

Evaluating Learning Style Personalization in Adaptive Systems: Quantitative Methods and Approaches

Elizabeth J. Brown, *Member, IEEE*, Timothy J. Brailsford, Tony Fisher, and Adam Moore

Abstract—It is a widely held assumption that learning style is a useful model for quantifying user characteristics for effective personalized learning. We set out to challenge this assumption by discussing the current state of the art in relation to quantitative evaluations of such systems and also the methodologies that should be employed in such evaluations. We present two case studies that provide rigorous and quantitative evaluations of learning-style-adapted e-learning environments. We believe that the null results of both these studies indicate a limited usefulness in terms of learning styles for user modeling and suggest that alternative characteristics or techniques might provide a more beneficial experience to users.

Index Terms—Evaluation/methodology, adaptive hypermedia, user issues, computer-assisted instruction, human information processing.

1 INTRODUCTION

THE use of personalized e-learning and adaptive educational hypermedia (AEH) has become increasingly important in recent years, with extensive research being devoted to finding different ways of tailoring the learning experience for individual students.

In addition to discovering techniques for adaptation, there is also the equally important issue of assessing the impact that such systems have upon their users. Unfortunately, there have previously been few rigorous evaluations carried out in this area. Most of the published works, when subjected to close scrutiny, possess inadequate experimental design and data analysis to provide a reliable appraisal of such systems.

As a consequence, there are some widely held beliefs that certain user models are effective in helping provide personalized e-learning, but these beliefs are not generally backed up with scientific evidence. One such approach to user modeling is that of basing the model upon students learning style preferences.

Previous work carried out by the authors has focused upon the use of visual/verbal and sequential/global learning style preferences in AEH. Despite rigorous quantitative evaluations, no statistically significant benefits for users have arisen from these approaches to personalization [7], [8]. The findings from these studies have led the

authors to produce a summary and an extension of the work done in this area, and thus, provide a crucial reference for those working in this field.

This paper focuses on the quantitative aspects of evaluations carried out into adaptive hypermedia systems used for education. It presents a survey of existing systems that use learning style as the basis of its user model, the quantitative methodologies appropriate for evaluating such systems, and also two case studies of empirical evaluations. Last, it discusses the issues that are most critical to this research and also presents some ideas for future work.

2 ADAPTIVE HYPERTEXT IN EDUCATIONAL SYSTEMS

There are several systems developed for educational purposes, commonly referred to as AEH systems. These systems base their user models largely on existing knowledge, and adaptation occurs at both content level (adaptive presentation) and/or link level (adaptive navigation) [11].

This section discusses the approaches employed across a variety of different AEH systems. Information provided in the following two tables gives a comprehensive overview of the most influential AEH systems of recent times. It is impossible to provide a truly exhaustive list of all AEH platforms, hence the systems mentioned here are those that have been developed and discussed most extensively within AEH literature.

2.1 Adaptive Navigation Systems

Adaptive navigation is used to implement adaptation in many AEH systems. They are summarized in Table 1, which gives information about the adaptation mechanism used in each system and also what characteristics are taken into account in terms of the user model.

- E.J. Brown, T.J. Brailsford, and A. Moore are with the School of Computer Science, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, UK. E-mail: {ejb, tjb, axm}@cs.nott.ac.uk.
- T. Fisher is with the School of Education, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, NG8 1BB, UK. E-mail: tony.fisher@nottingham.ac.uk.

Manuscript received 21 Nov. 2008; revised 23 Jan. 2009; accepted 8 Feb. 2009; published online 12 Feb. 2009.

For information on obtaining reprints of this article, please send e-mail to: lt@computer.org, and reference IEEECS Log Number TLTSI-2008-11-0101. Digital Object Identifier no. 10.1109/TLT.2009.11.

TABLE 1
Overview of Selected AEH Systems that Utilize Adaptive Navigation Techniques

| System: | Type of adaptive navigation: | User model addresses: | | | | |
|-----------------|--|-----------------------|-----------|------------|-----------------------|-------------|
| | | Goals | Knowledge | Background | Hyperspace experience | Preferences |
| ELM-ART [18] | Adaptive link annotation; adaptive link sorting | | ✓ | | | |
| ISIS-Tutor [15] | Direct guidance; adaptive link annotation; adaptive link hiding. | | ✓ | | | |
| ITEM/PG [17] | Adaptive link annotation | | ✓ | | | |
| HyperTutor [63] | Direct guidance; adaptive link hiding | | ✓ | | ✓ | |
| SHIVA [92] | Direct guidance | | ✓ | | | |
| Hypadapter [43] | Adaptive link hiding | | ✓ | | | ✓ |
| AHA! [29] | Adaptive link annotation; adaptive link hiding. | | ✓ | | | |
| Netcoach [84] | Direct guidance; adaptive link annotation. | ✓ | ✓ | | | ✓ |
| Interbook [19] | Direct guidance; adaptive link annotation. | | ✓ | | | |
| CHEOPS [34] | Link removal. | | ✓ | | | |
| TANGOW [20] | Link disabling; adaptive link sorting. | ✓ | | ✓ | | ✓ |

From the information contained in this table, it can be seen that most AEH systems that utilize adaptive navigation support do so by addressing the knowledge aspect of the user model. The most commonly used techniques are direct guidance and adaptive link annotation.

2.2 Adaptive Presentation Systems

In a similar fashion to Table 1, Table 2 gives an overview of AEH systems that implement adaptive presentation. It is worth noting that some systems appear in both Tables 1 and 2, indicating that they employ both forms of adaptation.

TABLE 2
Overview of Selected AEH Systems that Utilize Adaptive Presentation Techniques

| System: | Type of adaptive presentation: | User model addresses: | | | | |
|------------------|---|-----------------------|-----------|------------|-----------------------|-------------|
| | | Goals | Knowledge | Background | Hyperspace experience | Preferences |
| Anatom-Tutor [4] | Altering fragments. | | ✓ | ✓ | ✓ | |
| C-Book [45] | Inserting/removing fragments; altering fragments. | | | ✓ | | |
| Hypadapter [43] | Sorting fragments. | | ✓ | | | ✓ |
| AHA! [29] | Inserting/removing fragments; altering fragments. | | ✓ | | | |
| Metadoc [5] | Stretchtext. | | ✓ | | | |
| TANGOW [20] | Altering fragments. | ✓ | | ✓ | | ✓ |
| WHURLEHM [91] | Inserting/removing fragments. | | ✓ | | | |

In this type of modification, it can be seen that inserting/removing fragments and altering fragments are the preferred methods. User's knowledge remains the most popular aspect of the user model to be addressed. Background is a much wider distinction than knowledge since it encompasses factors such as motivation, users' memory capabilities, users' attitudes or beliefs, and social-economic status (SES). It is best exemplified in Beaumont's work [4], where he utilizes both knowledge and background as independent components of his user model.

3 ADAPTATION TO USER CHARACTERISTICS

3.1 Pedagogical Limitations of Current User Models

Teaching and learning in schools and colleges has traditionally been facilitated by streaming students into existing knowledge/ability groupings. It is therefore no surprise that a lot of educational software, online school-based resources, and AEH systems parallel this method of instruction and tend to accommodate students of a specific knowledge or ability level. (There are obviously some

differences, such as the more immediate temporal shift that can occur in AEH to move students between streams compared to school classes in the real world, allowing for more flexibility in differentiation.)

However, this model of contextual knowledge is relatively simplistic from a pedagogical perspective; it classifies students into a category but requires continual updating as the student gains knowledge. It also does not allow for other aspects of the learning process, such as the way a student approaches their learning from a psychological perspective.

Recent educational thinking has led toward a more direct cognitive approach that rationalizes the learning preferences of students. This is more satisfactory for addressing multiple learners' needs since it allows a large variety of preferences to be catered for, within specific knowledge domains, and may contribute to an enhanced learning experience [78]. Some recent work has been carried out by Bajraktarevic et al. [2], [3] that uses learning style theory to create the user profile, thus employing the "user preference" aspect in contrast to "existing knowledge." There is much literature in the fields of psychology and education

relating to learning styles, and this mode of adaptation may be more beneficial to learners than those simply based on domain knowledge. There is a large variety of learning style theories and tools that can be used to categorize learners. The next section (3.2) gives a brief overview of these main concepts and presents some additional AEH systems that have implemented learning style adaptation.

3.2 User Model Based on Learning Style

Keefe [46] states that learning styles are “*characteristic cognitive, affective, and psychological behaviors that serve as relatively stable indicators of how learners perceive, interact with, and respond to the learning environment.*” There are many published variants on this definition and there is also some debate on whether learning style preferences are contextual or not.

Examples of learning style include constructs such as field dependence/independence [89], reflexivity versus impulsivity [78], VAK (visual/auditory/kinesthetic), (w)holist/serialist [23], and models such as Dunn and Dunn [32] or Honey and Mumford [44].

A number of AEH systems utilize learning style as the basis of their user model. Examples include AES-CS [80], INSPIRE [40], Arthur [37], MANIC [77], AHA! [76], MOT [76], and CS-383 [21]. A detailed overview of these systems and their theoretical underpinnings can be found in Brown et al. [7]. Essentially, what these systems have in common is that users’ learning style forms an important part of the user profile; these learning style preferences are then used to inform how adaptation is performed in the AEH system. For example, only certain units of content or particular links will be displayed to those users who possess an appropriately matched learning style preference. In this way, the end document is one that is tailored to the individual preferences of the user but using a different user model compared to, e.g., domain knowledge.

A metaevaluation of these and other systems that have used learning style preferences for user profiling is presented in Section 5.2, but in order to interpret this, we first need to discuss the quantitative methods that are required to assess the relative benefits or disadvantages of such systems.

4 METHODOLOGY

The concept of user testing is an extremely important one. This section discusses the methodology that should be employed in order to achieve a quality-assured, rigorous set of quantitative and qualitative data that can be analyzed to investigate the effect of any interventions made by AEH systems.

4.1 Designing and Planning User Trials

Robson states that “*design is concerned with turning research questions into projects*” [69], so naturally it follows that the research question must be clear and unambiguous at the start of any experimental design. Once the research question has been clarified, the design of the project must be considered: is it to be an experiment, a survey, or an observational piece of work? A research project might combine aspects of all three. Whatever the methodology

chosen, there are several important factors to take into account. These include making sure that there are adequate levels or groupings in the design and ensuring sufficient “clean” sample sizes (including making allowances for participant unreliability or incomplete/inaccurate data); Mertens [51] recommends no less than 15 participants in the smallest grouping. There should also be care taken over choice of appropriate dependent variables. Lastly, it is strongly advised that initial pilot tests are carried out of any planned trials, so that potential problems with the study can be discovered ahead of time and planned around [59].

Data collection is an important aspect of any study but commonly used scales and measures are often not scrutinized as closely as they should be. Two essential components of such scales are their reliability and validity.

Reliability is a measure of how free the scale is from random error. It is judged by temporal stability (also known as test-retest reliability—high scores mean it is more reliable) and also internal consistency (the level at which different components of the scale are assessing the same traits, commonly measured by Cronbach’s alpha coefficient) [55], [59].

Validity refers to how accurately a scale measures what it is intended to measure, involving aspects of content validity, criterion validity, and construct validity. Content validity relates to how well a particular measure reflects the wider content domain. Criterion validity examines the connection between the scores of a scale and some other specific standard. Construct validity tests the scale against theoretically derived hypotheses that relate to the variable/s under examination. The way in which construct validity is explored is by examining associated constructs (convergent validity) and disparate ones (discriminant validity) [59], [73].

Reliability and validity therefore provide useful information about the appropriateness of selecting various scales or measurements for use within research projects. Other considerations include the preparation of questionnaires, such as response types and the wording of questions so as to avoid jargon, loaded or complex words and questions, and any cultural or emotional bias. Pallant suggests that, where possible, questionnaires should also include provisos for “don’t know” or “not applicable” [59].

Of course, scales and measurements form only one aspect of a research project. Historically, the scientific method of research involves observations and formulating/testing hypotheses by gathering and processing quantitative and qualitative data. Scientific objectivity is achieved by first of all presenting a “null hypothesis,” which states that there will be no statistically significant differences in quantitative data gleaned from disparate experimental groups or conditions. A number of alternate hypotheses can then follow, which suggest what these differences might be. These alternate hypotheses can be specific (for example, an increase in one factor might lead to an increase in the independent variable, i.e., what is being measured) or more general (for example, one particular group might have a higher score than another).

Empirical evidence is required to support or refute these hypotheses, which results from the statistical analysis of

quantitative data. The resulting statistics reveal the significance, or otherwise, of such data. Probability values are commonly utilized as a measure of significance, with a value of 0.05 or less (i.e., 5 percent or the probability of 1 in 20) required in order to declare something significant. Thus, if $p > 0.05$, the finding is not considered statistically significant because the chances are too great that the observed effect resulted from chance rather than the intervention under examination. Probability values are ascertained by statistical testing, which is discussed in Section 4.2.

There are also several different frameworks of experimental design, classified as fixed, flexible, and multiple design strategy [69]. Fixed designs tend to involve surveys and experiments; their principal characteristic is that much of the design specification is decided upon well in advance of the study. They are based upon well-developed theoretical frameworks, so that the researchers are aware of existing issues and how to control for various aspects of the experiment. Fixed designs are very much quantitative, whereas flexible designs tend to be qualitative. They are flexible in both in terms of the data that are typically collected from such work and also the approach used, where less advance preparation takes place and the research plan typically evolves and grows as the work is carried out.

Flexible designs can also accommodate quantitative methods in addition to qualitative, whereupon they tend to be referred to as mixed-method or multiple design strategies. There has traditionally been some dispute over which approach is more “scientific,” with scientific disciplines tending to be rooted in fixed design methodology and sociological domains preferring flexible procedures. Robson states that these techniques are not mutually exclusive in scientific studies, provided that they are carried out in a systematic and responsible manner [69]. The case studies presented in Section 6 utilize mostly quantitative designs with a small amount of qualitative methodology.

4.2 Data Analysis

Once a body of data has been gathered from the research study, it must often be cleaned and processed before it can be analyzed. Data cleaning involves screening the data set and resolving any problems with missing or incomplete data. Often, this means excluding such data from the overall analysis since they might not be a fair representation of that particular case or participant: for example, if a questionnaire is administered at two different time periods yet the user is only present at the first session, the second questionnaire could not have been completed, and thus, a comparison between the two would be meaningless.

Data may also need to be processed or transformed from one unit into another. For example, the number of Web pages seen by a user might need to be calculated as a percentage of the whole number of Web pages overall, or it might need to be converted into a unit of frequency.

Once the data set has been cleaned and processed, it can be analyzed statistically. This is usually done using one of a number of statistical software packages. The software used for data analysis throughout this research was SPSS.

The statistical tests that are carried out on the cleaned data to provide probability values vary according to what kind of data are being analyzed, how many groups or levels

there are, and also what the specific research question is. Generally speaking, statistical techniques either explore relationships between variables or differences between groups [59], [69]. Relationships between variables are often examined via correlations, multiple regression, or factor analysis [59]. A Pearson correlation assesses how strongly two variables relate to each other and whether this is a positive or negative association. A more sophisticated technique is multiple regression, where the relationship of a set of independent variables can be predicted against a continuous dependent variable. Factor analysis is a data reduction procedure, where a large grouping of variables can be condensed into a smaller, more manageable data set, which is then used for comparison.

Differences between groups can be investigated using t -tests, Analysis of Variance (ANOVA), Analysis of Covariance (ANCOVA), and Chi-square (χ^2) tests. If there are only two groups or two sets of data, the mean score of a continuous variable can be analyzed using t -tests. If there are two or more groups, a one-way ANOVA can be carried out: it studies the influence of a single independent variable on the dependent variable, although additional posthoc testing is required to determine in which of the groups the difference appears. If there are two independent variables, a two-way ANOVA can be used: This allows testing of an interaction effect between the two variables. Multivariate analysis of variance (MANOVA) testing can be used when group comparisons are required based on several unique (but associated) dependent variables. ANCOVA is a method by which an additional confounding variable needs controlling for, so that any differences between groups can be seen once this extra factor is taken into account [59]. Finally, Chi-square (χ^2) tests can determine if the distribution of a discontinuous variable is the same in two or more independent samples.

Further guidance on the empirical evaluation of user models and adaptation mechanisms can be found in the works of Brusilovsky et al. [14], Chin [22], Weibelzahl [86], Weibelzahl et al. [87], and Weibelzahl and Weber [88]. In particular, both Brusilovsky and Weibelzahl mention the idea of *layered evaluation*, where the success of adaptation is broken down and evaluated at different layers, to reflect the component parts of an adaptive system. Weibelzahl and Chin discuss in detail further aspects of reliability, validity, effect size, and power.

4.3 Other Considerations

Finally, there are two separate but somewhat related issues that need to be discussed—the Hawthorne effect and the confusion of correlation with causality. The Hawthorne effect is a phenomenon akin to a placebo effect, whereby the behavior of test subjects is temporarily altered because they are aware that they are participants in a study, and hence, expect to be given some special treatment that would help them perform certain tasks more effectively. It is named after a factory called the Hawthorne Works, where a series of productivity experiments were carried out on factory workers between 1924 and 1932. A variety of different interventions into physical working conditions were introduced in the factory, such as modifying the light levels. These interventions, regardless of what they were, had the

net result of increased productivity—but in both the experimental and the control groups [38], [83]. It seemed that any change perpetrated by the researchers caused this increase; it was not necessarily caused by any specific intervention. This has an impact upon the way in which any kind of user testing is carried out in other situations, since it has been shown that merely having been selected for some kind of different treatment can, in fact, have a positive effect on participants. In reality, this is very difficult to control for since for ethical reasons, taking part in a user trial must be voluntary and by informed consent. However, measures such as using a control group or a crossover design (where different groups all experience the same interventions but at different times in the research) help to reduce the potential consequences of the Hawthorne effect. It is also counteracted by using relatively homogeneous experimental groups, which neutralize the natural variability shown by individuals.

The other concern relates to causality. A strong correlation between variables does not signify that one causes—or is caused by—the other, but merely that the relationship exists. The reason for that association might be due to some other factor that has not been taken into account by the research.

It is thus critical to be cautious when statistically significant results are found as a result of a specific intervention and to show due regard for other factors that might contribute to these findings.

5 EXISTING EVALUATION STUDIES OF AEH

This section provides a survey of the state of the art in quantitative evaluation of adaptive hypermedia in education, focusing specifically on those that are driven by learning style adaptation.

There is a paucity of rigorous user evaluation in adaptive systems in the published literature. Studies tend to be fairly small in terms of sample sizes and rarely are statistical measures of significance used. Effect size is almost never published to the detriment of such studies. In most situations, systems are tested on a group of users who might possess some kind of bias (either implicit or explicit), such as the undergraduate students who engage with a system designed by the lecturer of their course. Such users might be expected to have favorable expectations of this software, which could result in a “self-fulfilling prophecy” whereby they perform well in the system because they expect to; the observer/novelty (a.k.a. Hawthorne) effect thus obscures any objective use of the system.

The survey of these studies is now divided into two parts: Section 5.1 investigates the use of evaluations in adaptive hypermedia from a broad perspective while Section 5.2 describes their application to systems that utilize learning style preferences as the predominant adaptation mechanism. Section 5.3 provides a summary of the information presented from these investigations.

5.1 Introduction

The AEH systems presented in Tables 1 and 2 were investigated further to determine which of them had

associated user trials based upon quantitative methodologies. Out of 15 systems, eight of them (ITEM/PG [17], SHIVA [92], Hypadapter [43], CHEOPS [34], Tangow [57], Anatom-Tutor [4], C-Book [45], WHURLE-HM [91]) did not appear to have any published data relating to empirical testing of the effectiveness of its adaptation mechanism.

Of the seven remaining systems, there was much variation in the depth of the empirical evaluations carried out. User trials have been carried out with four systems (ELM-ART [85], Hypertutor [81], Interbook [13], and AHA! [74], [75]), which resulted in the publication of some quantitative data; however, these are mostly related to how users interacted with the different aspects of the system and displayed as raw data or percentages rather than statistical results.

Only three evaluations actually used statistical testing to show the effectiveness of their systems: ISIS-Tutor, MetaDoc, and Netcoach. The research on ISIS-Tutor [16] presents ANOVA and T-test statistics but sample sizes were small and there were not enough participants for these results to be valid. User trials from MetaDoc [5], [10] show statistically significant results at the 1 percent level, using ANOVA and further post-hoc testing (Tukey) to determine where the differences were. However, sample sizes are not given, so it is unknown as to whether this evaluation was statistically valid. In addition, for both the ISIS-Tutor and MetaDoc user trials, there was no mention of effect size or how this might have been calculated. By stark contrast, the Chi-square (χ^2) and *t*-test analyses shown in the Netcoach evaluation [88] are extremely rigorous, using large statistically valid sample sizes and displaying careful experimental planning. The evaluation data show both probability values and effect size. This last study is the only general AEH user trial found to be of reasonably high quality with respect to quantitative experimental design and data analysis.

5.2 Learning Style Adaptation in AEH Systems

Further investigations were carried out into AEH systems that specifically utilize learning style as their adaptation mechanisms.

Out of 10 systems, there were six (ILASH [2], MOT [76], OPAL [25], AHA! [76], CS-383 [21], and Tangow [62]) which did not seem to have published any quantitative evaluations. Two systems (AES-CS [79] and INSPIRE [60]) presented some empirical data in the form of bar charts; pretest/posttest scores—together with the difference between them and also standard deviations—but no statistical testing was carried out and sample sizes were relatively small ($n = 10$ and $n = 23$, respectively).

Two systems did show statistical testing and this was done reasonably well. Bajraktarevic et al.’s empirical study [3] used 21 students, split into two groups and employing a crossover design. T-testing was used to generate test statistics and probability values. However, there was fairly high standard deviation shown throughout for most of the test scores and no effect size was calculated for any of the data analyses. The best example of statistical testing carried out in a learning-style-driven AEH is exemplified by Wolf with his work on the iWeaver system [90]. His research shows meticulous care and attention to experimental design and data analysis. Like Bajraktarevic et al., he also uses

paired *t*-tests to produce test statistics and probability values but additionally takes into account other factors such as effect size in considerable detail.

5.3 Summary of Quantitative Evaluation Studies Carried Out in AEH Systems

From a review of the published data concerning quantitative evaluation of adaptive educational hypermedia systems, in particular, those driven by learning style adaptation, it is clear that not enough is currently being done to test the effectiveness of adaptation mechanisms. Of the 24 different systems examined, only five of them had been tested with statistical techniques and of these, only two (those carried out with Netcoach and iWeaver) could be considered to be high-quality studies.

This apparent lack of quantitative studies in relation to adaptation mechanisms exemplifies a critical need to carry out further research in this area. It is clear that a systematic experimental approach, utilizing quantitative methods and statistical analysis of data, must be employed in order to find out objectively whether the adaptation technique is an effective means of providing personalization to users.

6 CASE STUDIES

The following case studies are examples of quantitative evaluations carried out by some of us at the University of Nottingham.

6.1 WHURLE-LS

Web-Based Hierarchical Universal Reactive Learning Environment—Learning Styles (WHURLE-LS) is an e-learning system that was used to support the teaching of Computer Science students at the University of Nottingham. Its user model is based on the Felder-Soloman Inventory of Learning Styles [33], focusing particularly on visual/verbal learning style. A quantitative user trial was carried out with over 200 students, who were assigned randomly into matched, mismatched, or “no preference” groups. Matched students were given content that corresponded to their learning style preference (e.g., a visual student receives primarily visual content). Mismatched students were given content that was contrary to their preferred learning style (e.g., a visual student would receive primarily verbal content). The “no preference” group consisted of students who either interacted with a “no preference” environment (a balance of visual and verbal content) or had a “no preference” learning style. The aim of the evaluation was to see if matching or mismatching students affected their academic performance after interacting with the system. It was hoped that this would provide some insights into the effectiveness of the user model as a means of improving student learning.

The study was carried out in a thorough and scientifically objective manner, with a null hypothesis that stated there would be no difference in markers of academic performance between the three groups. Statistical analysis of the data did not find any evidence to support the alternative hypotheses (that speculated on where the differences might be found) and so the null hypothesis was maintained. The foremost conclusion from this study

was that adaptation of e-learning materials to cater for differences in visual/verbal learning style could not be shown to be of any benefit, under these circumstances. Full details of the work can be found in Brown et al. [7].

6.2 DEUS

A further user trial was conducted with a similar system to WHURLE-LS. Digital Environment Utilizing Styles (DEUS) was an e-learning platform used by children aged 9-11 years old. The user trial was integrated with normal school teaching; the aim of the trial was, again, to see if adaptation to learning style preference would prove to be advantageous in terms of academic performance.

The learning style employed in this evaluation was sequential/global, another aspect of the Felder-Soloman model. Like the WHURLE-LS study, users were either matched or mismatched with respect to their learning style preference and the environment that they interacted with. Pretests and posttests were conducted to assess domain knowledge; this data was then analyzed statistically to see if there were any differences between these groups.

The findings were very similar to those from the WHURLE-LS trial; there were no significant differences between users from the matched or mismatched groups. Neither was there a difference between students with the same learning style preference. There was a very slight difference in the time it took to interact with the global environment compared with the sequential environment (irrespective of pupils’ learning style preferences) but this was very slight and not considered to be an important finding (the global environment took an average of 1.75 hours to work through; the sequential took an average of 2 hours to work through). This research was presented at the Hypertext 2007 conference and the exact details of the work can be found in those proceedings [8].

6.3 Conclusions

From these two case studies, it can be seen that adaptation to users’ learning style preference in AEH systems did not indicate any statistically significant benefit, under particular circumstances. In this respect, the findings concur with those from Kelly and Tangney [47].

However, the most crucial aspect of this work was not necessarily related to these findings. The authors consider that the approaches taken to evaluate these systems to be of equal, if not greater, importance. In particular, the way in which the user trials were designed in order to gather empirical data; the consideration given to appropriate intervention time and minimum group sizes; and also the scientific objectivity that has been a key focus of this work.

7 DISCUSSION

WHURLE-LS and DEUS have been presented as case studies of how we can evaluate the effect of using learning style as a means of personalization. Despite successful user trials, there were several limitations of the work that warrant some consideration. These can be subdivided into criticisms of the user trials themselves and also aspects of the research from a broader perspective.

7.1 Issues with the User Trials

The studies that were carried out were satisfactory from the perspective of sample size and amount of intervention, as discussed in Section 4. However, the Hawthorne effect could not be controlled for, and thus, any positive results would have required careful post-hoc analysis and possibly additional experimentation, in order to determine what had caused this beneficial effect. This did not prove to be an obstacle though, since the methodology employed in these two user trials would still have allowed for comparison between groups because the interventions were carried out under uniform conditions and with similar group compositions. However, it is useful to consider this issue here in greater detail, so that this potential problem might be resolved for future studies. In order to control for the Hawthorne effect, participants in the study cannot realize that they are part of a study where they might be being given special treatment. This is not always possible to do, and so the best way to deal with this is to use a control group or have a crossover design (where all participants experience the different interventions at different phases of the study). The Hawthorne effect tends to be more problematic when positive results are observed, since it is easy to attribute causality to the intervention rather than to the effect of being observed, resulting in a Type 1 error, otherwise referred to as “over-optimism” [59].

Another potential issue with the user trials was that the group compositions themselves might have proved problematic in terms of the granularity of the research. The way in which learning style itself was used could have been far too generalized: users were categorized into broad overarching groups reflecting the learning style under investigation, i.e., visual or verbal; sequential or global. For those who were actually in the “middle ground” of “little/no preference,” this was possibly not an ideal way to categorize them. With particular respect to DEUS, where a large number of pupils were in the central region of a normal distribution, it might be expected that they would not have benefited from any kind of adaptation. Thus, any positive results might have been obscured by a larger number of the negative effects as a consequence of this course-grained categorization of users (although this seems unlikely since the statistics showed no difference, even when these users were excluded from the data set). In addition, because the main factor under investigation was learning styles, other aspects of users were not considered as covariants. The groups tested in the user trials were as homogeneous as possible, so that the mix of gender, age, and generalized measures of intelligence were equivalent between groups. If the effect of learning style is a small one, this could easily have been masked by variance between these individual characteristics in the group as a whole. Hence, the homogeneity of groups could have contributed to the negative results and so further trials should be carried out with much stricter controls applied to the make up of user characteristics within groups. However, it should be noted that these group compositions were grounded in typical “real world” situations, and thus, provided truly authentic conditions for user testing. Given that schools in the real world rarely contain large groups of pupils that are exactly alike, the consequences resulting

from the user trials can thus be applied to schools and other educational institutions in general and reflect the highly variable assortment of learners seen in such situations.

Another factor to take into account was the way in which students were assessed, using the Felder-Soloman ILS model that was simplified down to a single axis (visual/verbal for the WHURLE-LS user trials and sequential/global in the DEUS study). Though necessary in the context of this work, it may not have been appropriate, since that axis is actually a component of the model and not the whole instrument. It is thus possible that students were not assessed as effectively as they might have been but this is beyond the scope of this work and would require input from additional detailed studies from an HCI/psychological perspective. In any case, the approach taken here is in keeping with previous studies that have also only used certain aspects of learning style and so the conclusions can be said to be comparable with those works.

7.2 Issues Related to the Broader Research Area

There seem to be two main problems that arise from using learning styles for user personalization. The first problem is concerned with learning style theory and its subsequent application, whilst the second addresses the complexity of learning as a process.

7.2.1 Learning Styles

The value of learning styles is a contentious issue with many psychologists and neuroscientists, who have questioned the scientific basis of learning styles and the theories upon which the models are based [24], [41]. Work by Coffield et al. [23], [24] suggests that many learning style models have low internal reliability and validity and do not measure that which was intended. It thus seems likely that many AEH researchers are employing a flawed user model for their studies (studies which in themselves may possess insufficient experimental design/evaluation). As a result, if positive results are found, it is all too easy to think that learning styles are making a positive impact upon learning. In truth, it may be almost impossible to determine the real reasons behind the experimental results.

There is also some controversy over the temporal stability of learning styles. It has yet to be established whether (and how often) learning styles change [27], [50], [64], hence any valid method of personalization that utilizes learning style may need to accommodate a flexible, dynamic model of user preferences rather than a fixed, static measurement.

There are some additional issues with respect to the particular learning style models employed in this research. The concept of visual and verbal preference is itself highly complex and very much more sophisticated than has been suggested by this work. For example, Paivio’s dual coding theory states that human cognition deals with visual and verbal processing simultaneously [58]. If this is true, then both of these types of representation should be catered for, for learning to be effective [53], [65]. Kozhevnikov et al. [48] state that although “verbalizers” tend to be a fairly uniform group, “visualizers” can be further categorized into those of high and low spatial ability. They found that these two groups (the “spatial” type and “iconic” type, respectively)

interpreted visual representations differently and this suggests that a further source of variance could have been introduced in the studies conducted with WHURLE-LS. There is also the issue of author bias: the representations may not have been suitably constructed for this wide range of visual and verbal users; even if they were appropriate for use in this system, the ways in which multiple representations affect learning have still not been fully explored [1], [71].

There is also the notion of visual literacy, a related but different concept to that of visual learning style. Visual literacy, a term first attributed by Debes [30] in 1968, is the “ability to construct meaning from visual images” [39] although subtly different definitions exist across different disciplines. It may be that learners with a strong visual learning preference are highly visually literate although there do not seem to be any published correlations between the two measures. Linguistic literacy, where meaning is derived from written or spoken language, is possibly related to verbal learning preference. However, researchers such as Kress [49] state that the integration of visual and linguistic literacies is essential to help students construct meaning, advocating a mixed media approach rather than dichotomous learning, reflecting the ideas of Paivio.

One of the overwhelming and insurmountable problems regarding learning style preferences is that it represents only one characteristic of the learner. Melis and Monthienvichienchai refer to eight other relevant criteria such as motivation, working memory capacity, and personality traits, amongst others [50]. Plass et al. suggest that there might be a link between learning styles and other features of the user, such as behavior or culture [65], whilst Germanakos et al. [36] refer to models of adaptation that involve emotional parameters. Thus, in addition to these factors impacting upon the learning experience as already discussed, it seems likely that they could—and should—make an important contribution to user models. Learning is clearly affected by a number of determinants and further work into these aforementioned aspects of users might provide crucial insights into what is effective in terms of providing personalization.

Another factor to take into account is the estimated effect size of learning styles. The low partial eta-squared scores in the aforementioned case studies suggest that effect size of learning style in these experiments was very small and this, in combination with numerous other variables (such as the limited gains in knowledge shown by pupils engaging with DEUS), would anticipate it being very difficult to find any results approaching statistical significance. The power of a particular phenomenon will determine what range of sample sizes and extent of the intervention needed to be able to find any statistically significant differences. Whilst the sample sizes were adequate in these experiments, it is likely that the amount of learning that took place and the time frames involved in both studies were not large enough to reveal any significant difference, if the effect size for learning styles is indeed so small.

Lastly, it is worth noting that learning style theory is not suggested as a replacement for user modeling based on knowledge but as a complementary improvement to these

existing user models. In addition, it is not suggested that any one particular learning style/preference is better than any other (regardless of the typology used), nor that learners should learn only in their preferred style. Indeed, it may be beneficial to a learner to study in a nonpreferred learning style for some of the time, since they will develop compensatory skills from this nonoptimal situation.

7.2.2 Learning as a Process

The second issue is a problem that is common to most educational research: the development of learning. The factors that affect learning are extremely complex; there are many different influences that, in combination or individually, can affect how people learn. Examples of such factors include IQ (itself influenced by several factors); socio-economic status; motivation; time and effort. In addition, there may be a variety of other distractions surrounding personal and social activities. All of these factors can affect the circumstances under which somebody learns [54], [67].

The complexity of learning can be described as a “wicked problem,” a term applied to problems that often have changing constraints and resources, where there is no straightforward solution [68]. Under such circumstances, it can be difficult to predict the outcome of certain interventions; it is possible to obtain a different result than before, even when the same intervention is carried out with in an identical study.

Somewhat related to “wicked problems” is the compounding phenomenon of the “butterfly effect.” This is an aspect of chaos theory that provides a metaphor for sensitive dependence on initial conditions; the name refers to the idea that the movement of a butterfly’s wings might create a slight disturbance in the atmosphere and indirectly cause a tornado to appear (or, conversely, prevent one from occurring) [42]. It is possible that projects involving educational research might well suffer from the butterfly effect, where similar starting conditions with very minor adjustments might result in widely differing outcomes. Complexity theory provides a more formal integration of the butterfly effect into educational research, where it has been used to study the adaptation of schools to their environment (i.e., the resources available, strategies employed, limitations imposed by the government or other educational bodies, etc.) [54]. Complexity theory, when applied to personalization for computer-based learning, suggests that user modeling is only one small component when taking into account the wider factors that might help provide for a more effective learning experience, such as making computer-based resources more widely available or examining the effect of computers in the classroom on pupil motivation, for example.

8 FURTHER WORK

There are several aspects of this research that merit further investigation, building on the studies carried out thus far and also introducing new ideas for consideration.

First, it seems sensible that before learning styles can be judged as inadequate for effective Personalized learning, further studies should be carried out with them to examine how they affect groups of truly homogeneous users. The

findings from this research are somewhat limited due to the simplicity of the user trials; it could be argued that this was a very narrow, “snapshot” view of what is a very broad and complex field. Future research should control for as many factors as possible, thus reducing the variance of initial conditions and lessening the “butterfly effect,” or the chance of a Type 2 error, where scientists are too cautious in their methodology or analysis and present negative findings when they should actually have been positive [59]. It is envisaged that age, gender, measures of intelligence/personality type, and motivation should all be as uniform as possible within a specific cultural group. This might cause practical problems where the number of potential test subjects is small although it does allow for much stricter management of controlled variables.

Second, it is suggested that further work be carried out with data analysis of Web log files in order to determine useful browsing patterns. This suggestion follows a later user trial carried out with WHURLE-LS and described in more detail in [6]. This study itself built upon work carried out by researchers who have elicited information for the user model from such browsing behavior: Hynecosum deduces the user’s experience level from their browsing patterns and HYPERCASE uses similar information to infer the user’s didactic goal [12], [52]. However, the information resulting from the amount of time spent on a node or the number of visits to that node is somewhat limited and does not yield reliable conclusions about the user’s intentions since there is no guarantee that the user has actively engaged with the content in each node [12]. Thus, the proposal is not to use this data to feed back into the user model, but rather to compare the patterns of navigation of one user to that of another and use commonly traversed paths as “recommended” routes for others, if guidance is requested by the user. In this way, no explicit user model is required and no causality is inferred as to which aspects of users might affect how the adaptivity evolves. It is a very simple idea that caters for the complexity of factors that contribute to the learning process without having to specifically state which of them is being addressed. It also removes the issue of asking the user for explicit information about their preferences (which might change at short notice and across different domains) and, hence, does not require continual updating. It is possible that this approach could be used to investigate which aspects of the user model result in certain navigational patterns, if this information were readily available. However, this would only establish correlations, which would then require further, more formal user trials to ascertain the relative benefits of using these user characteristics as models for adaptation. Nevertheless, this is potentially a very powerful method of providing adaptation for computer-based learning, once some established or suggested paths have been constructed. Brusilovsky suggests that these “nonsymbolic” approaches, which include case-based reasoning and neural networks, may help in providing adaptation decisions where no particular rules are available [9].

Similar recommender systems have already been used for providing personalization, typically seen in e-commerce situations such as Amazon.com [72], where users are shown

lists of other products bought by people who have just purchased the same item as them. Users might also be guided toward other products by the same author/musician or a best-seller in the same category of that product. It would be intriguing to investigate how recommender systems could be used to collect and analyze implicit data [56] to provide adaptation for educational content. Some interesting work has already been carried out by Plua and Jameson [66], Fok and Ip [35], Papanikolaou et al. [61], and Wang et al. [82] and these studies could prove a valuable starting point for future research.

Third, in this era of evolving technology and expanding networking capabilities, it seems logical to pursue the concept of device-based adaptation. Thus, users interacting with mobile devices such as laptops and “palmtops” or mobile phones would be expected to have different requirements and preferences from those interacting with materials on a desktop computer or a device with a bigger screen. Usability is already an important aspect of designing for small screens and it is possible that some of the modality (visual/verbal) or structure-based (sequential/global) aspects of learning style might influence the thinking behind the user interfaces for educational content displayed on such devices. The work by Dagger et al. [28] provide a fascinating introduction to this kind of adaptation and demonstrates its integration into the Adaptive Personalized eLearning Service (APeLS) system [26].

Fourth, a new framework is suggested to address an aspect of learning that has so far been largely neglected by AEH systems. The omission of social interaction in adaptive systems is possibly a critical one, since interpersonal relationships contribute a vital part of the learning process [31], [70]. Using the example of social networking Websites (which have augmented and enhanced communication and help construct distributed communities), it seems likely that similar techniques could be integrated into an AEH system, providing the social interaction that so many computer-based learning systems currently lack. Adaptation could then include interpersonal aspects, whereby users could prefer to interact with a small number of individuals or as part of a larger group, depending on their preferences.

9 CONCLUSIONS

It is clear that the field of learning style application in AEH is a highly complex and somewhat controversial area of research and one that has no quick answers. There does not seem to be any particular evidence to invalidate this area of research and any work carried out by others should not be dismissed out of hand; however, it does seem that personalization to show a statistically significant benefit in educational systems is much harder to create than first envisaged. A crucial aspect of this research was to exemplify a quality-assured, rigorous, and strongly scientific approach in the field of adaptive hypermedia so that the findings would be based on sound evidence gleaned from carefully controlled user trials and taking into account the many issues surrounding user variability and learning style theory. We would like to inspire continued debate amongst academics and practitioners, including readdressing the issue of learning styles for computer-based learning

and whether, in fact, they should be used for effective personalization.

However, until more evidence is acquired (for example, from more extensive user trials and/or user models), it is difficult to draw firm conclusions about the efficacy and validity of using learning styles as means of adaptation for computer-based learning. The lack of any kind of correlation seen in the case studies presented here might be a particular characteristic of those studies; however, it is possible (maybe even probable) that it is indicative of a universal pattern.

The issues raised in Section 7, along with many others, remain unanswered for the time being. It is evident though, that whatever future research is done, there is a clear and pressing need to produce quality-assured studies with as much quantitative and qualitative evidence as possible, in order to help us answer at least some of these questions.

ACKNOWLEDGMENTS

The authors wish to thank Helen Ashman, Shaaron Ainsworth, Vincent Wade, Peter Blanchfield, Craig Stewart, and Cees Van der Eijk for all their advice and useful comments at various stages of this research. They are also grateful to those who participated in the user trials. This research was supported by a PhD scholarship from the School of Computer Science at the University of Nottingham.

REFERENCES

- [1] S.E. Ainsworth, "DeFT: A Conceptual Framework for Learning with Multiple Representations," *Learning and Instruction*, vol. 16, pp. 183-198, 2006.
- [2] N. Bajraktarevic, W. Hall, and P. Fullick, "ILASH: Incorporating Learning Strategies in Hypermedia," *Proc. 14th ACM Conf. Hypertext and Hypermedia (HT '03)*, 2003.
- [3] N. Bajraktarevic, W. Hall, and P. Fullick, "Incorporating Learning Styles in Hypermedia Environment: Empirical Evaluation," *Proc. 14th ACM Conf. Hypertext and Hypermedia (HT '03)*, 2003.
- [4] I. Beaumont, "User Modelling in the Interactive Anatomy Tutoring System Anatom-Tutor," *User Modeling and User-Adapted Interaction*, vol. 4, pp. 21-45, 1994.
- [5] C. Boyle and A.O. Encarnacion, "Metadoc: An Adaptive Hypertext Reading System," *User Modeling and User-Adapted Interaction*, vol. 4, pp. 1-19, 1994.
- [6] E. Brown, T. Brailsford, T. Fisher, and C. van der Eijk, "Revealing the Hidden Rationality of User Browsing Behaviour," *Proc. 18th ACM Conf. Hypertext and Hypermedia (Hypertext '07)*, 2007.
- [7] E. Brown, T. Brailsford, T. Fisher, A. Moore, and H. Ashman, "Reappraising Cognitive Styles in Adaptive Web Applications," *Proc. 15th Int'l World Wide Web Conf. (WWW '06)*, 2006.
- [8] E. Brown, T. Fisher, and T. Brailsford, "Real Users, Real Results: Examining the Limitations of Learning Styles within AEH," *Proc. 18th ACM Conf. Hypertext and Hypermedia (Hypertext '07)*, 2007.
- [9] P. Brusilovsky, "Adaptive Hypermedia," *User Modeling and User Adapted Interaction*, Ten Year Anniversary Issue, A. Kobsa, ed., vol. 11, pp. 87-110, 2001.
- [10] P. Brusilovsky, "Efficient Techniques for Adaptive Hypermedia," *Intelligent Hypertext*, pp. 12-30, Springer, 1997.
- [11] P. Brusilovsky, "Intelligent Tutoring Systems for World-Wide Web," *Proc. Third Int'l World Wide Web Conf. (WWW '95)*, 1995.
- [12] P. Brusilovsky, "Methods and Techniques of Adaptive Hypermedia," *User Modeling and User Adapted Interaction*, vol. 6, pp. 87-129, 1996.
- [13] P. Brusilovsky and J. Eklund, "A Study of User Model Based Link Annotation in Educational Hypermedia," *J. Universal Computer Science*, vol. 4, pp. 429-448, 1998.
- [14] P. Brusilovsky, C. Karagiannidis, and D. Sampson, "The Benefits of Layered Evaluation of Adaptive Applications and Services," *Proc. Workshop Empirical Evaluations of Adaptive Systems at the Eighth Int'l Conf. User Modeling (UIM '01)*, 2001.
- [15] P. Brusilovsky and L. Pesin, "ISIS-Tutor: An Adaptive Hypertext Learning Environment," *Proc. Japan-CIS Symp. Knowledge-Based Software Eng. (JCKBSE '94)*, 1994.
- [16] P. Brusilovsky and L. Pesin, "Visual Annotation of Links in Adaptive Hypermedia," *Proc. Conf. Human Factors in Computing Systems (CHI '95)*, 1995.
- [17] P. Brusilovsky, L. Pesin, and M. Zyryanov, "Towards an Adaptive Hypermedia Component for an Intelligent Learning Environment," *Human-Computer Interaction*, L.J. Bass, J. Gornostae, and C. Unger, eds., pp. 348-358, Springer-Verlag, 1993.
- [18] P. Brusilovsky, E. Schwarz, and G. Weber, "Elm-Art: An Intelligent Tutoring System on World Wide Web," *Intelligent Tutoring Systems*, C. Frasson, G. Gauthier, and A. Lesgold, eds., pp. 261-269, Springer Verlag, 1996.
- [19] P. Brusilovsky, E. Schwarz, and G. Weber, "A Tool for Developing Adaptive Electronic Textbooks on WWW," *Proc. World Conf. Web Soc. (WebNet '96)*, 1996.
- [20] R.M. Carro, E. Pulido, and P. Rodríguez, "TANGOW: Task-Based Adaptive Learner Guidance on the WWW," *Proc. Second Workshop Adaptive Systems and User Modeling on the World Wide Web*, 1999.
- [21] C.A. Carver, R.A. Howard, and W.D. Lane, "Addressing Different Learning Styles through Course Hypermedia," *IEEE Trans. Education*, vol. 42, pp. 33-38, 1999.
- [22] D.N. Chin, "Empirical Evaluation of User Models and User-Adapted Systems," *User Modeling and User-Adapted Interaction*, vol. 11, pp. 181-194, 2001.
- [23] F. Coffield, D. Moseley, E. Hall, and K. Ecclestone, *Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review*. Learning & Skills Research Centre, 2004.
- [24] F. Coffield, D. Moseley, E. Hall, and K. Ecclestone, *Should We Be Using Learning Styles? What Research Has to Say to Practice*. Learning & Skills Research Centre, 2004.
- [25] O. Conlan, D. Dagger, and V. Wade, "Towards a Standards-Based Approach to E-Learning Personalization Using Reusable Learning Objects," *Proc. World Conf. E-Learning in Corporate, Gov't, Healthcare, and Higher Education (E-Learn '02)*, 2002.
- [26] O. Conlan, V. Wade, C. Bruen, and M. Gargan, "Multi-Model, Metadata Driven Approach to Adaptive Hypermedia Services for Personalized Elearning," *Proc. Second Int'l Conf. Adaptive Hypermedia and Adaptive Web-Based Systems*, 2002.
- [27] C.E. Cornett, *What You Should Know About Learning Styles*. Phi Delta Kappa, 1983.
- [28] D. Dagger, O. Conlan, and V. Wade, "Towards 'Anytime, Anywhere' Learning: The Role and Realization of Dynamic Terminal Personalization in Adaptive Elearning," *Proc. 15th World Conf. Educational Multimedia, Hypermedia, and Telecomm. (ED-MEDIA 2003)*, 2003.
- [29] P. De Bra and L. Calvi, "Aha! An Open Adaptive Hypermedia Architecture," *New Rev. Hypermedia and Multimedia*, vol. 4, pp. 115-139, 1998.
- [30] J. Debes, "Some Foundations of Visual Literacy," *Audio Visual Instruction*, vol. 13, pp. 961-964, 1968.
- [31] W. Doise, G. Mugny, and A.-N. Perret-Clermont, "Social Interaction and the Development of Cognitive Operations," *European J. Social Psychology*, vol. 5, pp. 367-383, 1975.
- [32] R. Dunn, "The Dunn and Dunn Learning Style Model and Its Theoretical Cornerstone," *Synthesis of the Dunn and Dunn Learning Styles Model Research: Who, What, When, Where and So What—The Dunn and Dunn Learning Styles Model and Its Theoretical Cornerstone*, R. Dunn and S. Griggs, eds., St John's Univ., 2003.
- [33] R.M. Felder and B.A. Soloman, "Index of Learning Styles," <http://www.ncsu.edu/felder-public/ILSpage.html>, 2009.
- [34] S. Ferrandino, A. Negro, and V. Scarano, "Cheops: Adaptive Hypermedia on the World Wide Web," *Proc. European Workshop Interactive Distributed Multimedia Systems and Telecomm. Services (IDMS '97)*, 1997.
- [35] A.W.P. Fok and H.H.S. Ip, "Personalized Education: An Exploratory Study of Learning Pedagogies in Relation to Personalization Technologies," *Advances in Web-Based Learning: ICWL 2004*, W. Liu, L. Qing, and Y. Shi, eds., pp. 407-415, Springer-Verlag, 2004.
- [36] P. Germanakos, N. Tsianos, Z. Lekkas, C. Mourlas, M. Belk, and G. Samaras, "A Semantic Approach of an Adaptive and Personalized Web-Based Learning Content—The Case of Adap-

- tive Web," *Proc. Second Int'l Workshop Semantic Media Adaptation and Personalization (SMAP '07)*, 2007.
- [37] J.E. Gilbert and C.Y. Han, "Adapting Instruction in Search of 'A Significant Difference,'" *J. Network and Computer Applications*, vol. 22, pp. 149-160, 1999.
- [38] R. Gillespie, *Manufacturing Knowledge: A History of the Hawthorne Experiments*. Cambridge Univ. Press, 1991.
- [39] C. Giorgis, N.J. Johnson, A. Bonomo, C. Colbert, A. Conner, G. Kauffman, and D. Kulesza, "Visual Literacy," *Reading Teacher*, vol. 53, pp. 146-153, 1999.
- [40] M. Grigoriadou, K. Papanikolaou, H. Kornilakis, and G. Magoulas, "Inspire: An Intelligent System for Personalized Instruction in a Remote Environment," *Proc. Eighth Int'l Conf. User Modeling (UM '01)*, 2001.
- [41] D. Hargreaves, *About Learning: Report of the Learning Working Group*. DEMOS, 1995.
- [42] R.C. Hilborn, "Sea Gulls, Butterflies, and Grasshoppers: A Brief History of the Butterfly Effect in Nonlinear Dynamics," *Am. J. Physics*, vol. 72, pp. 425-427, 2004.
- [43] H. Hohl, H.-D. Böcker, and R. Gunzenhäuser, "Hypadapter: An Adaptive Hypertext System for Exploratory Learning and Programming," *User Models and User Adapted Interaction*, vol. 6, pp. 131-156, 1996.
- [44] P. Honey and A. Mumford, *The Manual of Learning Styles*. Peter Honey Publications, 1992.
- [45] J. Kay and B. Kummerfeld, "Adaptive Hypertext for Individualised Instruction," *Proc. Workshop Adaptive Hypertext and Hypermedia at the Fourth Int'l Conf. User Modeling (UM '94)*, 1994.
- [46] J.W. Keefe, "Learning Style: An Overview," *Student Learning Styles: Diagnosing and Prescribing Programs*, J.W. Keefe, ed., NASSP, 1979.
- [47] D. Kelly and B. Tangney, "Adapting to Intelligence Profile in an Adaptive Educational System," *Interacting with Computers*, vol. 18, pp. 385-409, 2006.
- [48] M. Kozhevnikov, M. Hegarty, and R.E. Mayer, "Revising the Visualiser-Verbalizer Dimension: Evidence for Two Types of Visualizers," *Cognition and Instruction*, vol. 20, pp. 47-77, 2002.
- [49] G. Kress, *Literacy in the New Media Age*. Routledge, 2003.
- [50] E. Melis and R. Monthienvichienchai, "They Call It Learning Style but It's So Much More," *Proc. World Conf. E-Learning in Corporate, Gov't, Healthcare, and Higher Education (E-Learn '04)*, 2004.
- [51] D.M. Mertens, *Research Methods in Education and Psychology: Integrating Diversity with Quantitative and Qualitative Approaches*. Sage, 1998.
- [52] A. Micarelli and F. Sciarrone, "A Case-Based Toolbox for Guided Hypermedia Navigation," *Proc. Fifth Int'l Conf. User Modeling (UM '96)*, 1996.
- [53] R. Moreno and R.E. Mayer, "Cognitive Principles of Multimedia Learning: The Role of Modality and Contiguity," *J. Educational Psychology*, vol. 91, pp. 358-368, 1999.
- [54] K. Morrison, *School Leadership and Complexity Theory*. RoutledgeFalmer, 2002.
- [55] J.O. Nunnally, *Psychometric Theory*. McGraw-Hill, 1978.
- [56] D.W. Oard and J. Kim, "Implicit Feedback for Recommender Systems," *Proc. AAAI Workshop Recommender Systems*, 1998.
- [57] A. Ortigosa and R.M. Carro, "The Continuous Empirical Evaluation Approach: Evaluating Adaptive Web-Based Courses," *User Modeling 2003*, pp. 163-168, Springer, 2003.
- [58] A. Paivio, *Mental Representations*. Oxford Univ. Press, 1986.
- [59] J. Pallant, *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows (Version 12)*, second ed., Open Univ. Press (McGraw-Hill Education), 2005.
- [60] K.A. Papanikolaou, M. Grigoriadou, H. Kornilakis, and G.D. Magoulas, "Personalizing the Interaction in a Web-Based Educational Hypermedia System: The Case of Inspire," *User Modeling and User-Adapted Interaction*, vol. 13, pp. 213-267, 2003.
- [61] K.A. Papanikolaou, A. Mabbott, S. Bull, and M. Grigoriadou, "Designing Learner-Controlled Educational Interactions Based on Learning/Cognitive Style and Learner Behaviour," *Interacting with Computers*, vol. 18, pp. 356-384, 2006.
- [62] P. Paredes and P. Rodriguez, "A Mixed Approach to Modelling Learning Styles in Adaptive Educational Hypermedia," *Proc. Third IASTED Conf. Web-Based Education (WBE '04)*, 2004.
- [63] T. Pérez, P. Lopistéguy, J. Gutiérrez, and I. Usandizaga, "Hypertutor: From Hypermedia to Intelligent Adaptive Hypermedia," *Proc. Seventh World Conf. Educational Multimedia and Hypermedia (ED-MEDIA '95)*, 1995.
- [64] J.K. Pinto, M.A. Geiger, and E.J. Boyle, "A Three Year Longitudinal Study of Changes in Student Learning Styles," *J. College Student Development*, vol. 35, pp. 113-119, 1994.
- [65] J.L. Plass, D.M. Chun, R.E. Mayer, and D. Leutner, "Supporting Visual and Verbal Learning Preferences in a Second-Language Multimedia Learning Environment," *J. Educational Psychology*, vol. 90, pp. 25-36, 1998.
- [66] C. Plua and A. Jameson, "Collaborative Preference Elicitation in a Group Travel Recommender System," *Proc. AH 2002 Workshop Recommendation and Personalization in eCommerce*, 2002.
- [67] R.J. Riding, *School Learning and Cognitive Style*. David Fulton, 2002.
- [68] H. Rittel and M. Webber, *Dilemmas in a General Theory of Planning*, vol. 4. Elsevier Scientific Publishing Company, Inc., 1973.
- [69] C. Robson, *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Blackwell Publishing, 2002.
- [70] G. Salomon and D.N. Perkins, "Individual and Social Aspects of Learning," *Rev. Research in Education*, vol. 23, pp. 1-24, 1998.
- [71] M. Scaife and Y. Rogers, "External Cognition: How Do Graphical Representations Work?" *Int'l J. Human-Computer Studies*, vol. 45, pp. 185-213, 1996.
- [72] J.B. Schafer, J. Konstan, and J. Riedi, "Recommender Systems in E-Commerce," *Proc. First ACM Conf. Electronic Commerce*, 1999.
- [73] C. Stangor, *Research Methods for the Behavioral Sciences*. Houghton Mifflin, 1998.
- [74] N. Stash, A. Cristea, and P. De Bra, "Adaptation to Learning Styles in E-Learning: Approach Evaluation," *Proc. World Conf. E-Learning in Corporate, Gov't, Healthcare & Higher Education (E-Learn '06)*, 2006.
- [75] N. Stash, A. Cristea, and P. De Bra, "Learning Styles Adaptation Language for Adaptive Hypermedia," *Adaptive Hypermedia and Adaptive Web-Based Systems*, pp. 323-327, Springer, 2006.
- [76] N. Stash, A. Cristea, and P. De Bra, "Authoring of Learning Styles in Adaptive Hypermedia: Problems and Solutions," *Proc. 13th Int'l World Wide Web Conf.*, 2004.
- [77] M.K. Stern and B.P. Woolf, "Adaptive Content in an Online Lecture System," *Proc. Int'l Conf. Adaptive Hypermedia and Adaptive Web-Based Systems*, 2000.
- [78] J. Summerville, "Role of Awareness of Cognitive Style in Hypermedia," *Int'l J. Educational Technology*, vol. 1, 1999.
- [79] E. Triantafyllou, A. Pomportsis, and S. Demetriadis, "The Design and the Formative Evaluation of an Adaptive Educational System Based on Cognitive Styles," *Computers and Education*, vol. 41, pp. 87-103, 2003.
- [80] E. Triantafyllou, A. Pomportsis, and E. Georgiadou, "AES-CS: Adaptive Educational System Based on Cognitive Styles," *Proc. Second Int'l Conf. Adaptive Hypermedia and Adaptive Web-Based Systems (AH '02)*, 2002.
- [81] S.P. Tyerman, P.J. Woods, and J.R. Warren, "Loop Tutor and Hypertutor: Experiences with Adaptive Tutoring Systems," *Proc. Australian and New Zealand Conf. Intelligent Information Systems*, 1996.
- [82] D. Wang, Y. Bao, G. Yu, and G. Wang, "Using Page Classification and Association Rule Mining for Personalized Recommendation in Distance Learning," *Advances in Web-Based Learning: ICWL 2002*, J. Fong, C.T. Cheung, H.V. Leong, and Q. Li, eds., pp. 363-377, Springer-Verlag, 2002.
- [83] A. Weber, "The Hawthorne Works," *Assembly Magazine*, 2002.
- [84] G. Weber, H.-C. Kuhl, and S. Weibelzahl, "Developing Adaptive Internet Based Courses with the Authoring System Netcoach," *Hypermedia: Openness, Structural Awareness, and Adaptivity*, S. Reich, M. Tzagarakis, and P. De Bra, eds., pp. 226-238, Springer, 2001.
- [85] G. Weber and M. Specht, "User Modeling and Adaptive Navigation Support in WWW-Based Tutoring Systems," *Proc. Sixth Int'l Conf. User Modeling*, 1997.
- [86] S. Weibelzahl, "Problems and Pitfalls in the Evaluation of Adaptive Systems," *Adaptable and Adaptive Hypermedia Systems*, S.Y. Chen and G.D. Magoulas, eds., pp. 285-299, IRM Press, 2005.
- [87] S. Weibelzahl, S. Lippitsch, and G. Weber, "Advantages, Opportunities, and Limits of Empirical Evaluations: Evaluating Adaptive Systems," *Künstliche Intelligenz*, vol. 3, pp. 17-20, 2002.
- [88] S. Weibelzahl and G. Weber, "Evaluating the Inference Mechanism of Adaptive Learning Systems," *User Modeling 2003*, p. 148, Springer, 2003.

- [89] H.A. Witkin and D.R. Goodenough, *Cognitive Style—Essence and Origins: Field Dependence and Field Independence*. Int'l Univ., 1981.
- [90] C. Wolf, "Construction of an Adaptive E-Learning Environment to Address Learning Styles and an Investigation of the Effect of Media Choice," *School of Education*, p. 225, RMIT Univ., 2007.
- [91] M.R. Zakaria, A. Moore, H. Ashman, C. Stewart, and T. Brailsford, "The Hybrid Model for Adaptive Educational Hypermedia," *Proc. Second Int'l Conf. Adaptive Hypermedia and Adaptive Web-Based Systems (AH '02)*, 2002.
- [92] R. Zeiliger, "Adaptive Testing: Contribution of the Shiva Model," *Item Banking: Interactive Testing and Self-Assessment*, D. Leclercq and J. Bruno, eds., NATO ASF Series F, vol. 112, pp. 54-65, Springer-Verlag, 1993.



Elizabeth J. Brown received the PhD degree in computer science in 2008 for her thesis, "The Use of Learning Styles in Adaptive Hypermedia." Prior to this, she worked as a secondary school teacher and then at the University of Nottingham as a research assistant with an educational multimedia research group. She joined the School of Computer Science in 2001 as a teaching assistant and is currently a research fellow there. She is a reviewer for several conferences and journals and has published a number of papers, including two that have won Best Paper awards in 2006 and 2007. Her research interests lie predominantly in educational technology and she is passionate about exploring new ways of working and interacting with computers and other devices to inform learning and pedagogy. She is a member of the ACM, the IEEE, and IEEE Women in Engineering.



Timothy J. Brailsford received the PhD degree in biology in 1988 and subsequently spent several years doing biomedical research at the University of Nottingham. He is a lecturer in the School of Computer Science at the University of Nottingham and a member of the Learning Sciences Research Institute (LSRI). After learning to program mainframes in order to model biological systems and analyze data, he became intrigued by the potential of using computers to teach the fundamentals of science. This led him to become an early adopter of educational multimedia, and since the early 1990s, he has designed many educational hypermedia systems and experimented with a wide variety of approaches to technology-based learning. In 2007, he was conferred a "Lord Dearing Award" by the University of Nottingham in recognition of his services to teaching and learning within the University. He has research interests in various areas of educational technology—including adaptation and personalization, novel hyperstructures, and contextually-aware learning material.



Tony Fisher is a lecturer in education at the University of Nottingham, where he is also a member of the University's Learning Sciences Research Institute (LSRI). His research interests are in the use of new technologies in education and he has participated in several national research projects to evaluate the use and potential of new educational technologies, the most recent being an investigation of Web 2.0 technologies and learning in the secondary age-range. He teaches on higher degree courses in face-to-face, online, and blended modes. He joined the School of Education with 18 years of experience as a secondary school teacher, having held various positions of responsibility. He was also an advisory teacher for two years with Nottinghamshire Local Education Authority.



Adam Moore received the PhD degree in computational chemistry in 1998 and has since worked for the University of Nottingham on several projects in areas such as educational technologies, problem-based learning, e-learning, novel data structures, and bioinformatics. He is currently the project manager of where: now, a geospatial knowledge transfer hub at the University of Nottingham. He was previously a visiting fellow in e-learning at Hong Kong University. He has published in more than 30 peer reviewed publications and is a contributing author to the "HCI and the Web" chapter in *The Human-Computer Interaction Handbook*. He is currently interested in examining the role that the differing capabilities present in mobile devices play in the production of ad hoc experiential media.