Tech., Inst., Cognition and Learning, Vol. 5, pp. 69-89
Reprints available directly from the publisher
Photocopying permitted by license only

The Role of User Models in CAT: Exploring adaptivity variables

DR. EVANGELOS TRIANTAFILLOU[1*], DR. ELISABETH GEORGIADOU[2]
AND DR. ANASTASIOS ECONOMIDES[3]

[1]*Computer Science Department, Aristotle University of Thessaloniki, Dodekanisou 21 Thessaloniki 55131, Greece*

[2]*Department of Multimedia - Graphic Arts, Open University, Greece, Papakiriazi 12 Thessaloniki 54645, Greece, elisag@otenet.gr*

[3]*Department of Computer Networks, University of Macedonia, Greece, Egnatia 156,Thessaloniki 54006, Greece, economid@uom.gr*

This paper reviews all different variables that have been used in adaptive educational systems and then discuss their potential use to a hypothetical user model for Computer Adaptive Testing. From all the variables presented that triggers adaptation, modelling of '*knowledge on the domain presented'*, '*background-experience', 'preferences', 'personal data*' and *'mental models'* can produce more efficient CATs in terms of time, as fewer items will be needed to assess performance. Moreover, it could affect items' quality, since items can be more complex taking into account user characteristics, resulting in testing sessions that can contribute to the learning process and not merely arrange examinees on a problem complexity scale.

Keywords: CAT, user model, adaptivity variables, user modelling, computerised assessment, multiple modelling, measure performance.

1. INTRODUCTION

The use of computer-based testing has expanded rapidly the last two decades mainly due to advancements in communication and information technology that

*Corresponding author: vtrianta@csd.auth.gr
Electronic Link: http://www.edutech.gr/public/user_model_CAT.pdf

made computers with high power and speed affordable and effortlessly connected to broad bandwidth networks. Computer delivery of tests became feasible for licensure, certification and admission. Large-scale high-stake testing programs use, optionally or obligatory, computers for testing, such as MicroSoft-Certified Professional Program, Graduate Record Exam, Graduate Management Admission Test etc.

Computer-based delivery of tests has several advantages. Testing on demand can be facilitated, so as an examinee can take the test whenever and wherever s/he is ready. Multiple media can be used to create innovative item formats and more realistic testing environments. Other advantages are flexibility of test management; immediate availability of scores; increased test security; increased motivation etc. Moreover, with the use of computerized adaptive testing (CAT), a special case of computer-based testing, the statistical accuracy of test scores can be increased (Linden & Glass, 2003).

During the assessment, examinees might feel discouraged if the items are too difficult or, on the other hand, might lose interest if the items are too easy. The basic idea of an adaptive testing is to imitate automatically a good quality oral assessment where the examiner has the opportunity to change constantly the item's difficulty level according to examinee's performance.

In CAT the items presented are selected taking into account the examinee's individual performance during the test, or in other words how each examinee answered previous items. If the examinee answers the item presented correctly then a more difficult one follows. On the other hand, if the examinee answers the item incorrectly, an easier item appears. In this way low-ability examinees will be presented with relatively easy items, while high-ability ones with more difficult.

The computer continuously re-evaluates the ability of the examinee until the accuracy of the estimate reaches a statistically acceptable level or when some limit is reached such as a maximum number of test items is presented. All items in CAT are included in the item pool: a collection of test items with a full range of proficiency levels. The score is determined from the level of the difficulty, and as a result, while all examinees may answer the same percentage of questions correctly the high ability ones will get a better score as they answer correctly more difficult items.

The main advantage of CAT over any other computerized based test is efficiency. Since fewer questions are needed to achieve a statistically acceptable level of accuracy, significantly less time is needed to administer a CAT compared to a linear Computerized Based Test. CAT can reduce testing time by more than 50% while maintaining the same level of reliability.

Since the mid 80's when the first CAT became operational (Armed Services Vocational Aptitude Battery for US Department of Defense account) using adaptive techniques to administer multiple-choice items, much research and many technical challenges make possible new assessment tools. Dragsow (2002) classify them under the terms visualization, audition, and interaction.

Visualization

Technical abilities of monitors can present true color images with high resolution. Moreover 3D technology make possible for users to rotate and pan in and out objects in three dimensions. Ackerman et al. (1999) by exploiting these features developed a dermatological disorder examination for medical students. Skin disorders with photographic clarity are presented to examinees that can pan in and out to view them in detail. Case history is also presented and the students must select the matching one from a list of possible disorders. Many domains can benefit from improving visualization: identification of plants in dendrology course, of rocks in a geology course, of paintings in an art course etc.

Audition

Presentation of audio material is possible and inexpensive with today's computers. Use of audition facilitates the assessment of a wide variety of skills. For example, simulations of interactions with customers in telephone service centers can be presented to job applicants. Another example is Vispoel's (1999) musical aptitude test where musical sequences of varying complexity are presented to the examinees that have to identify a matching sequence.

Interaction

Collaboration is important in workplace and as a result social and interpersonal skills need to be assessed in many cases. Full motion video can be presented by the computer make possible the assessment of skills related to human interactions. For example, assess leadership skills that predict job performance of managers (Donovan et al.1998) or assess applicant's ability to deal with conflict in the workplace (Olson-Buchanan et al. 1998), or social skills of students to interact with their peers etc. Another example is the Test of Listening Communication Skills from Educational Testing Services (www.ets.org). This is an interactive multimedia test that presents real communication situations on video with actors and examinees are called to evaluate various ways to deal with the communication situations.

Current research explores the use of *Games and Simulations* as assessment tools for various constructs, such as learning ability; judgment and decision making; adaptability; emotional intelligence; metacognition; leadership; and attitudes, values, and beliefs (ETS, 2005).

The administration of computerized adaptive tests makes possible even more complex kinds of tests. Moreover, analysis of the results can go deeper than just calculate the right and wrong answers. Contemporary research in profile scoring involves the design and generation of enhanced score reports focus on the interpretation of score report components, feedback about skills, and educational advice (ETS, 2005).

Educational research constantly highlights the need for tools that broaden the understanding of individual differences and provides direction for fairness and equity initiatives. So far the computer has to select the items from a fixed item pool. However, item generation tools that will further increase the efficiency of test creation process is near to realization (ETS, 2005). Soon, items can be generated automatically taking into consideration several variables/specifications during the testing session. As a result new items with optimal properties will be presented to each examinee at each stage of the test assisting the potential for improved individual performance.

So far, most CAT programs merely model the level of problem complexity that an examinee can deal with and do not adapt to any other variable that defines the user as an individual. However, research on adaptive educational hypermedia systems has identified a number of variables that can prompt adaptivity. These variables are included in the user model that is embedded in such systems. The vast majority of CATs do not utilize a user model since their goal is to arrange examinees on a problem complexity scale that is relevant for graduation/admission decisions.

Current research in CAT is not limited to educational admissions, yet, focus on applications, in small and large scale, that address self-assessment, training, employment, teacher professional development for schools, industry, military etc. Moreover, as dynamic item generation tools are near to realization it is important to extend CAT's functionality to include as many variables as possible that refer to individual differences for improved individual performance and more efficient test delivery.

This paper is aiming to review all different variables that have been used in adaptive educational systems and then discuss their potential use to a hypothetical user model for Computer Adaptive Testing. The objective of this effort is to provide researchers designers, and developers of Computerized Adaptive Tests a perspective to utilize research outcomes from adaptive educational hypermedia to computerized adaptive testing. However, before that it is important to look at what a user model is and what is its purpose on a hypermedia system.

2. USER MODEL

The user model of an adaptive hypermedia system is defined as the system's representation of certain user characteristics and attitudes (Eklund & Zeilinger, 1996; Paiva, Self & Hartley, 1995). The life cycle of a user model includes collecting information about the user, constructing, updating, maintaining and exploiting it. This whole process is called 'user modelling' and its most important phase is the collection of information about the user.

Different techniques are used from systems to acquire such information (Kobsa, Müller, Nill, 1994). These include: (a) stereotypes (categorisation of user subgroups using stereotypes) are incorporated in the system by the system designer at design-time; then the system at runtime has to determine which stereotype is applicable to the current user; (b) user-supplied preferences through interviewing by the system at run-time aiming mainly to initialise the user model; (c) monitoring and analysis of user actions and plan recognition or inference (computed by the system at run-time): moves between pages, way of selecting the page (navigation/query), or sometimes explicit selection of tasks (Hook Karlgren, & Waern, 1995) can be used by the system to infer the user's intention.

User modelling plays an important role in adaptive hypermedia systems as, "adaptive hypermedia is a direction of research on the crossword of hypertext (hypermedia) and user modelling" (De Bra, Brusilovsky & Houben, 1999). A user model is formed by descriptions of what is considered relevant about the actual knowledge and/or aptitudes of a user, providing information for the system environment to adapt itself to the individual user. In case of learning applications, the user model is also called student model. The information kept in the user model of an adaptive hypermedia system belongs to the representation of user's preferences, tasks, goals or interests, domain knowledge, experience, and background (Brusilovsky, 1996a).

Student modelling differs from general user modelling in the diagnosis and representation of refinements on the user knowledge that are important to the adequate performance of an adaptive learning environment. It is more difficult to recognise changes in student knowledge, as there is usually little evidence and the learner's behaviour is difficult to interpret as it responds to creative, unexpected and novel actions (Greer and McCalla, 1994).

The main purposes of user modelling are: (a) to assist a user during learning of a given topic, (b) to offer information adjusted to the user, (c) to adapt the interface to the user, (d) to help a user find information, (e) to give to the user feedback about his/her knowledge, (f) to support collaborative work, and/or (g) to give assistance in the use of the system. (Koch, 2000)

User models can be adaptive, adaptable and dynamic. The user model is called adaptive when updated automatically by the system that monitors user's behaviour like mistakes, navigation patterns etc. In adaptable user models the user decides when the application will be adapted to the newest state of the user model and modify it accordingly. Allowing user to control the adaptivity and to alter the assumptions made by the system is very important as a problem with adaptive systems in general is that they might make wrong adaptations based on guesses they make about the user. "Modeling cannot be anything but a guess if it attempts to model the user's knowledge" (Kay, 1994). Finally, a dynamic user model results from the combination of adaptive and adaptable models. It is suitable in situations where users cannot customize effectively on their own and as a result they need assistance from the system. This approach looks the most promising because it can provide support for novices, help users to customize from the beginning of the program and be selective about what they customize, and finally help users maintain over time the interfaces that resulted from the customization.

Beyond this major division there are many categories of user models depending on several variables that concerns with the nature of the contents, the type of representation and the methods used to initialise, construct and exploit user models.

3. ADAPTIVITY VARIABLES

The variables that user models include can be classified to 'user dependent' that includes those directly related to the user and define him/her as an individual, and to 'user independent' that affect the user indirectly and are related mainly to the context of a user's work with a hypermedia application.

The user dependent variables are: (a) knowledge on the domain presented, (b) background - experience, (c) preferences, (d) interests, (e) individual traits, (f) user personality, (g) mental model, (h) personal data, (i) abilities/disabilities, (j) social-group. On the other hand, the user independent variables are: (a) current goal/task, (b) environment-work, and (c) situation variables. Next, the paper will proceed to examine what the above variables concern with and how they can be represented in the user model.

3.1 Dependent variables

3.1.1 Knowledge on the domain presented

In most of the existing adaptive hypermedia learning environments user's knowledge on the subject articulated appears to be the most used and the most important user characteristic. In reviewing adaptive hypermedia systems, Brusilovky (1996) argues that one third of the systems adapt their interface according to the perceived knowledge of the user. The use of user knowledge requires an understanding of the underlying structure of knowledge that can be defined as the structure of interrelationships between concepts and procedures in a particular domain, which is organized into a unified body of knowledge (Koubek and Salvendy 1991). These elements can be declarative (i.e. definitions, facts, or data) or procedural (i.e. how to). Adaptive hypermedia systems that rely on the user's knowledge variable have to identify the changes in the user's knowledge state and in view of that update the user model (Wu, 2002).

An overlay model or a stereotype user model usually represents user knowledge. Overlay model as a type of knowledge representation was initially developed in the area of intelligent tutoring systems and student modelling (Greer and McCalla, 1994). According to overlay model that is based on the structural model of the subject domain, user's knowledge of a subject is represented as an "overlay" of the domain model. For each concept in the domain model, an individual overlay model stores estimation of the user's knowledge degree of that concept. This estimation is usually presented by twofold concept-values (i.e., known or not known), qualitative states (good - average - poor), or a quantitative value (e.g., the probability that the user knows the concept, one for each domain concept of the task) (Brusilovsky, 1996).

Stereotype model is a more straightforward approach to classify the users. This model distinguishes several "stereotype" classes of users, which have preset values for the domain overlay (e.g. novice, expert). However, this model can be used several times simultaneously in the same system. For example, users can be classified for general concepts into a unified body of knowledge and then for each topic of the organised domains.

Overlay model can represent a user's knowledge of individual topics in a powerful way; however, it is not simple to initialise as stereotype model is as there are several variables to consider. Often, stereotype model is used to initialise overlay models in terms of classifying new users and set initial values for the overlay model.

3.1.2 Background - Experience

Another variable related to users previous general knowledge state is that of background-experience. This twofold variable is not concerned with the user's knowledge on the subject presented in the hypermedia system but it describes all the information related to the user's previous relevant experience outside the subject of the hypermedia system such as his/her familiarity with the information space and the ease of navigation within it. Moreover, it concerns with the user's profession, experience of work in related areas, and the user's point of view and perspective. Usually modelled using stereotype model (e.g. experience stereotype, background stereotype for profession).

3.1.3 Preferences

Preferences are user features that relate to the user's likes and dislikes. This variable describes that a user can prefer some types of nodes and links to others or some parts of a page over others. Moreover, preferences can indicate interface elements such as preferred colours, fonts, navigation ways, etc. User preferences are not assumed by the system; instead the user has to notify the system, directly or indirectly by providing feedback. Usually, the user through checklists can select preferred interface elements. Once the preferences are determined the system generalise the user's preferences and apply them for adaptation in new contexts (Brusilovsky, 1996).

In collaborative systems preferences are modelled numerically and several user preference models can be combined to form a group user model, which collects preferences of a specific group of users. A group model can be a good starting model for a new member of the group that could gradually evolve to user model. Moreover, group models are important in collaboration, as collaborators need to see the same adapted views (Brusilovsky, 1996).

3.1.4 Interests

Interests are a new adaptive variable that recently becomes popular in web-based information retrieval systems. It concerns with the user's long-term interests, and use these in parallel with the user's short-term search goal in order to improve the information filtering and recommendations. Interests can be modelled through navigation monitoring, for example, by observing which links the user visits more often.

3.1.5 Individuals Traits

Apart from the dimension 'background-experience' that defines a user as an individual, 'individual traits' can also prompt adaptation. Individual traits are stable features of a user as well (e.g. personality factors, cognitive factors, learning styles), however they are extracted by specially designed psychological tests unlike 'background-experience' where a simple interview is normally used. Importance of user traits acknowledged, but currently little agreement on which aspects can and should be used to drive adaptation (Brusilovsky, 2001).

3.1.6 Personal data

Personal data, such as gender, age, language, and culture should be taken into account when designing adaptive educational interfaces to optimise learner's potential to benefit from the system's design in terms of knowledge acquisition.

Males and females appear to have different preferences in terms of media presentation, navigation support, attitudes, and information seeking strategies (Magoulas and Dimakopoulos, 2005). For example, an empirical study into gender differences in collaborative web searching reveals that males formulate queries comprising fewer keywords, spent less time on individual pages, click more hypertext links per minute and in general were more active while online than females (Large, Beheshti and Rahman, 2001).

Research suggests that males significantly outperform females in navigating virtual environments. Special navigation techniques (Tan, Robertson, and Czerwinski, 2001) when combined with a large display and wide field of view, appeared to reduce that gender bias. That work has been extended with two navigation studies in order to understand the finding under carefully controlled conditions. The first study replicated the finding that a wide field of view coupled with a large display benefits both male and female users and reduces gender bias. The second study suggested that wide fields of view on a large display were useful to females despite a more densely populated virtual world (Czerwinski, Tan, and Robertson, 2002).

3.1.7 Abilities / Disabilities

People with disabilities often find difficulty to use computer-based systems, as the vast majority of these systems have no design considerations for them. These different users have varying needs regarding content and presentation of the information. For example, information for the blind should be presented in audio mode and a Braille display and speech synthesiser is needed so as to interact with the learning material; information for the deaf should never be in audio format.

3.1.8 User Personality

Murray and Bevan (1985) argue that human-computer interaction would improve if computers were assigned personalities, as the best way for a human to interact with a computer should closely resemble the interaction between two humans. On that view, Richter and Salvendy (1995) compared the performance of introverted and extroverted users using "extroverted" and "introverted" interfaces. The extroverted interface they design had more words, more "fun" pictures, more sounds, bold fonts and exclamation marks than the introverted interface. The subjects used in their empirical study were classified as introverted or extroverted according to the Eysenck Personality Inventory score. The main findings from this study suggested that users perceive the computer software as having personality attributes similar to those of humans and also using software designed with introverted personality results in general fastest performance for both individuals with extroverted and introverted personalities (Rothrock, Koubek, Fuchs, Haas, Salvendy, 2002).

3.1.9 Mental Model

Cognitive or learning styles refer to a user's information processing behaviour and have an effect on user's skills and abilities, such as preferred modes of perceiving and processing information, and problem solving. They can be used to personalise the presentation and organisation of the content, the navigation support, and search results (Magoulas & Dimakopoulos, 2005).

Learning style

Learning style is an important issue that affects the learning process and therefore the outcome. Many definitions and interpretations of learning styles appeared in literature the past decades. (Bedford, 2004). However, in general terms, learning styles is the individual preferences for how to learn (Sternberg, 1997). When designing instructional material, it is imperative to accommodate elements that reflect individual differences in learning as every learner has a unique way of learning. Learners can only benefit when specific characteristics of their learning style are embedded in adaptive educational hypermedia systems.

Papanikolaou and Grigoriadou (2004) suggest that important decisions underlying the incorporation of learning style characteristics in educational adaptive hypermedia systems demand the synergy of computer science and instructional science, such as: (i) the selection of proper categorizations, which are suitable for the task of adaptation, (ii) the design of adaptation, including the selection of appropriate adaptation technologies for different learning style categorizations and of apposite techniques for their implementation, (iii) the design of the knowledge representation of such a system in terms of the domain and the learner model, (iv) the development of intelligent techniques for the dynamic adaptation of the system and the diagnosis process of learners' learning style including also the selection of specific measurements of learners' observable behaviour, which are considered indicative of learners' learning style and studying attitude.

Several learning style theories have been applied to adaptive educational systems, such as Kolb's learning theory style, (Kolb, 1984), Felder-Silverman learning style theory (Felder and Silverman, 1988), Gardner's Multiple Intelligence theory (Gardner, 1993). Different systems collect student's learning styles using various techniques; interviews, questionnaires, monitoring of student's behaviour.

Cognitive Style

Cognitive style is the way individuals organize and structure information from their surroundings and its role is critically important associated with student success in any learning situation. Cognitive style is usually described as a personality dimension, which influences attitudes, values, and social interaction. It also refers to the preferred way an individual processes information. There are many different definitions of cognitive styles as different researchers emphasize on different aspects. However, Witkin's definition of field dependent (FD) and field independent (FI) is the most well known division of cognitive styles and is more relevant to hypermedia research than others (Witkin, Moore, Goodenough, Cox, 1977). Individuals have field dependent (FD) and field independent (FI) behavioural qualities that differentiate their cognitive style and classify them as more FD (global or undifferentiated) or more FI (analytic or differentiated)

According to Witkin, field dependence-independence has important implications for an individual's cognitive behavior and for his/her interpersonal behavior. While most learners fall on a range between these two cognitive processing approaches, each style is defined by certain characteristics. Specifically, field independent people tend to be more autonomous in relation to the development of cognitive restructuring skills and less autonomous in relation to the development of interpersonal skills. On the contrary, field dependent people tend to be more autonomous in relation to the development of high interpersonal skills and less autonomous in relation to the development of cognitive restructuring skills.

Furthermore, FD/FI dimension refers to a tendency to approach the environment in an analytical, as opposed to global, way. Field independent (FI) learners generally are analytical in their approach while Field Dependent (FD) learners are more global in their perceptions. Furthermore, FD learners have difficulty separating the part from the complex organization of the whole. In other words, FD individuals see things in the entire perceptual field (the forest than the trees). Additionally, FI individuals tend to be intrinsically motivated and enjoy individualized learning, while FD ones tend to be extrinsically motivated and enjoy cooperative learning. Specifically, FD individuals are more likely to require externally defined goals and reinforcements while the FI ones tend to develop self-defined goals and reinforcements (Witkin et al. 1977).

Many experimental studies have showed the impact of field dependence /independence on the learning process and academic achievement and identified a number of relationships between this cognitive style and learning, including the ability to learn from social environments, types of educational reinforcement needed to enhance learning, amount of structure preferred in an educational environment (Summerville, 1999, Ford & Chen, 2000, Weller, Repman & Rooze, 1994, Triantafillou, Demetriadis, Pombortsis, Georgiadou, 2004). The style of a user can be evaluated with the Educational Testing Service Hidden Figure Test (Ekstrom, French and Harman, 1976) or the Group Embedded Figures Test (Witkin, Ottman, Raskin, Karp, 1971). Moreover, there are ways to infer the cognitive style from the browsing strategy followed by the user (Stash & De Bra, 2004).

3.1.10 Social – group

Computer Supported Collaborative Learning (CSCL) and groupware applications are at the focus of educational research lately. Group models are important for collaborative work, since a standard group model should serve as a starting point for interaction for the new member that enters a group (Brusilovsky, 1996). While the new user starts to interact with the system, the user profile can be formed including those characteristics that are in common with, and are different from, the group profile.

To build the group profile, information from users can be acquired using similar techniques with those used for the individual user model: stereotypes, interviews, monitoring users' behaviour. However, these techniques take into account adaptivity variables such as mental models in order to select users for the group construction.

The group profile is quite important for web-based courses as web facilitates collaborative activities. The web browsing advisor called Broadway (Jaczinsky & Trousse, 1998) uses Case-Based Reasoning (CBR) to learn relevant cases from the navigation paths from a group of users in order to improve the recommendation process. CBR is based on the hypothesis that if two users went through a similar sequence of similar documents, they might have similar browsing objectives, and therefore enable us to recommend the same selection to both (Hinrichs and Kolodner, 1991).

3.2 Independent Variables

3.2.1 Current goal / task

The most changeable user feature that activates adaptation is the user's goal(s) or task(s). It is related to the context of a user's work with a hypermedia application rather than with the user as an individual. It informs what the user wants to accomplish by using the application. For example, in information retrieval systems, a user's goal is a search goal; in educational systems is a learning goal; in testing systems might be a problem-solving one (Wu, 2002).

User's goal or task is not firm but they constantly change from session to session and frequently change several times within a session. However, there can also be simultaneous goals i.e. simple, multiple, concurrent. General or high level goals are more stable than local or low-level goals. For example, in educational systems the learning goal is a high-level goal, while the problem solving goal is a low-level goal which changes from one educational problem to another several times within a session (Rothrock, Koubek, Fuchs, Haas, Salvendy, 2002).

In order to accommodate multiple user strategies, Rasmussen and Hurecon (2000) suggests that systems should be designed to adapt to the work contexts by supporting a set of possible user goals or tasks. A system for example can provide a set of possible user goals that users can recognise, and then the most suitable goal will be included in the user model. Vassileva (1996) argues, that the most advanced representation of possible user goals is a hierarchy of tasks. Alternatively, the user model may hold a probability value for each goal supported by the system, to determine the likelihood that a particular goal is the current user goal. This technique can be used also to refine the classification of a goal hierarchy.

3.2.2 Environment – Work

Adaptation to user's environment is a new kind of adaptation that was brought by web-based systems. Users of web-based systems can work irrespective of time and location using different equipment and as a result adaptation to the user's environment can result in better use of the system and yet better performance. Systems can adapt to the user platform, such as hardware, software and network bandwidth. Such adaptation usually involves selecting the type of material and media to present the content, for example, still image vs. movie, text vs. sound (Joerding, 1999).

Current Information and Communication Technologies developments focus on mobile information technology that allow for mobility in the physical space. Given the user and the information is connected to a network this technology facilitates accessibility of information from any point in the physical space. For communication purposes the user employs different devices that have, however, specific characteristics and limitations in terms of bandwidth and information presentation. For mobile information technology the particular challenge for adaptivity is the support of users at different locations. To achieve this, mobile information technology can be combined with technologies to identify the users' working environment and his or her position in the physical space such as infrared or General Positioning Systems (GPS) (Oppermann and Specht, 1999).

3.2.3 Situation Variables

In different situations the same user may have different requirements, and therefore, a system might need to take into account activities that are not expected from the user (Francisco-Revilla and Shipman, 2000). Situation variables that influence user abilities as well as task requirements include: time pressure, location in space and presence and location of targets; situation in time; weather conditions; visibility; and vibration and noise (Rothrock et al. 2002)

An example of how situation variables are examined is the 'Mars Medical Assistant' where three different situations are classified under time pressure: emergency, consultation and educational (Francisco-Revilla and Shipman, 2000). Time pressure is also the main characteristic of the user's profile used in Ready, an experimental system that adapts the type and the duration of advice given to people requesting for services over the telephone (Jameson, Schafer, Weis, Berthold, Weyrath, 1999).

4. DISCUSSION

Research in CAT moves beyond admission programs, on small and large scale applications that address many aspects of measuring performance in education and training. This, combined with new dynamic item generation tools can facilitate computerised assessments that take into consideration individual differences of the user resulting in improved individual performance and more efficient test delivery

What can a user model have to offer in computerised adaptive test delivery? A CAT in order to be more efficient than a linear computerised test, initially assess each individual's level by presenting first an item of moderate difficulty.

However, if the ‘*knowledge on the domain presented’* variable is modelled for each individual then this initial question could be more closer to the examinee’s ability estimation and this will result possibly in cutting down testing time as fewer items can be administered to evaluate the aptitude of the examinee.

Modelling ‘*background-experience’* variable could effect in simpler interfaces for examinees familiar with the information space and more explanatory ones for unfamiliar users. This, combined with the modelling of ‘*preferences’* variable that can basically indicate interface elements (preferred colours, fonts, navigation ways etc.) allows examinees to focus on the assessment process. More clear and self-explicit interfaces may result by taking into account the ‘*personal data*’ variable. For example, in examining gender, males and females appear to have different preferences in terms of media presentation, navigation support, attitudes, and information seeking strategies. Some examinees might feel frustrated or discouraged when they cannot work confidently with the assessment’s interface or when the interface is not designed to suit their individuality. In turn, this will result in poorer performance, since more time will be needed to process information. This is an important issue as in most assessments time is an important factor for measuring the overall performance.

Modelling ‘*interests’* variable for CAT can offer items closer to the long-term interests of each individual examinee. However, this may result in false measurement of performance as, in general, when examinees will be presented with items that always fall in their long term interests domain might perform better than having to deal with items other than that.

‘*Individual traits’* variable refers to stable features of the user such as personality factors, cognitive factors, and learning styles. It is acknowledged of importance, but currently little agreement on which aspects can and should be used to drive adaptation (Brusilovsky, 2001). In essence, it includes the variables of ‘*user personality*’ and ‘*mental models*’. Not much research exists, according to own knowledge, on user personality variable than this recorder in section 3.1.9. Richter and Salvendy (1995) suggested that users perceive the computer software as having personality attributes similar to those of humans. Interfaces designed with introverted personality can result in most cases fastest performance for extroverted and introverted individuals.

‘*Mental models*’ variable is acknowledged as the most significant in the literature of adaptive hypermedia. In interface design terms, with regards to cognitive style for example, a rigid structure should be provided for field dependent users as they need navigation and orientation support; they need to know where they are and where to go next; while a more flexible (or customisable) interface should be made available for field independent users. Furthermore, studies have shown that FD are holistic and require external help while FI people are serialistic and possess internal cues to help them solve problems. FD learners are more likely to require externally defined goals and reinforcements while FI tend to develop self-defined goals and reinforcements (Witkin et al. 1977). These implications of style characteristics in CAT design could result in clear, explicit directions, maximum amount of guidance and extensive feedback to FD examinees, and on the other hand minimal guidance and direction and least feedback to FI examinees.

The independent variables have an effect on the user indirectly, in terms that are not defining him/her as an individual. The most complicated variable to model is '*current goal or task*' as it change constantly from session to session and in many cases there are simultaneous goals within the same session. For example the main goal of taking a test is to pass it, however, simultaneously several goals exist, one for each item that is included in the test. In simple CATs modelling of '*current goal or task*' is not of a particular weight because it complicates the development of the test without any significant benefits for the examinee. However, in more sophisticated CATs, assessing for example leadership skills that predicts job performance of managers or social skills of students to interact with their peers etc. modelling of '*current goal or task*' variable is important as examinees will always face items that closely match their own individual goals.

A user is not tied to a particular hardware platform. S/he can work in one instance from a personal computer attached to a desk and on the other instance from a mobile device such as a Personal Digital Assistant (PDA). As a result dependent variables remain the same with regards to the user modelling. The independent variables of '*environment-work'* and '*situational*' cannot affect the content, yet it seriously affects the presentation mode. Systems can adapt to the user platform by selecting appropriate ways in terms of bandwidth, media etc. for presenting the information.

Summarising, from all the variables presented that triggers adaptation, modelling of '*knowledge on the domain presented', 'background-experience', 'preferences', 'personal data*' and *'mental models'* variables can produce more efficient CATs in terms of time, as fewer items will be needed to assess performance. Moreover, it could affect items' quality, since items can be more complex taking into account user characteristics. As a result, testing sessions would not be limited to measure performance but they can contribute to the learning process. In advanced CATs modelling of '*current goal or task*' can also affect CATs quality. Modelling of *'interests'* should be avoided as it may result in false measurements, as examinees will be presented with items that fall in their interests and not in the whole knowledge domain examined with a CAT. For educational courseware modelling of '*environment-work'* and '*situational*' variables may facilitate teaching and learning for disciplines related to outdoors activities such as zoology, botanology, sailing etc. However it is quite unusual to model '*environment-work'* and '*situational variables*' for testing purposes as there are not many situations when an examinee will need to be assessed for the same subject using a pc and a PDA.

During CAT development, mechanisms should be embedded in all programs to accommodate modelling of '*abilities/disabilities*' and '*social-group*' variables on demand. In examinee population almost always included people with disabilities. If a mechanism exist to assist such individuals on demand disable people will feel less disadvantaged as they could easily take part in any examination process. Modelling of '*social-group*' variable is also important as computer supported collaborative learning and consequently testing is currently at the focus of educational attention.

Modelling multiple variables is important, as users have complex characteristics that ultimate affect their performance. User models must incorporate multiple variables of the user; dependent and independent. A user model of a CAT could be in general stable during the assessment process as this usually lasts for a specific period of time. However, as complex CATs emerge that would not be tied in a specific time period developers should consider that a user model might vary over time as the examinee progresses through hyperpace and their goals and interests may change while they work with new concepts. In that case the user model must quickly adapt to these changes.

Adding additional variables will not always increase the accuracy of the user model but will always increase its complexity and the requirements to collect additional user information (Carver, Hill and Pooch, 1999). Moreover, multimedia adaptation adds additional complexity and requires a greater implementation effort. Media elements are difficult to generate and are not flexible to automatic recombination as text is. For example it is extremely difficult to automatically adapt video segments on the fly and present the results to users. There are many research questions related to multiple variables modelling such as 'what is the proper type and number of variables to measure?'; 'how the variables could be modelled?'; 'which dynamic techniques could be used to modify the weights associated with different variables to better represent the user?'; 'how can we maintain a balance between the number of variables, model complexity, and the accuracy of the model?' etc.

Kosba (2001) in reviewing the development of generic user modelling systems over the past twenty years concludes that predictions concerning the future of user modelling systems are fairly speculative, due to the rapidly changing nature of computing and computing devices. "Since personalization has already been demonstrated to benefit both the users and the providers of personalized services and since personalization is therefore going to stay, it is practically certain that generic tool systems that allow for the easy development and maintenance of personalized systems will be needed in the future as well. The exact form which user modelling systems of the future will take on is however likely to be strongly influenced by many characteristics of system usage that are difficult to predict" (Kosba, 2001, p.58).

Besides research questions the key issue remains; taking into account individual characteristics in interface design result in better user performance. The essence of testing is to measure performance and consequently user modelling for CAT must be the way ahead. The type and number of variables that each CAT would comprise in the user model depend heavily on the subject matter and the way that the test is implemented.

REFERENCES

Ackerman, T.A., Evans, J. Park, K-s., Tamassia, C., & Turner, R. (1999). Computer assessment using visual stimuli: A test of dermatological skin disorders. In F. Drasgow & J.B. Olson-Buchanan (Eds), *Innovations in computerized assessment,* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 137-150.

Bedford, T. A. (2004). *Learning styles: a review of literature* (first draft), Toowoomba, OPACS, The University of Southern Queensland.

Brusilovsky, P. (1996). Methods and Techniques of Adaptive Hypermedia. *Journal of user modeling and user-adapted interaction*.6 (2-3), 87-129.

Brusilovsky, P. (1996a) Adaptive hypermedia, an attempt to analyze and generalize. In P. Brusilovsky, P. Kommers, & N. Streitz (Eds.), Multimedia, Hypermedia, and Virtual Reality (Lecture Notes in Computer Science, Vol. 1077). Berlin: Springer-Verlag, 288-304.

Brusilovsky, P. (2001) Adaptive hypermedia. *User Modeling and User Adapted Interaction*, Ten Year Anniversary Issue (Alfred Kobsa, ed.) 11 (1/2), 87-110.

Carver, C., Hill, M. & Pooch, U. (1999). Third Generation Adaptive Hypermedia Systems. WebNet 99, Honolulu, Hawaii, 1999.

Czerwinski, M., Tan, D.S. & Robertson, G.G. (2002). Women take a wider view. *Proceedings of CHI 2002: Human Factors in Computing Systems*, ACM press, 95-202.

De Bra P., Brusilovsky P. & Houben G.J. (1999). Adaptive Hypermedia: From System to Framework. *ACM Computing Surveys*, 31 (4).

Donovan, M.A., Drasgow, F. & Bergman, M.E. (1998, April). Can a multimedia Computerised Assessment of leadership skills predicts job performance beyong g?. Paper presented at the annual meeting of the Society for Industrial and Organisational Psychology, Dallas, TX.

Drasgow, F. (2002). *The Work Ahead: A Psychometric Infrastructure for Computerized Adaptive Tests*. In Mills, C.N., Potenza, M.T., Fremer, J.J., Ward, W.C., *Computer-Based Testing: Building the Foundation for Future Assessments* (Eds). Mahway, NJ, London: Lawrence Erlbaum Associates, Publishers, 1-10.

Eklund J. & Zeiliger R. (1996). Navigating the Web: Possibilities and Practicalities for Adaptive Navigational Support. *Proceedings of Second Australian World Wide Web Conference*, AusWeb96.

Ekstrom, R.B., French , J.W., Harman, H.H., Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, N.J.: Educational Testing Service.

ETS. (2005). *Research Agenda 2006.* retrieved 1-12-2005 from www.ets.org.

Felder, R. M. & Silverman, L. K. (1988). Learning Styles and Teaching Styles in Engineering Education, Engr. Education, 78(7), 674-681.

Ford, N. & Chen, S. (2000). Individual Differences, Hypermedia Navigation, and Learning: An Empirical Study. *Journal of Educational Multimedia and Hypermedia*, 9 (4), 281-311.

Francisco-Revilla, L. and Shipman, F.M. (2000). Adaptive medical information delivery combining user, task, and situation model. *Proceedings of the International Conference on Intelligent User Interface 2000*, ACM Press.

Gardner, H. (1993). *Multiple Intelligences: The Theory in Practice*. New York: Basic books.

Greer, J. E. and McCalla, G. I., (1994). *Student Modeling: The Key to Individualized Knowledge-Based Instruction*, *NATO ASI, Series F*. Springer-Verlag.

Hinrichs, T.R. & Kolodner, J.L. (1991). The roles of adaptation in case-based design. *Proceedings of the Ninth National Conference on Artificial Intelligence*, Anaheim, CA, 28-33.

Jaczinsky, M. & Trousse, B. (1998). Broadway: a world wide web browsing advisor reusing past navigations from a group of users. *Proceedings of 3rd UK workshop on case-based reasoning*.

Jameson, A., Schafer, R., Weis, T., Berthold, A., Weyrath, T. (1999). Making systems sensitive to the user's time and working memory constraints. *Proceedings of the IUI 99*. ACM press.

Joerding, T (1999). A temporary user modeling approach for adaptive shopping on the Web. *Procceddings of Workshop on Adaptive Systems and User Modeling on World Wide Web at WWW '99 Conference*, Toronto, Canada, 75-79.

Kay, J, (1994). *The um toolkit for reusable long term user models*, SSRG Report 94/3/36.2.

Kobsa A., Müller D. & Nill A. (1994). KN-AHS: An Adaptive Hypertext Client of the User Modeling System BGP-MS. *Proceedings of Fourth International Conference on User Modeling*, Mitre Corporation.

Kobsa, A. (2001). Generic User Modeling Systems. *User Modeling and User-Adapted Interaction,* 11, 49-63. Netherlands: Kluwer Academic Publishers.

Koch, N., 2000. Software Engineering for Adaptive Hypermedia Systems. *Unpublished PhD thesis*, Mathematics and Informatics Department, University of Munich.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Koubek, R.J. and Salvendy, G. (1991). Cognitive performance of super-experts on computer program modification tasks. *Ergonomics*, 34(8), 1095-1112.

Large, A., Beheshti, J. & Rahman, T. (2002). Gender differences in collaborative Web searching behavior: an elementary school study. *Information Processing & Management*, 38(3), 427-443.

Large, J.A., J. Beheshti and T. Rahman. (2002). Design criteria for children's Web portals: the users speak out. *Journal of the American Society for Information Science and Technology.* 53(2), 79-94.

Magoulas G. D. & Dimakopoulos D. N. (2005). Designing Personalised Information Access to Structured Information Spaces. *Proceedings of the 1st International Workshop Workshop on New Technologies for Personalized Information Access*.

Murray, D. & Bevan, N. (1985). The social psychology of computer conversations. In B. Shackel (Ed.) *Human-Computer Interaction – INTERACT 84*. New York: Elsevier Science.

Olson-Buchanan, J.B., Drasgow, F. Moderg, P.J. Mead, A.D. Keenan, P.A., & Donovan, M. (1988). Conflict resolution skills assessment: A model-based multimedia approach). *Personnel Psychology,* 51, 1-24.

Oppermann, R. and Specht, M. (1999). Adaptive information for nomadic activity: a process oriented approach. *Proceedings of the Software-Ergonomie '99*, Stuttgart: Teubner, 255 –264.

Paiva A, Self J & Hartley R (1995). Externalizing Learner Models. In J Greer (ed.) *Proceedings of AIED95,* AACE publication, 509-519.

Papanikolaou, K.A. & Grigoriadou, M. (2004). Accommodating learning style characteristics in Adaptive Educational Hypermedia Systems. *Proceedings of the AH 2004 Workshop "Individual Differences in Adaptive Hypermedia"*.

Rasmussen, J. and Hurecon, S. (2000). Designing to support adaptation. In *Proceedings of the IEA 2000/HFES 2000 Congress*.

Richter, L.A. & Salvendy, G. (1995). Effects of Personality and Task Strength on Performance in Computerized Tasks. *Ergonomics,* 38(2), 281-291.

Rothrock L., Koubek R., Fuchs F., Haas M., Salvendy G. (2002). Review and Reappraisal of Adaptive Interfaces: Toward Biologically-Inspired Paradigms. *Theoretical Issues in Ergonomic Science*, 3(1), 47-84.

Stash, N. & De Bra, P. (2004). Incorporating Cognitive Styles In Aha! (The Adaptive Hypermedia Architecture, Web-based Education 2004, Innsbruck, Austria Editor(s): V. Uskov

Sternberg, R. J. (1997). *Thinking style*s. New York: Cambridge University Press.

Summerville, J. (1999). Role of awareness of cognitive style in hypermedia. *International Journal of Educational Technology*, 1(1).

Tan, D.S., Robertson, G.G. & Czerwinski, M. (2001). Exploring 3D Navigation: Combining Speed-Coupled Flying with Orbiting. In *Proceedings of CHI 2001, Human Factors in Computing Systems*, (Seattle, WA April 1-6, 2001), ACM press, 418-424.

Triantafillou, E., Demetriadis S., Pombortsis, A. & Georgiadou, E. (2004). The value of adaptivity based on cognitive style: an empirical study. British Journal of Educational Technology, 35 (1), 95-106.

van der Linden, W.J., Glas, C.A.W. 2003. Preface. In van der Linden, W.J., Glas, C.A.W (Eds). *Computerised Adaptive Testing: Theory and Practice*. Dordrecht, Boston, London: Kluwer Academic Publishers, vi-xii.

Vassileva, J. (1996). A task-centered approach for user modeling in a hypermedia –based information system. *In Proceedings of the 4th International Conference on User Modeling*, Hyannis, MA, 115-120.

Vispoel W.P. (1999). Creating computerized adaptive tests of music aptitude: Problems, solutions and future directions. In F. Drasgow & J.B. Olson-Buchanan (Eds), *Innovations in computerized assessment* (pp. 151-176). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Weber, G. and Brusilovsky, P. (2001). ELM-ART: An Adaptive Versatile System for Web-based Instruction. *International Journal of Artificial Intelligence in Education,* 12, 351-384.

Weller, H.G., Repman, J., & Rooze, G.E. (1994). The relationship of learning behaviour, and cognitive styles in hypermedia-based instruction: Implications for design of HBI. *Computers in the Schools*, 10, 401-420.

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*, 47(1), 1-64.

Witkin, H.A, Ottman, P.K., Raskin, E., Karp, S.A. (1971). *A manual for the Embedded Figures Tests*. Palo Alto, CA: Consulting Psychologists.

Wu H., 2002. A reference architecture for adaptive hypermedia applications. *Unpublished PhD thesis*, Technische Universiteit Eindhoven.