

Investigating Attributes Affecting the Performance of WBI Users

Rana A. Alhajri¹, Steve Counsell and XiaoHui Liu

Department of Information Systems and Computing, Brunel University,
Uxbridge, Middlesex, UB8 3PH, UK

Abstract Numerous research studies have explored the effect of hypermedia on learners' performance using Web Based Instruction (WBI). A learner's performance is determined by their varying skills and abilities as well as various differences such as gender, cognitive style and prior knowledge. In this paper, we investigate how differences between individuals influenced learner's performance using a hypermedia system to accommodate an individual's preferences. The effect of learning performance is investigated to explore relationships between measurement attributes including gain scores (post-test minus pre-test), number of pages visited in a WBI program, and time spent on such pages. A data mining approach was used to analyze the results by comparing two clustering algorithms (K-Means and Hierarchical) with two different numbers of clusters. Individual differences had a significant impact on learner behaviour in our WBI program. Additionally, we found that the relationship between attributes that measure performance played an influential role in exploring performance level; the relationship between such attributes induced rules in measuring level of learners' performance.

Keywords: Hypermedia systems, Cognitive style

1. Introduction

A learner's performance is determined by their varying skills and abilities as well as various personal features such as gender, preferences and background knowledge of the course content. Such differences, known as "individual differences of learners", have been found to be important factors in the development of non-linear learning systems (Calcaterra, et al., 2005; Mitchell, et al., 2005). In a hypermedia system, (a non-linear learning system) learners are permitted to learn in their own way and to decide on their own paths through the material (Large, et al., 2002). In this way, they learn at their own pace and construct their understanding of subject matter actively (Chen & Macredie, 2002; Littlejohn, 2002).

¹ Corresponding author. Tel.: +44-1895-274000; fax: +44-1895-251686.
E-mail addresses: rana.alhajri@brunel.ac.uk (r.alhajri), steve.counsell@brunel.ac.uk (s.counsell), xiaohui.liu@brunel.ac.uk (x.liu).

In this paper, we used a WBI program to accommodate preferences of individual differences using mechanisms provided in Chen, et al. (2006) and Chen & Liu (2008) where individual differences such as learner's prior knowledge and cognitive styles, more specifically field dependent and field independent were considered. In particular, we group the WBI users into clusters based on their characteristics using three important attributes in measuring their performance: gain score is defined as post-test minus pre-test (g-score), total number of topics pages visited by the participants (t-pages) and total time, in seconds, that each participant spent visiting the topic pages in the WBI program (t-time). Hierarchical and K-Means clustering algorithms were used to explore different numbers of clusters to strengthen our results. Investigation will focus on the following three key aspects. Firstly, learners were pre-identified using the intersection of the three individual differences (by combining gender, cognitive style and prior knowledge when identifying a learner). Secondly, we investigated the impact of the behaviour of individual difference' intersection on learner performance. Thirdly, we explored the relationship between attributes used to measure learner's performance to induce rules for performance level.

The paper is structured as follows. Section 2 presents related work. Section 3 describes the methodology used to conduct our study and the techniques applied to the analysis of the corresponding data. The findings of our analyses are then discussed in Section 4. A data mining approach provides the basis of our analyses exploring the relationship of attributes that affect the performance of the individual. Finally, conclusions are drawn and possibilities for future work are identified in Section 5.

2. Background

Many studies argue that no single style will result in better performance. Thus, learners may have different backgrounds, especially in terms of their knowledge skills and needs, so may show various levels of engagement in course content (Wang, 2007). However, learners whose browsing behaviour was consistent with their own favoured styles obtained the best performance results (Calcaterra, et al., 2005; Mitchell, et al., 2005). The lack of studies investigating the performance of individual after interacting with WBI programs accommodating user preferences is noteworthy.

2.1. Individual Differences

Previous studies have demonstrated the importance of individual differences as a factor in the design of web-based instruction. Such features can have a significant effect on user learning in web-based instruction and may affect the way in which they learn from, and interact with, hypermedia systems. These range from cognitive styles (Kim, 2001; Workman, 2004) to prior knowledge (Calisir & Gurel, 2003; Hölscher & Strube, 2000; Mitchell, et al., 2005) to gender differences (Beckwith, et al., 2005; Roy, et al., 2003; Schumacher & Morahan-Martin, 2001).

Gender: Most studies indicate that gender is a significant variable in the learning process. This implies that males and females might need different levels of support when they interact with the Web. Some studies have found that males are more actively engaged in browsing than females because they tended to perform more page jumps per minute (Large, et al., 2002; Roy, et al., 2003). They suggested that males out-perform females in their ability to retrieve information from the Web since they are more experienced users; they formulated queries with fewer keywords, spent less time on individual pages, clicked more links per minutes than females and have more positive attitudes towards online technology in general.

Prior Knowledge: Learners with different levels of prior knowledge, from experts to novices, benefit differently from hypermedia learning systems (Calisir & Gurel, 2003; Wildemuth, 2004). Many studies argue that there are different levels of perception in using hypermedia learning systems requiring different ways to navigate (Calisir & Gurel, 2003; McDonald & Stevenson, 1998; Shin, et al., 1994).

Cognitive Styles: Field dependence and field independence are probably the most well-known division of cognitive styles. The differences between field-dependent and field-independent learners are:

Field independent learners have an impersonal behavior. They are not interested in others and show both physical and psychological distance from people. They tend not to need external referencing methods to process information and are capable of restructuring their knowledge and developing their own internal referencing methods. Thus, field independent learners are generally analytical in their approach.

Field dependent learners have interpersonal behavior in that they show strong interest in others and prefer to be physically close to people. They make greater use of external social influences for structuring their information. Field dependent learners are more attentive to social cues than field independent learners. Thus, field dependent learners are more global in their perceptions (Witkin, et al., 1977).

Hypermedia and Program Design Elements

Hypermedia provides a flexible approach which helps users to work with the information from different viewpoints. *Additional support* can be provided to help novices in hypermedia learning. Thus, graphical overviews and structural cues are powerful and beneficial in providing navigation guidance to novices to ease potential disorientation problems (Chen, et al., 2006). Moreover, field dependent users look at examples, while field-independent users frequently examine detailed descriptions (Chen & Liu, 2008). As for the *content structure*, findings in Chen et al. (2006) indicate that experts focused on locating detailed information while novices tended to get an overview only. A field independent user performs well in terms of analytical thought; they tend to focus on information and browse fewer pages to directly get to relevant topics for completing their tasks. On the other hand, field dependent users have global perceptions to process information. They tend to build an overall picture by browsing more pages (Goodenough, 1976). For field dependent students, a global picture of the subject can be assisted with pop-up windows. In this case, a pop-up window can be used to show additional topics for field-dependent students who would like to get a global picture of the subject content (Chen & Liu, 2008). Thus, information that is related to tasks is put in the main window containing the topic instructions for field independent and field dependent users, while further information is displayed with a pop-up window for field dependent users.

As for *Navigation tools*, Chen et al. (2006) showed that index tools were helpful for experts. On the other hand, map and menu tools were beneficial for novice learners in hypermedia learning systems. Moreover, in the study of Lin and Chen (2008), 101 individuals were examined and their cognitive styles identified by the Cognitive Styles Analysis (CSA) by Riding (1991). Results showed that field independent individuals favoured an alphabetical index and a search engine whereas field dependent individuals preferred to use a map to build the entire perceptual fields. Thus field independent users often prefer an alphabetical index to locate specific

information, whereas field dependent users often use a hierarchical map to get a global picture of the subject content (Chen & Liu, 2008; Chen, 2010; Farrell & Moore, 2001; Chen & Macredie, 2010).

Lee et al. (2009) investigated the relationships between cognitive styles and users' learning behaviour in web-based learning programs. They found that a cognitive style, more specifically field dependent and field independent, could be reached by some rules. "These rules can be applied to replace the CSA or other cognitive style tests and work as criteria for automatic identification of the students' cognitive styles" (Lee, et al., 2009). In other words, they found that field independent learners prefer non-linear navigation which we provide it by index approach because field independent learners are tend to be more analytical (Ford , et al., 1994) and they are very task oriented (Witkin, et al., 1977). Such a finding is in line with those of Lee et al. (2005). On the other hand, they found that field dependent learners prefer linear presentation of learning material and have difficulties in non-linear learning which we provide it by hierarchical map approach. Thus, field-dependent learners often use the hierarchical map to illustrate the relationships among different concepts (Turns, et al., 2000), which reflects the conceptual structure of the subject content (Nilsson & Mayer, 2002)

Measuring Learners' Performance in Existing Studies

McDonald and Stevenson (1998) measured navigation performance in terms of speed and accuracy in answering questions and locating particular nodes. Results showed that the performance of knowledgeable participants was better than that of non-knowledgeable participants. Additionally, in Mitchell et al. (2005), performance was measured by a *gain score* (henceforward 'g-score'), calculated as post-test score minus pre-test score. Those subjects that performed poorly on the pre-test made a greater improvement in the post-test.

The study of Kim (2001) investigated how differences in cognitive style and online search experience influenced the search. They used the time spent for retrieving information and the number of nodes visited for retrieving information as two different indicators for measuring search performance.

As for the gain score, results indicate that novices show a greater improvement in their learning performance than experts. More specifically, those who performed poorly on the pre-test make a greater improvement on the post-test. As for number of visited pages in the WBI programs, studies have found that male, field dependent, experts browse more pages than female, field independent, novices (Chen & Liu, 2008; Ford & Chen, 2000; Large, et al., 2002; Roy, et al., 2003). As for time spent in browsing the WBI programs, some studies have found that male, field independent users spent less time than female field dependent ones (Chen & Liu, 2008; Lee, et al., 2009; Roy, et al., 2003).

The literature on the effects of hypermedia systems on user performance focuses extensively on measurement attributes such as time spent using the system by a user, g-score and number of pages visited in the system. However, there is a dearth of studies which explore the relationship between such attributes in measuring performance level. There is also a lack of studies demonstrating the influence of the behaviour of individual differences' intersection on their performance using such measurements and after interacting with a WBI system.

In this paper, we attempt to answer the following research questions:

RQ1: "What are the relationships between the attributes values in measuring the performance level of the individual differences?"

RQ2: "How does the behaviour of individual differences influence learner's performance using three performance measurement attributes?"

2.2. Data Mining

Data mining is the process of discovering interesting, unexpected or valuable information from large datasets (Hand, 2007). It uses data to find unexpected relationships and patterns (Wang, et al., 2002). By doing so, hidden relationships and interdependencies can be discovered and predictive rules generated (Hedberg, 1995; Gargano & Raggad, 1999).

Data mining can be divided into clustering, classification and association rules (Witten, et al., 2011). Among these three approaches, clustering is selected for analyzing data in our study because it can form groups that share similar characteristics where each group consists of objects that are similar amongst themselves and dissimilar to objects of other groups (Roussinov & Zhao, 2003).

Clustering methods may be grouped into the following two categories: hierarchical and non-hierarchical clustering (Jain & Dubes, 1999). A hierarchical clustering procedure involves the construction of a hierarchy or tree-like structure, a nested sequence of partitions (Fraley & Raftery, 1998), while non-hierarchical or partitioned procedures end with a particular number of clusters at a single step. Commonly used non-hierarchical clustering algorithms include the K-means algorithm, graph-theoretic approaches via the use of minimum spanning trees, evolutionary clustering algorithms, simulated annealing based methods as well as competitive neural networks such as Self-Organizing Maps. In this paper, we have used both hierarchical clustering and the widely used non-hierarchical clustering method, K-means, to group users into clusters based on their characteristic in measuring their performance using three attributes, ‘g-score’, ‘t-pages’ and ‘t-time’.

3. Methodology Design

In this paper, we used quantitative data obtained from the pre-test, post-test and the log file of the WBI program for our analysis. We define our attribute values: a) gain scores (g-score), b) pages visited in WBI program (t-pages) and c) time spent in browsing the WBI program (t-time) as performance attributes and the independent variables as: a) cognitive style (field dependent and field independent), b) prior knowledge (expert and novices) and c) gender (male and female).

3.1. Research Instruments

The WBI program presents an introduction to PowerPoint and provides participants with links within the text, navigation tools, including a hierarchical map and an alphabetical index (Figure 1). Chen and Liu (2008) state that: “field-dependent students rely more heavily on external cues, thus, they prefer to get concrete guidance from examples. One of the possible ways to address their different needs is to show both of the display options, detailed description and concrete examples, within a table. By using a table, all of the relevant information about a particular case can put together in one place. For example, one column can be used to present the detailed descriptions of a particular topic, while the other column provides the illustration with examples for that topic”. In our WBI, each topic will be presented in two display options, description details and illustrated examples (Chen, et al., 2006). Figure 2 shows the design of a topic page presenting the same structure (description and examples).

There are two types of overview, a general overall picture and specific information on a topic. Specific information is displayed as a pop-up window in our WBI program named ‘further details page’. In this way, learners are given control of deciding their own learning paths and choosing their favored navigation tools and display options. Additionally, we logged the display of each page when a participant clicked on any link in the WBI program; the chosen hyperlink was either from an index or from a map frame. The students were handed out a set of tasks to complete on PowerPoint while utilizing the WBI. The tasks sheets contained different main tasks used to cover the questions that are provided in the pre-test and post-test. All of their interactions with the WBI were logged by the system.

A pre-test was given to the participants to identify their prior knowledge of using PowerPoint in order to decide whether they were novices or experts. A post-test was provided to measure the g-score of each participant (post-test score minus pre-test score). The pre-test and post-test included 20 multiple-choice questions, each with four different answers and an "I don't know" option, from which the students chose one response. The questions were matched on the pre-test and post-test so that each question on the pre-test had a similar (but not the same) question on the post-test. Creating similar questions on the post-test was achieved by re-phrasing the question. For example, a question in the pre-test was stated as: "The following are some views of the PowerPoint interface: a) Normal, Show, Action, b) Sorter, Show, design, c) Normal, Sorter, Show, d) None of the above and e) I don't know"; the similar question in the post-test was "Normal and Sorter can be known as a: a) Slide Design, b) Slide View, c) Slide Layout, d) None of the above and e) I don't know"). For each of the 20 questions, subjects received 1 mark towards their score if the answer was correct. If the answer was wrong or the "I don't know" option was chosen, the participant received 0.

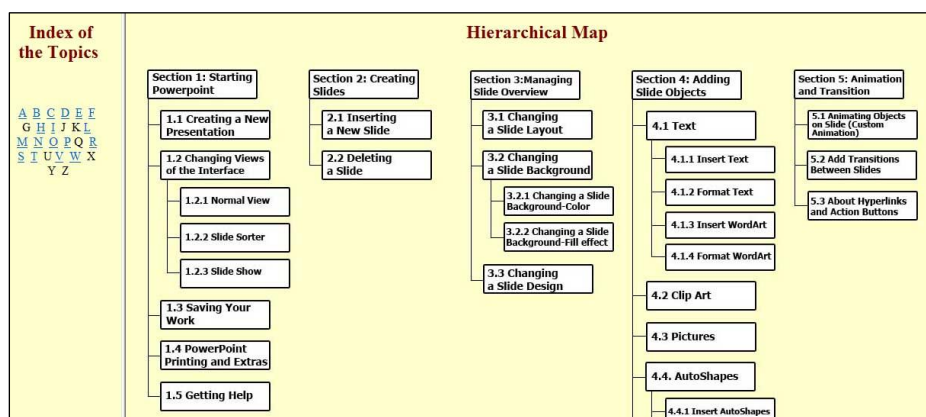


Figure 1: The main page of the WBI program.

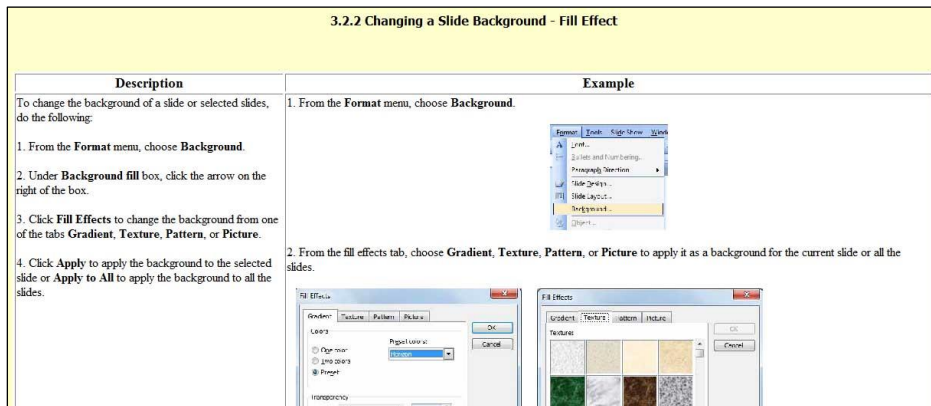


Figure 2: Popup window displaying the topic contents in details description and example.

3.2. Participants

We conducted the experiment at the Higher Institute of Telecommunication and Navigation (HITN) in Kuwait. The total number of participants was 91 and their ages ranged between 18 and 25 years. Participants had different computing and internet skills and were classified in terms of cognitive style, gender, and prior knowledge. Firstly, males and females were placed in different groups. Secondly, in keeping with findings from previous studies the field independent learner favored using the index. Conversely, field dependent learner preferred to use the map (Chen & Liu, 2008; Chen & Macredie, 2002; Ford & Chen, 2000). We used these findings to identify the field dependent and field independent using our WBI program. This was done by inspecting the log file of each participant; we calculated the number of Map and Index pages that each user had visited. Using a hierarchical clustering test, to identify learners as field dependent and field independent learners, we found that if the number of map pages was more than 50% of the pages they had visited, the participant was identified as field dependent. If the number of index pages was greater than 50%, the participant was identified as a field independent user. We chose the 50% to show which were the most pages visited by each learner (Map or Index pages).

Finally, for prior knowledge, we calculated the mean of the pre-test scores of all participants. The calculated mean was 8.5 (SD=3.45) out of a possible 20. If the participant's score in the pre-test was less than or equal to this mean value, the participant was identified as a novice. If the score was greater than the mean, the participant was identified as an expert. Table 1 shows the number of participants after identifying them in their appropriate individual difference groups.

Individual differences classes	Cognitive style		Gender		Prior knowledge	
	FD	FI	Male	Female	Experts	Novices
Number of participants	51	40	45	46	47	44

Table 1: Number of participants in each class; FD: Field dependent, FI: Field independent

3.3. Procedures

The experiment consisted of four phases. Firstly, participants were asked to practice 30 minutes using the PowerPoint application; this helped to refresh their prior knowledge of PowerPoint. However, 3 hours was given to the participants in working on the whole experiment including pretest, task, posttest and survey. Pre-test and post-test values were measured at the beginning and at the end of the experiment. Participants individually completed a pre-test and a post-test and no time limit was set for the completion of these tests. Secondly, a pre-test was introduced to each participant to identify their prior knowledge of subject content and to clarify whether they were novices or experts. The results of the pre-test were later used in the final phase. Thirdly, all participants were given an introduction to the use of the WBI programs. They were then asked to spend 2 hours (maximum) interacting with the WBI program using a task. In this way, participants were free to choose preferred navigation tools, display options and content structure by themselves; their interactions with the WBI were stored in a log file. The log file recorded participant movement and registered visited pages as well as the time they spent visiting such pages. The time was used to calculate the time spent on each page to examine how the time and number of page attributes influenced their performance. Each of the 20 questions provided in the pre-test and the post-test was covered as a specific task in the task sheet to ensure that the participants practiced such a topic before the post-test. Finally, in this phase, a post-test was given to the participants to check whether they performed positively by achieving a higher g-score in the post-test than those in the pre-test.

4. Results and Discussion

In our study, we compared the results of two clustering algorithms, namely K-Means and Hierarchical clustering algorithms. These algorithms were used to study the behaviour of three individual differences: gender, prior knowledge and cognitive style. The cases of individual differences and their intersection defined in our study are as follows:

1. FFIE: Female who is known as Field Independent and Expert learner.
2. FFIN: Female who is known as Field Independent and Novice learner.
3. FFDE: Female who is known as Field Dependent and Expert learner.
4. FFDN: Female who is known as Field Dependent and Novice learner.
5. MFIE: Male who is known as Field Independent and Expert learner.
6. MFIN: Male who is known as Field Independent and Novice learner.
7. MFDE: Male who is known as Field Dependent and Expert learner.
8. MFDN: Male who is known as Field Dependent and Novice learner.

Table 2 shows the cases of individual differences and number of participants and the percentage for each case in our study.

	FFIE	FFIN	FFDE	FFDN	MFIE	MFIN	MFD E	MFD N	Total
Frequencies	8	3	17	18	14	15	9	7	91
Percentage	8.79	3.30	18.68	19.78	15.38	16.48	9.89	7.69	100

Table 2: Intersection of individual differences' frequencies

An ANOVA test was computed to explore any significance between individual differences as an independent variable with performance measurements attributes (g-score, t-pages and t-time) as dependent variables. We found significant differences at the 5% level for the g-score value. However, there were no significant differences with the t-pages and t-time attributes (significance was greater than 5%). Thus, g-score will be used as the first measuring attribute to compare between learner's performance levels.

For each participant interacting with the WBI program, we used three attribute values; 'g-score' (post-test minus pre-test, where mean of pre-test of the 91 participants is 8.5 and mean of post-test is 11.30), 't-pages' and 't-time'.

Table 3 shows the overall mean values for each attribute. These overall mean values are calculated by using the attribute values for each participant.

		g-score	t-pages	t-time
N	Valid	91	91	91
	Missing	0	0	0
Std. Deviation		2.785	7.268	1105.759
Mean		2.77	15.35	2015.36

Table 3: Overall mean values of attribute using for clustering

4.1. Results of Four Clusters

In this section, we study the clustering algorithms using four clusters (C1, C2, C3, C4) starting with the K-Means algorithm in Analysis One and the Hierarchical algorithm in Analysis Two. The mean values of g-score, t-pages and t-time of the clusters will be used to study the characteristics of each cluster by comparing these values with the overall mean values shown in Table 3.

Analysis One:

In Analysis One, we began with the K-Means algorithm using K=4; the attributes that we used in this algorithm are shown in Table 3. Additionally, we labeled the cases in the used algorithm of each one of the individual differences as shown in Table 2.

Table 4 shows that the highest number of individual differences was located in C1 (N=37). The lowest number was located in C2 (N=4). In C1, all the individual differences were allocated into this cluster, where the highest number of individual differences is in the MFIN category. In C2, we see that FFIN, FFDE and FFDN are allocated into this cluster where the highest number of individual differences in C2 is FFDN.

		Cases of individual differences							Total	
		FFIE	FFIN	FFDE	FFDN	MFIE	MFIN	MFDE		MFDN
Cluster Number	C1	1	1	5	8	6	11	4	1	37
	C2	0	1	1	2	0	0	0	0	4
	C3	3	0	7	5	4	0	2	5	26
	C4	4	1	4	3	4	4	3	1	24
Total		8	3	17	18	14	15	9	7	91

Table 4: Cluster distribution of individual differences of K-Means algorithm (4 clusters)

Table 5 shows the K-Means clustering results. We used these attribute values to compare the mean values in each cluster (using the words High and Low) with the overall mean value of all participants; the overall mean value is given in the last row of Table 5 ('Total' row). From this comparison, we found the following:

Cluster Number		g-score		t-pages		t-time	
C1	Mean	3.05	High	15.38	Slightly	1810.30	Low
	N	37		37	High	37	
	Std. Dev	2.79		5.88		266.42	
C2	Mean	4.00	High	23.00	High	5130.75	High
	N	4		4		4	
	Std. Dev	4.97		14.85		658.40	
C3	Mean	2.08	Low	17.73	High	2954.27	High
	N	26		26		26	
	Std. Dev	2.48		5.76		409.93	
C4	Mean	2.88	High	11.46	Low	795.13	Low
	N	24		24		24	
	Std. Dev	2.71		7.44		358.28	
Total (overall values)	Mean	2.77		15.35		2015.36	
	N	91		91		91	
	Std. Dev	2.79		7.27		1105.76	

Table 5: Cluster Centroids of K-Means algorithm (4 clusters)

1. Participants allocated into clusters C1, C2 and C4 had a higher g-score than the overall mean. Those allocated into C3 had a lower g-score than the overall mean value.
2. Participants allocated into clusters C1, C2 and C3 visited more t-pages than the overall mean value; however, C1 was slightly higher. Those allocated into C4 visited fewer t-pages than the overall mean value.
3. Participants allocated into clusters C2 and C3 spent a higher t-time than the overall mean value. Those who were allocated into clusters C1 and C4 spent less t-time than the overall mean value on visiting topic pages (t-pages).

Analysis Two:

In this analysis, we used a Hierarchical Clustering Algorithm. We set the number of clusters = 4. The used attributes are shown in Table 3 and we labeled cases in the used algorithm of each one of the individual differences as shown in Table 2.

Table 6 shows that the highest number of individual differences was located in C1 (N=54). On the other hand, the lowest number was located in C4 (N=3). In C1, all the individual differences were allocated into this cluster, where the highest number of individual differences is MFIN. In C4, we can see that FFIN, FFDE and FF DN are allocated into this cluster (one participant for each of the individual differences).

		Cases of individual differences							Total	
		FFIE	FFIN	FFDE	FFDN	MFIE	MFIN	MFDE		MFDN
Cluster Number	C1	3	2	7	10	10	15	5	2	54
	C2	3	0	7	6	4	0	2	5	27
	C3	2	0	2	1	0	0	2	0	7
	C4	0	1	1	1	0	0	0	0	3
Total		8	3	17	18	14	15	9	7	91

Table 6: Cluster distribution of individual differences of Hierarchical algorithm (4 clusters)

From Table 7, we see the report of the Hierarchical clustering results. We used these attribute values to compare the mean values in each cluster (using the words High and Low) with the overall mean value of all participants; the overall mean value is given in the last row of Table 7 (Total row). From this comparison, we found the following:

1. Participants allocated into clusters C1 and C4 had a higher g-score than the overall mean. Those allocated into C2 and C3 had a lower g-score than the overall mean value.
2. Participants allocated into clusters C2 and C4 visited more t-pages than the overall mean value. Those allocated into C1 and C3 visited fewer t-pages than the overall mean value.
3. Participants allocated into clusters C2 and C4 spent higher t-time than the overall mean value. Those allocated into clusters C1 and C3 spent less t-time than the overall mean value.

Cluster Number		g-score		t-pages		t-time	
C1	Mean	3.11	High	14.61	Low	1551.56	Low
	N	54		54		54	
	Std. Dev	2.81		6.28		454.72	
C2	Mean	2.07	Low	17.59	High	3006.11	High
	N	27		27		27	
	Std. Dev	2.43		5.69		483.88	
C3	Mean	2.00	Low	7.86	Low	325.71	Low
	N	7		7		7	
	Std. Dev	2.00		7.84		199.76	
C4	Mean	4.67	High	26.00	High	5389.67	High
	N	3		3		3	
	Std. Dev	5.86		16.64		498.00	
Total (overall values)	Mean	2.77		15.35		2015.36	
	N	91		91		91	
	Std. Dev	2.79		7.27		1105.76	

Table 7: Cluster Centroids of Hierarchical algorithm (4 clusters)

Discussion of Analysis One and Analysis Two:

The bar charts in Figure 3 show the comparison between the four clusters with the three attribute values (t-pages, t-time, and g-score) using the K-Means algorithm. From these charts and the results of Analysis One, we conclude that participants who achieved a higher g-score after they visited fewer t-pages and spent less t-time in visiting these pages, demonstrate better performance, as shown from the results of C1 and C4, although the t-pages value of C1 is similar to the overall mean value.

We can also conclude that those participants who had achieved a lower g-score after they visited more t-pages and spent higher t-time in visiting these pages did not perform well, as shown from the results of C3. We ignored cluster C2 because it has a low number of participants (4 out of 91 participants); the majority of participants are located in C1, C3 and C4 (numbering 87).

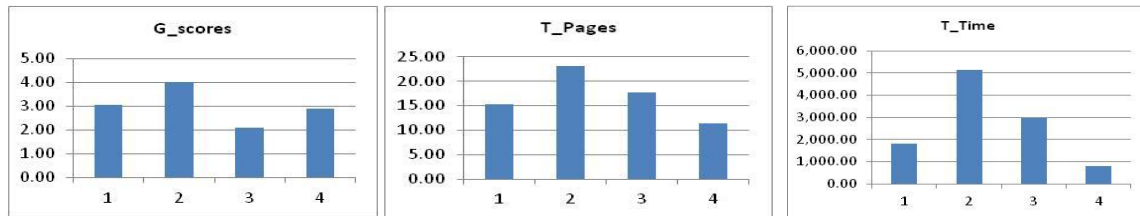


Figure 3: K-Means algorithm for four clusters and the three attribute values

The bar charts in Figure 4 show the comparison between the four clusters with the three attribute values, t-pages, t-time and g-score. From these charts and the results of Analysis Two, when comparing C1 with C2, we conclude that participants, who had achieved higher g-score after they visited fewer t-pages and spent less t-time in visiting these pages, improved their performance (C1).

We can also conclude that those participants who had achieved lower g-score after they visited more t-pages and spent higher t-time in visiting these pages did not perform well (C2). We ignored clusters C3 and C4 because of their low number of participants (C3=7 and C4=3); the majority number of participants were located in C1 and C2 (81 out of 91 participants).

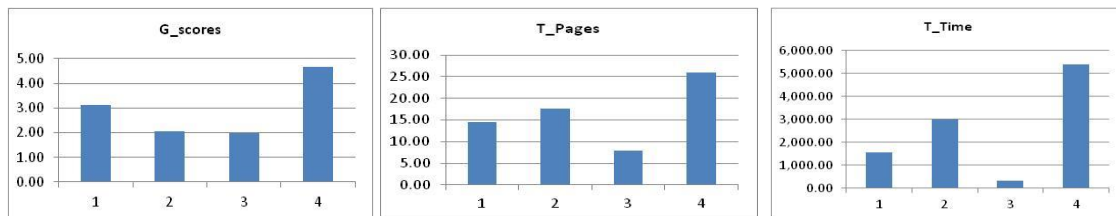


Figure 4: Hierarchical algorithm for four clusters and the three attribute values

4.2. Results of Five Clusters

In this section, we study the clustering algorithms using five clusters (C1, C2, C3, C4, C5) starting with the K-Means algorithm in Analysis Three and the Hierarchical algorithm in Analysis Four. The mean values of g-score, t-pages and t-time of the clusters will be used to study the characteristics of each cluster by comparing these values with the overall mean values shown in Table 3.

Analysis Three:

In this analysis, we started with the K-Means algorithm using K=5. The attributes used in this algorithm are shown in Table 3. Additionally, we labeled the cases in the algorithm we used for each of the individual differences (Table 2).

From Table 8, results show that the highest number of individual differences was located in C5 (N=37). The lowest number was located in C1 (N=3). In C5, all individual differences were allocated into this cluster, where the highest number of individual differences is MFIN. In C1, we see that FFIN, FFDE and FF DN are allocated into this cluster (one participant for each of the individual differences).

		Cases of individual differences								Total
		FFIE	FFIN	FFDE	FFDN	MFIE	MFIN	MFDE	MFDN	
Cluster Number	1	0	1	1	1	0	0	0	0	3
	2	4	1	4	3	4	3	3	1	23
	3	1	0	2	3	1	0	0	2	9
	4	2	0	5	3	3	1	2	3	19
	5	1	1	5	8	6	11	4	1	37
Total		8	3	17	18	14	15	9	7	91

Table 8: Cluster distribution of individual differences of K-Means algorithm (5 clusters)

Table 9 presents the K-Means clustering results. We used these attribute values to compare the mean values in each cluster (using the words High and Low) with the overall mean value of all participants; the overall mean value is given in the last row of Table 9 (Total row). From this comparison, we conclude the following:

Cluster Number		g-score		t-pages		t-time	
C1	Mean	4.67	High	26.00	High	5389.67	High
	N	3		3		3	
	Std. Dev	5.86		16.64		498.00	
C2	Mean	2.70	Slightly Low	10.96	Low	773.65	Low
	N	23		23		23	
	Std. Dev	2.62		7.18		350.18	
C3	Mean	2.11	Low	16.22	High	3569.33	High
	N	9		9		9	
	Std. Dev	1.76		6.96		387.36	
C4	Mean	2.26	Low	18.05	High	2702.26	High
	N	19		19		19	
	Std. Dev	2.83		4.98		207.36	
C5	Mean	3.08	High	15.62	High	1782.92	Low
	N	37		37		37	
	Std. Dev	2.82		6.01		266.53	
Total (overall values)	Mean	2.77		15.35		2015.36	
	N	91		91		91	
	Std. Dev	2.79		7.27		1105.76	

Table 9: Cluster Centroids of K-Means algorithm (5 clusters)

1. Participants allocated into clusters C1 and C5 achieved a higher g-score than the overall mean. Those allocated into C3 and C4 had a lower g-score than the overall mean value. However, C2 is slightly lower.
2. Participants allocated into clusters C1, C3, C4 and C5 had visited more t-pages than the overall mean value. Those allocated into C2 had visited fewer t-pages than the overall mean value.
3. Participants allocated into clusters C1, C3 and C4 spent a higher t-time in visiting these pages than the overall mean value. Those who were allocated into clusters C2 and C5 had less t-time in visiting pages than the overall mean value.

Analysis Four:

We set the number of clusters = 5; the attributes used are shown in Table 3 and we labeled the cases in the algorithm we used of each of the individual differences as per Table 2. From Table 10, results show that the highest number of individual differences was located in C1 (N=54). The lowest number was located in C4 (N=3). In C1, all the individual differences were allocated into this cluster, where the highest number of individual differences is for the MFIN category. In C4, we can see that FFIN, FFDE and FFDN are allocated into this cluster (one participant for each of the individual differences).

		Cases of individual differences								Total
		FFIE	FFIN	FFDE	FFDN	MFIE	MFIN	MFDE	MFDN	
Cluster Number	C1	3	2	7	10	10	15	5	2	54
	C2	3	0	5	4	4	0	2	4	22
	C3	2	0	2	1	0	0	2	0	7
	C4	0	1	1	1	0	0	0	0	3
	C5	0	0	2	2	0	0	0	1	5
Total		8	3	17	18	14	15	9	7	91

Table 10: Cluster distribution of individual differences of Hierarchical algorithm (5 clusters)

We used Table 11 values to compare the mean values in each cluster (again using the words High and Low) with the overall mean value of all participants; similarly, the overall mean value is given in the final row of Table 11 (Total row). From this comparison, we conclude the following:

1. Participants allocated into clusters C1 and C4 achieved a higher g-score than the overall mean. Those allocated into C2, C3 and C5 had a lower g-score than the overall mean value.
2. Participants allocated into clusters C2 and C4 had more value for t-pages than the overall mean value. Those allocated into C1, C3 and C5 had fewer values of t-pages than the overall mean value.

Cluster Number		g-score		t-pages		t-time	
C1	Mean	3.11	High	14.61	Low	1551.56	Low
	N	54		54		54	
	Std. Dev	2.81		6.28		454.72	
C2	Mean	2.09	Low	18.77	High	2820.73	High
	N	22		22		22	
	Std. Dev	2.58		5.58		270.21	
C3	Mean	2.00	Low	7.86	Low	325.71	Low
	N	7		7		7	
	Std. Dev	2.00		7.84		199.76	
C4	Mean	4.67	High	26.00	High	5389.67	High
	N	3		3		3	
	Std. Dev	5.86		16.64		498.00	
C5	Mean	2.00	Low	12.40	Low	3821.80	High
	N	5		5		5	
	Std. Dev	1.87		2.41		343.36	
Total (overall values)	Mean	2.77		15.35		2015.36	
	N	91		91		91	
	Std. Dev	2.79		7.27		1105.76	

Table 11: Cluster Centroids of Hierarchical algorithm (5 clusters)

- Participants allocated into clusters C2, C4 and C5 had a higher t-time in visiting pages than the overall mean value. Those who were allocated into clusters C1 and C3 had less t-time in visiting pages than the overall mean value.

Discussion of Analysis Three and Analysis Four:

The bar charts in Figure 5 show the comparison between the five clusters with the three attribute values, t-pages, t-time and g-score using the K-Means algorithm. From these charts and the results of Analysis Three, we conclude that participants, who had achieved higher g-score after they visited fewer t-pages and spent less t-time in visiting these pages, improved their performance, as shown from the results of C2 and C5, although the t-pages value of C5 is equal to the overall mean value. We can also conclude that those participants, who had achieved lower g-score after they visited more t-pages and spent higher t-time in visiting these pages did not perform well, as shown from the results of C3 and C4. We ignore the cluster C1 because it has a low number of participants (3 out of 91 participants); where the majority of participants are located in C2, C3, C4 and C5 (88).

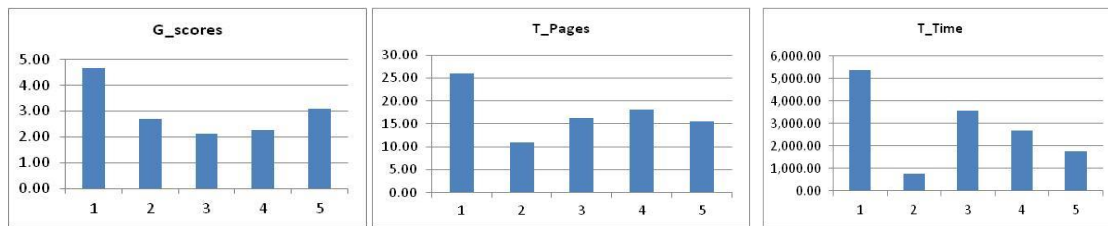


Figure 5: K-Means algorithm for four clusters and the three attribute values

The bar charts in Figure 6 show a comparison between the five clusters with the three attribute values, t-pages, t-time and g-score using the Hierarchical algorithm. From these charts and the results of Analysis Four, by comparing C1 with C2, we conclude that participants who had achieved higher g-score after they visited fewer t-pages and spent less t-time in visiting these pages performed better. We can also conclude that those participants who had achieved lower g-score after they visited more t-pages and spent higher t-time in visiting these pages did not perform well. We ignore clusters C3, C4 and C5 because of their low number of participants (C3=7 and C4=3, C5=5); the majority number of participants are located in C1 and C2 (76 out of 91 participants).

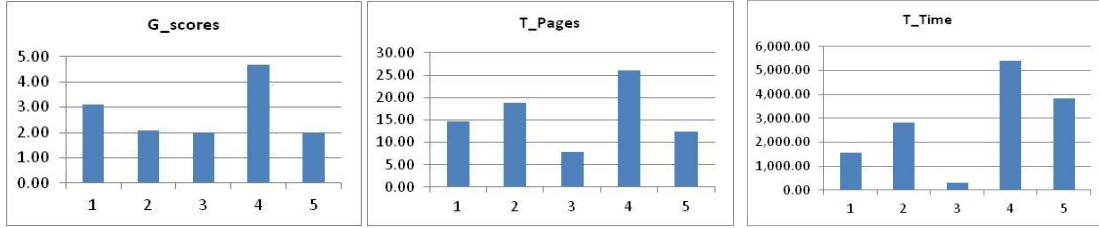


Figure 6: Hierarchical algorithm for four clusters and the three attribute values

4.3. Discussion

From the previous discussion, we can conclude that the g-score, t-pages and t-time attributes had a great effect on measuring the performance level of the individual difference intersection. Additionally, there is a significant relationship between such attributes. These relationships can be encapsulated in the following rules:

1. Rule 1: this rule was established for the relationship “Participants who achieve higher g-score after visiting fewer t-pages and spending less t-time in visiting these pages compared to the global mean value”. This relationship considered a learner’s best performance. Thus, the best performance is if learner achieved a higher g-score than the mean of all the participants’ attribute values after they spent lower time browsing the WBI pages and visited more pages than the mean of all the participants’ attribute values.
2. Rule 2: this rule was established for the relationship “The participant who achieves a lower g-score after visiting more t-pages and spending higher t-time in visiting these pages compared to the global mean value”. This relationship considered a learner’s worst performance. Thus, the worst performance is if learner achieved lower g-score than the overall mean value, after they spent higher time browsing the WBI pages, and visited more pages than the mean of all the participants’ attribute values.

To investigate the performance level (High/Low) of the individual differences intersection using Rule 1 and Rule 2, we compared the means of the attribute values (t-pages, t-time and g-score) of the individual difference intersection in each of our four analyses with the mean values of the three attributes shown in Table 3 (t-pages = 15.35, t-time = 2015.36 and g-score = 2.77).

Analyses	Clusters	Individual differences	g-score	t-pages	t-time
Analysis One	C1	FFDN	3.63	12.75	1,941.38
		MFIE	2.67	13.00	1,810.50
	C4	FFIN	5.00	1.00	804.00
		FFDN	3.33	6.00	676.67
Analysis Two	C1	FFIN	3.00	11.00	1,407.50
		FFDN	3.60	11.40	1,710.50
		MFIE	3.10	14.10	1,520.10
Analysis Three	C2	FFIN	5.00	1.00	804.00
		FFDN	3.33	6.00	676.67
	C5	FFDN	3.63	12.75	1,941.38
Analysis Four	C1	FFIN	3.000	11.000	1,407.500
		FFDN	3.60	11.40	1,710.50
		MFIE	3.10	14.10	1,520.10

Table 12: Comparison means of individual difference intersection in clusters of our analyses (high performance)

1. According to Rule 2, we found that some of the individual difference intersection, those who had shown in the four analyses that they had low performance level are female-field dependent-expert (FFDE), male-field independent-expert (MFIE) and male-field dependent-expert (MFDE). These findings are shown in Table 13.

We conclude that the intersection of individual differences had a great effect on the performance of the learner.

Analyses	Clusters	Individual differences	g-score	t-pages	t-time
Analysis One	C3	FFDE	1.00	16.00	2,974.14
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00
Analysis Two	C2	FFDE	1.00	16.00	2,974.14
		FFDN	2.67	18.33	3,421.83
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00
Analysis Three	C3	FFIE	1.00	17.00	3,225.00
		MFIE	2.00	19.00	3,180.00
	C4	FFDE	1.00	17.60	2,716.60

		MFIE	1.67	22.33	2,733.00
		MFDE	0.50	16.50	2,489.00
Analysis Four	C2	FFDE	1.00	17.60	2,716.60
		FFDN	2.25	20.75	3,051.25
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00

and Table 13 show the results of performance level in our four analyses. We can conclude the following:

2. According to Rule 1, we observe that, of the individual difference intersection, those who had shown in our four analyses that they had high performance level are female-field dependent-novice (FFDN). These findings are shown in

Analyses	Clusters	Individual differences	g-score	t-pages	t-time
Analysis One	C1	FFDN	3.63	12.75	1,941.38
		MFIE	2.67	13.00	1,810.50
	C4	FFIN	5.00	1.00	804.00
		FFDN	3.33	6.00	676.67
Analysis Two	C1	FFIN	3.00	11.00	1,407.50
		FFDN	3.60	11.40	1,710.50
		MFIE	3.10	14.10	1,520.10
Analysis Three	C2	FFIN	5.00	1.00	804.00
		FFDN	3.33	6.00	676.67
	C5	FFDN	3.63	12.75	1,941.38
Analysis Four	C1	FFIN	3.000	11.000	1,407.500
		FFDN	3.60	11.40	1,710.50
		MFIE	3.10	14.10	1,520.10

Table 12: Comparison means of individual difference intersection in clusters of our analyses (high performance)

3. According to Rule 2, we found that some of the individual difference intersection, those who had shown in the four analyses that they had low performance level are female-field dependent-expert (FFDE), male-field independent-expert (MFIE) and male-field dependent-expert (MFDE). These findings are shown in Table 13.

We conclude that the intersection of individual differences had a great effect on the performance of the learner.

Analyses	Clusters	Individual differences	g-score	t-pages	t-time
Analysis One	C3	FFDE	1.00	16.00	2,974.14
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00
Analysis Two	C2	FFDE	1.00	16.00	2,974.14
		FFDN	2.67	18.33	3,421.83
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00
Analysis Three	C3	FFIE	1.00	17.00	3,225.00
		MFIE	2.00	19.00	3,180.00
	C4	FFDE	1.00	17.60	2,716.60
		MFIE	1.67	22.33	2,733.00
		MFDE	0.50	16.50	2,489.00
Analysis Four	C2	FFDE	1.00	17.60	2,716.60
		FFDN	2.25	20.75	3,051.25
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00

4. .

Analyses	Clusters	Individual differences	g-score	t-pages	t-time
Analysis One	C1	FFDN	3.63	12.75	1,941.38
		MFIE	2.67	13.00	1,810.50
	C4	FFIN	5.00	1.00	804.00
		FFDN	3.33	6.00	676.67
Analysis Two	C1	FFIN	3.00	11.00	1,407.50
		FFDN	3.60	11.40	1,710.50
		MFIE	3.10	14.10	1,520.10
Analysis Three	C2	FFIN	5.00	1.00	804.00
		FFDN	3.33	6.00	676.67
	C5	FFDN	3.63	12.75	1,941.38
Analysis Four	C1	FFIN	3.000	11.000	1,407.500
		FFDN	3.60	11.40	1,710.50
		MFIE	3.10	14.10	1,520.10

Table 12: Comparison means of individual difference intersection in clusters of our analyses (high performance)

- According to Rule 2, we found that some of the individual difference intersection, those who had shown in the four analyses that they had low performance level are female-field dependent-expert (FFDE), male-field independent-expert (MFIE) and male-field dependent-expert (MFDE). These findings are shown in Table 13.

We conclude that the intersection of individual differences had a great effect on the performance of the learner.

Analyses	Clusters	Individual differences	g-score	t-pages	t-time
Analysis One	C3	FFDE	1.00	16.00	2,974.14
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00
Analysis Two	C2	FFDE	1.00	16.00	2,974.14
		FFDN	2.67	18.33	3,421.83
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00
Analysis Three	C3	FFIE	1.00	17.00	3,225.00
		MFIE	2.00	19.00	3,180.00
	C4	FFDE	1.00	17.60	2,716.60
		MFIE	1.67	22.33	2,733.00
		MFDE	0.50	16.50	2,489.00
Analysis Four	C2	FFDE	1.00	17.60	2,716.60
		FFDN	2.25	20.75	3,051.25
		MFIE	1.75	21.50	2,844.75
		MFDE	0.50	16.50	2,489.00

Table 13: Comparison of means of individual differences' intersection in clusters of our analyses (low performance)

5. Conclusions

There has been a notable lack of studies investigating the performance of different individual differences after interacting with WBI programs accommodating user preferences. In this paper, we used a WBI program which accommodated preferences of individual differences such as learner's prior knowledge, gender and cognitive styles. In particular, we make advances in grouping the WBI users into clusters based on three important attributes using both hierarchical clustering and K-Means algorithms. Our investigation has been focused on three key aspects.

Firstly, learners were defined using the intersection of the three individual differences (gender, cognitive style and prior knowledge). The concern of this intersection is to identify each learner by each of such individual differences (e.g., MFIE known as Male who is Field Independent and an Expert learner). Secondly, we investigated the impact of individual difference intersection on learner performance; learners were pre-identified using the intersection of three individual differences to understand the impact of the individual differences on learners' performance. Thirdly, the combined performance measurement attributes give a better understanding of how learners performed. In this paper, we explored the relationship between attributes that had been used to measure learner's performance in order to induce rules for performance level.

Results showed that attributes relationships had an impact on measuring learners' performance level. Those learners were defined using the intersection of the three individual differences. Additionally, a suggested relationship of such attributes was provided for optimal performance. The results obtained using clustering were compared to investigate the attributes' relationships that explore the performance level. The first research question related to "What are the relationships between the attributes values in measuring the performance level of the individual differences' intersection?" We demonstrated that the relationship of the three attributes had a significant effect on the performance level of the individual differences' intersection. Moreover, we demonstrated two different rules in measuring the optimal and the worst level of performance. We also found that the intersection of the individual differences female-field independent-novice (FFIN) and the intersection female-field independent-expert (FFIE) had the best performance, whereas the intersection of the individual differences male-field independent-expert (MFIE) had the worst performance.

The second research question was "How the behavior of individual differences' intersection influenced learner's performance using three performance measurement attributes?" we found that Learners achieve optimal performance when they gain a higher score (post-test minus pre-test) after spending lower time browsing the WBI program, and browsing fewer pages compared to the overall mean values of the all learners for each of such attributes. Learners exhibit worst performance when they gain a lower score after they spent more time browsing the WBI program and browse more pages compared to the overall mean values for each of attributes. From these findings, we can acknowledge that those learners who have better performance are those who

improved better after using our WBI program but they are not necessarily known as better learners (those who are identified as experts). This implies that a learner may use specific preferences accommodated in a WBI program although it may not be helpful for improving their learning performance. These findings imply that “*what learners like may not be what they need*” (Minetou, et al., 2008). The other explanation is that performance and preferences are two different things (Minetou, et al., 2008).

Few previous studies have been carried out to investigate these three attributes (g-score, t-pages and t-time) and what relationships between such attributes may affect learners’ performance level using the intersection of three individual differences (gender, cognitive style and prior knowledge). As future work, there is a need to analyze learners’ performance using other data-mining approaches (e.g., classification and association rules) as well as to consider more subjects and a larger sample to provide additional evidence.

References

- Beckwith, L., Burnett, M., Wiedenbeck, S. & Cook, C., 2005. Effectiveness of end-user debugging software features: are there gender issues?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI’05)*. Portland, OR. ACM Press, New York, NY, p. 869–878.
- Calcaterra, A., Antonietti, A. & Underwood, J., 2005. Cognitive style, hypermedia navigation and learning. *Computers & Education*, pp. 44(4), 441-457.
- Calisir, F. & Gurel, Z., 2003. Influence of text structure and prior knowledge of the learner on reading comprehension, browsing and perceived control. *Computers in Human Behavior*, pp. 19, 135–145.
- Chen, L.-H., 2010. Web-based learning programs: Use by learners with various cognitive styles. *Computers & Education*, pp. 54, 1028–1035.
- Chen, S. Y., Fan, J. & Macredie, R. D., 2006. Navigation in hypermedia learning systems: experts vs. novices. *Computers in Human Behavior*, pp. 22, 251–266.
- Chen, S. Y. & Liu, X., 2008. An Integrated Approach for Modeling Learning Patterns of Students in Web-Based Instruction: A Cognitive Style Perspective. *ACM Transactions on Computer-Human Interaction*, p. 15 (1).
- Chen, S. Y. & Macredie, R. D., 2002. Cognitive styles and hypermedia navigation: Development of a learning model. *Journal of the American Society for Information Science and Technology*, pp. 53(1), 3-15.
- Chen, S. Y. & Macredie, R. D., 2010. Web-based interaction: A review of three important human factors. *International Journal of Information Management*, Volume 3, p. 379–387.
- Farrell, I. H. & Moore, D. M., 2001. The effect of navigation tools on learners’ achievement and attitude in a hypermedia environment. *Journal of Educational Technology Systems*, pp. 29(2), 169–181.
- Ford, N. & Chen, S., 2000. Individual differences, hypermedia navigation and learning: an empirical study. *Journal of Educational Multimedia and Hypermedia*, pp. 9, 281–311.
- Ford, N., Wood, F. & Walsh, C., 1994. Cognitive styles and online searching. *Online and CD_ROM Review*, 18(2), p. 79–86.
- Fraley, C. & Raftery, A. E., 1998. How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis. *The Computer Journal*, pp. 41(8), 578-588.

- Gargano, M. L. & Raggad, B. G., 1999. Data mining—A powerful information creating tool. *OCLC Systems Services*, pp. 15(2), 81–90.
- Goodenough, D., 1976. The role of individual differences in field dependence as a factor in learning and memory. *Psychol. Bull.*, pp. 83, 675–694.
- Hand, D. J., 2007. Principles of Data Mining. *Drug Safety*, pp. 30 (7), 621-622.
- Hedberg, S. R., 1995. The data gold rush. *Byte*, pp. 20(10), 83–88.
- Hölscher, C. & Strube, G., 2000. Web search behavior of Internet experts and newbies. *Computer Networks*, pp. 33, 337-346.
- Jain, A. & Dubes, R., 1999. Data clustering. *ACM Computing Survey*, pp. 31, 264-323.
- Kim, K. S., 2001. Implications of user characteristics in information seeking on the world wide web. *International Journal of Human-Computer Interaction*, pp. 13(3), 323-340.
- Large, A., Beheshti, J. & Rahman, T., 2002. Gender differences in collaborative web searching behavior: an elementary school study. *Information Processing and Management: an International Journal*, pp. 38(3), 427–443.
- Lee, C., Cheng, Y. W., Rai, S. & Depickere, A., 2005. What affect student cognitive style in the development of hypermedia learning system?. *Computers and Education*, 44(1), p. 1–19.
- Lee, M. W., Chen, S. Y., Chrysostomou, K. & Liu, X., 2009. Mining students' behavior in web-based learning programs. *Expert Systems with Applications*, pp. 36, 3459–3464.
- Lin, H.-J. & Chen, L.-H., 2008. *Discovering learning pattern in different cognitive style of learners*. s.l., The Third International Conference on Convergence and Hybrid Information Technology.
- Littlejohn, A., 2002. New lessons from past experiences: recommendations for improving continuing professional development in the use of ICT. *Journal of Computer Assisted Learning*, pp. 18, 166–174.
- McDonald, S. & Stevenson, R., 1998. The effects of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors*, pp. 40, 18 –27.
- Minetou, C. G., Chen, S. Y. & Liu, X., 2008. Investigation of the Use of Navigation Tools in Web-Based Learning: A Data Mining Approach. *International Journal of Human-Computer Interaction*, pp. 24(1), 48–67.
- Mitchell, T. J. F., Chen, S. Y. & Macredie, R. D., 2005. Hypermedia learning and prior knowledge: domain expertise vs. system expertise. *Journal of Computer Assisted Learning*, pp. 21, 53–64.
- Nilsson, R. M. & Mayer, R. E., 2002. The effects of graphic organizers giving cues to the structure of a hypermedia document on users' navigation strategies and performance. *International Journal of Human-Computer Studies*, 57(1), pp. 1-26.
- Riding, R. J., 1991. *Cognitive Styles Analysis. Learning and Training Technology*. Birmingham: s.n.
- Roussinov, D. & Zhao, J. L., 2003. Automatic discovery of similarity relationships through Web mining. *Decision Support Systems*, pp. 35, 1, 149–166.
- Roy, M., Taylor, R. & Chi, M. T. H., 2003. Gender differences in patterns of searching the web. *Journal of Educational Computing Research*, pp. 29, 335–348.
- Schumacher, P. & Morahan-Martin, J., 2001. Gender, internet and computer attitudes and experiences. *Computers in Human Behavior*, pp. 17(1), 95-110.
- Shin, E., Schallert, D. & Savenye, C., 1994. Effects of learner control, advisement, and prior knowledge on young students' learning in a hypertext environment. *Educational Technology Research and Development*, pp. 42(1), 33-46.
- Turns, J., Atman, C. J. & Adams, R., 2000. Concept maps for engineering education: A cognitively motivated tool supporting varied assessment Functions. *IEEE Trans. Educ.*, Volume 43, p. 164–173.
- Wang, M., 2007. Designing online courses that effectively engage learners from diverse cultural backgrounds British. *Journal of Educational Technology*, pp. 38(2), 294–311.
- Wang, M., Rees, S. J. & Liao, S. Y., 2002. *Building an online purchasing behavior analytical system with neural network*. s.l., WIT Press.
- Wildemuth, B. M., 2004. The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, pp. 55(3), 246-258.
- Witkin, H. A., Moore, C. A., Goodenough, D. R. & Cox, P. W., 1977. Field-dependent and field independent cognitive styles and their educational implications. *Review of Educational Research*, pp. 47, 1, 1–64.
- Witten, I. H., Frank, E. & Hall, M. A., 2011. *Data Mining: Practical Machine Learning Tools and Techniques. 3rd edition*. s.l.:Morgan Kaufmann Publishers.
- Workman, M., 2004. Performance and perceived effectiveness in computer-based and computeraided education: do cognitive styles make a difference?. *Computers in Human Behavior*, pp. 20, 4, 517–534.

