formats and the task types, as well as individual differences of students, would influence a student's perception of the visualization. In order to understanding the relationships between these aspects, we constructed several hypotheses. The hypotheses will be discussed in the next section.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

Chapter 3 RESEARCH MODEL

From the literature review, we have seen an abundance of studies working on building visualization tools with class room orchestration technology. The graph comprehension enables us to get a grasp of how individual perform on different visualization formats, but it does not elaborate on the individuals differences in the various aspects of academic or learning performance. This study adapts the concept of graph comprehension in regard to learning analytics. Multiple factors have been considered for the process of using online visualization learning analytics tool to review paper-based programming exams. Learning goal orientation was adopted as a personality trait to measure the individual differences regarding to learning and academic performance. Two types of tasks were also identified: "search-fact" and "inference-generation". As for the graph visualization formats, we present three types of graph visualization formats in our study: line graphs, bar graphs and radar graphs.

In order to further explore the effects on individual perceptions via graph visualization, two types of constructs were proposed: "Learning Comprehension" and "Understanding of Visualization". We adapted the concept of graph comprehension (Shah & Freedman, 2011) to proposed the construct of learning comprehension, which refers to an evaluation of how students can integrate their prior knowledge and information embedded in the graph visualization to perform learning. Shah, Mayer, & Hegarty (1999) extended studies of bar versus line graphs work to characterize how Gestalt principles might affect comprehension of common graphs depicted in high school social studies textbooks. The results demonstrated that viewers' descriptions, if based on the visual pieces, would differ depending on format. Originating from this study, we focused on the user perceptions on different formats and proposed the constructs of understanding of visualizations. To measure how useful the visualization

DOI:10.6814/NCCU201901137

was to provide knowledge for students, we used the concept of perceived learning (Wu, Hiltz, Roxanne, & Bieber, 2010), which was adapted from the concept of perceived usefulness in TAM (Davis, 1989).

Table 3.1 summarizes the representative literature for the involved factors. These constructs are proposed to capture how students perceive the proposed orchestration technology platform with different formats, task types and personality. The influences of these factors on a user's perception and the respective research hypotheses are discussed in the next subsection.

Table 3.1 Proposed constructs.

Perceptions	Reference Studies
Learning Comprehension	(Shah & Freedman, 2011)
Understanding of Visualization	(Shah et al., 1999)
Perceived Learning	(Wu et al., 2010)

3-1 Learning Goal Orientation, Format and Task

The construct of goal orientation has recently received increasing attention due to the increase in web-based distance learning out of the IS and HCI domains. The goal orientation concept was first proposed to compare the orientations of students who approached college to acquire new skills and knowledge verses those who approached college to obtain high grades (Eison, 1979). Individuals with a high learning goal orientation tend to have a learning motivation to understand something new or to enhance their level of competence. Klein et al. (2006) examined how learning goal orientation relates to motivation to learn and course outcomes. The results suggest that learners high in learning goal orientation have significantly higher motivation to learn within a blended learning condition. In our study, we theorized that students with different personality types would have different behavior when reviewing their exam

DOI:10.6814/NCCU201901137

政 治 大 學 National Chengchi University

results. Students who prioritize the pursuit of new knowledge or skills for their own personal development tend to prefer exploring visualization analytics over students who are merely in pursuit of scores. This assumption matches the concept of learning goal orientation. Thus, to measure the effectiveness and learning outcome of students who review the paper-based programming exam in the online visualization learning analytics tool, we adopted learning goal orientation as a factor.

Several formats were investigated throughout previous studies of graph comprehension. The study conducted by Shah & Freedman (2011) reported that viewers are more likely to describe the interactions between variables on the X axis and Y axis when viewing line graphs, and they are more likely to describe main effects and the interactions between the variables in the legend and Y axis when viewing bar graphs. In the study of Toker et al. (2013), radar graphs were chosen to compare with bar graphs given that radar graphs are widely used for multivariate data. In this study, we focus on the visual characteristics of these common graphs and their influence on comprehension. We extended the work of Shah and Toker and adopted three types of common graphs for this study: line graphs, bar graphs and radar graphs. Although radar graphs are often considered inferior to bar graphs for common information seeking tasks (Few, 2005), there are indications that radar graphs may be just as effective as bar graphs for more complex tasks or integrated information (Toker et al., 2012).

This study utilizes two tasks that touching on two broad genres of learning: procedural and inferential. Both genres have broad records in the human behavioral literatures, and represent two very different types of knowledge: creation and use. Procedural learning is the learning composed of a sequence of iterative tasks (Sun, Merrill, & Peterson, 2001). Inference learning, on the other hand, is the learning that comes to a conclusion or a concept from available data. Several transmutations are used in the process of inference learning, including induction, deduction, generalization and

DOI:10.6814/NCCU201901137

comparison (Michalski, 1993). In the study of Green et al. (2010), both procedural tasks and inferential tasks are applied to explore the impact of individual differences in personality factors on different interfaces. Shah & Freedman's work (2011) uses both fact-retrieval and inference generation tasks to investigate the graph comprehension on interaction of top-down and bottom-up processes. Through these works, two types of learning tasks are identified in this study: "search-fact" and "inference-generation".

3-2 Learning Comprehension

In the theory of graph comprehension, the process of comprehension starts as the visual elements such as nodes, lines, and colors are identified and grouped together into clusters by viewers. Then these visual clusters influence a viewer's interpretations of the data. Specifically, the display is clustered based on the Gestalt principles of proximity, good continuity, and similarity(Pinker, 1990). Graph comprehension can evaluate the effect of individual differences on the information visualization (Okan, Garcia-Retamero, Cokely, & Maldonado, 2011; Okan, Garcia-Retamero, Galesic, & Cokely, 2012) and can also be used to depict how does an individual's prior knowledge (topic familiarity and graphical literacy skills) interact with format to influence s viewer's interpretations of graph visualization. Shah & Freedman (2011) investigated the effects of format (line vs. bar), viewers' familiarity of topics, and viewers' graphical literacy skills on the comprehension of multivariate data presented in graphs. The results showed that high-skilled graph viewers were able to make main effect inferences when viewing bar graphs that supported their ability to make the necessary mental computations, but not when viewing line graphs. Low-skilled graph viewers, however, could not make such inferences, even when viewing bar graphs. It may be useful to present different formats of graph visualization for users with different graphical literacy skills. The study also showed that skill may correspond to greater

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

differentiation between formats in open-ended tasks than fact-retrieval tasks.

As for personality trait factors, learning goal orientation was used intensively in education research as an indicator of learning performance or academic achievement. Debicki et al. (2016) developed a model to test potential mediating effect of learning goal orientation, prove performance orientation and avoid performance goal orientation between core self-evaluations and academic performance. The results showed that students with high core self-evaluations and learning goal orientation might utilize their perceived high capability to gain new experiences and increase their knowledge in search for personal development, thus creating a positive relationship with academic performance.

Therefore, in regards to our study, we propose learning comprehension as a factor because of its origination from the concept of graph comprehension. The degree of learning comprehension represents how well students review and learn the Java programming topics from the graph visualizations. If the students are motivated by increasing their competence through learning programming rather than just motivated by passing the course, they are willing to explore more topics which they are not familiar with when reviewing exam results in the web-based learning environment. The students who are enthusiastic for learning could benefit from the visualization analytics and uncover more knowledge, thus resulting in a better learning performance. Also, students may perform different degrees of comprehension due to the different presented graph visualization formats and the different types of reviewing tasks. Therefore we proposed the following hypotheses.

H1a. Learning goal orientation would have a positive influence on the degree of learning comprehension.

DOI:10.6814/NCCU201901137

H1b. The graph visualization formats would have different influences on the degree of the learning comprehension.

H1c. The type of reviewing tasks would have different influences on the degree of the learning comprehension.

3-3 Understanding of Visualization

The study of Toker et al., (2013) suggested that visualization types should be taken into account from the interaction effects found in user's cognitive abilities. More specifically, we looked into Gestalt principles (Koffka, 2013), which described how humans typically see objects by grouping similar elements, recognizing patterns and simplifying complex images. Gestalt principles are widely used in data visualization applications (Nesbitt & Friedrich, 2002; Patel et al., 2010) Shah et al. (1999) characterize how Gestalt principles might affect comprehension of common graphs. In the bar graph, the proximity principle predicts that for bar graphs a viewer would encode the grouped sets of bars representing levels of word familiarity (low, medium, and high). In the line graph, the principle of good continuity suggests that individuals would encode three visual clusters formed by the lines representing reading skill (low, medium, and high). The results showed that viewers' descriptions, if based on the visual clusters, would differ depending on format. We adapted this idea and proposed the constructs for the understanding of visualization, which refers to the level of how well students can interpret the visualization in the generation and meaning of the graph. According to Gestalt principles, the embedded information would be different corresponding to the format of graph. Line graphs are useful for displaying smaller changes in a trend over time according to the law of continuity. Bar graphs are easy to compare sets of data between different groups at a glance according to the law of

DOI:10.6814/NCCU201901137

proximity. Though there is no significant law on radar graphs, radar graphs may contain integrated information which is useful for complex data. We argued that such user perceptions on graphs will be different in line graphs, bar graphs and radar graphs. We also wanted to investigate within our study if personality traits induces different effects in terms of understanding of visualization. We assumed that a student with high learning goal orientation would show a greater willingness to depict graph visualization comprehensively, thus, accordingly they would have a better understanding of visualization in our system. Therefore, we proposed the following hypotheses.

H2a. Learning goal orientation would have a positive influence influences on the degree of understanding of visualization.

H2b. The graph visualization formats would have different influences on the degree of understanding of visualization.

3-4 Perceived Learning

According to Caspi & Blau (2008), perceived learning is "the set of beliefs and feelings one has regarding the learning that has occurred". Perceived learning was used extensively in educational researches, including game-based learning systems (Barzilai & Blau, 2014), asynchronous online courses (Swan, 2001), personal differences (Rovai & Baker, 2005), and visualization-based learning environments (Wang, Wu, Kinshuk, Chen, & Spector, 2013). In our context, perceived learning refers to the students' self-evaluation of their learning experience while using our system, indicating the degree of knowledge gained from the visualization. Therefore, to measure how useful the visualization was to provide knowledge for the students, we proposed the construct of perceived learning (Wu et al., 2010), which was adapted from the concept of perceived usefulness in TAM (Davis, 1989). We assumed that a student with high learning goal

DOI:10.6814/NCCU201901137

orientation would learn better in our system because of their aggressive motivation to learning programming in order to increase their competence, a trait for higher skills in perceived learning. In our study, we also wanted to know if the formats of graph visualization induces different effects in terms of perceived learning. Thus, we proposed the following hypotheses:

H3a. Learning goal orientation would have a positive influence on the degree of perceived learning.

H3b. The graph visualization formats would have different influences on the degree of perceived learning.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

Chapter 4 METHODOLOGY

4-1 Dataset

In this study we used grading data exported from Web Programming Grading Assistant (WPGA), a system from previous studies[1], to conduct an exploratory analysis on students' behavioral patterns toward topics in the programming exam. WPGA was a web-based system that facilitates grading and feedback delivery of paper-based programming assessments. We intended to provide students with a visualization tool which is independent from their existing learning environment. We named it Topic Combination Analysis Visualization (TCAV). TCAV aimed to support learners in exploring cognitive, topic-based, and behavioral insights of students' performance in exams.

WPGA was first introduced to the Java Programming class in the first semester of 107 academic year. There was a total of three exams during the semester. Before each exam, an instructor can set the grading rubrics related to the current learning material by inputting involved topics and corresponding scores. Figure 4.1 shows the grading interface of WPGA and topic rubrics of the questions. At the end of each exam, grading data was exported from WPGA. By analyzing the grading data, we can get an initial insight into how familiar an individual was with a particular topic, as well as discover potential correlation between topics through peer comparison. Extremely detailed grading data is available after applying exploratory data analysis on grading data from WPGA. This grading data can then be used to make various types of user-adaptive visualization, which is the core function of TCAV.

[1] https://cidsewpga.fulton.asu.edu/login/

DOI:10.6814/NCCU201901137

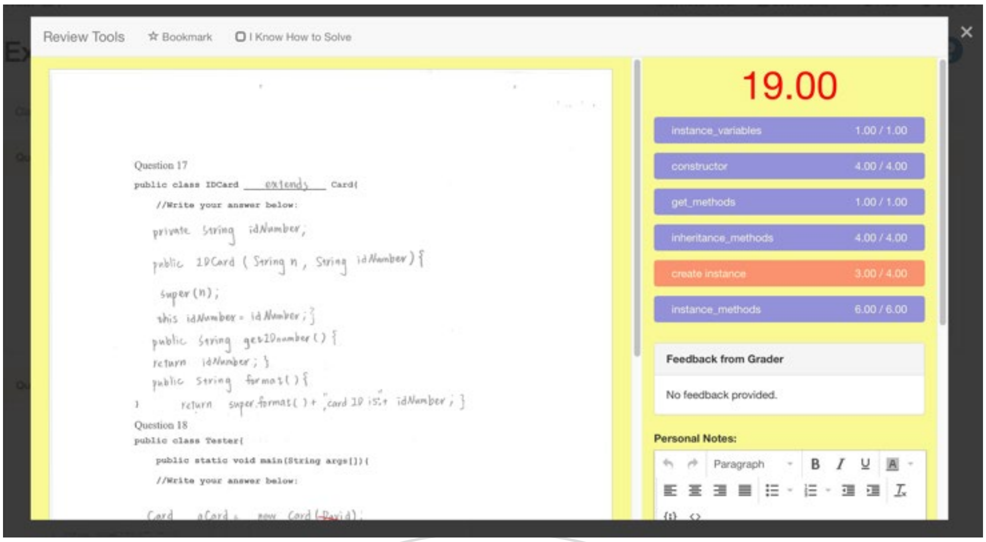


Figure 4.1 WPGA student interface detail view

We also conducted a lab study in the second semester of the 107 academic year. Participants were students who had enrolled in the Java Programming class in the Department of Management Information System of National Chengchi University. All participants joined the study voluntarily and acknowledged their right to decline their participation with a consent form. The data collection schedule is as figure 4.2.

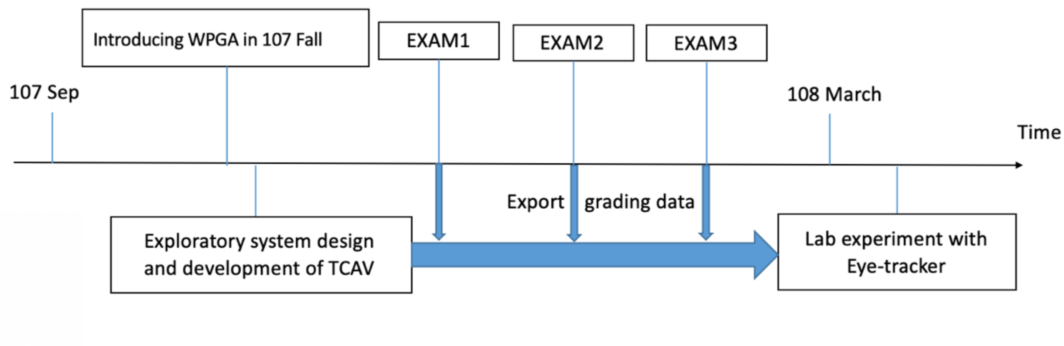


Figure 4.2 Experiment Schedule

4-2 System Development and Interface

We designed TCAV based on the grading data collected from participating students who attended the 107 Fall Java Programming Language I in the Department of Management Information System of National Chengchi University. It is a mandatory

DOI:10.6814/NCCU201901137

course for all first year students of MIS, however, it is also comprised by nearly one-third of students from other colleges.

The visualizations were created through the following steps (Figure 4.3). First, the instructors determined several related Java programming conceptual topics for each exam according to the lectures and text books used in the class. They then assigned several topics and grading rubrics for each question. The topics in the present study cover basic java programming concepts, including loops, control statements, objects, interfaces, etc. After students finish the exam, TA's (teaching assistants) graded the exam online through WPGA based on the rubrics. Personalized visualizations were generated from the detailed grading data. We calculated the scores students got for each topic of question. Then we correlated the visualization analytics containing topics involved in the question with percentage of correct answers.

We particularly focused on two kinds of rates of correct answers: individual answer rates versus the class and individual averages. Class average was the mean correct answer rates of this question calculated by the whole class for each topic involved in the current question (green area of Figure 4.4). Individual average, on the other hand, was the mean correct answer rate of other questions of a current exam calculated per individual for each topic involved the in current question (orange area of Figure 4.5). Students could discover the topics in which they performed poorly compared to the class average. These topics could be crucial fundamental concepts they needed to focus on. Students could also discover the topics in which they performed not poorly compared to the individual average. These topics could be the drilling concepts of the specific question. Finally, we applied the semantic results to visualizations with different types of graphs including bar graphs, line graphs and radar graphs (Figure 4.6-4.8).

DOI:10.6814/NCCU201901137

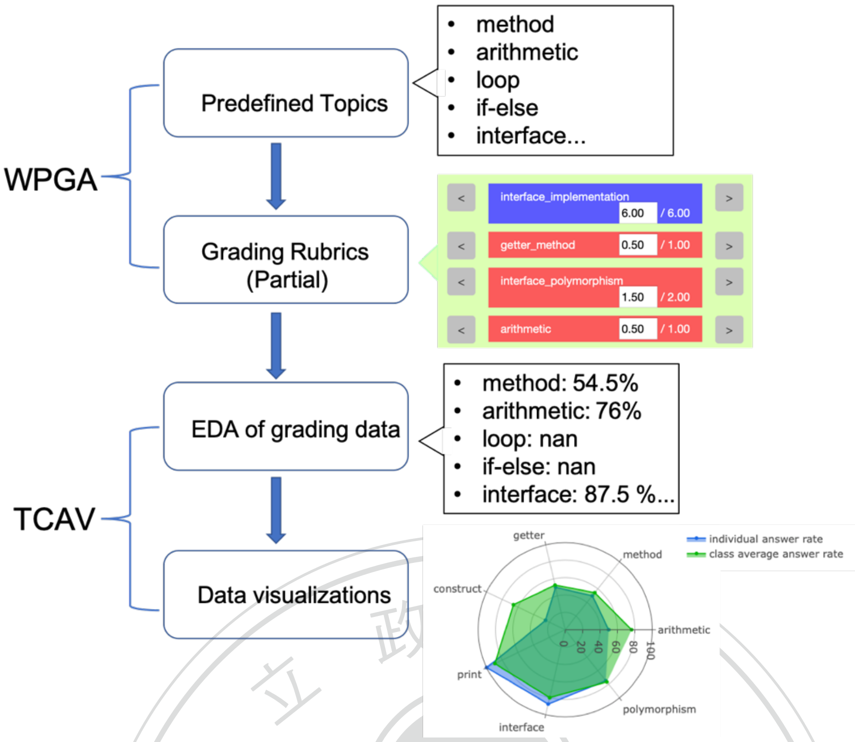


Figure 4.3 Data processing schema

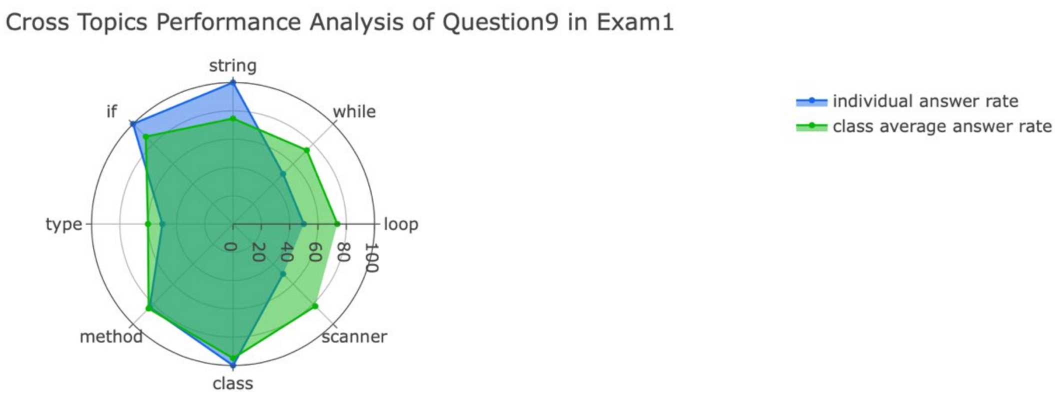


Figure 4.4 Visualization analytics: individual answer rate comparing to the class average

DOI:10.6814/NCCU201901137

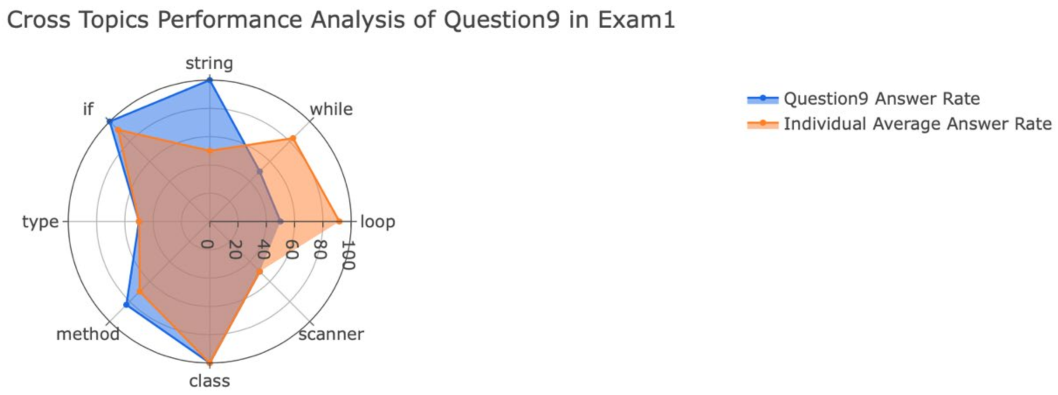


Figure 4.5 Visualization analytics: individual answer rate comparing to the average

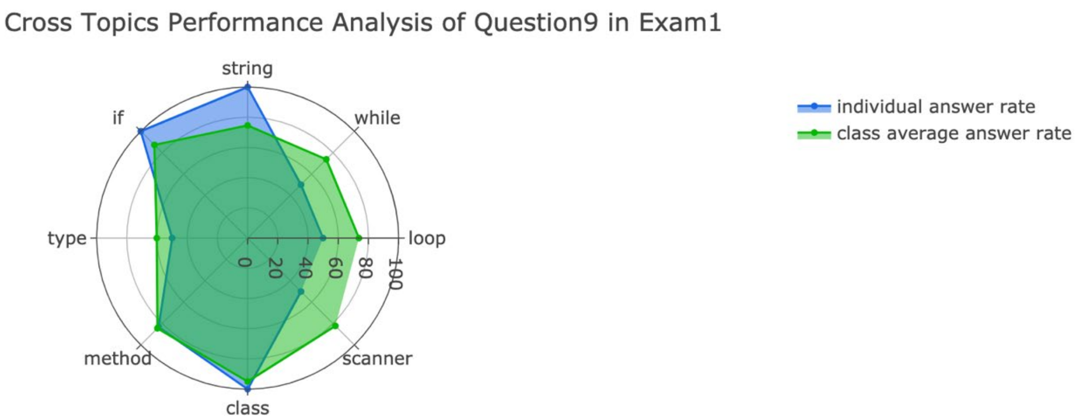


Figure 4.6 Graph visualization: radar graph

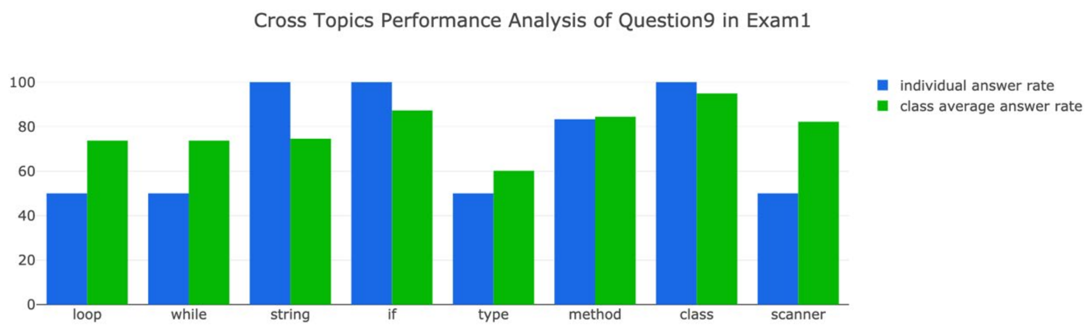


Figure 4.7 Graph visualization: bar graph

DOI:10.6814/NCCU201901137

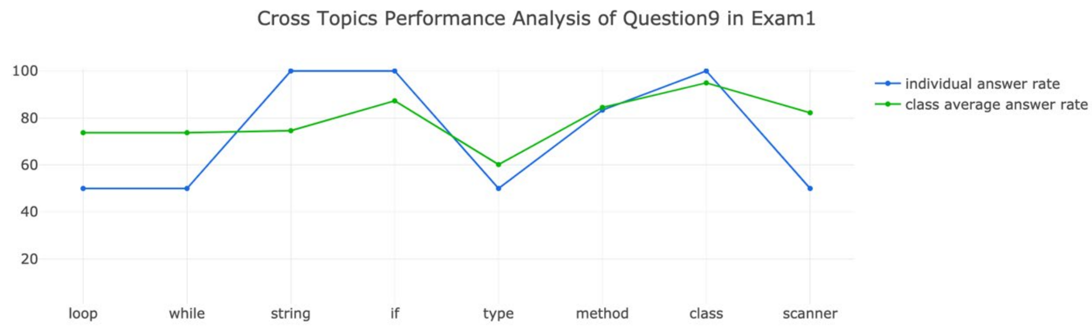


Figure 4.8 Graph Visualization: line graph

4-3 Search Fact Tasks and Inference Generation Tasks

To answer our research question 1 (RQ1), which explored how the visualization format and task type influence student's learning comprehension, we implemented two interfaces: a "search fact" interface with only a topic correct answer rate comparison between the individual and class averages, and an "inference generation" interface with a topic correct answer rate comparison between individual scores and both the class and individual average. Each interface included two main areas: the visualizations and question areas.

Figure 4.9 shows the interface for search fact questions. The upper half of the page was the graph visualizations. The graph visualizations compared the correct answer rates of topics involved in the current question between the individual and the class average. The lower half of the page was the question representing a search fact task involving Java conceptual concepts with a checkbox of possible answers. We established search fact questions through the instruction "retrieve the information provided in the visualizations". Thus, participants needed to go through the visualization analytics and check the topics corresponding to the different search fact questions. There are a total of three search fact questions in this study:

1. Please check all the topics which are involved in this question.

DOI:10.6814/NCCU201901137

2. Please check the topics in which the individual answer rate is higher than the class average answer rate.
3. Please check the topics in which the individual answer rate is lower than the class average answer rate, and is lower than 60%.

Figure 4.10 shows the interface for inference generation questions. The upper half of the page is the graph visualizations. The difference between the interface for search fact questions and inference generation questions is that we applied graph visualizations to compare both individual average and the class average. Participants could switch between these visualizations while answering the inference generation questions. The reason for providing two kinds of visualization analytics is that we created instructions for the inference generation questions as "compare and integrate multiple information sources to generate inference toward specific question." The process of inference generation may contain information retrieval, mapping, comparing, classifications etc. To answer the inference generation questions, the participants not only needed to compare the different visualizations, but also needed to rely on subjective prior knowledge of Java programming. Therefore, we designed a total of three inference generation questions for this study:

1. Please check the topics which an individual needs to review and strengthen for this question.
2. Given the chapter contents, please check the chapters which are needed to be reviewed for this question.
3. Following the previous question, please sort the chapters you checked according to their priority when you review the exam.

The corresponding interface of Q2 and Q3 were showed in figure 4.11-4.12

DOI:10.6814/NCCU201901137

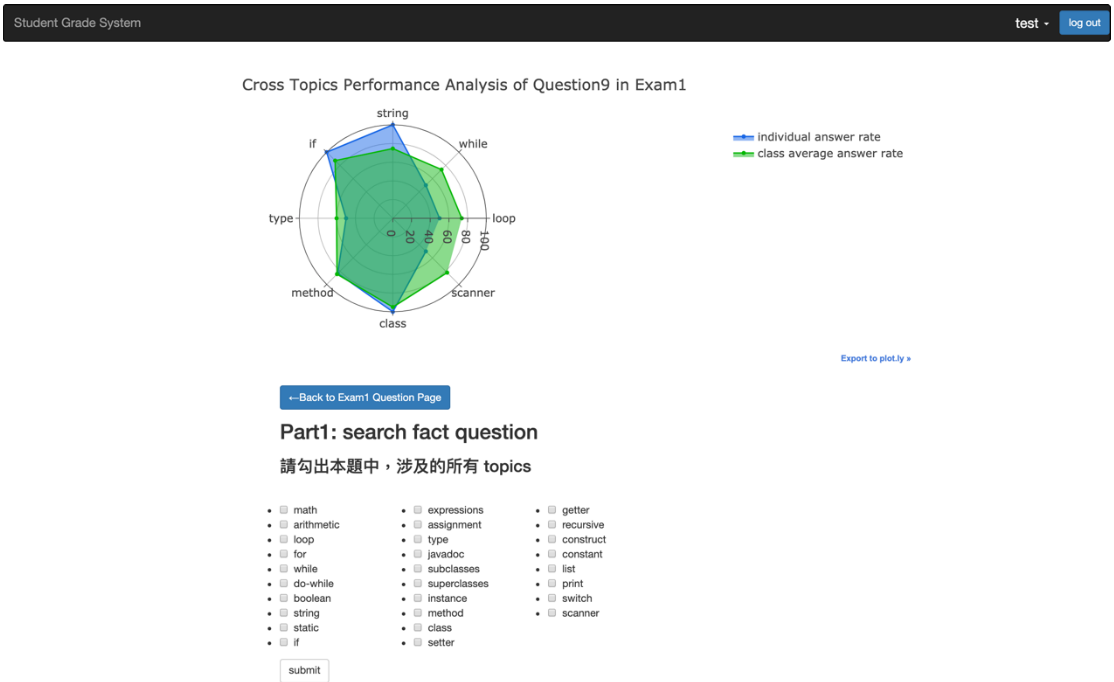


Figure 4.9 Visualization interface: search fact question

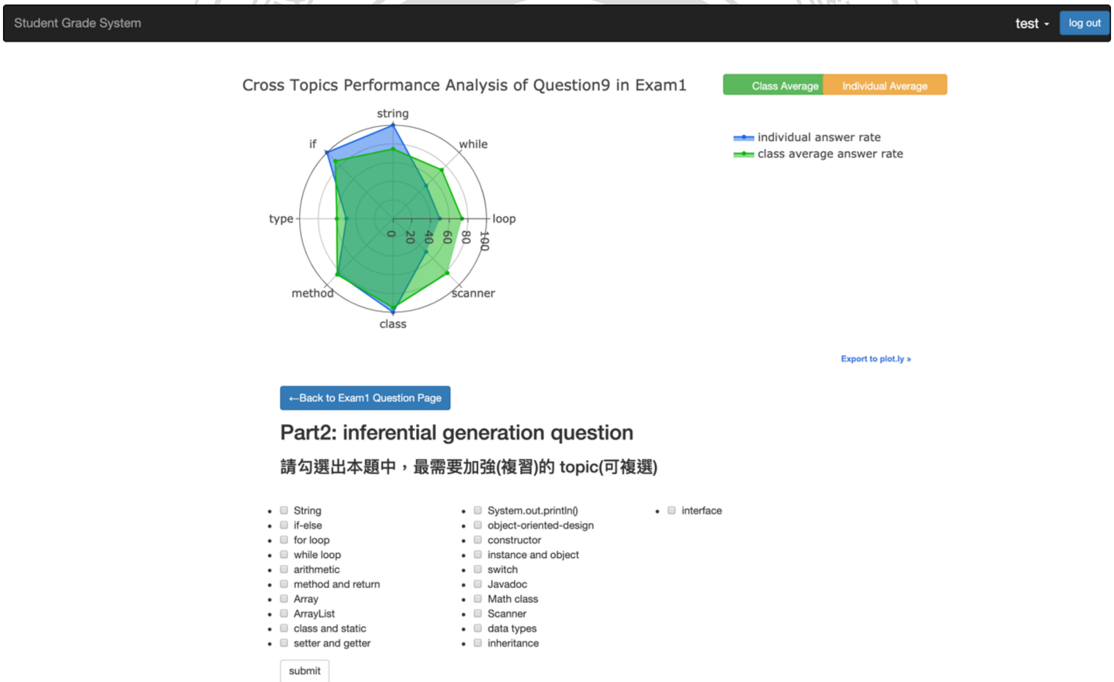


Figure 4.10 Visualization interface: inference generation question

DOI:10.6814/NCCU201901137

Figure 4.11 Question interface: inference generation question 2

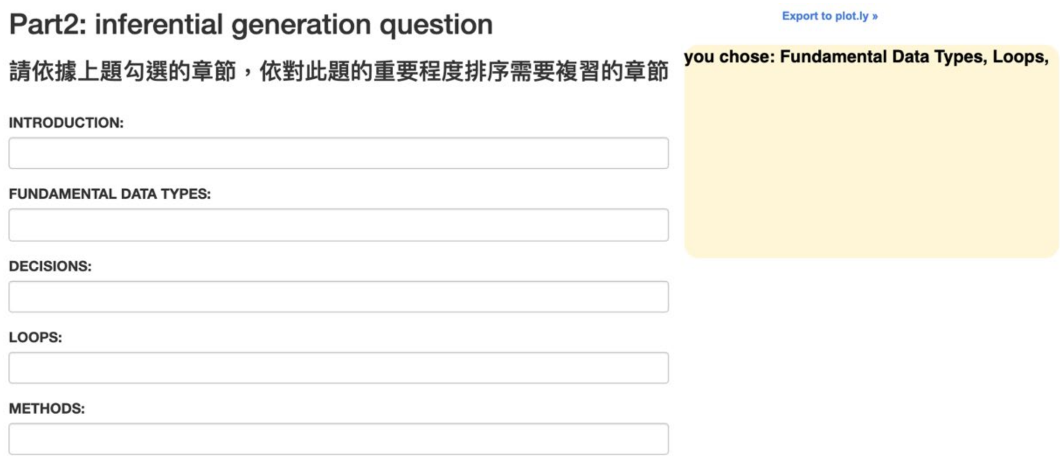


Figure 4.12 Question interface: inference generation question 3

4-4 Apparatus

A 24-inch computer screen with a resolution of 1280 X 1024 was used to display the system, and a Tobii X2-60 eye tracker with a sampling rate of 60Hz was implemented to collect the participants' eye-movements.

4-5 Subjects and Experiment Procedure

We conducted a user study at the Department of Management Information System of National Chengchi University. The experiment simulated the process of reviewing the paper-based Java programming exam on TCAV. Limited to the nature of our visualizations and research system, the participants could only be recruited from those who had taken at least one Java programming courses prior this research. For those who participated, a worth of NTD100 gift card and extra credit were given. In

National Chengchi University

DOI:10.6814/NCCU201901137

total 34 student participants (17 females and 17 males) were recruited. Their average age was 21.03 years old (min=19, max=24, SD=1.1 years).

The participants were informed that this experiment was based on a simulated exam and they had finished their exams and were prepared to review the exam results on TCAV. We adopted within-subjects design in this user study, hence the participants would need to review three exams, which were from 2018 Fall Java programming course, in the experiment. Each exam was paired with one kind of the format of graph visualizations described in the previous section, and the order of each participant was decided by Latin square order (Table 4.1). The participants would review starting from exam 1 and ending with exam 3 because the difficulty of exams was incremental, and we believed that this is the normal learning process for students. While the participants were viewing the exam questions, an eye-tracker was utilized to collect their eye movement.

In each exam, the task of participants was to answer the search fact questions and inference generation questions according to the given visualizations. For the consistency of the visualizations, we used the same grading data for each participant so that every participant would view the same visualization analytics based on the same grading score in each exam. For simplicity of the study, the participants were required to review only one question appointed beforehand in each exam. To control the difficulty consistency of the questions between the exams, we selected the questions which had a close class average correct rate in the range of 70% to 80% (Table 4.2).

Table 4.1 Sequence of exam number and format pairing.

Exam number / User ID	Exam1	Exam2	Exam3
test1	Radar	Bar	Line
test2	Radar	Line	Bar
test3	Bar	Radar	Line

DOI:10.6814/NCCU201901137

Table 4.2 Selected exam questions and corresponding answer rate

Exam number	Exam1	Exam2	Exam3
Selected question	Question 9	Question 10	Question 21
Class average *	77%	76.6%	76.2%
Number of topics			

* Class average correct answer rate in 2018 Fall Java programming course

The procedure of the experiment was as follows:

1. The participants were required to fill in a demographic questionnaire.
2. The participants were required to complete a cognitive test to test their perceptual speed (Ekstrom, French, & Harman, 1976).
3. The participants were required to browse the questions of the three exams in the paper. The purpose was to let the participants get familiar with the paper-based Java programming exams and recall the basic Java knowledge.
4. An eye-tracker was set up and calibration was conducted for each participant to make sure the eye movement was successfully collected.
5. The participants were then introduced to the system interfaces and their task. The experimenter would focus on introducing the functions of the system.
6. The participants would review questions from exam 1 to exam 3, each of which was paired with one of three designed visualization formats (see Appendix A). Meanwhile, an eye-tracker would collect their eye movement, this step would take roughly 30 to 45 minutes and the participants would be encouraged to take their time during this step instead of rushing to finish.
7. The experiment had three iterations, which are corresponded to exam 1 to exam 3. In each iteration, the participants were given a specific visualization format, three search fact questions and three inference generation questions. The participants would first review the exam question. Then they could check the

DOI:10.6814/NCCU201901137

exam solution (Figure 4.13). After the participants had reviewed the exam question and solution, they could move to the main interface to check visualization analytics with search fact and inference generation questions of the current question. There would be 3 search fact questions first, followed by 3 inference generation question. The procedure in each iteration was showed in Figure 4.14.

8. After finishing each iteration, the participants would be asked to fill in a questionnaire. The questionnaire was in regard to the participants' perceptions of the assigned visualizations, specifically its presentation format in specific (Figure 4.15).
9. After finishing reviewing all three exams, the participants would answer a post questionnaire and a short interview regarding their whole experience of the system and experiment.

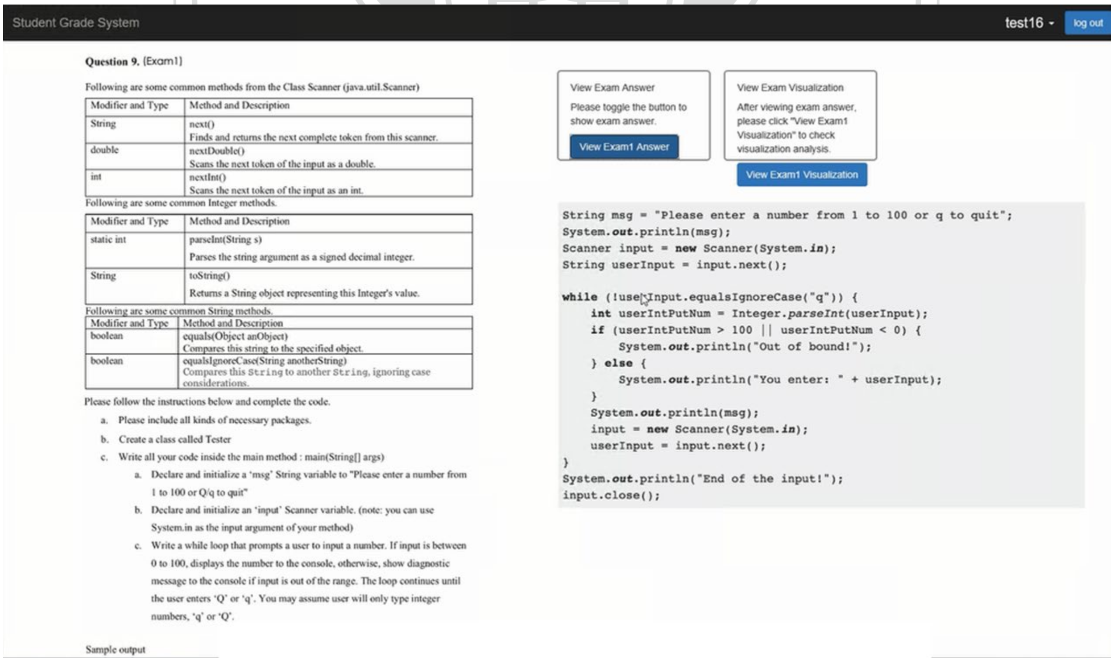


Figure 4.13 Original exam question interface: exam question and exam solution

DOI:10.6814/NCCU201901137

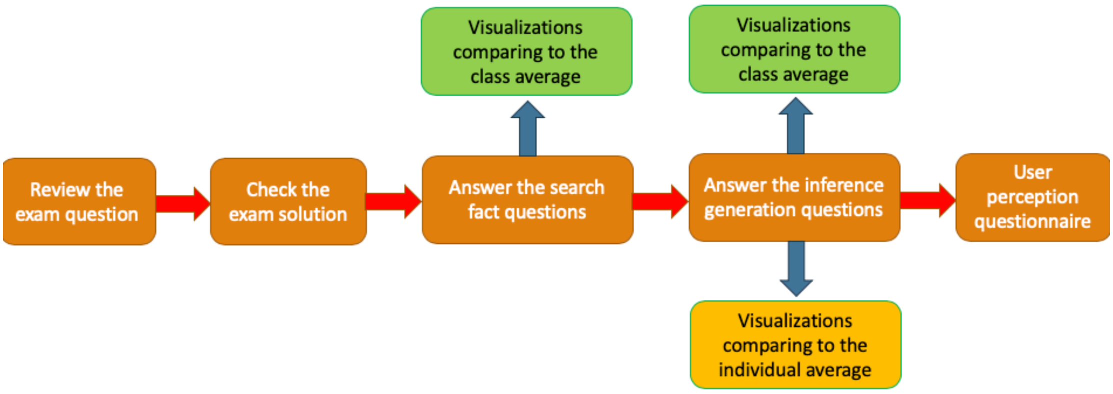


Figure 4.14 Experiment Procedure in one iteration.

Figure 4.15 User perception questionnaire.

4-6 Analysis Method

This study focused on the interaction between users and learning materials and more detailed learning analytics based on the paper-based exams to enhance user's awareness and self-monitoring. To explore the relationship between students' learning perceptions and the visualization formats, even though a questionnaire analysis would be able to tell us students' perceptions towards the formats, employing an eye tracker was able to give us additional information to gain a deeper understanding. As a result, we employed an eye-tracker to collect students' eye movement during the experiment.

DOI:10.6814/NCCU201901137

Totally, we adopted questionnaire, log and eye-tracking analysis to depict user's pattern and reflect their perceptions toward our system. We wanted to evaluate the influence of goal orientation, format and task on participants' perceptions despite the influence of gender and the familiarity of programming skills. Thus the regression analysis was applied to depict the dependency between these factors and user perceptions under the control of several conditions.

The construct of perceived learning was collected through the questionnaire after reviewing each exam as the subjective measurement of learning efficiency on the proposed system. The construct of learning comprehension was calculated based on the correct answer rate of the search fact questions and inference generation questions from system log as the objective measurement of learning performance on the proposed system. We were also interested in how visualization formats would influence the students' behavior. Thus, we measured the understanding of visualization after reviewing each exam through and calculated the correct answer rate from system log. An eye-tracking analysis was also conducted to obtain a deeper insight of the effects of visualizations in an exam review context. The result of the eye-tracking analysis could objectively evaluated the meaningfulness of the formats and the placement of specific interface elements, hence reveal some implications to improve the design of the interface. The collected data and corresponding measurements were depicted in the chapter 5.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

Chapter 5 DATA AND MEASUREMENTS

An experiment of within subject analysis was conducted to test the hypotheses. A lab study with eye-tracking analysis was conducted to collect data, including TCAV system logs, questionnaire, and eye-tracking system logs.

We conducted a lab study to collect data from students who have taken at least one Java programming course. In total, data of 34 participants were analyzed, 17 of them (50%) were male, the average age was 21.03 (std = 1.1), 30 of them were from IS background, others included finance and sociology. 26 of them had experience of programming for more than 1 year. Subjective learning perception data were also collected through questionnaire. And both system log and eye-tracking data were collected for objective user behavior analysis.

Our research question could be answered by analyzing the data from user system log and questionnaire. But we also wanted to explore more on user's behavior on the visualization analytics. As a result, objective eye-tracking data were applied to support our hypothesizes.

5-1 User Behavior and Perception Data

To measure the learning goal orientation, we used the 8-items measurement proposed by Zajac, Button, & Mathieu (1996). A sample item reads, "The opportunity to do challenging work is important to me." Response were made on a 7-point scale (1=Strongly Disagree, 7 = Strongly Agree). Higher scores indicate higher learning goal orientation. One factor analysis was conducted to test the internal validity among the items. Item quality and factor correlation were satisfied and no item was needed to be dropped. The Cronbach's alpha is 0.9, which support the internal consistency reliability of the factor. Appendix B provides the complete list of these items. The format of the graph visualization and the type of the learning task were recorded as participants

DOI:10.6814/NCCU201901137

finishing each question in the lab study. To measure the learning comprehension, we calculated the average answer rate of search fact questions or inference generation questions for each user. For search fact questions, it was the rate that the participants checked the correct topics which met the query of the question. For inference generation questions, the same calculation was applied on Q1. For Q2 and Q3 of inference generation questions, we took them as a set questions and used the normalized discounted cumulative gain (NDCG) to measure the ranking quality. To measure the understanding of visualization, we asked the participants to answer what is the information the prior visualization mainly imply after finishing each iteration. The question list multiple choice of the possible information that the proposed visualization could imply. To calculate the correct answer rate, we set default answer for each visualization format (table 5.3). Finally the measurement of perceived learning is the 3-items measurement with 5-point scale (1 = Strongly Disagree, 5 = Strongly Agree). The participants need to reply their perceived learning of the prior visualization analytics after finishing each iteration in the lab experiment. One factor analysis was conducted to test the internal validity among the items. Item quality and factor correlation were satisfied and no item was needed to be dropped. The Cronbach's alpha is 0.92, which support the internal consistency reliability of the factor. The description of variables are summarized in table 5.1 and 5.2.

DOI:10.6814/NCCU201901137

Table 5.1 Description of user perceptions variables

Research variables	Measurement items and description
Learning Comprehension	Average correct answer rate that the participants checked the correct topics for search fact questions 1-3 and inference generation question 1. NDCG score for inference generation question 2-3.
Understanding of Visualization	Average correct answer rate that the participants reply of the following multiple selection question. What is the information mainly imply in the prior visualization analytics (you may choose more than one response)? Chapter 1 Show the trend of the answer rate. Chapter 2 Compare the difference between individual and the class average Chapter 3 Emphasize the extreme value belongs to specific topic of the individual or class average Chapter 4 Show the Java topics involved in the question Chapter 5 Show the correlations of difficulty between the Java topics.
Perceived Learning	The average score of the following 3 measurement items with 5-point scale (strongly disagree to strongly agree): 1. I learned knowledge of Java programming from the visualization analytics. 2. The visualization analytics helped me learn Java programming. 3. The visualization analytics improved my familiarity with Java programming.

DOI:10.6814/NCCU201901137

National Chengchi University

Table 5.2 Description of learning goal orientation, format and task variables

Research variables	Measurement items and description
Learning Goal Orientation	The average score of the following 8 items measurement with 7-point scale(1=Strongly Disagree, 7 = Strongly Agree). (Zajac et al., 1996) A sample item reads, "The opportunity to do challenging work is important to me."
Format	Recorded as participants finishing each question in the lab study The value is either Radar, Bar or Line
Task	Recorded as participants finishing each question in the lab study The value is either SearchFact or InferenceGeneration

Table 5.3 Default answer of understanding of visualization for each format

Format type	Information mainly imply
Line graphs	1. Show the trend of the answer rate. 3. Emphasize the extreme value belongs to specific topic of the individual or class average 4. Show the Java topics involved in the question
Bar graphs	2. Compare the difference between individual and the class average 4. Show the Java topics involved in the question
Radar graphs	2. Compare the difference between individual and the class average 3. Emphasize the extreme value belongs to specific topic of the individual or class average 4. Show the Java topics involved in the question 5. Show the correlations of difficulty between the Java topics.

DOI:10.6814/NCCU201901137

To answer our research question, regression analysis was adopted to explore the dependency between learning goal orientation, format, task and user perceptions. We integrated the collected data of these variables. In total, we collected data from 34 participants. Each participant would view three exam questions with each kind of visualization formats. Each exam question had both search fact task and inference generation task. Hence, the sample contained 204 observations.

Dependent variables represent the user perceptions. The variable learning comprehension is the percentage of the correct answer rate respectively in search fact questions and inference generation questions. The variable understanding of visualization is the correct answer rate percentage of the multiple selection question to test how participants know the meaning of the visualization format. Finally the variable perceived learning is the average score of the measurement items with 5-point scale (strongly disagree to strongly agree).

Independent variables reflect the factors which would have influences on the user perceptions. The variable goal orientation represents the degree of individual is motivated by the opportunity to develop and master new skills rather than desire to demonstrate their abilities or to avoid failure. We transform the original average score of 8-items measurement with 7-point scale to categorical variable which have three values: low, middle, high. The three values represent the score which is under the 25%, between 25% and 75%, above 75% respectively. The variable format is a categorical variable, which takes on values of the format for the graph visualization (line, bar and radar). The variable format is also a categorical variable, which represents the types of the learning task (search-fact and inference-generation).

The control variables include differences between programming-experienced students and beginners. Experience is a dummy variable reflecting whether the students is experienced in programming. We define the students who have more than one year

DOI:10.6814/NCCU201901137

experience in programming as programming-experienced students. Even though we select the question which have the close class average correct answer rate for each exam, we still control for the effects of exam number, which may has effects on the perceptions of the students. The gender is also controlled. Table 5.4 lists the description and summary statistics of variables. Table 5.5 reports the pair-wise Pearson correlation coefficients of our dependent and independent variables. Table 5.6 shows the frequency table for the categorical control variables.

Table 5.4 Descriptive statistics of variables (N=204)

Variable	Mean	Std. Dev.	Min	Max
Dependent variables				
Learning comprehension	0.846	0.217	0	1.000
Understanding of Visualization	0.611	0.249	0	1.000
Perceived learning	3.69	0.799	2.00	5.00
Independent variables				
Goal orientation	Categorical variable: high(N=54), middle(N=96), low(N=54)			
format	Categorical variable: radar(N=68), bar(N=68), line(N=68)			
task	Categorical variable: searchfact(N=102), inference(N=102)			

Table 5.5 Correlation of variables

	(1)	(2)	(3)
Learning comprehension	1		
Understanding of Visualization	-0.0023	1	
Perceived learning	0.0278	0.2373	1

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

Table 5.6 Frequency table of categorical variables

Variable	Count
Programming-experienced	204
Experienced	156
Beginner	48
Exam number	204
1	68
2	68
3	68
Gender	204
Male	102
Female	102

5-2 Eye-tracking Data

When the participants were reviewed questions on TCAV in lab experiment, meanwhile, an eye-tracker were adopted to collect data. A Tobii X2-60 eye tracker with a sampling rate of 60Hz was used to collect eye movement data throughout the experiment. We focused on the following aspects to analyze the eye tracking data:

- *Area of Interest (AOI)*: The predefined region in the specific interface where user looked at while performing task. It helps to extract metrics from the selected region.
- *Fixation*: The moment that the eyes are relatively motionless and fix on a portion of the interface.
- *Saccade*: The eye movements occur between fixations.
- *Transition*: The saccade between two AOIs.

In the present study, we defined 4 major AOIs: Visualization, Question, Legend and Title (Figure 5.1) on the visualization interface of TCAV. Two metrics from fixation data, fixation duration and fixation count, were used to show how much time and attention the participants spent on the AOIs. Fixations could reflect user's attention

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

allocation during the experiment. However, longer fixation duration and more number of fixation count might indicate difficulty in extracting information and less efficient search (Goldberg & Kotval, 1999; Just & Carpenter, 1976). The transition data were also calculated to depict user patterns between two AOIs.

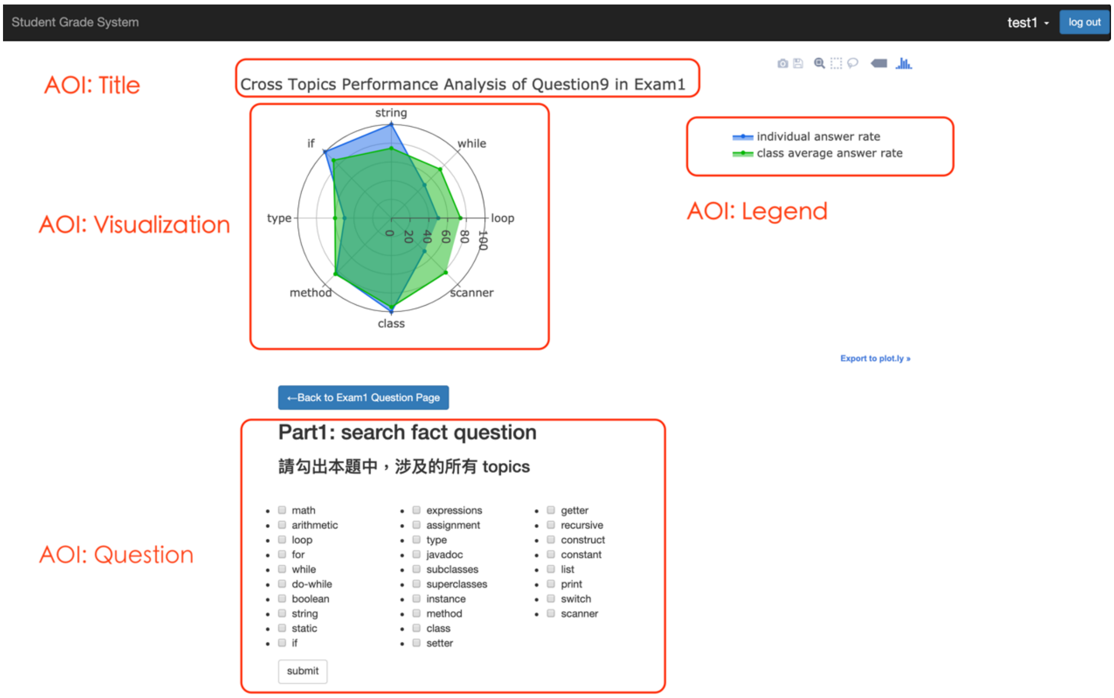


Figure 5.1 Predefined AOIs: Visualization, Question, Legend and Title

Regression analysis was also used to explore the dependency between the eye-tracking metrics and learning goal orientation, format, task. For each AOI, we labeled its visualization format (radar, bar or line) and the types of the learning task (search fact or inference generation). Totally 28 out of the 34 participants were viable for eye-tracking analysis. Four of them were missing data due to the malfunction of the eye-tracker. Two of them were incompletely recorded due to technical problems of eye-tracker during recording of the eye movements so we decided to drop them. Each participants would view three interfaces with different formats, and each interface contained two tasks. In each task, we defined two AOIs. To investigate the effect of visualization format and the learning type of task on different AOI, we integrated the collected data of these variables. We only selected the data which the total fixation time

DOI:10.6814/NCCU201901137

was not equal to 0. In total, the sample contained 555 observations. Table 5.7 and 5.8 show the descriptive statistic of total fixation duration time and fixation count.

Table 5.7 The descriptive statistic of Total Fixation Duration in seconds.

	Search Fact (N=285)			Inference Generation (N=270)		
	Mean	**Std**	**N**	**Mean**	**Std**	**N**
Radar						
Question	30.43	27.28	28	29.18	23.0	28
Visualization	26.58	18.65	28	25.52	24.15	28
Legend	2.06	1.88	24	2.50	2.39	22
Title	0.70	0.84	18	0.82	1.78	18
Bar						
Question	26.97	22.95	28	26.19	23.0	28
Visualization	16.47	14.62	28	20.31	21.82	28
Legend	1.31	1.40	21	2.852	2.46	20
Title	0.61	0.73	14	0.73	0.89	11
Line						
Question	27.91	21.68	28	24.87	15.65	28
Visualization	16.62	56.86	28	21.17	22.81	28
Legend	2.20	1.58	22	4.57	7.26	20
Title	0.44	0.65	18	0.72	0.89	11

DOI:10.6814/NCCU201901137

Table 5.8 The descriptive statistic of Fixation Count.

	Search Fact (N=285)			Inference Generation (N=270)		
	Mean	**Std**	**N**	**Mean**	**Std**	**N**
Radar						
Question	154.79	97.46	28	175.71	121.43	28
Visualization	129.89	72.85	28	129.43	106.33	28
Legend	11.29	8.33	24	13.91	13.31	22
Title	4.78	5.25	18	5.22	10.0	18
Bar						
Question	147.32	97.59	28	158.5	117.13	28
Visualization	94.71	70.7	28	115.89	115.14	28
Legend	7.76	7.99	21	16.7	14.0	20
Title	3.57	3.23	14	5.36	5.45	11
Line						
Question	151.54	90.37	28	154.04	82.18	28
Visualization	104.18	56.86	28	118.46	117.05	28
Legend	12.14	8.29	22	24.15	36.69	20
Title	2.67	3.09	18	4.82	5.33	11

The eye tracker software captured a complete sequence of fixations for both search fact task and inference generation task performed by each participant with different formats. The transition is participant's eye movement from one AOI to another. In order to keep the readability of gaze transitions, we focused on Visualization (V) and Question (Q) AOIs. And classified the remaining AOIs, Legend and Title, as Others

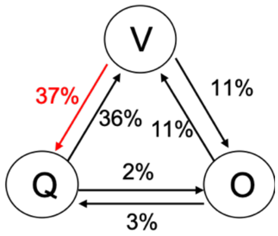
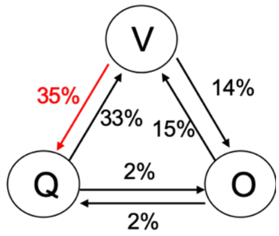
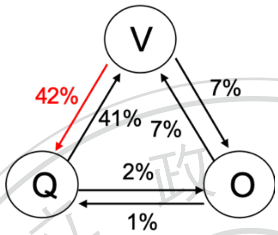
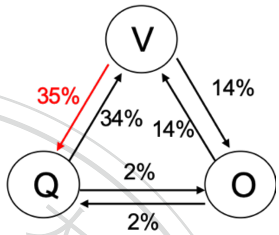
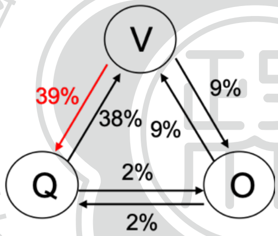
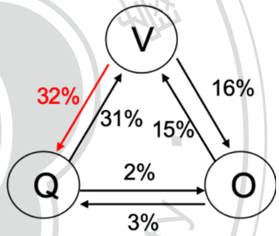
DOI:10.6814/NCCU201901137

(O). Table 5.9 shows two-state transition diagrams—one for the search fact task and another for the inference generation task of the three visualization formats. The state diagrams display the relative transition frequencies (in percentages), i.e., the frequency of each sum of the frequencies of VQ, QV, VO, OV, QO, OQ. The dominated eye movement transition for both search fact and inference generation task is VQ and QV, which is fairly reasonable because Visualization and Question AOIs are two major area of the visualization interface of TCAV. It indicates that users attempt to reference to the visualization when they answer the questions. To better understand difference in patterns of user behavior, we also conduct a regression analysis to explore the dependency between the goal orientation, formats, task and the transition rate between Question AOI and Visualization AOI. The results were depicted in chapter 6.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

Table 5.9 The state diagrams of gaze transitions among AOIs.

	Search Fact	Inference Generation
Radar		
Bar		
Line		

National Chengchi University

DOI:10.6814/NCCU201901137

Chapter 6 MODEL SPECIFICATIONS

6-1 Log-based User Behavior and Perception Data Analysis

We specify the following model to test our research hypotheses, and the linear regression model was used for estimation.

- Learning comprehension (H1)

$$LearningComprehension_i = \beta_0 + \sum_{j=1}^{2} \beta_j GoalOrientation_i + \sum_{j=3}^{4} \beta_j Format_i + \beta_5 Task_i + \beta_6 ProgrammingExperienced_i + \sum_{j=7}^{8} \beta_j ExamNumber_i + \beta_9 Gender_i + \varepsilon_i \quad (1)$$

- Understanding of visualization (H2)

$$UnderstandingOfVisualization_i = \beta_0 + \sum_{j=1}^{2} \beta_j GoalOrientation_i + \sum_{j=3}^{4} \beta_j Format_i + \beta_5 ProgrammingExperienced_i + \sum_{j=6}^{7} \beta_j ExamNumber_i + \beta_8 Gender_i + \varepsilon_i \quad (2)$$

- Perceived learning (H3)

$$PerceivedLearning_i = \beta_0 + \sum_{j=1}^{2} \beta_j GoalOrientation_i + \sum_{j=3}^{4} \beta_j Format_i + \beta_5 ProgrammingExperienced_i + \sum_{j=6}^{7} \beta_j ExamNumber_i + \beta_8 Gender_i + \varepsilon_i \quad (3)$$

The equation (1) test how goal orientation impact the learning comprehension. Also, it test the influences of the format of graph visualization and the type of learning tasks on the learning comprehension. The index i stands for each observation in our integrated collected data. The effects of dummy variable ***GoalOrientation*** on learning

DOI:10.6814/NCCU201901137

comprehension are denoted by the parameters β_1 and β_2, which represents the difference level of the high degree observation and the low degree observation of the goal orientation, and the difference level of the high degree observation and the middle degree of the goal orientation. β_3 and β_4 are coefficients for the effect of dummy variable ***Format*** on learning comprehension, which represents the difference level of the line format and the bar format, and the difference level of the radar format and the bar format. β_5 reflects the effect of dummy variable ***Task*** on learning comprehension, which represents the difference level between search fact task and inference generation task on learning comprehension. β_6, β_7, β_8, β_9 are coefficients for the effect of control variable. β_6 reflects the effect of dummy variable ***ProgrammingExperienced***, which represents the difference level between the participants experienced in programming and the beginner. β_7 and β_8 are estimate of dummy variable ***ExamNumber***, which represents the difference level of the exam 2, exam 3 and the exam 1 as the baseline. β_9 is estimate of dummy variable ***Gender***, which reflects the effect of difference level between the male and female. Finally the ε_i is the level-1 error term.

In sum, the coefficient estimates of *GoalOrientation_Low* is statistically significant at the level of 0.05 and *Task_SearchFact* is statistically significant at the level of 0.001. The results show that there is a significant difference between the high degree observation and the low degree observation of the goal orientation, which strongly support for H1c and support for H1a. However, the estimates of *Format_Line* and *Format_Radar* provide no statistical support for H1b. The estimated results for learning comprehension are summarized in **Table 6.1.** Also we have observed significant difference of control variables exam number 1 and exam number 2, which show that there are difference between the three exams. We discuss this in chapter 7.

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

Table 6.1 Estimated results for learning comprehension.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	0.75939	0.054018	14.031	< 2e-16 ***
Goal orientation (Low)	-0.088151	0.036983	-2.384	0.0181 **
Goal orientation (Middle)	-0.031032	0.034210	-0.907	0.3655
Format (Line)	-0.002351	0.032971	-0.071	0.9432
Format (Radar)	-0.037530	0.032971	-1.138	0.2564
Task (SearchFact)	0.128123	0.026909	4.761	3.76e-06 ***
Programming-experienced (Experienced)	-0.053657	0.033959	-1.580	0.1157
Exam number (2)	0.163285	0.032971	4.952	1.59e-06 ***
Exam number (3)	0.147371	0.032971	4.470	1.33e-05 ***
Gender (Female)	0.025574	0.027578	0.927	0.3549

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

The equation (2) test how goal orientation impact the understanding of visualization. Also, it test the influences of the format of graph visualization on the understanding of visualization. The index i stands for each observation in our integrated collected data. The effects of dummy variable ***GoalOrientation*** on learning comprehension are denoted by the parameters β_1 and β_2 , which represents the difference level of the high degree observation and the low degree observation of the goal orientation, and the difference level of the high degree observation and the middle degree of the goal orientation. β_3 and β_4 are coefficients for the effect of dummy variable ***Format*** on learning comprehension, which represents the difference level of the line format and the bar format, and the difference level of the radar format and the bar format. β_5, β_6, β_7, β_8 are coefficients for the effect of control variable. β_5

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

reflects the effect of dummy variable ***ProgrammingExperienced***, which represents the difference level between the participants experienced in programming and the beginner. β_6 *and* β_7 are estimate of dummy variable ***ExamNumber***, which represents the difference level of the exam 2, exam 3 and the exam 1 as the baseline. β_8 is estimate of dummy variable ***Gender***, which reflects the effect of difference level between the male and female. Finally the ε_i is the level-1 error term.

In sum, the coefficient estimates of *Format_Line* and *Format_Radar* are statistically significant at the level of 0.01. The results show that there is a difference between three formats on the understanding of visualization, which fairly strong support for H2b. However, the estimates of *GoalOrientation_Low* and *GoalOrientation_Middle* provide no statistical support for H2a. The estimated results for understanding of visualization are summarized in **Table 6.2**.

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201901137

Table 6.2 Estimated results for understanding of visualization.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	0.71088	0.06361	11.176	< 2e-16 ***
Goal orientation (Low)	-0.05556	0.04497	-1.235	0.21815
Goal orientation (Middle)	-0.04556	0.04160	-1.095	0.27473
Format (Line)	-0.12367	0.04009	-3.085	0.00233 ***
Format (Radar)	-0.11703	0.04009	-2.919	0.00392 ***
Programming-experienced (Experienced)	0.12021	0.04129	2.911	0.00402 ***
Exam number (2)	-0.03795	0.04009	-0.947	0.34499
Exam number (3)	-0.05902	0.04009	-1.472	0.14259
Gender (Female)	-0.08599	0.03353	-2.564	0.01109 *

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

The equation (3) test how goal orientation impact the perceived learning. Also, it test the influences of the format of graph visualization on the perceived learning. The index i stands for each observation in our integrated collected data. The effects of dummy variable ***GoalOrientation*** on learning comprehension are denoted by the parameters β_1 *and* β_2, which represents the difference level of the high degree observation and the low degree observation of the goal orientation, and the difference level of the high degree observation and the middle degree of the goal orientation. β_3 *and* β_4 are coefficients for the effect of dummy variable ***Format*** on learning comprehension, which represents the difference level of the line format and the bar format, and the difference level of the radar format and the bar format. β_5, β_6, β_7, β_8 are coefficients for the effect of control variable. β_5 reflects the effect of dummy variable ***ProgrammingExperienced***, which represents the difference level between the

DOI:10.6814/NCCU201901137

participants experienced in programming and the beginner. β_6 *and* β_7 are estimate of dummy variable ***ExamNumber***, which represents the difference level of the exam 2, exam 3 and the exam 1 as the baseline. β_8 is estimate of dummy variable ***Gender***, which reflects the effect of difference level between the male and female. Finally the ε_i is the level-1 error term.

In sum, the coefficient estimates of *GoalOrientation_Low* is statistically significant at the level of 0.001. The results show strong support for H3a. However, the estimates of *Format_Line* and *Format_Radar* provide no statistical support for H3b. The estimated results for perceived learning are summarized in **Table 6.3.**

Table 6.3 Estimated results for perceived learning.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	3.96456	0.20965	18.910	< 2e-16 ***
Goal orientation (Low)	-0.56790	0.14821	-3.832	0.000172 ***
Goal orientation (Middle)	-0.14607	0.13709	-1.065	0.287977
Format (Line)	0.01183	0.13213	0.090	0.928735
Format (Radar)	-0.04473	0.13213	-0.339	0.735313
Programming-experienced (Experienced)	-0.16053	0.13609	-1.180	0.239610
Exam number (2)	0.06898	0.13213	0.522	0.602247
Exam number (3)	-0.07677	0.13213	-0.581	0.561914
Gender (Female)	0.16727	0.11052	1.514	0.131764

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

DOI:10.6814/NCCU201901137

6-2 Eye-tracking Data - Fixation Analysis

The log-based and questionnaire could provide both subjective and objective analysis of users' learning performance on our system. However, the users' pattern of depicting the visualizations could not be observed through the system log. Eye movement data could support such analysis and acted as another objective measurement of how users engage in the visualization.

To further investigate the effect of format, task and goal orientation on the participants' eye movement behavior, we specify the following model to estimate the total fixation duration time in seconds and fixation count. Same with previous section, the linear regression model was used for estimation.

- Total fixation duration time

$$FixationDurationTime_i = \beta_0 + \sum_{j=1}^{2} \beta_j GoalOrientation_i + \sum_{j=3}^{4} \beta_j Format_i + \beta_5 Task_i + \sum_{j=6}^{8} \beta_j AOI_i + \beta_9 ProgrammingExperienced_i + \sum_{j=10}^{11} \beta_j ExamNumber_i + \beta_{12} Gender_i + \varepsilon_i \quad (4)$$

- Fixation count

$$FixationCount_i = \beta_0 + \sum_{j=1}^{2} \beta_j GoalOrientation_i + \sum_{j=3}^{4} \beta_j Format_i + \beta_5 Task_i + + \sum_{j=6}^{8} \beta_j AOI_i + \beta_9 ProgrammingExperienced_i + \sum_{j=10}^{11} \beta_j ExamNumber_i + \beta_{12} Gender_i + \varepsilon_i \quad (5)$$

The equation (4) test how goal orientation impact the total fixation duration time. Also, it test the influences of the format of graph visualization and the type of learning tasks on the total fixation duration time. The index i stands for each observation in our integrated collected data. The effects of dummy variable ***GoalOrientation*** on fixation duration are denoted by the parameters β_1 *and* β_2, which represents the difference level

DOI:10.6814/NCCU201901137

of the high degree observation and the low degree observation of the goal orientation, and the difference level of the high degree observation and the middle degree of the goal orientation. β_3 *and* β_4 are coefficients for the effect of dummy variable ***Format*** on fixation duration, which represents the difference level of the line format and the bar format, and the difference level of the radar format and the bar format. β_5 reflects the effect of dummy variable ***Task*** on fixation duration, which represents the difference level between search fact task and inference generation task on fixation duration. $\beta_6, \beta_7, \beta_8, \beta_9, \beta_{10}, \beta_{11}, \beta_{12}$ are coefficients for the effect of control variable. β_6, β_7 *and* β_8 are estimate of dummy variable ***AOI*** on fixation duration, which represents the difference level of the AOI Question, AOI Title, AOI Visualization and the AOI Legend as the baseline. β_9 is estimate of dummy variable ***ProgrammingExperienced***, which represents the difference level between the participants experienced in programming and the beginner. β_{10} *and* β_{11} are estimate of dummy variable ***ExamNumber***, which represents the difference level of the exam 2, exam 3 and the exam 1 as the baseline. β_{12} is estimate of dummy variable ***Gender***, which reflects the effect of difference level between the male and female. Finally the ε_i is the level-1 error term.

In sum, the effect of goal orientation on fixation duration time was found between high and middle degree of the goal orientation at the level of 0.01. The effect of format was found between radar and bar at the level of 0.05. The results show goal orientation and format has impact on the total fixation duration time. The estimated results for total fixation duration time in ALL page (without considering AOIs) are summarized in **Table 6.4.** Also we have found significant difference of control variables ***AOI*** between AOI Visualization, AOI Question, and the baseline. Thus we further estimate the same regression model on AOI Visualization and AOI Question. The results are summarized in **Table 6.5-6.6.**

DOI:10.6814/NCCU201901137

Table 6.4 Estimated results for total fixation duration time in All page.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	5.178	2.917	1.775	0.07643 *
Goal orientation (Low)	-1.405	1.787	-0.786	0.43218
Goal orientation (Middle)	-4.506	1.631	-2.763	0.00592 ***
Format (Line)	0.634	1.581	0.401	0.68858
Format (Radar)	3.948	1.562	2.527	0.01177 **
Task (SearchFact)	-0.398	1.274	-0.312	0.75487
AOI (Question)	26.271	1.758	14.942	< 2e-16 ***
AOI (Title)	-2.667	2.065	-1.292	0.19697
AOI (Visualization)	19.791	1.758	11.256	< 2e-16 ***
Programming-experienced	4.538	1.769	2.565	0.01058 **
Exam number (2)	-11.040	1.545	-7.148	2.87e-12 ***
Exam number (3)	-12.441	1.554	-8.005	7.30e-15 ***
Gender (Female)	3.035	1.294	2.345	0.01940 **

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201901137

Table 6.5 Estimated results for total fixation duration time in AOI Visualization.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	30.4693	5.1134	5.959	1.60e-08 ***
Goal orientation (Low)	-3.8283	3.4788	-1.100	0.2728
Goal orientation (Middle)	-6.0958	3.1475	-1.937	0.0546 *
Format (Line)	-0.2194	2.9898	-0.073	0.9416
Format (Radar)	7.6618	2.9841	2.568	0.0112 **
Task (SearchFact)	-2.4405	2.4365	-1.002	0.3181
Programming-experienced	6.7904	3.3503	2.027	0.0444 **
Exam number (2)	-21.7752	2.9841	-7.297	1.35e-11 ***
Exam number (3)	-20.9936	2.9898	-7.022	6.13e-11 ***
Gender (Female)	4.1293	2.4657	1.675	0.0960 .

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

Table 6.6 Estimated results for total fixation duration time in AOI Question.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	33.696	6.581	5.120	8.77e-07 ***
Goal orientation (Low)	-1.860	4.477	-0.415	0.678384
Goal orientation (Middle)	-8.227	4.051	-2.031	0.043935 **
Format (Line)	-1.054	3.848	-0.274	0.784430
Format (Radar)	3.223	3.841	0.839	0.402613
Task (SearchFact)	1.689	3.136	0.539	0.590867
Programming-experienced	6.100	4.312	1.415	0.159116
Exam number (2)	-13.916	3.841	-3.623	0.000391 ***
Exam number (3)	-18.984	3.848	-4.934	2.03e-06 ***
Gender (Female)	5.726	3.173	1.804	0.073072 *

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

The equation (5) test how goal orientation, format and task impact the fixation count. The index i stands for each observation in our integrated collected data. The

DOI:10.6814/NCCU201901137

effects of dummy variable ***GoalOrientation*** on fixation duration are denoted by the parameters β_1 *and* β_2, which represents the difference level of the high degree observation and the low degree observation of the goal orientation, and the difference level of the high degree observation and the middle degree of the goal orientation. β_3 *and* β_4 are coefficients for the effect of dummy variable ***Format*** on fixation count, which represents the difference level of the line format and the bar format, and the difference level of the radar format and the bar format. β_5 reflects the effect of dummy variable ***Task*** on fixation count, which represents the difference level between search fact task and inference generation task. β_6, β_7, β_8, β_9, β_{10}, β_{11}, β_{12} are coefficients for the effect of control variable. β_6, β_7 *and* β_8 are estimate of dummy variable ***AOI*** on fixation duration, which represents the difference level of the AOI Question, AOI Title, AOI Visualization and the AOI Legend as the baseline. β_9 is estimate of dummy variable ***ProgrammingExperienced***, which represents the difference level between the participants experienced in programming and the beginner. β_{10} *and* β_{11} are estimate of dummy variable ***ExamNumber***, which represents the difference level of the exam 2, exam 3 and the exam 1 as the baseline. β_{12} is estimate of dummy variable ***Gender***, which reflects the effect of difference level between the male and female. Finally the ε_i is the level-1 error term.

In sum, the results are similar to the results of fixation duration time. The effect of goal orientation on fixation count was found between high and middle degree of the goal orientation at the level of 0.05. The effect of format was found between radar and bar at the level of 0.05. The results show goal orientation and format has impact on the total fixation duration time. The estimated results for total fixation count are summarized in Table 6.7. We also found significant difference of control variables AOI between AOI Visualization, AOI Question, and the baseline. Thus we further

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

estimate the same regression model on AOI Visualization and AOI Question. The results are summarized in Table 6.8-6.9.

Table 6.7 Estimated results for total fixation count in All page .

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	34.548	13.208	2.616	0.00915 ***
Goal orientation (Low)	-5.762	8.090	-0.712	0.47660
Goal orientation (Middle)	-14.642	7.383	-1.983	0.04786 **
Format (Line)	3.893	7.159	0.544	0.58685
Format (Radar)	14.155	7.072	2.001	0.04584 **
Task (SearchFact)	-8.163	5.768	-1.415	0.15758
AOI (Question)	148.581	7.961	18.665	< 2e-16 ***
AOI (Title)	-13.764	9.349	-1.472	0.14154
AOI (Visualization)	107.028	7.961	13.445	< 2e-16 ***
Programming-experienced	21.319	8.010	2.662	0.00801 ***
Exam number (2)	-59.067	6.993	-8.446	2.78e-16 ***
Exam number (3)	-70.921	7.037	-10.079	< 2e-16 ***
Gender (Female)	12.994	5.860	2.217	0.02701 *

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

Table 6.8 Estimated results for total fixation count in AOI Visualization.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	165.999	23.792	6.977	7.82e-11 ***
Goal orientation (Low)	-13.322	16.187	-0.823	0.4117
Goal orientation (Middle)	-18.860	14.645	-1.288	0.1997
Format (Line)	1.913	13.912	0.138	0.8908
Format (Radar)	24.357	13.885	1.754	0.0813 .
Task (SearchFact)	-11.667	11.337	-1.029	0.3050
Programming-experienced	27.244	15.589	1.748	0.0825 .
Exam number (2)	-112.821	13.885	-8.125	1.21e-13 ***
Exam number (3)	-113.876	13.912	-8.186	8.50e-14 ***
Gender (Female)	24.906	11.473	2.171	0.0314 **

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

Table 6.9 Estimated results for total fixation count in AOI Question.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	203.829	28.301	7.202	2.27e-11 ***
Goal orientation (Low)	-9.669	19.254	-0.502	0.6163
Goal orientation (Middle)	-26.157	17.420	-1.501	0.1352
Format (Line)	-5.494	16.548	-0.332	0.7403
Format (Radar)	12.339	16.516	0.747	0.4561
Task (SearchFact)	-11.536	13.485	-0.855	0.3936
Programming-experienced	35.153	18.542	1.896	0.0598 *
Exam number (2)	-80.804	16.516	-4.892	2.43e-06 ***
Exam number (3)	-115.565	16.548	-6.984	7.53e-11 ***
Gender (Female)	17.431	13.647	1.277	0.2034

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

DOI:10.6814/NCCU201901137

6-3 Eye-tracking Data - Transition Analysis

- Transition

$$Transition_i = \beta_0 + \sum_{j=1}^{2} \beta_j GoalOrientation_i + \sum_{j=3}^{4} \beta_j Format_i + \beta_5 Task_i + \beta_6 ProgrammingExperienced_i + \sum_{j=7}^{8} \beta_j ExamNumber_i + \beta_9 Gender_i + \varepsilon_i \quad (6)$$

From the state diagrams of gaze transitions, we found that the transition between AOI Visualization and AOI Question is the chief transition of the visualization interface on TCAV. Hence, we focused on this transition and use regression model to explore the dependency between the transition rate and goal orientation, format and learning task.

The equation (6) test how goal orientation, format and learning task impact the transition rate between AOI Visualization and AOI Question. The index i stands for each observation in our integrated collected data. The effects of dummy variable ***GoalOrientation*** on learning comprehension are denoted by the parameters β_1 *and* β_2, which represents the difference level of the high degree observation and the low degree observation of the goal orientation, and the difference level of the high degree observation and the middle degree of the goal orientation. β_3 *and* β_4 are coefficients for the effect of dummy variable ***Format*** on learning comprehension, which represents the difference level of the line format and the bar format, and the difference level of the radar format and the bar format. β_5 reflects the effect of dummy variable ***Task*** on learning comprehension, which represents the difference level between search fact task and inference generation task on learning comprehension. β_6, β_7, β_8, β_9 are coefficients for the effect of control variable. β_6 reflects the effect of dummy variable ***ProgrammingExperienced***, which represents the difference level between the participants experienced in programming and the beginner. β_7 *and* β_8 are estimate

DOI:10.6814/NCCU201901137

of dummy variable ***ExamNumber***, which represents the difference level of the exam 2, exam 3 and the exam 1 as the baseline. β_9 is estimate of dummy variable ***Gender***, which reflects the effect of difference level between the male and female. Finally the ε_i is the level-1 error term.

In sum, the effect of task on transition rate was found between search fact task and inference generation task at the level of 0.05. The effect of format was found between radar and bar at the level of 0.1. The results show goal orientation and format has impact on the transition rate between AOI Visualization and AOI Question. The estimated results for transition rate are summarized in **Table 6.7.**

Table 6.10 Estimated results for transition rate between AOI Visualization and Question.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	0.63853	0.05826	10.959	< 2e-16 ***
Goal orientation (Low)	0.02566	0.03964	0.647	0.51830
Goal orientation (Middle)	0.05091	0.03586	1.419	0.15773
Format (Line)	-0.03527	0.03407	-1.035	0.30207
Format (Radar)	-0.05841	0.03400	-1.718	0.08777 *
Task (SearchFact)	0.06629	0.02776	2.388	0.01813 **
Programming-experienced	0.01973	0.03817	0.517	0.60593
Exam number (2)	0.09721	0.03400	2.859	0.00482 ***
Exam number (3)	0.07609	0.03407	2.234	0.02692 **
Gender (Female)	0.02843	0.02810	1.012	0.31307

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

DOI:10.6814/NCCU201901137

Chapter 7 DISCUSSIONS

The main objective of this study is to investigate the influence of learning goal orientation, visualization format, and type of learning task on students' learning perception and learning performance of our system in the context of Java programming. To fulfill the research goal, three research questions and corresponding hypotheses were proposed.

7-1 The Influence on User Behavior and Perception

The first research question aims to investigate the influence of the visualization format, type of learning task and individual differences on the learning comprehension, since these factors have been proved to be the indicators of graph comprehension (Shah & Freedman, 2011) and learning performance (Debicki et al., 2016). Based on the results from the linear regression analyses, H1a and H1c are supported, indicating that the learning goal orientation and task type could influence the student's learning comprehension. Students with a relatively high learning goal orientation would have a better degree of learning comprehension. The results also show that with the assistance of visualizations, students perform search fact task better than inference generation task in our system. However, H1b is not supported, indicating that visualization format has no significant influence on the student's learning comprehension. Similar results are found in the third research question, which aims to investigate the influence of the visualization format and individual differences on the student's perceived learning. From the linear regression analyses, H3a is supported, indicating that the learning goal orientation had an influence on the student's perceived learning. However, H3b is not supported, which means that visualization format has no significant influence on the student's perceived learning. These results show that the learning goal orientation is an important factor of learning performance, which is consistent with the previous studies.

DOI:10.6814/NCCU201901137

The students who are high in learning goal orientation would have higher motivation to learn in the blended learning condition, thus have a better comprehension and learning outcome to the programming learning contexts. Also, they will perceive a better learning performance when learning programming in the proposed learning system. We don't find any significant effect of visualization format on the student's learning comprehension and perceived learning, which does not meet our expectation (e.g. the radar graphs will be more effective for reviewing integrated information of exams). The possible reason is that we measure learning comprehension through the reviewing questions designed by ourselves, which are not various enough for users to demonstrate the difference of visualization formats.

The second research question aims to investigate the influence of the visualization format and individual differences on the understanding of visualization. H2b is supported, indicating that the format do have an influence on the student's understanding of visualization. However, H2a is not supported, which meant that learning goal orientation has no significant influence on the student's understanding of visualization. These results show that the users would depict the information visualization in different ways according to a different graph format is given. We measure the understanding of visualization by asking users a question with multiple choice based on Gestalt principles (e.g. The proposed visualization imply the trend of the correct answer rate). The difference between the bar and line graphs is consistent with the law of proximity and continuity of Gestalt principles. Also, comparing to the bar chart, students have a relatively worse understanding on line chart and radar chart. It indicates that the radar chart and line chart may not be suitable in our context because our tasks are mainly about comparing the difference between individual and class average. The results are consistent with prior study that radar graphs are considered inferior to bar graphs on common information seeking tasks (Few, 2005).

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

According to the interview, some participants reported the radar graph was more intuitive as it was widely used to display core competencies of school course, and some preferred bar graph and thought radar graph was hard to understand. The user feedback is consistent with the results of understanding of visualization. However, we didn't find any significant influence of visualization format on students' learning performance in our system. The possible reason is that each format convey a part of information which the students need or lack. Although there is difference between formats, our questions are not complicated enough for students to demonstrate the difference. They can retrieve the information from each kind of format. In sum, each kind of format may somehow be helpful for students to review exams. As a result, we can't find significant difference on learning performance between formats.

The results of learning comprehension also showed that there are significant difference between the three exams. Even though we control the effects of the three different exams, it indicate that there may be learning effect between the three exams. Thus we further estimated the same regression model on each exam. The results are summarized in **Table 7.1-7.3.** The possible reason is that we asked the same search fact questions and inference generation questions in each exam. Participants may tend to answer the question in purpose of achieving as high correct answer rate as possible. As a result, at the first exam, participants would reference to the visualization and answer the question step by step. But after the first exam, they learned to reference to the visualization in a more efficient way to answer the same question. Thus the format itself is not as important as the first iteration. Though the system log could support the objective results of users' learning performance, the results were also influenced by how we design the question and experiment procedure. Hence, the eye-tracking data could make up the deficiency of log analysis results.

DOI:10.6814/NCCU201901137

Table 7.1 Estimated results for learning comprehension in exam 1.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	0.72079	0.09732	7.406	5.06e-10 ***
Goal orientation (Low)	-0.09155	0.06920	-1.323	0.1908
Goal orientation (Middle)	-0.03229	0.06012	-0.537	0.5932
Format (Line)	-0.06695	0.06277	-1.066	0.2905
Format (Radar)	-0.12884	0.06274	-2.054	0.0444 **
Task (SearchFact)	0.31893	0.04683	6.810	5.26e-09 ***
Programming-experienced (Experienced)	-0.02056	0.06599	-0.312	0.7565
Gender (Female)	-0.03388	0.04834	-0.701	0.4861

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

Table 7.2 Estimated results for learning comprehension in exam 2.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	0.94122	0.07044	13.361	< 2e-16 ***
Goal orientation (Low)	-0.10939	0.05494	-1.991	0.0510 *
Goal orientation (Middle)	-0.08213	0.05143	-1.597	0.1155
Format (Line)	0.07466	0.04960	1.505	0.1375
Format (Radar)	0.08060	0.04941	1.631	0.1081
Task (SearchFact)	0.03074	0.03932	0.782	0.4374
Programming-experienced (Experienced)	-0.08643	0.05123	-1.687	0.0968 *
Gender (Female)	0.06617	0.04134	1.601	0.1147

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

DOI:10.6814/NCCU201901137

Table 7.3 Estimated results for learning comprehension in exam 3.

	Estimate	Std. Error	t value	Pr(>|t|)
Constant	0.947248	0.083509	11.343	< 2e-16 ***
Goal orientation (Low)	-0.113797	0.061708	-1.844	0.0701 *
Goal orientation (Middle)	-0.013169	0.056197	-0.234	0.8155
Format (Line)	0.006314	0.058335	0.108	0.9142
Format (Radar)	-0.041197	0.057088	-0.722	0.4733
Task (SearchFact)	0.034706	0.043940	0.790	0.4327
Programming-experienced (Experienced)	-0.062621	0.056753	-1.103	0.2743
Gender (Female)	0.042573	0.046107	0.923	0.3595

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

7-2 Eye Movement and User Behavior

In order to gain a better understanding of the effect of presentation format, we included eye movement analyses in the present study. The linear regression analysis was also adopted to analyze the effects.

- **Fixation analysis**

Goal orientation - Main effects on fixation duration time

We found main effects of goal orientation on fixation duration in regression model of All page, AOI Question and AOI Visualization. The results are summarized in Table 7.4. The results showed that the users with the middle degree of the goal orientation had a lower fixation duration time than high goal orientation users. It indicated that the users with high degree of the goal orientation tended to pay more attention on the format

DOI:10.6814/NCCU201901137

立 政 治 大 學 國 National Chengchi University

and focus on answering the review questions with the assistance of visualizations (see Figure 7.1, 7.2, 7.3). The results were consistent with the results of learning comprehension and perceived learning, which provided a valid evidence of H1a and H3a.

We also observed that users with low degree of goal orientation tended to have longer fixation duration with bar graphs from Figure 7.1, 7.2, 7.3. However, we didn't find the main effects in regression analysis. The possible reason was that these users were allocated with bar graphs in exam1. And they tended to spend more time in their first iteration rather than the following two iterations. This bias might influence the results of fixation analysis. As a result, we could only observe the difference between the middle goal orientation users and the high goal orientation users in the fixation analysis despite the fact that we observed the similar results between the low goal orientation users and the high goal orientation users in line and radar graphs (see Figure 7.1). It could also explain why the regression results of log analysis and eye-tracking analysis were not completely consistent.

Table 7.4 Main effects of goal orientation on fixation duration.

AOI	Coefficient	Estimate	Pr(>|t|)
All page	Goal orientation (Middle)	-4.506	0.00592 ***
Visualization	Goal orientation (Middle)	-6.0958	0.0546 *
Question	Goal orientation (Middle)	-8.227	0.043935 **

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

DOI:10.6814/NCCU201901137

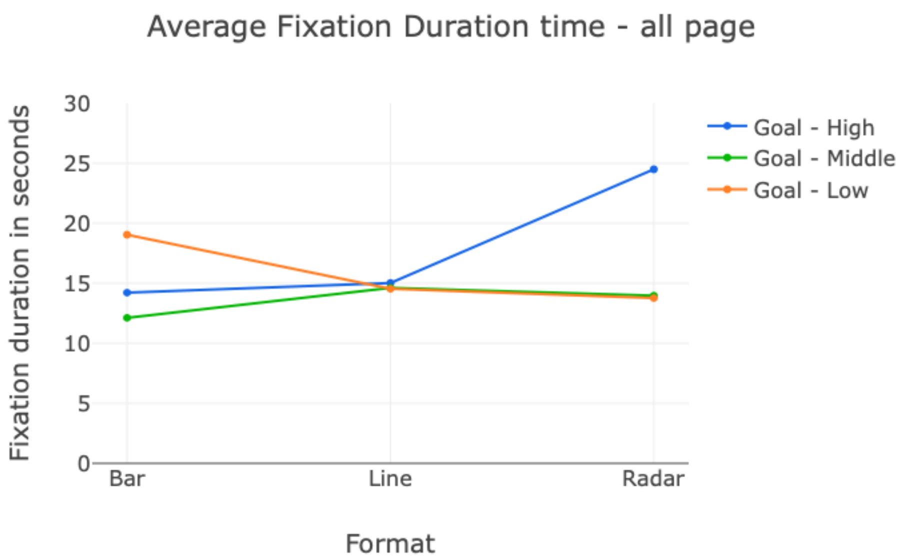


Figure 7.1 Main effect of goal orientation on fixation duration in All page.

* Y-axis is the average fixation duration of four AOIs.

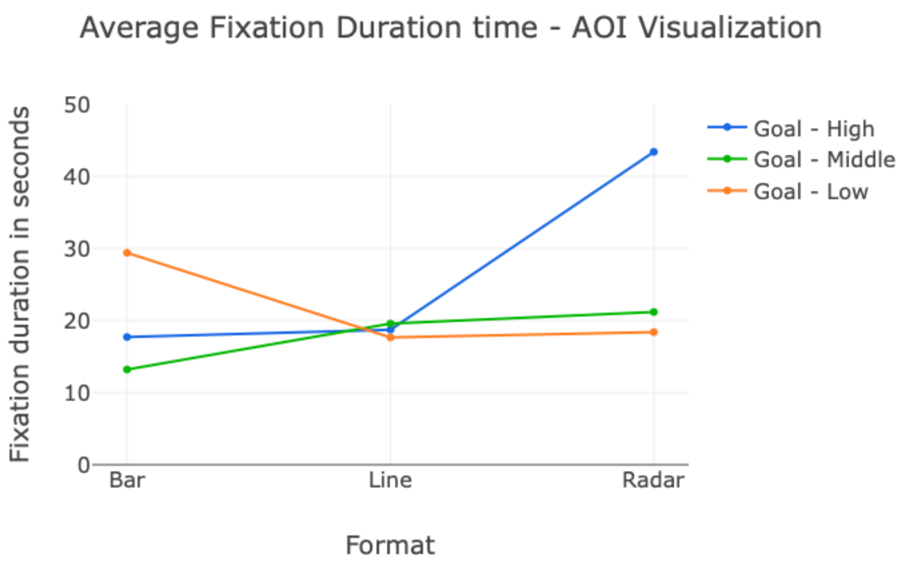


Figure 7.2 Main effect of goal orientation on fixation duration in AOI Visualization.

DOI:10.6814/NCCU201901137

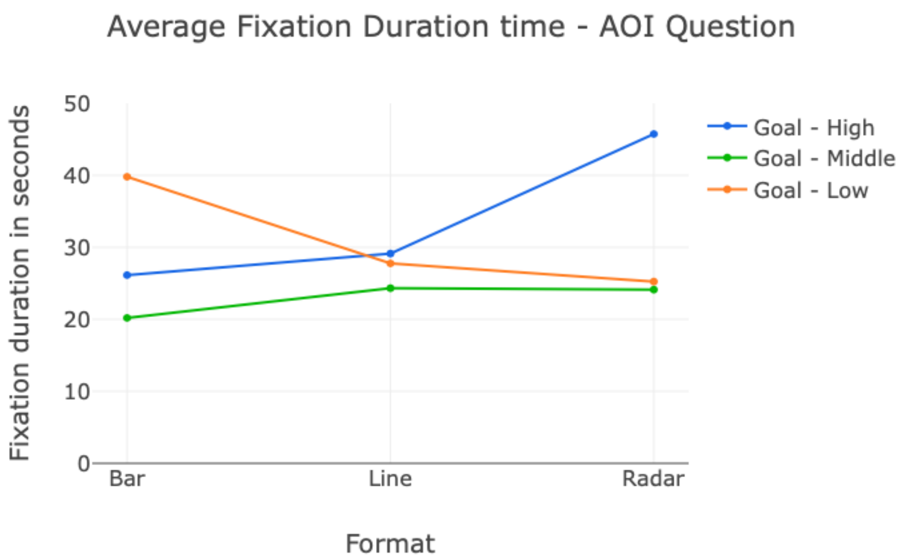


Figure 7.3 Main effect of goal orientation on fixation duration in AOI Question.

Format - Main effects on fixation duration

We found main effects of goal orientation on fixation duration in regression model of All page and AOI Visualization. The results are summarized in Table 7.5. The results showed that users had a high average of fixation duration on radar chart than bar chart in AOI Visualization (see Figure 7.4, 7.5). It indicated that the users spent more time on retrieving the information of radar chart. The possible reason was that radar chart had a more complex composition than bar chart. Hence users may be confused when doing information seeking tasks. The results were consistent with the results of understanding of visualization, which provided a valid evidence of H2b.

Same with the results of goal orientation on fixation duration, Figure 7.4 and 7.5 also showed that users with low degree of goal orientation tended to have longer fixation duration with bar graphs. However, we didn't find the main effects in regression analysis either.

DOI:10.6814/NCCU201901137

Table 7.5 Main effects of format on fixation duration.

AOI	Coefficient	Estimate	Pr(>|t|)
All page	Format (Radar)	3.948	0.001177 **
Visualization	Format (Radar)	7.6618	0.0112 **

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

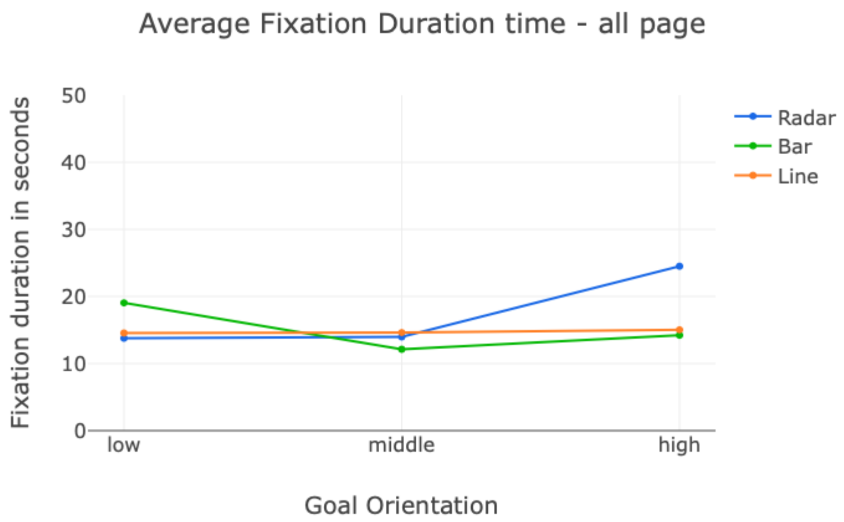


Figure 7.4 Main effect of format on fixation duration in All page.

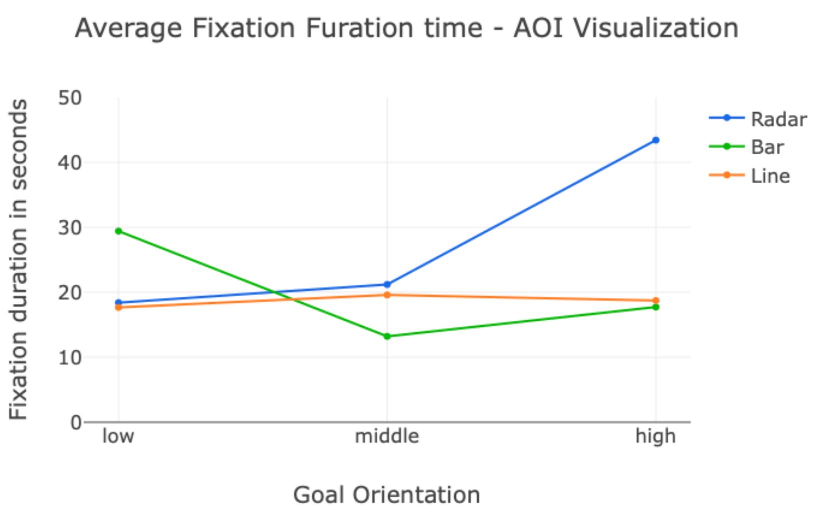


Figure 7.5 Main effect of format on fixation duration in AOI Visualization.

DOI:10.6814/NCCU201901137

Goal orientation - Main effects on fixation count

Similar to the results of fixation duration, we found main effects of goal orientation on fixation count in All page regression model. The results are summarized in Table 7.6. Users with a middle degree of goal orientation had a relatively low average of fixation count than the users with a high degree of goal orientation (see Figure 7.6). Comparing to the results of fixation duration time, though there was no significance of main effects between middle degree of goal orientation and high degree of goal orientation on fixation count in AOI Question (Pr(>|t|) = 0.1352), we still observed the similar trend between these two groups (see Figure 7.7). The possible reason that the fixation duration time result was not completely agree with the fixation count result was that we counted fixation duration time as the average of the duration for all fixations within an AOI, and the fixation count as the number of times the user fixated on an AOI. Hence the amount of fixation count was not completely correlated to the average fixation duration time due to each fixation could be just a few seconds or dozens of seconds.

Table 7.6 Main effects of format on fixation count.

AOI	Coefficient	Estimate	Pr(>|t|)
All page	Goal orientation (Middle)	-14.642	0.04786 **

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

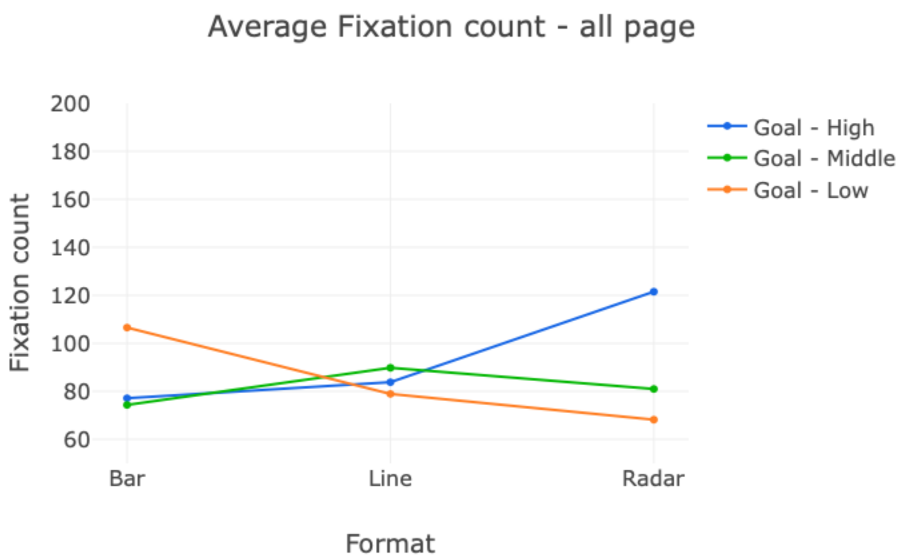


Figure 7.6 Goal orientation on fixation count in All page.

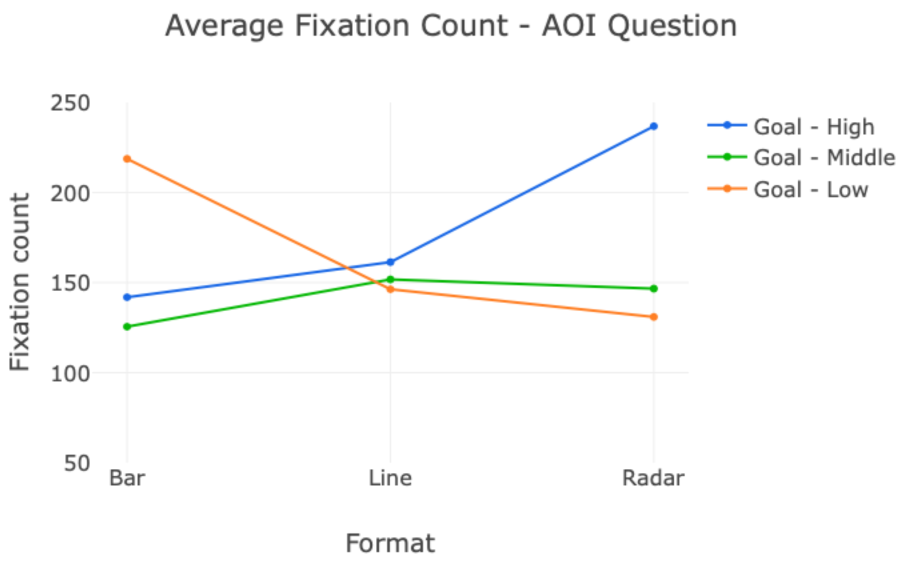


Figure 7.7 Goal orientation on fixation count in AOI Question.

Format - Main effects on fixation count

Similar to the results of fixation duration, we found main effects of format on fixation count in All page and AOI Visualization regression model. The results are summarized in Table 7.7. The results showed that users had a high average of fixation count on radar chart than bar chart in AOI Visualization (see Figure 7.8, 7.9). It

DOI:10.6814/NCCU201901137

indicated that users used more fixations on retrieving the information of radar chart, which indicated a less efficient search.

Figure 7.8 and 7.9 also showed that users with low degree of goal orientation tended to have more fixation count with bar graphs. However, we didn't find the main effects in regression analysis either.

Table 7.7 Main effects of format on fixation count.

AOI	Coefficient	Estimate	Pr(>|t|)
All page	Format (Radar)	14.155	0.04584 **
Visualization	Format (Radar)	24.357	0.0813 *

* Significance (Sig.) at 0.1 level, ** Sig. at 0.05 level, *** Sig. at 0.01 level.

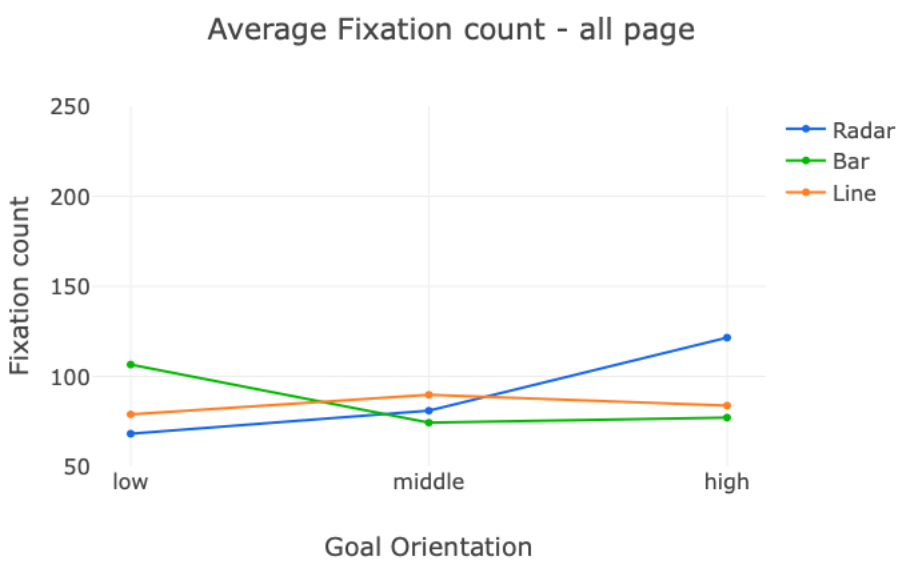


Figure 7.8 Main effect of format on fixation count in All page.

DOI:10.6814/NCCU201901137

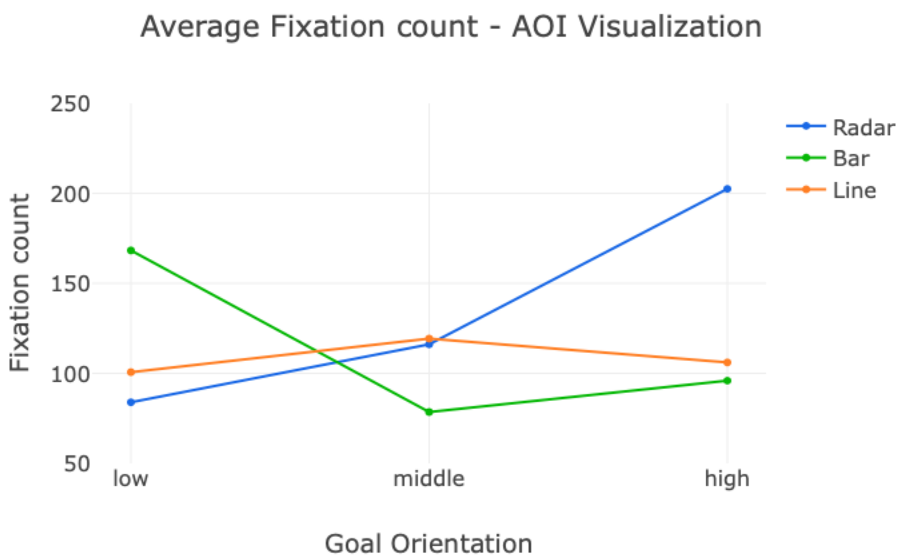


Figure 7.9 Main effect of format on fixation count in AOI Visualization.

- **Transition analysis**

To explore the dependency of the transition between areas of interest (AOIs), we focused on AOI Visualization, AOI Question and the corresponding factors because the dominated eye movement transition for both search fact and inference generation task is VQ and QV. We found the main effects of format and task on the transition rate between AOI Visualization and AOI Question. The results showed that the users performed more transitions between AOI Visualization and AOI Question during answering search fact questions than inference generation questions. It indicated that visualizations could facilitate the information search fact tasks. The results were consistent with the results of learning comprehension, which provided a valid evidence of H1c. We also found that users had a relatively less transition rate between AOI Visualization and AOI Question on radar chart than bar chart. It indicated that radar graph was not suitable for the information search task due to its complicated composition. The results were consistent with the results of understanding of visualization and fixation analysis, which provided a valid evidence of H2b.

DOI:10.6814/NCCU201901137

In sum, the results of fixation analysis and transition analysis show that learning goal orientation and format have impact on the fixation duration and fixation count, while format and type of learning task have impact on the transition rate. In terms of the learning goal orientation, participants who have a middle degree of learning goal orientation tend to have less fixation duration and fixation count than who have a high degree of learning goal orientation. This result suggest that the students with high learning goal orientation are more detailed when they review the visualization analytics. They are willing to spend more time to explore the information embedded in the visualization for their own personal development. The results also show that participants have a relatively higher fixation duration and fixation count on radar graph then other formats, which meant that participants tend to spend more time on radar graph. The possible explanation is that radar graph contained more information than the other format, so that participants needed to spend more time on it to decode information from radar graph. It is consistent with the feature that radar graph was inferior on information retrieving. Finally, the transition analysis results show that participants have relatively higher transition rate between AOI Visualization and AOI Question on search fact task than inference generation task. As Visualization and Question AOIs are two major area of the visualization interface of TCAV, VQ and QV are two transitions which dominate user's eye movement transition. The higher frequency of using these two transitions in the visualization interface might indicate that these processes were better supported by the design of visualizations analytics.

DOI:10.6814/NCCU201901137

Chapter 8 CONCLUSION

This present study investigated the effect of the learning goal orientation, visualization format and type of learning task on the students' perception of our visualization analytics system. We conducted a within-subject model experiment and utilized questionnaires and an eye-tracker to collect survey, user log and eye movement data. The results revealed that the students with higher learning goal orientation would have a better learning comprehension and perceived learning while the different visualization format induced different levels of understanding of visualization. Further, from the eye-tracking analysis, we have found that students who had high level of learning goal orientation would had more motivation to explore the system to review the exam and learn Java programming. The results of fixation analysis and transition analysis showed that users had a relatively higher fixation duration time and fixation count but lower transition rate between AOI Visualization and AOI Question in radar graphs than other graphs. It indicated that users tended to perform less efficient search using radar graph, which support the results of understanding of visualization from user log analysis. Finally, we explore the transition rate between AOI Visualization and AOI Question and found that search fact task is better reinforced than inference generation task by visualization analytics.

Based on our results, several implications for practitioners can be drawn from them. First, the type of learning task should be taken into consideration while designing learning analytics tool and providing visualization. The pattern of performing search fact tasks and inference generation tasks are different. Users tend to rely more on the visualization analytics than their own prior knowledge when asked to answer specific search fact question. In other words, when asked to answer inference generation questions, which are more open-ended, the viewers would select what they consider to

DOI:10.6814/NCCU201901137

be the important information and used their knowledge about format to make the judgments. This finding also support the previous studies (Shah & Freedman, 2011). Hence, such design of visualization analytics support search fact task better. Therefore, the difference of learning task would in turn influence the learning comprehension of the context. One should always keep in mind that whether the target learning task is suitable for going with visualization analytics. Although there is no significant correlation between formats and learning performance in log analysis, we still found main effect of format on understanding of visualization and fixation analyses. Users not only have low degree of understanding of visualization on radar graphs, but also perform less transition between AOI Visualization and AOI Question during answering the review task questions. They interpreted the information less efficiently through radar graphs. As a result, visual format should be chosen carefully in order to let the users understand and interpret the information effortlessly.

The main contribution of our work is to propose a learning analytic tool on an orchestration technology which integrates the visual analytic dashboard into the online virtual exam review system. Existing research in the field of learning analytics focuses on depicting the effects of individual difference on learning performance or academic performance. Most of the learning analytics tools are developed with innovative approaches, but lack of using theoretically established approaches (Papamitsiou & Economides, 2015). Some psychological indicators are involved including self-efficacy, locus of control, learning engagement and learning goal orientation (Albert & Dahling, 2016; Joo et al., 2013). However, as the analytic dashboards become increasingly aware, the user performance of utilizing the visual analytics should also be considered in the learning analytics research. Therefore, we dug into the graph theory and combined graph comprehension into learning analytics. We extended the literature by combining graph comprehension theory and well-developed learning goal orientation into the

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

development of an operating dashboard platform, which aims to support students to review paper-based programming exam. Further, we empirically establish instructional strategies of developing and investigating under the consideration of psychological indicators, graph comprehension and learning genres types using regression model.

The second contribution of our work is to utilize the objective trace data and subjective feedback from users to investigate the usability and perception of the proposed system. Existing research in the field of information visualization and interface focuses on depicting the personality factors on interface action. Some of them collected and decoded user's comments, then depicted with the personality factors and graphical skills (Shah & Freedman, 2011). Others used eye-tracking data and discussed the effect of user's cognitive abilities on the user gaze behavior and pattern (Toker et al., 2013). In this study, by combining questionnaire, log and eye-tracking data analysis, we used the various data sources to measure student's learning performance and perception on our visualization interface. Our study demonstrated the added value of the interfaces with visualization analytics in the information search tasks of reviewing programming exams. Questionnaire and system log analysis report both subjective and objective results for the observed difference of learning performance of visualization analytics. Eye-tracking analysis further highlight differences between learning goal orientation, visualization formats and learning task type and connected them with user learning perceptions of visualization analytics.

Some limitations existed in the present study. First, the participants were not general enough due to our system limitation. We focused on reviewing Java programming paper-based exam on the virtual online platform. So the participants only be recruited from students who took Java programming course. Some of the results might be biased in regression analysis due to the insufficient numbers of participants. Therefore, to promote our system, further experiment and analysis are required to

DOI:10.6814/NCCU201901137

confirm the effect of the visualization formats and personality factors. Second, the experiment question design was rather simple. Because we used the same question set of search fact task and inference generation task in all three iterations, there may be a learning effect during answering the experiment questions. There was difference of time spent between the first exam and the rest. Users could be easy to retrieve the answers from the visualization regardless of the format of the visualization after finishing the first iteration. So the results of fixation duration and fixation count analysis may be influenced. The insignificant effect of format of log analysis might result from the simplicity of our question design. Nonetheless, some useful insights were drawn from our analysis results. These findings shed light on the future works of information visualization.

In the future, we should increase the number of participants if we need to perform the user study again. A detailed illuminated manuscript and operational test before the formal experiment should be also provided. More participants and manuscript will decrease the bias between the iterations and brings a more reliable results. Also, both the search fact questions and inference generation questions should be redesign for the reason that there may be learning effect between three iterations. The questions should be various enough to decrease the learning effect of participants. Also, the questions, especially the search fact questions, should be more challenging so that participants will more rely on the visualizations and demonstrate the difference between the formats.

Second, further study of an interface with more deluxe AOIs may be helpful to learn more about the interaction between the personality factors and the interface. The transition between the AOI Visualization and AOI Question is dominated among the user eye movements in the present study. With more AOIs involved, we hope to find more sophisticated results between user patterns and learning performance from fixation analysis and transition analysis.

DOI:10.6814/NCCU201901137

Another part we should add in the future is the field study conducted in the real programming course classroom. As it is the classroom orchestration system aims to enhance learning awareness to programming learners by providing elaborated visualization results, the system should be implemented in the real classroom. The results of our user study should be reproduce in the field study. Without collecting the eye-tracking data, we can still measure the user learning perceptions through questionnaire and system user log. At the beginning of the new semester, we can collect student's learning goal orientation through questionnaire in pretest. After each exam, students are given visualization analytics with a specific format. Students can get additional credits if they have reviewed the exam on the platform. To finish reviewing exams, they should go through several questions which are labeled as search fact questions and inference generation questions, just as the review tasks practiced in the user study. The results of the field study should support a more reliable and general results to the user study due to it has more participants join in the research.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

REFERENCE

Albert, M. A., & Dahling, J. J. (2016). Learning goal orientation and locus of control interact to predict academic self-concept and academic performance in college students. *Personality and Individual Differences*, *97*, 245–248.

Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers and Education*, *58*(1), 470–489.

Ali, N., & Peebles, D. (2013). The effect of Gestalt laws of perceptual organization on the comprehension of three-variable bar and line graphs. *Human Factors*, *55*(1), 183–203.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*(2), 117–148.

Barzilai, S., & Blau, I. (2014). Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers & Education*, *70*, 65–79.

Brusilovsky, P., Hsiao, I. H., & Folajimi, Y. (2011). QuizMap: Open social student modeling and adaptive navigation support with TreeMaps. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6964 LNCS*, 71–82.

Cascio, M. I., Botta, V. C., & Anzaldi, V. E. (2013). The role of self efficacy and internal locus of control in online learning. *Journal of E-Learning and Knowledge Society*, *9*(3), 95–106.

Caspi, A., & Blau, I. (2008). Social presence in online discussion groups: Testing three conceptions and their relations to perceived learning. *Social Psychology of Education*, *11*(3), 323–346.

DOI:10.6814/NCCU201901137

Chang, M. M. (2005). Applying self-regulated learning strategies in a web-based instruction - An investigation of motivation perception. *Computer Assisted Language Learning*, *18*(3), 217–230.

Chen, I. S. (2017). Computer self-efficacy, learning performance, and the mediating role of learning engagement. *Computers in Human Behavior*, *72*, 362–370.

Conati, C., & Maclaren, H. (2010). Exploring the role of individual differences in information visualization. *Proceedings of the Working Conference on Advanced Visual Interfaces*, 199–206.

Crawford, E. R., LePine, J. A., & Rich, B. L. (2010). Linking job demands and resources to employee engagement and burnout: A theoretical extension and meta-analytic test. *Journal of Applied Psychology*, *95(5)*, 834–848.

Cui, W., Zhou, H., Qu, H., Wong, P. C., & Li, X. (2008). Geometry-based edge clustering for graph visualization. *IEEE Transactions on Visualization and Computer Graphics*, *14*(6), 1277–1284.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319.

Debicki, B. J., Kellermanns, F. W., Barnett, T., Pearson, A. W., & Pearson, R. A. (2016). Beyond the Big Five: The mediating role of goal orientation in the relationship between core self-evaluations and academic performance. *International Journal of Management Education*, *14*(3), 273–285.

Denny, P., Luxton-Reilly, A., & Hamer, J. (2008). The PeerWise system of student contributed assessment questions. *Conferences in Research and Practice in Information Technology Series*, *78*, 69–74.

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers and Education*, *69*, 485–492.

Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. *New*

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201901137

Science of Learning: Cognition, Computers and Collaboration in Education, 525–552.

Eison, J. A. (1979). The development and validation of a scale to assess different student orientations towards grades and learning.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for Kit of factor-referenced cognitive tests*. Educational testing service Princeton, NJ.

Few, S. (2005). Keep Radar Graphs Below the Radar–Far Below. *Perceptual Edge*, (May), 1–5.

Freedman, E. G., Shah, P., & Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. *The Cambridge Handbook of Visuospatial Thinking*, 426–476.

Gansner, E. R., & North, S. C. (1999). An Open Graph Visualization System and Its Applications. *Software - Practice and Experience*, *30*(May 1999), 1203–1233.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let ' s not forget : Learning analytics are about learning. *TechTrends*, *59*(1), 64–71.

Gibson, D., & de Freitas, S. (2016). Exploratory Analysis in Learning Analytics. *Technology, Knowledge and Learning*, *21*(1), 5–19.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs Joseph. *International Journal of Industrial Ergonomics*, *24*(6), 631–645.

Govaerts, S., Verbert, K., Duval, E., & Pardo, A. (2012). The student activity meter for awareness and self-reflection. *Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*, 869–884.

Green, Tear Marie, & Fisher, B. (2010). Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction. *VAST*

DOI:10.6814/NCCU201901137

10 - IEEE Conference on Visual Analytics Science and Technology 2010, Proceedings, (November), 203–210.

Green, Tera Marie, Jeong, D. H., & Fisher, B. (2010). Using personality factors to predict interface learning performance. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 1–10.

Helfman, J. I., & Goldberg, J. H. (2007). Selecting the Best Graph Based on Data, Tasks, and User Roles. *Usability Professionals' Association, UPA Conference on Patterns: Blueprints for Usability, Austin, TX*, (July).

Hsiao, I-Han. (2016). Mobile Grading Paper-Based Programming Exams: Automatic Semantic Partial Credit Assignment Approach. In K. Verbert, M. Sharples, & T. Klobučar (Eds.), *Adaptive and Adaptable Learning* (pp. 110–123). Cham: Springer International Publishing.

Hsiao, I-Han, Pandhalkudi Govindarajan, S. K., & Lin, Y.-L. (2016). Semantic visual analytics for today's programming courses. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 48–53.

Hsiao, I. -Han, & Brusilovsky, P. (2012). Motivational Social Visualizations for Personalized E-Learning (pp. 153–165).

Hsiao, I. H., Huang, P. K., & Murphy, H. (2017). Uncovering reviewing and reflecting behaviors from paper-based formal assessment. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 319–328.

Hsiao, I. Han, Bakalov, F., Brusilovsky, P., & König-Ries, B. (2013). Progressor: Social navigation support through open social student modeling. *New Review of Hypermedia and Multimedia*, *19*(2), 112–131.

Hsiao, I. Han, & Lin, Y. L. (2017). Enriching programming content semantics: An evaluation of visual analytics approach. *Computers in Human Behavior*, *72*, 771–782.

DOI:10.6814/NCCU201901137

Hsiao, I., Sosnovsky, S., & Brusilovsky, P. (2008). Adaptive Navigation Support in an E-Learning System for Java Programming. *Journal of Computer Assisted Learning*, *26*(4), 270–283.

Hu, P. J. H., & Hui, W. (2012). Examining the role of learning engagement in technology-mediated learning and its effects on learning effectiveness and satisfaction. *Decision Support Systems*, *53*(4), 782–792.

Joo, Y. J., Lim, K. Y., & Kim, J. (2013). Locus of control, self-efficacy, and task value as predictors of learning outcome in an online university context. *Computers & Education*, *5*(4), 149–158.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*(4), 441–480.

Klein, H. J., Noe, R. A., & Wang, C. (2006). Motivation To Learn and Course Outcomes: the Impact of Delivery Mode, Learning Goal Orientation, and Perceived Barriers and Enablers. *Personnel Psychology*, *59*(3), 665–702.

Koffka, K. (2013). *Principles of Gestalt psychology*. *Routledge*.

Li, R., Pelz, J., Shi, P., & Haake, A. R. (2012). Learning image-derived eye movement patterns to characterize perceptual expertise. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *34*(34), 1900–1905.

Lu, Y., & Hsiao, I.-H. (2016). Seeking Programming-related Information from Large Scaled Discussion Forums, Help or Harm? *Edm '16*, 442–447.

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, *9*(11), 8–8.

Mazur, D. J., & Hickam, D. H. (1993). Patients' and Physicians' Interpretations of Graphic Data Displays. *Medical Decision Making*, *13*(1), 59–63.

Michalski, R. S. (1993). Inferential Theory of Learning as a Conceptual Basis for

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

Multistrategy Learning. *Machine Learning*, *11*(2), 111–151.

Nesbitt, K. V., & Friedrich, C. (2002). Applying Gestalt principles to animated visualizations of network data. *Proceedings of the International Conference on Information Visualisation*, *2002-Janua*(January), 737–743.

Okan, Yashmina, Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2011). Individual Differences in Graph Literacy: Overcoming Denominator Neglect in Risk Comprehension. *Journal of Behavioral Decision Making*, *21*, 453–456.

Okan, Yasmina, Garcia-Retamero, R., Galesic, M., & Cokely, E. T. (2012). When Higher Bars Are Not Larger Quantities: On Individual Differences in the Use of Spatial Information in Graph Comprehension. *Spatial Cognition & Computation*, *12*(2–3), 195–218.

Papamitsiou, Z., & Economides, A. A. (2015). Temporal learning analytics visualizations for increasing awareness during assessment. *RUSC. Universities and Knowledge Society Journal*, *12*(3), 129.

Paredes, Y. V., Huang, P. K., Murphy, H., & Hsiao, I. H. (2017). A Subjective Evaluation of Web-based Programming Grading Assistant: Harnessing digital footprints from paper-based assessments. *CEUR Workshop Proceedings*, *1828*, 23–30.

Patel, K., Bancroft, N., Drucker, S. M., Fogarty, J., Ko, A. J., & Landay, J. A. (2010). Gestalt: Integrated Support for Implementation and Analysis in Machine Learning. *Uist*, 37–46.

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology*, *92*(1), 128–150.

Pinker, S. (1990). A theory of graph comprehension. *Artificial Intelligence and the Future of Testing*, 73–126.

DOI:10.6814/NCCU201901137

Ratwani, R. M., & Boehm-davis, D. A. (2008). Thinking graphically: Connecting vision and cognition during graph comprehension. *Journal of Experimental Psychology*, (703).

Ratwani, R. M., & Trafton, J. G. (2008). Shedding light on the graph schema: Perceptual features versus invariant structure. *Psychonomic Bulletin and Review*, *15*(4), 757–762.

Rovai, A. P., & Baker, J. D. (2005). Gender Differences in Online Learning: Sense of Community, Perceived Learning, and Interpersonal Interactions. *Quarterly Review of Distance Education*, *6*(1), 31–44.

Shah, P., & Freedman, E. G. (2011). Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Topics in Cognitive Science*, *3*(3), 560–578.

Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as Aids to Knowledge Construction: Signaling Techniques for Guiding the Process of Graph Comprehension. *Journal of Educational Psychology*, *91*(4), 690–702.

Siemens, G., & Baker, R. S. J. de. (2012). Learning analytics and educational data mining: towards communication and collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 252–254.

Siemens, George, & Gasevic, D. (2012). Guest editorial-learning and knowledge analytics. *Journal of Educational Technology and Society*, *15*(3).

Steichen, B., Carenini, G., & Conati, C. (2013). User-adaptive information visualization: using eye gaze data to infer visualization tasks and user cognitive abilities. *Iui*, 317–328.

Sun, R., Merrill, E., & Peterson, T. (2001). From implicit skills to explicit knowledge_ a bottom-up model of skill learning. *Cognitive Science*, *25*, 203–244.

DOI:10.6814/NCCU201901137

Swan, K. (2001). Virtual interaction : Design factors affecting student satisfaction and perce ... *Education*, *22*(2), 306–331. Retrieved from

Toker, D., Conati, C., Carenini, G., & Haraty, M. (2012). Towards adaptive information visualization: On the influence of user characteristics. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7379 LNCS*, 274–285.

Toker, D., Conati, C., Steichen, B., & Carenini, G. (2013). Individual user characteristics and information visualization: connecting the dots through eye tracking. In *proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 295–304).

VandeWalle, D. M. (1997). Development and validation of a work domain achievement goals instrument. *Educational and Psychological Measurement*, *8*(6), 995–1015.

Venkatesh, V. (2000). Determinants of perceived ease of use: integrating perceived behavioral control, computer anxiety and enjoyment into the Technology Acceptance Mode. *Information Systems Research*, *11*(4), 342–365.

Wang, C., Shannon, D. M., & Ross, M. E. (2010). Students ' characteristics , self-regulated learning , technology self-ef fi cacy , and course outcomes in online learning. *Dissertation*, *34*(3), 302–323.

Wang, M., Wu, B., Kinshuk, Chen, N. S., & Spector, J. M. (2013). Connecting problem-solving and knowledge-construction processes in a visualization-based learning environment. *Computers and Education*, *68*, 293–306.

Wu, D., Hiltz, Roxanne, S., & Bieber, M. (2010). Acceptance of educational technology: field studies of asynchronous participatory examinations. *Communications of the Association for Information Systems*, *26*(1), 21.

Xanthopoulou, D., Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2009).

DOI:10.6814/NCCU201901137

國立政治大學 National Chengchi University

Reciprocal relationships between job resources, personal resources, and work engagement. *Journal of Vocational Behavior*, *74*(3), 235–244.

Yi, M. Y., & Hwang, Y. (2003). Predicting the use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human Computer Studies*, *59*(4), 431–449.

Yigitbasioglu, O. M., & Velcu, O. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, *13*(1), 41–59.

Zajac, D. M., Button, S. B., & Mathieu, J. E. (1996). Goal Orientation in Organizational Research : A Conceptual and Empirical Foundation. *Organizational Behavior and Human Decision Processes*, *67*(1), 26–48.

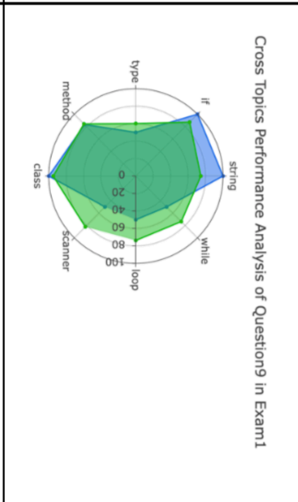
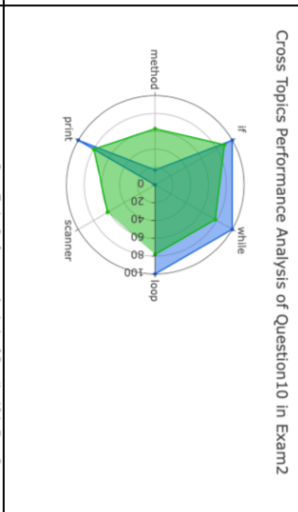
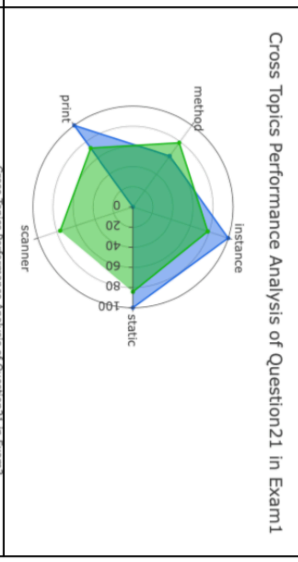
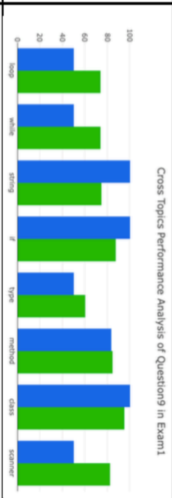
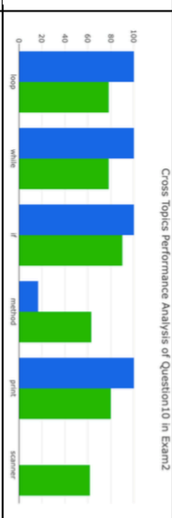
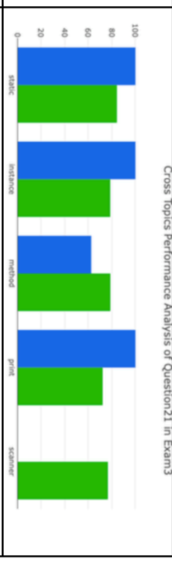
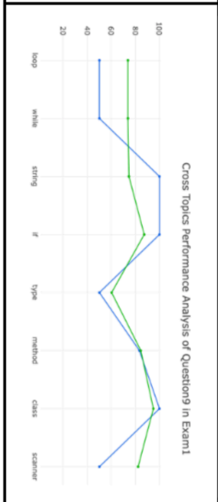
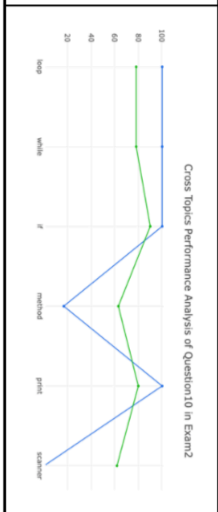
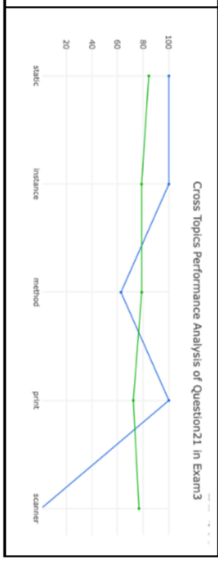
Ziemkiewicz, C., Crouser, R. J., Yauilla, A. R., Su, S. L., Ribarsky, W., & Chang, R. (2011). How locus of control influences compatibility with visualization style. *VAST 2011 - IEEE Conference on Visual Analytics Science and Technology 2011, Proceedings*, 81–90.

Ziemkiewicz, C., & Kosara, R. (2009). Preconceptions and individual differences in understanding visual metaphors. *Computer Graphics Forum*, *28*(3), 911–918.

立 政 治 大 學 國 · · National Chengchi University

DOI:10.6814/NCCU201901137

Appendix A: Visualizations with different format

	Exam1	Exam2	Exam3
Radar			
Bar			
Line			

DOI:10.6814/NCCU201901137

Appendix B: Learning Goal Orientation Measurement Items

1. The opportunity to do challenging work is important to me.
2. When I fail to complete a difficult task, I plan to try harder the next time I work on it.
3. I prefer to work on tasks that force me to learn new things.
4. The opportunity to learn new things is important to me.
5. I do my best when I’m working on a fairly difficult task.
6. I try hard to improve on my past performance.
7. The opportunity to extend the range of my abilities is important to me.
8. When I have difficulty solving a problem, I enjoy trying different approaches to see which one will work.

DOI:10.6814/NCCU201901137