Creating Bridges: The Role of Exploratory Design Research in an Intelligent Tutoring System

1st Author Name, 2nd Author Name, 3rd Author Name, 4th Author Name, 5th Author Name

Affiliation	Affiliation	Affiliation
Address	Address	Address
e-mail address	e-mail address	e-mail address
Optional phone number	Optional phone number	Optional phone number

ABSTRACT

Designers of Intelligent Tutoring Systems (ITS) have long been interested in delivering personalised teaching to individual students, typically by ensuring that the student receives content appropriate to their skills and knowledge. Nonetheless, a more holistic view on what constitutes good teaching practice has challenged whether this approach to user modelling is sufficient. Teaching is not only defined by what is taught, but also by how it is taught. In this paper, we demonstrate that exploratory design research can support this view by generating a more inclusive set of user attributes for purposes of user modelling. Through a case study, we show that design research for user modelling can function as a boundary object serving three important roles, that underpin more specifically the design of user modelling and more broadly ITS design. First, design research can establish common ground by encapsulating domain knowledge in an accessible form. This can support diverse project stakeholders to make decisions on what is to be modelled. Second, design research can reveal a wide range of teaching and learning perspectives that in turn introduce transparency to the decision-making process of user modelling and provoke a sense of criticality and accountability amongst project stakeholders. Third, design research can build new bridges between the design of the technology and the user model that underpins it. To this end, user attributes deemed important, yet too complex or cumbersome to develop, can become design principles in the context of the overall ITS design.

INTRODUCTION

Author Pre-Print Copy

Design thinking begins with the understanding and exploration of a problem. This exploratory, early stage of design enables designers to break free from existing ways of thinking and knowing (Dorst, 2011). Design research – encompassing any research activity that supports the process of design – has been instrumental both in this early stage of design and in refining technology design once a direction has been decided (Vines et al., 2013). Even though this view on design has in many ways become ingrained in mainstream HCI, it has not always been adopted by communities that have overlapping interests with the HCI community. A prominent example is user modelling, an area of interest for researchers working on Artificial Intelligence (AI) and HCI. Within the field of user modelling, AI researchers have largely occupied themselves with the production of better algorithms through which to understand and model human intelligence. Design research has typically featured as a way to enrich researchers' theories and perspectives, rather than to define or prioritise them. Meanwhile, HCI research has typically conversed with AI in improving *the use* of AI algorithms after they have been developed (Grudin, 2009).

In cases where HCI researchers have sought to influence AI in the design stage, the process of doing so has not been straightforward. For example, reporting on their experience of designing an adaptive serious game, Vasalou and Khaled (2013) point out the difficulties they encountered when introducing generative design research in the project. Rooted in a tradition of data-driven design methods, AI researchers were hesitant to embrace the subjectivity that design methods invited. Thus, even though both HCI and AI researchers recognized the reciprocal interaction between user modeling and design, their divergent goals, epistemology, and related practices became a barrier to their collaboration. As the authors explain, the consequence of this was that design decisions between teams were frequently disjointed. While acknowledging these challenges, in this paper we advocate the view that design thinking can play an important role in the design of user modelling. In particular, we propose that exploratory design research for user modelling can act as a 'boundary object' (Star and Griesemer, 1989). In this capacity, it can help frame, inform and justify user modelling decisions, while at the same time it can influence a broader cohort of design decisions involved in the development of intelligent technology. This proposal is explored through a case study of an intelligent tutoring system (ITS) designed to teach reading (word decoding) skills to primary school children with dyslexia.

The work we report contributes to our knowledge of design by demonstrating that design research can play three different roles that directly support design decision-making processes for user modelling and ITS. *Firstly*, design research can encapsulate domain knowledge in an accessible form that in turn creates common ground between diverse project stakeholders, providing all team members with the knowledge and confidence to make decisions. *Secondly*, given its inductive nature, design research can reveal a wide range of teaching and learning perspectives that introduce transparency to the design process and as a consequence provoke a sense of criticality and accountability amongst project stakeholders. *Thirdly*, design research for user modelling captures core teaching and learning knowledge that is of relevance more broadly to the design of ITS and can thus directly inform other design decisions and activities.

In the following section we give an overview of current user modelling approaches used within ITS and their relationship to theories of learning. Next, we elaborate on why the introduction of design research in the exploratory stages of design can inspire new ways of approaching user modelling, as well as inform broader design decisions within ITS. We go on to present the application of design research within a case study of a project whose aim is to develop an ITS to support children with dyslexia in acquiring reading skills. The implications of this case study are discussed, focussing on how our design research acted as a boundary object within our multidisciplinary team collaboration to inform user modelling and ITS design.

BACKGROUND

It has long been recognised that personalised education is more effective than a one-size-fits-all teaching approach. Bloom showed that an average student who received one-to-one tutoring from an expert tutor scored two standard deviations higher than an average student taught in a traditional group-based instructional setting (Bloom, 1984). Cohen and colleagues (1982) found a similar result (though with a weaker effect size) based on a meta-analysis of tutoring more broadly. While personalised education has recently been forecast as one of the global challenges we face in coming years (LaFontaine, 2013), for decades it has been a mainstay objective in the field of ITS, which has aspired to construct personalised learning experiences that "make the benefits of individualized instruction available to all students at affordable costs" (Anderson, 1992, p. 3). The main approach to delivering personalised education with technology is through a *user model*. User modelling is based on a simple idea: By having information about a specific individual, technology can make decisions that are best suited to that individual. Any technology that behaves differently for different users employs a user model. User models can be big or small, complex or simple, rich or sparse. Orwant (1996) captures the essence of user modelling by explaining that, "user modelling is nothing more than a fancy term for automated personalisation. Humans model each other all the time. I am modelling *you* as I write; my topics, presentation, and language are all aimed at a hypothetical, average reader of this journal. If I have guessed well, you will enjoy this essay. If not, you will skip to the next one. That is what user modelling systems do – they make guesses, and hopefully educated ones, about their users" (Orwant, 1996, p. 398).

Author Pre-Print Copy

ITS researchers have focused their modelling efforts on a number of key areas. Some researchers have modelled student *knowledge*. For example, Anderson (1992) developed a high school mathematics tutor using pre-existing maths curriculum as a foundation for progressing students through the software. Taking a different approach to modelling knowledge, Radlinski and McKendree (1992) created the COBOL tutor based on an understanding of the level of difficulty required to solve particular problems. COBOL was thus based on a production system model of an "ideal student", suggesting what coding errors a student has made and what type of code should follow what they have written. A second area of research has been concerned with the modelling of student *affect*. Conati, for example, worked on detecting students' affect during the use of educational game, using a combination of physical sensors and log files to detect their emotions while playing the game. Depending on which emotions were detected, a game character appeared and changed the form of feedback it gave to the player (Conati et al., 2003; Conati and Maclaren, 2009). A second representative example is work by Mota and Picard (2003) who developed a model that relied on a student's posture to infer their interest during a learning task, which they proposed could be used in turn to provide assistance or encouragement at the right time. A third area of development has focused on modelling motivation. Costagliola et al. (2010) detailed a method for modelling attention from body posture. Given the student's motivation, the system changed the topic and level of content being presented.

These examples indicate consensus regarding the requirement to tailor learning experiences to students' individual characteristics and needs. However, despite this common ground, the diversity of these applications provides evidence to the different approaches taken to solving the same design problem of how to best personalise an educational experience through technology. For example, when developing a user model for a maths tutor, a fundamental difference exists between modelling the acquisition of knowledge (e.g. the student knows that $1 + 2 = 3$), abilities (e.g. the student can perform simple addition) or strategies (e.g. the student knows how to use a number line). The student's subsequent learning experience will depend on the user modelling approach taken.

The roots of these differences can be traced to prevailing theories used by AI researchers, as well as contextual domain requirements, both of which encourage particular ways of framing the design problem (Selwyn, 2011; Laurillard, 2012). For example, those tutoring systems inspired by behaviourist approaches – that consider behaviour to be conditioned by a series of responses to various stimuli – have tended to rely on the selection of appropriate activities and content, using reinforcement through game-like rewards or punishments to encourage the desired behaviour (e.g., Conati et al., 2003).

Conversely, researchers grounded in cognitivist traditions have been concerned with the internal process of mental action and the complex descriptions of how knowledge is stored and processed. This has inspired modelling efforts that map student errors against an optimal way of solving a problem (e.g., Radlinski and McKendree, 1992). A further, related consideration arises from the post method era, which has created a division between those basing themselves entirely within the remit of theory and evidence-based interventions, and researchers in favour of prioritising practitioners' tacit knowledge (Akbari, 2007). In designing their user models based on tacit knowledge, some AI researchers have heavily relied on domain expert opinions, while others have been informed by contextual real world observations of teaching practices (Ragnemalm, 1995). Highlighting the importance of understanding practitioners' tacit knowledge, Porayska-Pomsta and colleagues additionally argue that research into the cognitive and affective components of learning is still on-going; it is thus often necessary to conduct exploratory studies aimed at eliciting teacher's tacit practices and learner experiences within particular domains of inquiry that in turn address researchers' knowledge gaps (Porayska-Pomsta et al., 2013). Alongside the formative role of theory or practice on user modelling, AI researchers must also address pragmatic concerns, and focus on those requirements that are most pressing in their respective problem domains. For instance, software that aims to foster skills in STEM subjects such as mathematics may require information about students' knowledge, while a game that instigates conflict in order to teach conflict resolution skills calls attention to a deeper understanding of players' coping with negative emotions.

MOTIVATION AND RESEARCH OBJECTIVES

Traditionally, AI researchers working on ITS created user models that supplied appropriate content to each student. Despite the importance of this work, good teaching involves decisions on both *what* and *how* to teach. Thus, teachers frequently adapt their strategies in-situ based on an understanding of their students' knowledge or progress, alongside their motivation levels or frustration. ITS researchers have become increasingly interested in how both knowledge and affect can be integrated into user modelling. Given the evolving nature of ITS, this context provides an ideal exemplar for highlighting how design research can help identify new perspectives on teaching and learning by revealing a wider range of user characteristics than those that might be typically pursued.

In reviewing the design practices involved in user modelling, we observe that the primary role of design research has been to support the development of solutions to defined design problems. In other words, AI researchers have tended to conduct design research only after deciding which knowledge or psychological constructs (e.g. motivation) necessitate adaptation, with research used to refine and enrich the perspective taken. A recent review by Porayska-Pomsta and colleagues (2013) illustrate this by identifying a number of methodologies – such as logs of student actions, verbal protocol, interviews and walkthroughs – typically used by AI researchers to study tutor interactions with students toward building predictive models of affect (e.g., Porayska-Pomsta et al., 2008). Design problems, nonetheless, such as that of user modelling, are 'wicked' by nature (Buchanan, 1992). In other words, different perspectives on the education problem will motivate different approaches to user modelling. Before choosing a perspective through which the problem can be seen, designers must reach a deeper understanding of the situation (Gaver, 2012; Dorst, 2011; Sengers et al., 2005). This design exploration is idealistic in its aim to "reveal alternatives to the expected and the traditional, aspiring to transcend accepted paradigms and bring matters to a head" (Fallman and Stolerman, 2010). In keeping with this position, we propose that the introduction of research in the early stages of design stands to make an important contribution to the field of user modelling. *Our first aim is to demonstrate that design research can reveal overlooked user attributes and directions for adaptation that advance approaches taken within a particular domain of inquiry.*

Author Pre-Print Copy

While design research may open up new avenues for user modelling, as Gaver (2012) has pointed out "design involves specifying many different decisions of an artefact". User modelling forms one component within a highly complex web of technologies. Yet the centrality of the user model means that it interacts with and affects many different design decisions across the development of an ITS. Large-scale ITS development projects involve AI researchers, designers and domain experts (among others) who must ultimately reach consensus regarding how a user model interacts with other design decisions and constraints. This task is in no way trivial, as different stakeholders must reconcile conflicting values and priorities, not least of all arising as a result of different epistemologies evidenced through the reflective, ethnographic, and participatory methods of design and the data-driven methodologies preferred in the field AI (Khaled and Ingram, 2012; Vasalou et al., 2012; Benton and Vasalou, 2013). Recognising the importance to reconcile such differences, Star and Griesemer (1989) observed the role of 'boundary objects'. As they explained, "boundary objects are objects which are both plastic enough to adapt to local needs and the constraints of the several parties employing them, yet robust enough to maintain a common identity across sites… they have different meanings in different social worlds but their structure is common enough to more than one world to make them recognisable, a means of translation". They went on to propose that the form of boundary objects can be abstract or concrete, physical or virtual, as long as they provide diverse communities – such as those involved in ITS projects – with a focussed point from which decisions can be made. Thus, boundary objects serve to support a form of distributed cognition both by mediating cognition and shifting decisions from a single individual to multiple actors (Walenstein, 2002; Perry, 2003; Wright et al., 2000). To our knowledge, AI researchers have not examined if and how design research for user modelling has supported decisions within the broader design context of ITS. *Our second aim is to explore the role of design research as a boundary object with a particular focus on its impact on decision-making by diverse project stakeholders.*

These two research aims were pursued in the context of an on-going ITS project centred around improving the reading skills of children with dyslexia. Adopting an inductive approach to understanding individualised teaching for children with dyslexia, we identified a broad range of student characteristics and teaching strategies, which we treated as a boundary object. The following section describes our project, including the project's need for user modelling as well as the stakeholders involved.

CASE STUDY – AN INTELLIGENT TUTORING SYSTEM FOR TEACHING LITERACY TO CHILDREN WITH DYSLEXIA

Developmental dyslexia, or specific reading disability, has been defined as "an unexpected, specific, and persistent failure to acquire efficient reading skills despite conventional instruction, adequate intelligence, and sociocultural opportunity" (Demonet et al., 2004). A recent UK government report into teaching literacy to children with dyslexia proposed a similar definition identifying difficulties in phonological awareness, verbal memory and verbal processing speed (Rose, 2009). Dyslexia is considered as a continuum rather than a distinct category (Snowling, 2008; Goswami, 2008). Individuals are likely to have a sub-set of the many difficulties associated with dyslexia. As a consequence of this, it is vital that each child with dyslexia benefits from an individual intervention plan (Rose, 2009) and best practice teaching requires customised lesson plans to support a child in addressing specific areas of difficulty (e.g. suffixing, syllable division).

Our research was aimed at developing an ITS to help children with dyslexia between the ages of 9-11 improve their reading skills. Supporting this age group was particularly important as these children will shortly begin their transition from primary to secondary education, resulting in substantial changes to their current support model and often reductions in the level of individualised support they receive. The continuum nature of dyslexia suggested the relevance of including a user modelling component within the technology. By tracking the individual difficulties a given child may have, we aspired to provide targeted support. To give an illustrative example of targeted literacy support, while one child may be supported with their difficulty regarding suffixes and prefixes, another may be given support with syllables and vowel sounds. Our ITS takes the form of a tablet-based application consisting of two main components: A *reader* with supportive reading features and an *educational game* integrating a series of learning activities. While children were the primary users of the ITS, the software was envisioned as a teaching tool. Thus, teachers would be able to observe and approve a child's progress through various means including reports produced by the software. Our project stakeholders included two researchers focusing on user modelling, four domain experts in linguistics, developmental psychology and policy, nine developers, and four designers undertaking the design research and game design.

Author Pre-Print Copy

Design Research

Methods and Procedure

Teaching is a multifaceted practice, and it is important to deploy a variety of methodological techniques to capture a holistic portrayal of teaching practices and reflections on practices. As such, we adopted a mixed methods approach, consisting of observations, interviews and online surveys. Table 1 presents a summary of the methods we used and the participants who were involved in our research.

Table 1 about here

User modelling requires the definition of measurable user attributes based on which the software responds; thus neither a user attribute nor a teaching strategy is useful in isolation. Given this design requirement, our aim was to identify user attributes that teachers perceived and responded to through targeted teaching strategies. Our methodology was designed to elicit teaching strategies from teachers, from which we subsequently inferred the user attributes that teachers perceived. This decision was made as it was deemed more natural for teachers to reflect and articulate on what they do in practice. Our methodology was developed to further measure teaching strategies at two levels. First, we examined teaching strategies that were applicable to *all* children with dyslexia, enabling us to capture common and shared issues. Second, we collected teaching strategies that were used by teachers with specific children, allowing us to distinguish individualized strategies from those broadly provided to all children with dyslexia. Our data collection consisted of three stages, which included observations of teaching sessions, interviews with teachers and an online survey targeted at specialist dyslexia teachers.

The *first* stage of our design research was to observe a number of teaching sessions both at specialist dyslexia centres and mainstream schools. By undertaking observations at different centres as well as within mainstream schools which use different approaches to their teaching, we hoped to capture a rounded view of existing teaching practices as well as to understand how specialist dyslexia teachers structure and adapt their lessons. The first set of observations were undertaken at two specialist dyslexia teaching centres in two separate cities, one in the South West (Centre A) and the other in the South East (Centre B) of the UK. Two different researchers undertook the observations, but followed the same protocol agreed in advance. Each researcher spent approximately five hours observing specialist intervention sessions. These sessions differed between centres: At Centre A all sessions involved pairs of children and at Centre B the sessions involved both individual children as well as groups ranging from two to four children. The second set of observations were undertaken at a mainstream primary school in South East London (School A). One researcher observed two year six literacy classes, one low-ability class and one high-ability class, of children aged 10-11. Each lesson lasted approximately one hour. The researchers passively observed the sessions and took written notes of what happened, particularly focusing on the teacher's strategies for structuring the sessions, adapting content, providing help, building confidence and engaging the children.

We followed up our observations, in the *second* stage of our design research, with a series of interviews with one specialist teacher from Centre A and two specialist teachers from Centre B as well as a special educational needs (SEN) teacher and a mainstream year six teacher from School A. One of the researchers also interviewed a SEN teacher at a further mainstream primary school, School B, in South East London (it had not been possible to undertake observations of lessons in conjunction with this). Each teacher was asked specific questions about the way they work, their specific strategies for teaching children with dyslexia (where relevant) and also how they thought personalised technology could benefit their teaching. Where appropriate during the observations, critical moments were flagged within the researchers' notes such that they could be followed up during the interview. Similar interviews were carried out with SEN coordinators at mainstream schools such that we could combine the specialist approach with the approach taken within mainstream schools. Each of the interviews lasted approximately 30 minutes and were audio recorded where consent was given (and later transcribed) or alternatively where there were issues with audio recording the researcher took detailed notes of the interview.

As a *third* and final stage, to ensure that sufficient consideration was given to the needs of specific children with dyslexia, an online survey was sent out to dyslexia specialist teachers to find out how they adapt their lessons for a given student. The surveys asked the teachers to consider individual students, potentially allowing more detailed characteristics to emerge that might be otherwise overlooked if attempting to generalise across a wider dyslexic population. Three teachers reported on their intervention sessions, reporting on a total of five students. The survey included 22 questions, the majority of which encouraged the teacher to reflect on the intervention session they had recently undertaken with a specific child. The survey included questions that asked about the specifics of their last teaching session with that child, focusing on the difficulties covered, activities used, and the child's response to those activities. It also asked teachers to reflect on what did or did not work within the session and why. Finally, it asked teachers to reflect on the individual aspects of the specific child and how they changed their standard teaching approach to account for those differences. The survey was accessible for two weeks and teachers received the equivalent of one hour of pay for each response they made.

Author Pre-Print Copy

We note that our methodological approach intentionally focused on exploring teachers' practices rather than seeking to involve children as active participants in this part of the design process. Given our focus to understand children's learning needs and the characteristics that underpin these needs, we focused on what currently occurs in classrooms; a teacher relies on their subjective assessment of children's behaviour and based on their extensive training and experience determines how to react or adapt the learning intervention. To accomplish our research aim, we explored teachers' expert pedagogical approaches from different perspectives in order to select the most significant attributes to model, a process we do not believe a child could meaningfully contribute to. This is in line with the work of Druin (2002) who argues that there are inappropriate roles for children within the design process, for instance "we cannot expect them to know what educational goals need to be covered in a school curriculum as well as a teacher does". Good and Robertson (2006) also observe that children are not capable of designing their own learning goals, which are best articulated by a teacher or an educational technologist. In the case of young children with dyslexia, we further note their particular challenges in knowing how they best learn (Rose, 2009).

Analysis

Our analysis was inductive, in that our data was coded to capture the factors that teachers took into consideration when putting into action individualised teaching strategies. Specifically, we conducted a thematic analysis (Braun and Clarke, 2006) on the notes and transcripts from the observations and interviews, as well as on the survey results, which resulted in a set of codes. In each case, the researcher who had originally collected the data performed the analysis. In those cases where there were two investigators involved (i.e. in the first and second stage), the researchers compared their codes and applied them back across their individual transcripts. Triangulation was employed during the analysis comparing data from several different sources to help us better understand the phenomenon. This process established 19 distinct codes, each representing an attribute that could be modelled. In order to ensure that these codes could be consistently applied across the data, a series of statistical tests were performed for interrater agreement (Tinsley and Weiss, 2000). Based on the recommendations in (Lombard et. al, 2002), 25% of each data set was randomly selected and subsequently coded by two independent coders. Prior to the coding, the judges received a codebook with definitions for each code to establish a shared understanding of the codes (Lombard et. al, 2002). Cohen's kappa is a measure used to examine interrater agreement on the assignment of a variable on a nominal scale and was thus chosen to establish the degree of agreement between the two coders (Tinsley and Weiss, 2000). Interrater agreement (reported in Table 2) ranged from moderate to almost perfect agreement (Landis and Koch, 1977).

Table 2 about here

The 19 codes yielded from our analysis were next organised into six high-level themes: *linguistic ability, cognitive style, motivation and interests, goals, personality and behaviour* and *co-occurring characteristics and difficulties*. These themes are discussed in detail below with excerpts from the data, illustrating how well the data fits within these themes.

Results

Linguistic ability

Given our focus on literacy, it was not surprising that teachers concentrated on children's linguistic abilities. The teachers we interviewed needed to know a child's specific *linguistic level* to ensure the content they provided during the lessons was appropriate for the child. Furthermore one of the specialist teachers at Centre B suggested storing "a word bank personalised to the child", which would provide more personalised content than simply assigning the child to a generic set of words that might not target their difficulties.

Some of the teachers also discussed assessing the *linguistic strategies* that individual children were using when tackling particular difficulties, to ensure that the strategy they were using was effective. For instance, the Centre B principle argued that it was important to "assess the strategies they are using" to solve particular problems. Similarly, the SEN teacher from School A noted helping the student learn was "partly to do with showing them that there are support strategies that they can use themselves".

Author Pre-Print Copy

Cognitive style

Cognitive style was highlighted by both the principle of Centre B and the SEN teacher at School B who argued that it was important for teaching sessions to be tailored to use the "best way the child learns and adapt accordingly, for example through muscle memory or visuals". The exact form this *multisensory* approach can take was made clear by the responses of our online survey, with four of the responses highlighting the various approaches content can be made multisensory. For instance, one of the teachers highlighted how their student was a visual learner with excellent drawing skills and therefore would include tasks such as creating an annotated picture, which encourages their drawing skills and improves their written work at the same time. As another example, a different teacher mentioned one of their students who had done a project on World War II where he "worked independently to create some great work which included audio files at various points to indicate the sounds of the guns with explanations", thus reducing the time and effort that would have been required if he had to share everything in written form.

Many teachers highlighted broad teaching strategies they used across their teaching which they found useful. One of the most common strategies was *overlearning*, which the SEN teacher from School A explained requires "a lot of practicing of things that have already been done". This is a particularly important consideration for students with dyslexia who typically have poor memory skills – if activities and areas of difficulties are not revisited then new skills that a student has previously demonstrated can be easily forgotten.

In addition to these broad strategies, we also identified a number of specific teaching strategies deployed within individual activities. For example, a number of teachers discussed the *introduction of new content* and highlighted how it was important to make clear what content was new and what content the child had seen before. For example, the principle of Centre B argued that in order "to build self esteem, make sure that 99% of the lesson they already know and 1% is new". The majority of the specialist dyslexia teachers noted how important *self-checking* was as a motivation tool but also as a means of highlighting what the student still needed to work on. This was coupled with a general approach ensuring that students did not move onto new activities without understanding what they did wrong. In situations where interventions needed to be made, one of the teachers from Centre A made clear that this should be done at the end of an activity so as to not disrupt the child’s flow. Finally, a survey respondent noted that one of her students needed *instructions* to be kept short and the student would then be asked to repeat back the instructions to ensure the instructions had been correctly understood.

Motivation and interests

The teachers discussed motivational strategies, and a key component within these strategies was the identification and integration of the child’s *interests* within the teaching. There could also be differences between the specific motivations of individual children. The mainstream teacher from School A pointed out that what is motivating to a low ability class (giving out lots of merits) is not necessarily motivating for higher ability groups.

Our analysis highlighted the potential for leveraging a child’s interests as a motivational strategy. This was discussed by many of the teachers from both the specialist dyslexia teaching centres and mainstream schools, with one of the specialist teachers who completed the survey explicitly stating, “personal interests are important”. The mainstream teacher from School A attempted to tailor his lessons to “something the children can relate to… what interests the children” while the SEN teacher from School B noted how “tailoring the sessions to the children’s interests” can increase a child’s motivation. One specific strategy suggested by the SEN teacher from School B was to have “a conversation with the child about their interests and make games linked in with that, e.g. ghosts, that can be used for multiple purposes”.

We also identified several ways in which teachers attempt to maintain motivation within their students whilst presenting challenging material. These motivational techniques fell into two categories. The first are concerned with broad level approaches, distinct from individual activities. The *construction of the learning environment* in which the children are taught appeared to have been carefully considered at Centre A. Many books of various lengths, aimed at various ages and with different levels of content were clearly displayed within the classroom and posters of famous people with dyslexia were displayed on the walls. This suggests that the construction of the learning environment to encourage children to feel more comfortable with books and reading as well as with their dyslexia diagnosis could potentially motivate the children with their learning. The larger group classes within Centre B provide an opportunity for *peer learning*, where the younger children work alongside older children. The younger students get to see older children having the same problems (decreasing the sense of isolation) and the older students get to help the younger students (improving the sense that they can achieve things), which can help to boost motivation on both sides. Common to both dyslexia centres is that the teachers noted the importance of a consistent *lesson structure*. This means that even if the content is unfamiliar to the child, the familiar structure helps the child feel more comfortable with attempting the exercise. Many teachers across both mainstream and specialist schools noted the importance of using a quick succession of short activities to keep the students engaged which prevents the child becoming stuck on an activity for an extended period and potentially decreasing their motivation.

Author Pre-Print Copy

The second category includes motivational strategies deployed within individual activities, which primarily focuses on how they *present feedback* to the children during and after these activities. The SEN teacher from School B observed that when correcting a student, you have to make it clear that the mistake “is not a reflection on their intelligence and point out what they are good at. However, don’t pretend that it is not an issue”. Similarly, many teachers made it clear that you should not give praise for praise sake as the students will begin to see the praise as worthless. The strategy used at Centre A was to correct mistakes as the activity is progressing, giving a “congratulations” or “well done” at the end of an activity, along with a comment as to how they might improve. Furthermore, the teachers at Centre A made it clear that they highlight if the student gets something right that most people get wrong. Finally, as discussed under learning strategies, self-checking during activities and getting students to record their own progress can be highly motivating, allowing the student to see the progress they have made and giving them ownership over their work.

Goals

A few of the SEN and specialist teachers attempted to establish what each child’s personal learning aims were. The SEN teacher from School B did this through “talking with the children about what they want to be able to do”, to incorporate this into their teaching. This could be a *general aim* such as acquiring a certain literacy skill or a more *specific goal*. For example, one of the teachers at Centre A mentioned a particular student who really wanted to be able to read some of Jacqueline Wilson’s books. One of the surveyed teachers noted that one of her students hates being below the rest of his classmates in class and wants to be on the same reading books. This recognises that although not all children will be able to overcome all of their dyslexia difficulties, teaching could be tailored towards helping a child to achieve the goals that are personally important for them.

Personality and behaviour

A further theme identified within both the interview and survey data was effectively *managing a child’s behaviour* during a session, with the SEN teacher from School A saying “you might have someone who is being naughty or not trying and then I would say you need to concentrate” and the mainstream teacher from School A stating that a child might have an attitude problem, e.g. “[if a child is] not applying themselves then we work very closely with the child… basically if it is a behaviour thing and we have a lot of strategies to bring that round”. The structure of the lesson, as well as being educationally important, can also be used to moderate behaviour. One of our survey respondents discussed how her lesson plan will clearly state on the board which student's turn it is to sit on the 'red' chair and who will be choosing the game at the end of the lesson. This helps to mitigate disagreements within the class.

Being able to *deal with failure* was also highlighted as a specifically important issue, with the SEN teacher from School A stating that failure is "something a lot of children have got a big problem with". The specialist teachers who completed the survey discussed particular strategies they had based on their knowledge of the child's personality, for instance "activities are very carefully selected as he is unable to work independently" or knowing that a particular child "responds to challenging work". One of the surveyed teachers stated that "We do not tick or cross work as too many incorrect answers will make him upset. He [the student] would try to change the answer and then try to convince me that it was right all along". That said, a different respondent said that their student relished the challenge of trying to get 10 out of 10 questions correct.

Many behavioural aspects, such as those related to *social and emotional issues*, were discussed as being important. This was particularly clear from the survey responses as the teachers were considering the behaviour of individual children they taught. These aspects included health issues, as well as excitability and fatigue, all of which can further exaggerate a child's specific difficulties. Fatigue was a particular issue which, although a behavioural aspect, was not raised during the observations and interviews but was captured in the online surveys, e.g. "this session was good as the boys were less excitable than is often the case", and "tired today, close to end of term, feeling unwell". Two survey respondents noted how important their students' emotional needs were. One stated that their emotional state is important as he often needs 'perking up' when arriving at the centre. This had led to the teacher opting to run less challenging difficulties with that student. In short, if the student is experiencing some form of emotional distress, whatever the cause, the difficulty of the lesson is typically moderated.

Co-occurring characteristics and difficulties

The last theme that we identified, predominantly through the survey responses, was that of co-occurring characteristics and difficulties linked with dyslexia. *Specific co-existing conditions* such as dyspraxia and attention deficit (hyperactivity) disorder (ADD/ADHD) were both mentioned by the specialist teachers from both centres, with suggestions of more frequent changes of activity as one strategy to use with a child who additionally has an ADD/ADHD diagnosis. A number of further *associated difficulties* were also mentioned including working memory and memory load, organisational issues, and problems reading moving text. Whilst some of these can be considered as personality traits, they are also difficulties associated with dyslexia more generally and as such need to be accounted for.

Author Pre-Print Copy

Low self-esteem and low self-confidence are commonly found personality characteristics of children with dyslexia. Many of the specific attributes we have identified are, to a greater or lesser extent, related to trying to boost the student's confidence or at least not to damage it. Examples include the self-checking of work, the motivational strategies used, the construction of the learning environment, the proportion of new content being introduced and the ability to deal with failure. There are many complex social and emotional issues that are related to building self-esteem and self-confidence, making it worthwhile to identify them as attributes in their own right.

DISCUSSION

Our design research revealed that teachers adapted their teaching style by taking into account their students' *linguistic ability,* their *cognitive styles* and *behaviour,* as well as the *co-occurring difficulties* their students' encountered. They aligned their interventions to students' *individual goals* and used numerous strategies to strengthen their *motivation*. Table 3 presents the 19 potential user modelling attributes we identified grouped under the aforementioned themes, alongside example teaching strategies used by teachers. We note that the themes are not mutually exclusive and that certain attributes might belong to multiple themes. While the attributes we identified may be particular to the group of teachers and students we worked with, our design philosophy with respect to user modelling was based on conscious reflection on the role of subjectivity. Importantly, the breadth of our findings exemplifies the complexity of identifying and ascertaining user modelling attributes. The aim of this paper was to employ design research as a way to reveal new avenues for user modelling, while concurrently investigating the value of using such research as a boundary object within the design context of ITS more broadly. To this end, Table 3 summarises how each of the potential attributes contained within the boundary object was either implemented in the user model, deemed out of scope, applied as a design principle or generated new research directions. We next explore how different stakeholders within our project interpreted the research, the dynamics it introduced, and the role of the boundary object in moderating design decisions regarding user modelling and beyond.

Table 3 about here

Establishing Common Ground

When tackling new design problems, designers must become quasi-domain experts (Buchanan, 1992). The outcomes of our design research served as a springboard for designers and technologists to develop this expertise. The knowledge they acquired helped create common ground with the domain experts of the project, mainly by providing them with a foundation for interpreting conversations and recognising design opportunities. For instance, while the AI researchers of the project were aware of the centrality of linguistic knowledge, it was not until they had engaged with the design research that they understood what these linguistic difficulties entailed and the potentially vast scope of the linguistic knowledge requiring attention. More specifically, prior to us undertaking the design research, the domain experts had presented other project stakeholders with an exemplar teaching session in an attempt to unpack the rationale and decision-making involved. The highly technical and complex nature of teaching had made it difficult to discern discrete strategies that were computationally feasible. The boundary object reduced this complexity to a set of higher-level dimensions that were accessible to everyone's understanding. Using these dimensions as a shared point of reference allowed the domain experts to drill down to the particulars. For example, the domain experts had often spoken of derivational and inflectional suffixes, which was a challenging concept for other project stakeholders. Basing themselves on the *linguistic level* attribute, domain experts stopped using such technical language and instead explained the underlying four rules of suffixing in English (add, drop, double and change). These rules refer to how the stem word, particularly the last letter, is affected by suffixing. Along with examples, this helped communicate the linguistic difficulty of suffixing in a form all the project stakeholders could understand (e.g. for the double rule, run becomes running, the "n" being doubled). Therefore, the foundational and coarse-grained knowledge expressed in the boundary object made communicating across stakeholder expertise easier as it provided a common ground that could be built upon and represented a neutral language through which the teams could communicate. This understanding gave design and AI researchers the confidence to organise a series of collaborative activities with the domain experts in which they determined the project's scope of linguistic knowledge. While these activities did not use the boundary object per se (they instead branched out from one of its attributes), they were triggered by the design research.

Establishing common ground through the boundary object additionally helped inform concrete design decisions. For example, the domain experts had highlighted early on the importance of considering children's *interests*. However, given the multiple interpretations that this user attribute invites, AI researchers had not implemented this recommendation with a corresponding design decision. Indeed, the complexity of *motivation* and *interests* was elucidated through the specific teaching strategies the design research found with regards to embedding a child's interests into lessons. The concrete examples of teaching strategies suggested by our research were discussed in a project meeting. These strategies assisted the AI researchers in considering how to both model *interests* and embed this attribute within the system design, subsequently integrating it within the user model with a view to adapt the reader through the selection of books and texts by genre. The boundary object was thus used to translate broad teaching concepts that the domain experts understood and intuitively practiced into a form our technology experts could understand and implement.

Author Pre-Print Copy

In summary, the boundary object mediated our project discussions, it allowed diverse project stakeholders to access and discuss complex issues and it supported the process of decision-making. The difficulties our project has had in terms of establishing common ground are by no means unique. However, our work suggests that in approaching findings from design research as a boundary object, diverse stakeholders can be brought together to the benefit of projects.

Transparency as a Consequence of Possible Alternatives

Designers will often unconsciously introduce their own values, perspectives, and biases when designing new technologies. In critically appraising their role and responsibility in shaping human behavior, it is necessary that designers bring these to light (Sengers et al., 2005). In multi-disciplinary projects where design decisions are negotiated and shared, we argue that reflection on design choices becomes even more important. Our project's focus on reading, alongside the salient representation of domain expertise within our team, meant that there was a strong incentive to model students' linguistic abilities following a linear pathway for their progress. Our design research presented new alternatives, which enhanced our critical awareness. In recognising the different teaching strategies for dyslexia, our design and development team became aware of the holistic teaching approach often adopted in dyslexia teaching centres and school contexts. Therefore, even though our design research supported the primarily behavioural approach to learning advocated by the domain experts of the project, it also highlighted the importance of supporting the development of strategies that foster independence, cognitive styles and motivation. This encouraged us to consider other possibilities, in turn supporting recent arguments that a rounded understanding of how people learn may be best suited to developing learning technologies (Laurillard, 2012). Indeed, the breadth of these alternatives prompted the domain experts of the project to articulate additional teaching strategies – namely those supporting *motivation* and *self-esteem* – which they had neglected to acknowledge as a consequence of focusing too closely on children's phonological deficits. After the design research had been completed, the domain experts responded to our identification of *motivation* as a potential attribute by presenting the many strategies they use within their teaching sessions (such as never stating that an answer is wrong but focussing on teaching how to make the answer more correct).

In other cases, a reflective, alongside a pragmatic, lens on our design process was applied to consciously choose *not* to follow particular directions due to the complexity they introduced. For example, while *peer learning* was identified to be pedagogically beneficial, its associated technical difficulties were deemed to be too complex. *Peer learning* necessitated a complete re-design of the software, as the networking system and the login functionality would require additional functionality. To overcome these difficulties, a social solution would be to have pairs of students working on a single tablet. However, such a decision would undermine the operation and accuracy of the user model, and indeed the individualised teaching approach of the ITS. In balancing these complexities with the anticipated benefit *peer learning* would accrue, *peer learning* was not further developed. A second example of a direction we chose to pursue only in part was the student's *interests* within the software. In the intervention sessions we observed, teachers would often tailor the content to an extremely fine grain level, for example by using dinosaurs or birds to appeal to children's interests. Recording such fine grain information within the user model would mean that any texts input into the reader would have to be automatically processed to establish whether they corresponded with these detailed *interests*. Additionally, providing example words within learning game activities of the software would necessitate a large set of technical vocabulary to capture each *interest*, resulting in a significant amount of additional work. Recognising the importance of interests, and considering the alternatives, a compromise was reached whereby the selection of books and texts would be guided by a child's preference for a particular genre of books, stored within the user model.

Therefore, the boundary object revealed a wide range of possibilities for our user model. In broadening the dominant perspective in the project, it centred the discussion on other important considerations such as *motivation* that had not been discussed. It also supported the process of decision-making: in considering all pedagogical alternatives, we had to balance our understanding of the pedagogical benefit of a given attribute against the complexity of extracting and processing information on that attribute and the ways in which the software would adapt based on that attribute. This assisted our design and development team to make decisions from a vantage point of understanding, reflecting and considering the technical complexity on the alternatives rather than relying on a prescribed direction. Having developed an initial set of recommendations for the user model, the boundary object was then used as a means of achieving consensus across the wider team. Circulated beforehand, it acted as a focal point, allowing discussion and consensus to be achieved in a way that ensured that the entire team understood that decisions were taken on the basis of evidence, and that alternatives had been considered in light of their technical demands while assessing where the most impact could be created.

Author Pre-Print Copy

A Springboard For New Activities

Our design research became an end in itself as well as a foundation for further activity, both for user modelling and design. As an end, in line with our discussion above, it inspired the introduction of a further attribute to our user model, *interest*. However, it also revealed critical user attributes that were too time intensive to develop within the framework of a user model, such as, for example, children's *co-occurring difficulties*. Designers within the project drew upon these attributes to develop a set of baseline pedagogical and accessibility design principles, as presented in Table 4, that were subsequently applied throughout the design of the software. An example principle was to avoid moving text due to some children's difficulty in reading it. In generating follow up research, the designers built upon the importance of *motivation* and *self-esteem* through involving children within an informant design process, where we involved children at various stages of the design process to which they could make important design contributions (Scaife et al. 1997). Aiming at ensuring that children would be motivated to interact with the learning game embedded in the software, designers held a series of focus groups with children to identify engaging game genres. This led to an initial game concept, which was followed up in an art workshop where children within our target age group used art materials to develop motivating narratives and characters for the game (Benton et al., 2014). Moreover, with a view to strengthen children's *self-esteem*, a further series of workshops was organised comprising of children with reading difficulties, where ideas for appropriately pitched in-game feedback were elicited and developed. The design methodology developed to support children's ideation was in part based on the teaching strategies observed within the present research; for instance, we provided children with clear instructions and cues regarding their progress (Benton et al., 2014).

Table 4 about here

Future Work

The aim of this paper has been to explore the role of design research in the context of ITS. We note that while research in the early phase of design can introduce alternative perspectives of a design problem, methodologies such as knowledge elicitation (Porayska-Pomsta et al, 2013) will continue to play a vital role in informing later phases through the collection of fine-grained annotated data. While our approach can be viewed as a first step to designing a user model, future work can also explore whether boundary objects can make a more direct methodological contribution to user modeling. Researchers have begun to move towards using Distributed Cognition as a methodological and operational frame of reference for user modeling. Such an approach "drives design by elements of human-oriented theory rather than by technological-oriented assumptions" (Eberle et al., 2011). In alignment with this perspective, boundary objects such as our own could be viewed as a form of human-oriented theory. Our boundary object encapsulates both theory and practice, further enriching the user model design. Such a framework may offer many advantages but is particularly powerful when coupled with processes underpinned by models of the software such as the Model-Driven Architecture (MDA) (Eberle et al., 2011; Brown, 2004). While a detailed description of MDA is outside the scope of this paper, the essential concept is that as models of the software are refined, code can automatically be generated to perform the functions that have been modeled. In these circumstances, our boundary object can offer a starting point; by further refining the model of the user model, and defining its interaction with other software components, MDA can allow the automatic generation of the user modeling code based on the design research undertaken.

CONCLUSION

The field of HCI has typically conversed with AI in improving the use of AI algorithms after they have been developed (Grudin, 2009). The aim of this paper was to show that if introduced early in the design process, design research can contribute new directions for user modelling as well as inform the design of ITS more broadly. Specifically, the design research we conducted aimed at informing a user model for an ITS designed to help develop the reading skills of children with dyslexia. We adopted a mixed methods approach to capture information about teachers' existing practices with a focus on how they tailor their teaching to the needs and characteristics of individual students with dyslexia. By undertaking an inductive, thematic analysis on the data collected, we identified 19 potential attributes for the user model. These attributes were organised across a number of themes, namely linguistic ability, cognitive style, motivation and interests, goals, personality and behaviour, and co-occurring characteristics and difficulties.

Author Pre-Print Copy

In line with our research objective, we treated the attributes, alongside example teaching strategies, as a boundary object within our multidisciplinary development team. We discovered that design research assisted in creating common ground between domain experts and other stakeholders leading to further collaborative work and supporting decision-making. Furthermore, the breadth of perspectives captured within the design research supported more criticality and accountability amongst our team, leading to the introduction and recognition of previously overlooked teaching strategies. Finally, one of the strongest functions of our design research was to support new avenues of research and development, both for user modelling and software design. Taking this approach in our project has created synergy and cohesion between user modelling and other design decisions. While the roles of our boundary object may have been in part bound to the politics and personalities of our project stakeholders, we hope that future research will continue to build on this work toward accumulating a deeper understanding of boundary objects in the context of ITS and other multidisciplinary development projects.

Our design research also contributes a rich dataset of different teachers' approaches to personalising lessons to match the needs of children with dyslexia. This work can be of value to teachers, assisting them to reflect upon their own practices and extend them. It additionally provides technologists wanting to develop new tools and technologies to support children with dyslexia in reading with a much-needed inclusive foundation. This speaks to a bigger issue. User models tend to be developed within the confines of specific projects. Without any means of communicating the basis for the design of that user model, work is frequently repeated in developing new user models for similar projects. This means that the user model itself, and the processes used to create it, are lost. What is needed is some form of externalised representation of the user model that can be extended and adapted for future projects. A boundary object created through the process of design research is ideal for this purpose.

ACKNOWLEDGEMENTS

This work was funded by the ILearnRW (project no: 318803) FP7 ICT EU project. We thank all of the teachers and children who took part in this research and graciously donated their time. We also acknowledge the three reviewers whose comments improved this work.

REFERENCES

1. Akbari, R., 2007. Reflections on reflection: A critical appraisal of reflective practices in L2 teacher education. *System*, *35*(2), pp. 192-207.
2. Anderson, J. R., 1992. Intelligent tutoring and high school mathematics. In *Intelligent tutoring systems*, pp. 1-10. Springer Berlin Heidelberg.
3. Benton, L., and Vasalou, M., 2013. The Wicked Problem of Undertaking Responsible and Ethical Research in Multidisciplinary and Geographically Diverse. Observatory for Responsible Research and Innovation in ICT, accessed from http://responsible-innovation.org.uk/torrii/resource-detail/1081.
4. Benton, L., Vasalou, A., Khaled, R., Johnson, H., & Gooch, D., 2014. Diversity for Design: A Framework for Involving Neurodiverse Children in the Technology Design Process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press.
5. Bloom, B. S., 1984. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, *13*(6), pp. 4-16.
6. Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), pp. 77-101.
7. Brown, A., 2004. An Introduction to Model Driven Architecture.

   http://www.ibm.com/developerworks/rational/library/3100.html
8. Brusilovsky, P., 2003. Developing adaptive educational hypermedia systems: From design models to authoring tools. In *Authoring Tools for Advanced Technology Learning Environments* (pp. 377-409). Springer Netherlands.
9. Buchanan, R., 1992. Wicked Problems in Design Thinking. *Design Issues*, 8(2), pp. 5-21.
10. Cohen, P. A., Kulik, J. A., and Kulik, C. L. C., 1982. Educational outcomes of tutoring: A meta-analysis of findings. *American educational research journal*, 19(2), pp. 237-248.
11. Conati, C., and Maclaren, H., 2009. Empirically building and evaluating a probabilistic model of user affect. *User Modelling and User-Adapted Interaction*, 19(3), pp. 267-303.
12. Conati, C., Chabbal, R., & Maclaren, H., 2003. A study on using biometric sensors for monitoring user emotions in educational games. In *Workshop on Assessing and Adapting to User Attitudes and Affect: Why, When and How*.
13. Costagliola, G., De Rosa, M., Fuccella, V., Capuano, N., and Ritrovato, P., 2010. A Novel Approach for Attention Management in E-learning Systems. In *DMS*, pp. 222-227.
14. Demonet, J., Taylor, M. and Chaix, Y., 2004. Developmental dyslexia. *Lancet*, 363, pp. 1451-1460.
15. Dorst, K., 2011. The core of 'design thinking' and its application. *Design studies*, *32*(6), pp. 521-532.
16. Druin, A., 2002. The role of children in the design of new technology. Behaviour and IT, 21 (1), pp. 1-25.
17. Eberle, P., Schwarzinger, C. and Stary, C., 2011. User modelling and cognitive user support: towards structured development, Universal Access in the Information Society, 10(3), pp. 275-293.
18. Fallman, D. and Stolterman, E., 2010. Establishing criteria of rigour and relevance in interaction design research. In *Digital Creativity*, 21(4), pp. 265-272.
19. Gaver, W., 2012. What should we expect from research through design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 937-946. ACM Press.
20. Good, J. and Robertson, J., 2006. CARSS: A Framework for Learner-Centred Design with Children International Journal of Artificial Intelligence in Education, 16 (4), 381-413.
21. Goswami, U., 2008. Learning difficulties: Future challenges. A paper prepared as part of the Foresight Review on Mental Capital and Wellbeing, accessed from http://www.bis.gov.uk/assets/foresight/docs/mental-capital/learning_difficulties.pdf
22. Grudin, J., 2009. AI and HCI: Two fields divided by a common focus. *AI Magazine* 30(4), pp. 48-57.
23. Khaled, R. and Ingram, G., 2012. Tales from the front lines of a large-scale serious game project. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 69-78. ACM Press.
24. LaFontaine, W., 2013. Global Technology Outlook 2013. IBM report, accessed from http://www.zurich.ibm.com/pdf/isl/infoportal/Global_Technology_Outlook_2013.pdf
25. Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.

Author Pre-Print Copy

26. Laurillard, D., 2012. Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology. Routledge, Taylor & Francis Group.
27. Lombard, M., Snyder‐Duch, J., and Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human communication research*, *28*(4), pp. 587-604.
28. Mota, S., and Picard, R. W., 2003. Automated posture analysis for detecting learner's interest level. In *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03.* Vol. 5, pp. 49. IEEE.
29. Orwant, J., 1996. For want of a bit the user was lost: Cheap user modelling. IBM Systems Journal, 35(3/4), pp. 398-416.
30. Perry, M., 2003. Distributed Cognition. HCI, Models, Theories, and Frameworks, London: Morgan Kaufman.
31. Porayska-Pomsta, K., Mavrikis, M., and Pain, H., 2008. Diagnosing and acting on student affect: the tutor's perspective. *User Modelling and User-Adapted Interaction*, *18*(1-2), pp. 125-173.
32. Porayska-Pomsta, K., Mavrikis, M., D'Mello, S., Conati, C., & Baker, R. S., 2013. Knowledge Elicitation Methods for Affect Modelling in Education. In *International Journal of Artificial Intelligence in Education* 22, pp. 107-140.
33. Radlinski, R., and McKendree, J., 1992. Grace meets the "real world": tutoring COBOL as a second language. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 343-350. ACM Press.
34. Ragnemalm, E. L., 1995. Student diagnosis in practice; bridging a gap. *User Modelling and User-Adapted Interaction*, *5*(2), pp. 93-116.
35. Rose, J., 2009. Identifying and Teaching Children and Young People with Dyslexia and Literacy Difficulties. Report for the Department of Children, Schools and Families, June 2009.
36. Scaife, M., Rogers, Y., Aldrich, F. and Davies, M., 1997. Designing for or designing with? Informant design for interactive learning environments. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (CHI '97). ACM, New York, NY, USA, pp. 343-350.
37. Selwyn, N., 2011. Education and technology: Key issues and debates. Continuum International Publishing Group.
38. Sengers, P., Boehner, K., David, S. and Kaye, J., 2005. Reflective design. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility* (CC '05), pp. 49-58. ACM Press.
39. Snowling, M.J., 2008. Dyslexia. A paper prepared as part of the Foresight Review on Mental Capital and Wellbeing, accessed from http://www.bis.gov.uk/assets/foresight/docs/mental-capital/sr-d2_mcw_v2.pdf
40. Star, S. L., and Griesemer, J. R., 1989. Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social studies of science*, 19(3), pp. 387-420.
41. Tinsley, H. E. and Weiss, D. J., Interrater Reliability and Agreement. *In:* Tinsley, H. E., & Brown, S. D. (Eds.). (2000). *Handbook of applied multivariate statistics and mathematical modeling*. Academic Press.
42. Vasalou, A., Ingram, G., and Khaled, R., 2012. User-centered research in the early stages of a learning game. In *Proceedings of the Designing Interactive Systems Conference*, pp. 116-125. ACM Press.
43. Vasalou, A. and Khaled, R., 2013. Designing from the Sidelines: Design in a Technology-Centered Serious Game Project. Proceedings of CHI Workshop "Let's talk about Failures: Why was the Game for Children not a Success?"
44. Vines, J., Clarke, R., Wright, P., McCarthy, J., & Olivier, P., 2013. Configuring participation: on how we involve people in design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 429-438. ACM Press.
45. Walenstein, A., 2002. Foundations of Cognitive Support: Toward Abstract Patterns of Usefulness. In *Proceedings of the 14th Annual Conference on Design, Specification, and Verification of Interactive Systems,* pp. 133-147, Springer Berlin Heidelberg.
46. Wright, P., Fields, R. and Harrison, M., 2000. Analyzing Human-Computer Interaction as Distributed Cognition: The Resources Model. Human-Computer Interaction, 15, pp. 1-41.

Author Pre-Print Copy

Method	Participants
Observations	Dyslexia Teaching Centre A: 2 dyslexia specialist teachers and 10 students (6 male, ages 7-12). Dyslexia Teaching Centre B: 1 dyslexia specialist teacher and 11 students (7 male, ages 8-16). Mainstream Primary School A: 1 mainstream teacher in a low ability class with 14 students (mixed gender, ages 10-11). Mainstream Primary School A: 1 mainstream teacher in a high-ability class with 20 students (mixed gender, ages 10-11).
Interviews	4 dyslexia specialist teachers (1 male) 2 SENCO – special education needs coordinators (2 female) 1 mainstream teacher (1 male)
Online Survey	3 dyslexia specialist teachers reporting on 5 students (4 male, ages 9-12).

Table 1: Summary of the methods and participants deployed across the study.

Author Pre-Print Copy

Code for Potential User Modelling Attribute	**Questionnaire Data**	**Interview Data**	**Observation Data**
Linguistic level	$\kappa = 0.64$ (p<0.01), 95% CI (0.58 to 0.70)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.78$ (p<0.001), 95% CI (0.76 to 0.81)
Linguistic strategies	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.74$ (p<0.001), 95% CI (0.54 to 0.94)
Multisensory approach	$\kappa = 0.64$ (p<0.01), 95% CI (0.58 to 0.70)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Overlearning	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Introduction of new content	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.85$ (p<0.001), 95% CI (0.82 to 0.88)
Self checking	$\kappa = 0.64$ (p<0.01), 95% CI (0.58 to 0.70)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.80$ (p<0.001), 95% CI (0.76 to 0.83)
Presentation of instructions	$\kappa = 0.77$ (p<0.001), 95% CI (0.73 to 0.81)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.74$ (p<0.001), 95% CI (0.71 to 0.77)
Interests	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Construction of learning environment	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.79$ (p<0.001), 95% CI (0.75 to 0.83)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Peer learning	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Lesson structure	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.78$ (p<0.001), 95% CI (0.76 to 0.80)
Presentation of feedback	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.91$ (p<0.001), 95% CI (0.89 to 0.93)	$\kappa = 0.91$ (p<0.001), 95% CI (0.90 to 0.92)
General aim	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Specific goals	$\kappa = 0.87$ (p<0.001), 95% CI (0.85 to 0.89)	$\kappa = 0.79$ (p<0.001), 95% CI (0.75 to 0.83)	$\kappa = 0.92$ (p<0.001), 95% CI (0.90 to 0.93)
Managing behaviour	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.77$ (p<0.001), 95% CI (0.74 to 0.80)	$\kappa = 0.80$ (p<0.001), 95% CI (0.76 to 0.83)
Dealing with failure	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.77$ (p<0.001), 95% CI (0.74 to 0.80)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Social and emotional issues	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.66$ (p<0.001), 95% CI (0.59 to 0.72)	$\kappa = 0.66$ (p<0.001), 95% CI (0.60 to 0.72)
Specific co-existing conditions (e.g. Dyspraxia, ADHD, ADD etc.)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)
Associated difficulties (e.g. self-esteem, memory load, organisational ability, moving text)	$\kappa = 1.00$ (p<0.001), 95% CI (1.00 to 1.00)	$\kappa = 0.81$ (p<0.001), 95% CI (0.78 to 0.83)	$\kappa = 0.80$ (p<0.001), 95% CI (0.76 to 0.83)

Table 2: Cohen Kappa Coefficient for Interrater Agreement (κ), Significance Levels (p) and Confidence Intervals (95% CI)

Author Pre-Print Copy

Potential User Modelling Attributes	Example Teaching Strategy	Design Implication
Linguistic Ability		
Linguistic level	Assess the child's current level of linguistic knowledge and give the student appropriate exercises	✔
Linguistic strategies	Equip students with effective strategies that they can use independently to solve problems	★
Cognitive Style		
Multisensory approach	Present learning materials or allow the student to answer questions in a variety of forms	★
Overlearning	Previously acquired skills need to be re-practiced. The selection of which skills to work on can be adjusted for recent skills or older skills	✔
Introduction of new content	Generally 99% familiar content to 1% new content. This proportion can be adjusted according to how well the student can cope with challenging content	★
Self checking	Allow the student to self-check their own work where appropriate	✖ ○
Presentation of instructions	Provide short, precise spoken instructions	★
Motivation and Interests		
Interests	Use content related to the student's interests	✔
Construction of learning environment	Present images of famous people with dyslexia	✖
Peer learning	Allow older children to assist younger children on certain exercises	✖
Lesson structure	Always keep the lesson structure consistent	★
Presentation of feedback	Combine praise with a discussion of what areas still need to be worked on	★
Goals		
General aim	Tailor exercises towards specific linguistic skills	✖
Specific goals	Tailor exercises towards a student's specific goals (e.g. reading a particular book)	✖
Personality and Behaviour		
Managing behaviour	Make clear in the lesson plan which students get to perform which actions that day (e.g. select the game at the end of the session)	✖
Dealing with failure	Select activities that provide an appropriate level of challenge	★ ○
Social and emotional issues	If overly tired, moderate the difficulty of the activities, perhaps focus on overlearning skills	✖
Co-occurring Characteristics and Difficulties		
Specific co-existing conditions (e.g. Dyspraxia, ADHD, ADD etc.)	Use a sequence of fast moving activities	★
Associated difficulties (e.g. self-esteem, memory load, organisational ability, moving text)	Do not utilise games including moving text	★ ○

Author Pre-Print Copy

Table 3: A summary of user attributes, example teaching strategies and design implications. Design implications are annotated as follows: Implemented in user model (✔), out of scope (✖), implemented as a static design principle (★), new direction for investigation (○)

Design Principle	Linguistic Ability	Cognitive Style	Motivation and Interests	Goals	Personality and Behaviour	Co-occurring Characteristics and Difficulties
Provide a succession of short activities		x	x		x	x
Never say something is wrong but rather focus on teaching the strategy to correct it	x		x			
Provide the child with tips and hints for discovering the right answer	x					
Children are not allowed to move on without understanding what they did wrong		x				
Allow a child to beat their previous scores			x			
Start with presenting progress and then present areas of weakness					x	
Allow children to decide when they are ready to go up to the next level			x	x	x	
Don't give praise for praise sake		x	x			
Present 99% old content to 1% new		x				
Ensuring all activities are multisensory		x				
Provide dyslexia-friendly presentation of text. No moving text.						x
Provide an audio option for all text instructions						x
Provide clear navigation throughout the system						x
Do not increase cognitive load by playing audio in the background for phonological tasks						x

Table 4: Design principles for dyslexia friendly software

Author Pre-Print Copy