Santos, P., Cook, J., & Hernández-Leo, D. (2015). M-AssIST: Interaction and Scaffolding Matters in Authentic Assessment. *Educational Technology & Society, 18* (2), 33–45.

M-AssIST: Interaction and Scaffolding Matters in Authentic Assessment

Patricia Santos[1*], John Cook[1] and Davinia Hernández-Leo[2]
[1]Department of Arts and Cultural Industries, University of the West of England, Bristol, UK // [2]Department of Technologies and Information Technologies, University Pompeu Fabra, Barcelona, Spain // Patricia.Santosrodriguez@uwe.ac.uk // John2.Cook@uwe.ac.uk // davinia.hernandez@upf.edu
[*]Corresponding author

ABSTRACT

Authentic assessment is important in formal and informal learning. Technology has the potential to be used to support the assessment of higher order skills particularly with respect to real life tasks. In particular, the use of mobile devices allows the learner to increase her interactions with physical objects, various environments (indoors and outdoors spaces), augmented digital information and with peers. Those interactions can be monitored and automatically assessed in a way that is similar to traditional objective tests. However, in order to facilitate a meaningful interaction with formative purposes, we propose that the assessment process can be assisted through scaffolding mechanisms that transform the mobile system into a 'more capable peer'. In this context, this paper presents the m-AssIST model which captures the necessary emergent properties to design and analyse m-assessment activities. The model is used to analyse the benefits and limitations of existing m-test based systems. This paper discusses the importance of meaningful interactions, and the provision of scaffolding mechanisms to support formative and authentic assessment.

Keywords

Mobile-assessment, Authentic assessment, Meaningful interactions, Scaffolding

Introduction

Authenticity in assessment has been widely discussed by several authors in this research field. As Brown, et al., (1989) assert, authentic or situated learning can be seen from the perspective of social constructivism, it occurs when the learner takes part in relevant activities to her real life and which take place within a culture similar to an applied context. Gulikers et al., (2004) state that current educational and professional curriculums are more focused on the development of competences than on a simple knowledge acquisition. In this sense, ILA-practices (Instruction, Learning and Assessment) are characterized as following an instructional-approach focused on learning and competence development, supporting reflective-active knowledge construction, contextualized, interpretative and performance assessment (Birenbaum, 2003). The goal of ILA practices is the acquisition of higher-order thinking skills instead of factual knowledge and basic skills. As a consequence the assessment has a formative goal. However, assessment in formal learning is commonly held to contribute to feedback to students on their learning and the certification of their achievement (Boud & Falchikov, 2006).

The term "higher-order thinking skills" (HOTS) is used to delineate cognitive activities that are beyond the stage of understanding and lower levels of application (Bloom & Krathwohl, 1956). In this paper, we use the definition provided by Kaminsky et al., (1997) who define these skills as the ones where the level of thinking depends upon the context, with a real-world situation, where the individual has to apply, reorganize and embellish knowledge in the context of the thinking situation. Assessing learners' HOTS is important in formal learning but is particularly important in informal learning contexts such as the workplace. One key proponent of an empirical tradition of work-based learning research is Eraut. In Eraut (2007) the author proposes a typology of early career learning, where he emphasizes the importance of consolidating, extending and refining skills in workplace environments. In this context, self-assessment has an important role in order to support staff and management to develop their skills (Boud, 2013). Considering the future of younger learners, some authors have stated that new ways of thinking about assessment in formal education have to be proposed (for example Boud, 2006; Davies, 2010). Brown et al. (1989) identifies: "*many of the activities students undertake (in formal education) are simply not the activities of practitioners and would not make sense or be endorsed by the cultures to which they are attributed… When authentic activities are transferred to the classroom, their context is inevitably transmuted; they become classroom tasks and part of the school culture... Classroom tasks, therefore, can completely fail to provide the contextual features that allow authentic activity*."

In this context *authentic* means that learners should be able to demonstrate and practice their skills like in real life tasks. One of the main problems in formal learning occurs when some of the considered authentic skills are included in educational curriculums or training programmes, but the methods to assess them are not appropriate (Benett, 1993; Boyle & Hutchison, 2009). The main aim of this approach is to provide meaningful learning/assessment and promoting the transfer of knowledge to real-life situations (Choi & Hannafin, 1995). Choi and Hannafin state that knowledge is *situated*, when it is product of an activity, context and culture in which it is developed and used. We agree with the vision proposed by these authors and for this reason this paper claims that mobile technology is the most suitable instrument for supporting the assessment of authentic/situated tasks due to its "situational" facilities.

The European Network of Excellence in Technology Enhanced Learning (TEL), STELLAR published a report with the title "Education in the wild: contextual and location-based mobile learning in action" (Brown, 2010) claiming that a re-conceptualization of educational theories have to be undertaken in order to develop an approach for the present and future of learning designs. The use of mobile devices enables technicians and practitioners to rethink learning, teaching and assessment strategies. As Engeström (1999) states: "*human activity is endlessly multifaceted, mobile, and rich in variations of content and form…the theory of activity should reflect that richness and mobility*."
In related research, Cook (2010) proposes the term "*Augmented Contexts for Development" (ACD) where: learners can use mobile devices (e.g., smartphones) in combination with augmented digital information to interact with each other and interact with the physical and a virtual environment with the goal of creating their own ACD. The smartphone and the location-triggered learning activity it holds are considered here as the "more capable peer*" (Vygotsky, 1978, p. 86) because the device provides guidance to the learner in order to solve a problem and as a consequence the learner augments her context of development. The main elements to develop the ACD are: (a) the physical environment, (b) a pedagogical plan (e.g., an assessment activity), (c) tools/devices for an augmented oriented approach, (d) learner co-constructed "temporal context for development" and (e) collaborative learners' interpersonal interactions using tools. We observe here the importance of interacting with the environment and with other peers, for this reason we use the definition provided by Woo and Reeves (2007) where they present the term "meaningful interaction" as those "*interactions that have direct influence on learners' intellectual growt."* The ACD approach is also supported by the definition of cognitive systems proposed by Hollan et al., (2000) which claim that the organization of cognitive systems is extended to cover the interactions between people and with the resources and materials of the environment distributed in time.

Taking into account the above approaches we propose that a *pedagogical plan* can consist in an authentic assessment activity performed in a real-world environment. But we claim that this plan has to be supported by scaffolding mechanisms in order to help the learner to solve the problem by doing meaningful interactions through her mobile device. We use the description of scaffolding provided by (Wood et al., 1976): "*process that enables a child or a novice to solve a problem, carry out a task, or achieve a goal which would be beyond his unassisted efforts*." As Lieberman & Linn (1991) state, scaffolding is one of several ways to encourage self-directed learning in students, especially in terms of TEL scaffolding it is useful to support self-assessment and formative assessment methods and provide automatic assistance. This paper claims that the monitoring of the interactions done by the learner during the assessment activity through her mobile device can be used to *assist* the learner in real time and provide helpful assistance in order to achieve the adequate meaningful interactions and corresponding knowledge. Our view is that mobile devices have the potential to support more authentic types of assessment integrated within real locations/objects in ways that stimulate and motivate learners and support their meaning making.

From these approaches, we assert that: the Environment, Time, Interactions and Scaffolding mechanisms matter when designing an authentic assessment activity for a corresponding *context*. We use the term "*emergent properties"* proposed by Dourish (2004) to refer to the elements that allow us to define the context of an authentic assessment activity. In particular we show how the meditation of those interactions through mobile devices is a key issue to ensure the authenticity of the activity in a specific yet emergent context (i.e., time + location). We capture these emergent properties in a model called m-AssIST. The m-AssIST model can be used to analyse existing m-assessment scenarios in order to identify limitations related to the assessment of HOTS. The model can also be used during the design process of an assessment activity in order to reflect about the importance of these properties for the success of the authenticity of the activity.

From the wider number of different assessment instruments, this paper focuses its analysis on the use of automatic methods of assessment (i.e., objective tests commonly used to support self-assessment). We agree with the description provided by Boud (2013) that self-assessment has to "*stress the importance of learners constructing rather than receiving knowledge, of promoting the taking of responsibility for learning."* So, this previous sentence stresses the formative strength of self-assessment methods, and the importance of changing the role of the learner from a passive one to an active one though the use of mobile devices.

In the following section, we discuss concepts for assessment in TEL and scaffolding mechanisms (we define these below) to support the assessment and development of authentic tasks and HOTS. After the discussion of the literature, the emergent properties of the m-AssISt model are described. In order to show the application of the m-AssIST model, this paper also contains a section reviewing three existing m-assessment scenarios where the goal was to assess situated HOTS, this review will be followed by an analysis using the m-AssIST model in order to identify those scaffolding mechanisms and strategies to support meaningful interactions which could be applied in the selected scenarios. The scenarios are significant because they have taken into consideration scaffolding strategies to support an authentic m-assessment activity. The three scenarios are sufficiently different to illustrate diverse ways of interacting with authentic environments (i.e., a natural park, an old city district, and a botanical garden) in order to assess different HOTS assigned to different levels of learning and subjects. Finally, we finish with a final reflection of the benefits of the emergent properties captured in the m-AssISt model.

Relevant literature

In this section we elaborate on our claim that mobile devices have the potential to augment and support authentic assessment. In order to do that, we first need to understand how technology has been employed to support traditional assessment, and how the analysis of the limitations of traditional assessment has allowed the identification of new ways of supporting more authentic assessment in TEL.

Computer Assisted Assessment

Most of researchers in the Computer Assisted Assessment (CAA) field have concentrated their efforts in creating innovative question-items or using complex computational solutions to represent complex questions and tests. They claim that the creation of new types of questions can improve the measuring of learners' HOTS. Creating advanced assessment scenarios should involve the use of technology to perform test-activities which would be impossible or very expensive to reproduce using traditional methods of assessment (e.g., paper & pencil). Computers offer the possibility of using simulations, managing a big quantity of updated and enriched information, increasing the interaction with the information and making the learner more participative in the assessment process (Conole & Warburton, 2005).

Bull & McKenna (2003) show in their blueprint for CAA that the use of objective tests (or automatic tests) provides advantages such as: improving the interactivity with the learning content, presenting a question, obtaining a response, evaluating a response, providing a mark and answering with feedback automatically. Despite the benefits of using automatic tests, the traditional transcription from paper to computer is not enough to assess authentic tasks and HOTS such as: problem solving, problem exploration, applying, collaboration, analysing, discovering rules, spatial or time perception, among others (Mayotte, 2010).

Previous research projects have studied how to use objective tests in lifelong learning contexts (considered to mean more "authentic" contexts than the classroom). On the one hand, the EU FP7 research project 'TEN-Competence' developed an Open Source environment to support self-directed lifelong learning based on IMS standards (Louys et al., 2009). The system included web-based tools for self-assessment and embedded-assessment in competence oriented scenarios. The self-assessment outcomes (obtained after answering an automatic test) were used to help the learner to identify her competence gaps and generate her learning path for competence development. Furthermore, the ICOPER project (Crespo et al., 2010) was focused on finding ways that can effectively bring together educational achievements that should results in outcome-based assessment and units of learning. One of the contributions of this project was a conceptual map for learning assessment. This model comprises the whole process, including the activity of the learners taking a computational exam, as well as the phase where instructors include the corresponding feedback/grades. The project showed how the adoption of IMS standards (IMS, 2006) for learning outcomes definitions facilitates to some extent the alignment between assessment outcomes and the educational curriculums.
Both projects contributed significantly to the research field of CAA, however with the fast evolution of the technology in the last 5 years we find that some of the contributions are outdated for the following reason: IMS does

not consider complex representations of questions and tests, instead it is typically used to represent traditional types of questions such as Multiple Choice or Multiple Response in order to assess knowledge. As a consequence, the tests developed by these computational-based systems do not support the assessment of situated knowledge, where the assessment activity is developed at least in a similar context (i.e., location) than the real associated one.

By drawing on our analysis of the previous approaches, we can now identify the main limitation of using computers to assess authentic tasks. First of all, its staticity: although complex simulations of the real-world can be pre-defined, this kind of approach is limited in terms of being used in any situational (uncontrolled) activity where the *location* and specific *time* where the action is performed can have an impact over the solution that has to be applied by the learner. Second, it is not possible to use this approach to *interact* with physical objects or people in an outdoors environment like the ACD example given above.

Boyle & Hutchison (2009) stated that the more technological resources you use, the higher skills and more sophisticated tasks can be evaluated. The JISC report "Effective assessment in digital age" (Davies, 2010) claimed that technology has to be used to create authentic assessment. For this reason some researchers state that the use of handheld devices will transform the learning and assessment practices (Thompson, 2005), opening up the possibilities of using other technologies to support assessment based on tests. In particular, the wider use of mobile devices has come up with a positive transformation in the assessment of authentic practices.

Computing – Based testing: Adding authenticity to Computer Assisted Assessment

As proposed in a previous publication (Santos et al., 2012) we consider the term "Computing" instead of "Computer" as more appropriate when referring to assessment scenarios mediated with a wider range of learning technologies and devices. In particular, our interest is focused on the use of different mobile devices such as smart-phones, PDA's, consoles or supporting technologies such as NFC or RFID tags, GPS and Bluetooth to support more authentic assessment (Santos et al., 2013, de-Marcos et al., 2010; Hwang & Chang, 2011; Susono & Shimomura, 2006; among others). The use of appropriate technologies increases the possibilities of creating new types of questions, tests and assessment activities based on tests more adapted to real life interactions, situations and locations. The word *computing* covers the different ways of technologically representing and manipulating a test. See Figure 1.

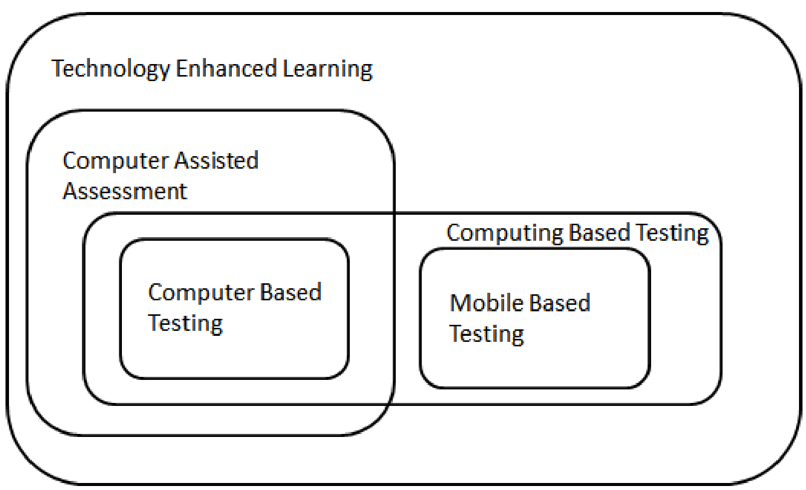


*Figure 1***.** Relations between the assessment research contexts in TEL

Mobile devices include a variety of technology, apps and features: wearable devices, tablets, smartphones, GPS, smart watches, google glasses, camera, pedometer, etc. The use of mobile technologies to enhance learning has been identified as beneficial for encouraging the development of meta-level thinking skills (Facer et al., 2004; Wegerif, 2002). These devices allow the user to interact with physical objects, physical locations, digital information, and with other users. We found really significant the statement made by Sharples et al., (2005): "*Context is constructed by learners through interaction: To explore the complexity of mobile learning it is necessary to understand the contexts* *in which it occurs. Context should be seen not as a shell that surrounds the learner at a given time and location, but as a dynamic entity, constructed by the interactions between learners and their environment.*"

Considering this reflection, we agree with Sharples that mobile devices can be adequately employed to allow users to interact with the physical/virtual context in a more realistic way. But mobile technologies are not enough, as we have discussed in the introduction; an adequate pedagogical plan is needed in order to guarantee the success of the assessment activity. We state that there are two important issues have to be considered when designing the pedagogical plan: First of all, the learning/assessment objectives have to be aligned with the performance of meaningful interactions. This means that the plan of the activity itself has to consider the kind of actions (interactions) done by the learner in order to solve the problem. Second, monitoring all the actions done by the learner can be difficult, but the higher the monitoring is the higher support can be provided to scaffold the learner during (and after) to understand which meaningful interactions has to perform in order to solve the activity.

Supporting meaningful interactions through scaffolding mechanisms

In this paper we want to place an emphasis on the importance of supporting "meaningful interactions", this means to understand the potential of mobile devices to help learners to interact adequately with the environment in order to support more realistic tasks. As Dourish (2004) states, computational systems can be made sensitive and responsive to the setting in which they are harnessed and used. Mobile and sensor technologies play an important role monitoring interactions in context-aware activities. We agree with Dourish's approach that the context can be seen as an interactional problem. Specifically, we analyse the importance of the context as interactional problem when designing an m-assessment activity.

As Woo & Reeves (2007) state: "*Interaction in learning is a necessary and fundamental process for knowledge acquisition and the development of both cognitive and physical skills*." But these authors also identify in their paper the notion that the use of authentic tasks is not a guarantee of the adequate performance of meaningful interactions. One of the factors that help to guarantee the success of the activity is the application of scaffolding strategies. Hillman et al. (1994) insisted that all interaction in TEL is mediated via a medium. For this reason, it is necessary to understand the potential of mobile technologies (as a medium) to improve the interaction possibilities for assessment purposes. Engeström (based on Vygotsky's idea of mediation) also discusses the importance of mediation as the use of artifacts, and/or concepts used by an actor, and how tools and the social context influence the actor-structure interactions (Engeström, 1999). Furthermore, as Cook (2010) states, mobile devices can be used as mediators in an ACD using them as the more capable peer can guide and scaffold the learner to find the adequate solutions. As Muirhead and Juwah (2004) identified, interaction is a dialogue between two or more participants and objects (i.e., physical or virtual) which occurs synchronously (or not) mediated by response or feedback and interfaced by technology. If we analyse the sentence, it can be applied to a process of answering questions in a test, where learners have to interact with information (e.g., physical, virtual or both) in order to provide an answer, and s/he receive feedback of this action by obtaining a response or a grade. Some researchers (e.g., Chen et al., 2003; Hwang and Chang 2011; Melero et al., in press; described with more detail below) have demonstrated how capturing information about location, dialogues among peers, time, position, interaction with certain objects, context, etc; can be used to analyse in real time the actions done by the learner and scaffold her steps during the activity.

We don't forget that in an authentic real-world situation it would be unthinkable to have scaffolding assistance, however in this case we are talking about formative assessment situations where it is important to support the acquisition of the necessary learner's skills.

Below we discuss which elements have to be considered to allow the learners to increase their *interaction* possibilities within a real learning *context*, and how by providing certain *scaffolding strategies* the user can be guided to engage with certain authentic tasks (by interacting with real objects, digital informal and other peers) and obtain specific results that can be monitored and assessed. In the following section we propose a model, called m-AssIST, that captures the necessary emergent properties to design an m-assessment context.

m-AssIST: Mobile - Interaction Scaffolding Temporal model

We claim that the use of mobile devices allows the learner to provide more authentic ways of answering questions by undertaking specific actions/interactions (not the traditional selection of a choice in a multiple choice question) which are associated with specific environments (i.e., locations and objects) and people (i.e., peers or people related with the context). These interactions with specific *physical objects* in an environment (i.e., measure the different angles of a room, run a distance in a given time, find and analyse a specific plant in a botanical garden, etc.) can be automatically monitored and assessed by the m-assessment system. The monitored actions can be used in favour of formative assessment by supporting the learner with scaffolding strategies during and after the assessment activity. However, in order to provide adequate scaffolding strategies for supporting the performance of meaningful interactions it is necessary to design adequately the *pedagogical plan* of the activity before its execution. We state that by using assistance and scaffolding mechanisms the dialogue between the learner and the m-assessment system can be mediated but some emergent properties have to be considered in order to guarantee the success of the activity.

In Figure 2 we present the mobile-Assessment Interaction Scaffolding Temporal Model (m-AssIST), this is an extension of the Context Model proposed by Tarasewich (2003). M-AssIst considers the different interaction possibilities during the *temporal axis* of an assessment activity (i.e., before, during and after the activity): (a) interactions with other *learners*, (b) interactions with the *environment* (i.e., indoors/outdoors spaces), (c) interactions with *physical objects* encountered in an environment, (d) interactions with the assessment *activity*, and finally (e) interactions with *augmented digital information* shown to support the activity.

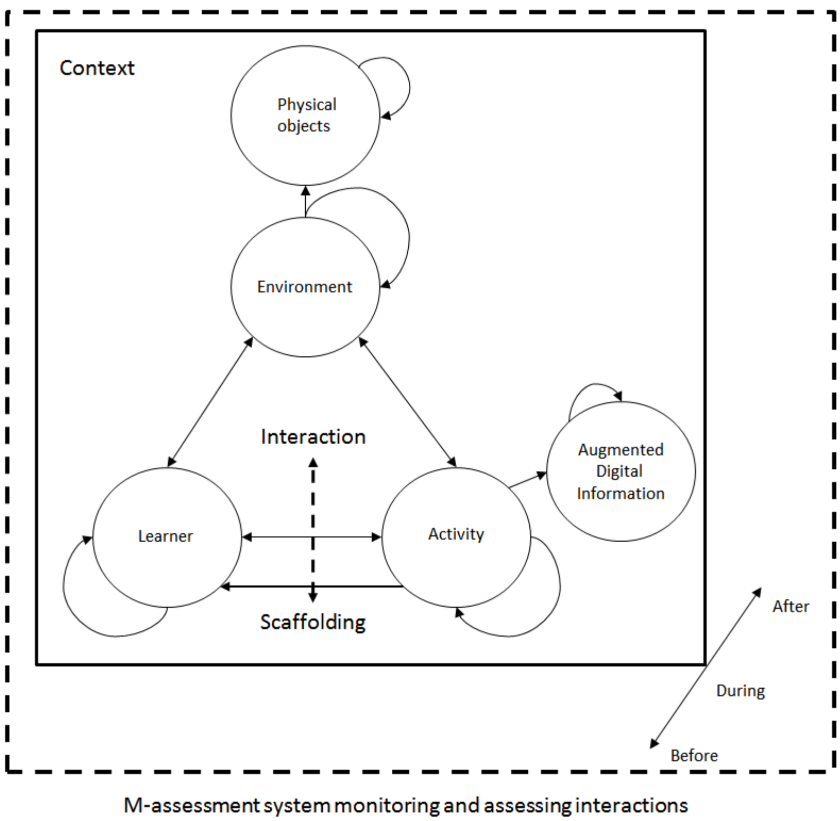


*Figure 2***.** m-Assessment Interaction Scaffolding Temporal Model (m-AssIST)

As we illustrate in the model, the performance of an assessment activity has a temporal axis, this means that *before*, *during* and *after* the activity the m-assessment system has to use the learning objectives of the activity and the outcomes from the monitoring process and scaffold the learner to ensure the performance of meaningful interactions. The m-AssIST model has the potential to be useful especially when practitioners and technicians are in the phase of designing an m-assessment activity (this means, *Before the Activity*). It is at this moment that the *pedagogical plan* of the activity has to be completed. From our understanding the pedagogical plan has to consider the learning/assessment objectives of the activity (i.e., HOTS to be assessed). Gulikers (2004) in his model reflects the need of having an alignment between the authentic learning tasks, the context (physical and social) and the learning goals (skills to be assessed). We recommend using Bloom's Taxonomy (Bloom & Krathwohl, 1956) to indicate the type of HOTS that practitioners can assess with respect to a specific learning objective. Level 1: Knowledge; Level 2: Comprehension; Level 3: Application; Level 4: Analysis; Level 5: Synthesis; and Level 6: Evaluation. The higher a level is the most sophisticated, and this is the task (set of interactions) that the user has to do in order to answer the question. As we have already discussed in the previous sections, the feature of mobile devices allows the learner to be more active and this can be used in benefit of the assessment process. However it depends on the practitioner to take profit of this benefit for assessing the learners' HOTS. In relation to these approaches, our model makes emphasis on the importance of supporting the learners' interaction with the context of the activity in order to provide the adequate assistance to scaffold her learning.

In this paper we claim that the pedagogical plan has also to include an analysis of the emergent properties (see Figure 2) that have an effect over the context of the designed activity. An important question to be answered is: *which meaningful interactions the learner has to perform within the environment selected (i.e., the space where the activity is performed, for instance a Botanical Garden), with specific physical objects (i.e., plants, tools, people, etc.), with other learners (if collaboration is necessary) and with the augmented digital information provided by the system in order to solve the problem adequately?*

By observing the m-AssIST model we see that the *Interaction* factor is in the middle of the figure. In this sense, for us the Interaction context is seen as the influential *physical and social context* discussed by Vygotsky (1978) or the relation between the *physical environment and the interpersonal interactions* proposed by Cook (2010). The adequate interaction of the learner with all the different factors considered in the context of the activity is essential to achieve the learning/assessment objectives of the activity. When the necessary interactions are identified by the practitioner these can be considered by the system in order to provide *scaffolding* assistance during the activity if some interaction is missed. For this reason a bold arrow goes from the *Activity* to the *Learner*, showing that a scaffolding assistance have to be provided to the learner if necessary. The missing interactions can be identified by the mobile system by monitoring the actions done by the student in real time or by considering the incorrect answers. In this case, the m-assessment system will be acting as the more capable peer.

m-AssIST is especially useful during the design phase of a m-assessment activity, but the model can also be used to analyse existing scenarios. In the following section we analyse existing m-assessment scenarios which accomplish the following characteristics: they are framed in a realistic context to assess higher order skills related to a subject matter, learners use mobile devices to interact with physical objects, digital information and peers, the students' interactions are monitored and assessed in real time, and finally some guidance or scaffolding mechanism is provided to support the student. Specifically, we apply our m-AssIST model with the goal of analysing the examples selected identifying the methods (or limitations) used to guide and scaffold the learner in order to support meaningful interactions with assessment purposes. The corresponding analysis can be used by other practitioners to replicate strategies of an existing scenario, or to improve the identified limitations.

Authentic tasks, meaningful interaction and scaffolding in m-assessment

Three illustrative real examples have been selected in order to validate the fitness-for-purpose of the m-AssIST model from the point of view of using the properties to reflect about the necessary meaningful interactions in the authentic context and the adequateness of the scaffolding mechanisms used to provide assistance before, during and/or after the activity. The first two examples have been selected by using the combination of tags 'Mobile, Assessment, Scaffolding, Authentic' in Google scholar, then we selected two papers by taking into consideration the significance of the papers in terms of their scholarly impact (higher number of citations), different authors and designers, sufficiently different authentic environments (a natural park and an oldest part of a city), subject matters (natural science and local culture) and different skills assessed (further detail below). The third example also accomplishes the main characteristics of the first examples, but has been selected because some of the authors of this paper were involved in the design, creation and analysis of the m-assessment system and scenarios described and, therefore, are especially familiar with its definition and details. We use the emergent properties captured in the m-AssIST model to analyse the benefits and limitations of the selected examples in terms of: authentic tasks, meaningful interaction and scaffolding mechanisms. We claim that if we show the value/utility and "usability" of the model for three significant and sufficiently different scenarios we can also discuss its transferability to other cases.

Analysis of existing examples

The first example is illustrated through the Bird-Watching Learning (BWL) (Chen et al., 2003). BWL was designed to assess students of natural science in preliminary school level when they are in an outdoors *environment* observing birds (in this case birds acts as *physical objects* and the natural park is the environment). PDAs and a Wi-Fi based wireless network are used as mediator technology during the activity. The assessment *activity* consists of the following tasks: while the students are observing the birds, a specific bird is selected by the instructor and students have to find the real specie selected by themselves. So, the learning objectives of the activity include assessing the following HOTS: *analysis* in situ of the information provided by the system, taking information apart (*inspect, examine*) from the exploration of the environment, and *compare* (differentiate, identify, resolve) this information in order to find the correct birds. As shown in Figure 3, HOTS (in form of verbs) are associated to the Activity property, this classification can be used to understand how to relate the activity with the other properties of the model. The BWL system has a testing subsystem that evaluates each learner's level of learning/skills and attitudes. The instructor delivers information from the birds to the students in real time, so all the students have the same chance to see the related information (introduction texts are used as *scaffolding* mechanism *before* stating the main activity). The instructor can select 10 birds' questions from the records, and transmit them to the students. By taking into consideration the responses, the system provides 'different levels of assistance' (manipulated by the instructor as *scaffolding during* the activity). The system provides 'static images and hints', and the students use this information to *analyse, and explore* the environment and discover the birds. The scaffolding level is determined according to the level of ability (application of HOTS) of each learner. The authors of the BWL system apply the following scaffolding strategy: the task is decomposed into (1) hierarchical tasks according to the skills. According to the step 1, the instructor (2) classifies the amount of support in decreasing levels and sets up a (2) repetitive authentic practice. The 3 steps are evaluated through an (4) ongoing assessment process (i.e., test questions are delivered to assess the accomplishment of tasks). This case is a very good example that shows the benefits of using scaffolding mechanisms to support meaningful interaction with the environment and physical objects (birds) whilst making possible a formative assessment activity. However, the application of the m-AssIST model allows us to detect some limitations in this scenario (see Figure 3).

Due to the technological limitations in terms of mobile devices in 2003, some mechanisms to monitor the activity in real time and support automatic assessment and scaffolding *during the activity* could be improved nowadays. For instance, instead of textual information, *augmented simulations* of the birds could be shown in the mobile device to help learners to achieve the HOTS, for instance identifying the correct bird. As used in Savannah (Facer, 2004) audio simulation of birds singing can be used to guide the students to find the adequate positions where they have to find the species. The *interaction among learners* is only referred to pairs of learners; however data from other peers could be used to improve the scaffolding assistance. For instance, the information from successful peers (pairs of leaners who answered correctly the questions) could be used to scaffold other unsuccessful peers by sharing the location of those learners where they have found the correct birds. Finally, it would be useful to save the interactions undertaken by the students during the activity in a log and provide statistics *after* the activity to help the learners to reflect in their performance.

In the second example selected, Hwang and Chang (2011) propose a Formative Assessment-based Learning guiding Mechanism (FAML) to learn in situ about local culture (the Chin-An temple in southern Taiwan). In this scenario, fifth grade students used PDAs in a wireless network *outdoors and indoors environment*, they had to interact with the temple environment in order to answer related questions. In this case, the required HOTS were: students have to apply the theory learnt at classroom, *classify* the provided information and *interpret* this in a new situation (i.e., the real location) in order to solve the different questions (see Figure 4). In this case the system plays the role of the "more capable peer" guiding the students *during the activity*. In terms of a scaffolding strategy, the system provides "hints" and other 'extra multimedia content' to the students instead of giving an assessment of the validity of answers. One of the main benefits of the FAML system is that it was designed to encourage the student to interact with the environment instead of being focused on providing correct answers. When a question is answered incorrectly, instead of providing the correct solution the system encourages the student to make further interactions by providing more hints.

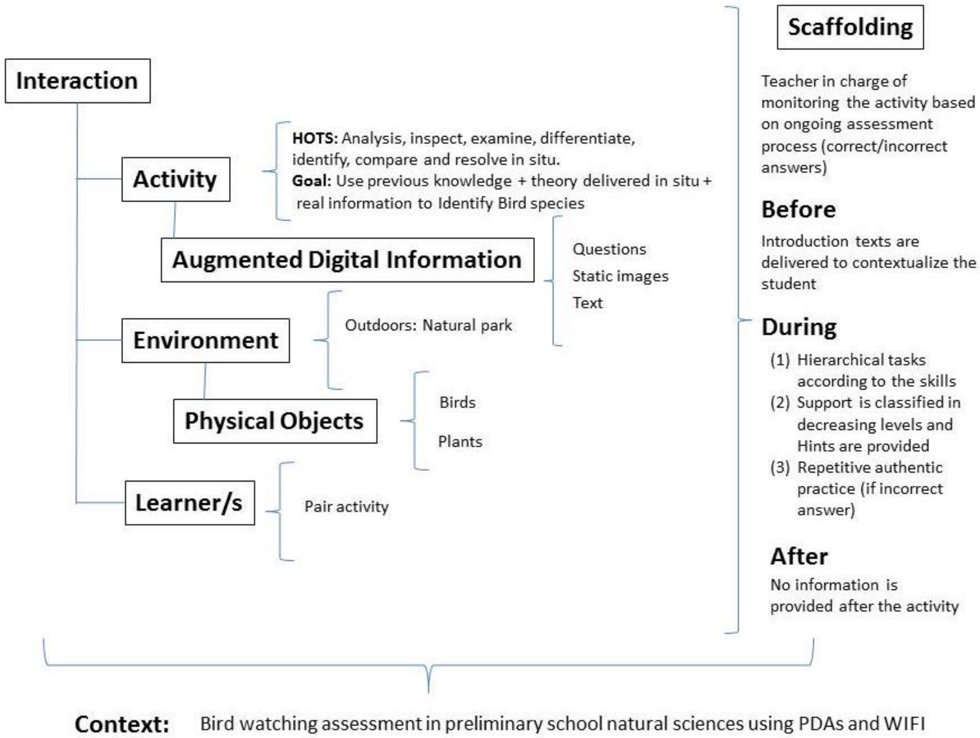
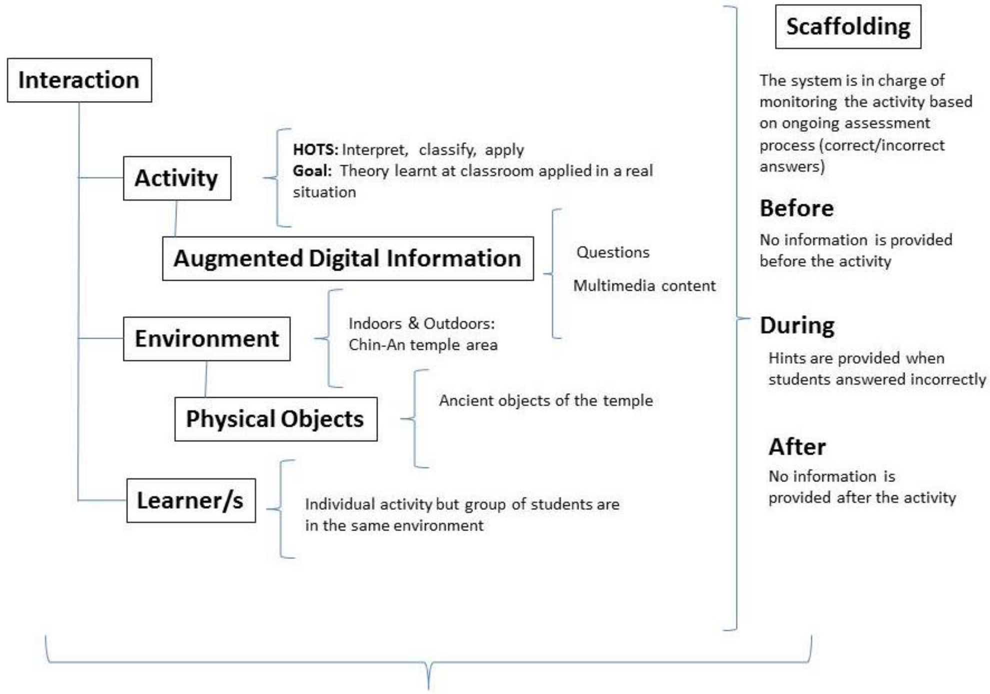


Figure 3. m-AssIST model applied to the BWL scenario

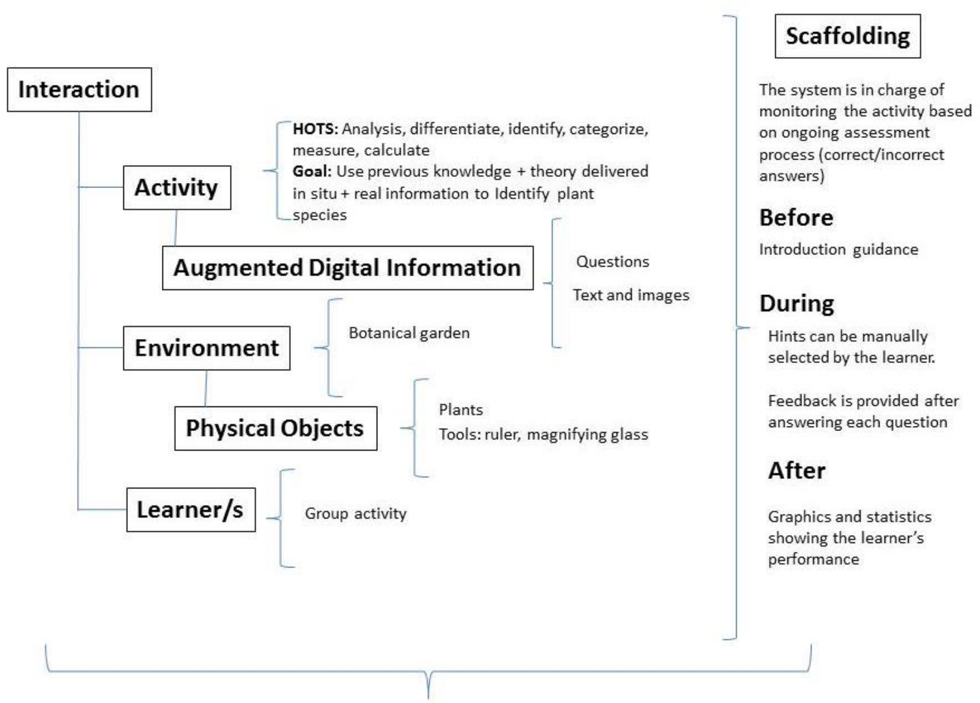
Using the emergent properties of the m-AssIST model (see Figure 4), we would like to highlight as limitation the fact of using the answers provided by the learners (correct or incorrect ones) as the unique mechanism to provide hints and guidance to the students. Nowadays, mobile devices allow the tracking of actions/interactions done by the user in spaces and with objects. As Zhou et al., (2008) show, tracking techniques can be used to detect changes in the viewer's position and reflect these changes in rendered augmented digital information. We envisage that calibration and registration tools could be used to align the real world (i.e., position of a statue in the temple) with the virtual one (i.e., activity and augmented digital information) when the user view is fixed. Other limitations of this example are observed regarding the scaffolding strategies, in particular before and after the activity where any information is provided to help the learner to understand her goal at the beginning of the activity, or to reflect about her performance at the end.

As the final example, we use the QuesTInSitu m-assessment system developed by (Santos et al., 2011, Melero et al., in press). We have selected a particular scenario to analyse the features of the tool to design m-assessment activities. QuesTInSitu allows learners to answer geolocated questions in situ (in an *outdoors environment* using GPS tracker, and in *indoors environments* using manual selection of questions over digital maps). When learners are in the correct position, a question appears and they have to *apply their knowledge* by interacting with the environment and physical objects in order to solve the question. For instance, one example of activity was done in a botanical garden where learners were asked to find specific plants in the garden (by using only the descriptions of the plants). Students put in practice the following HOTS by using tools to *measure, calculate and analyse* the plant (i.e., the average length of its leaves), by previously *differentiate* and *identify* the plant from others in the garden (see Figure 5). As in previous examples, the level of interaction complexity (and the application of HOTS) depends on the task asked by the question. In the first version of the tool (Santos et al., 2011) automatic feedback delivery was used after answering a question as a scaffolding mechanism *during the activity*. However, after performing several activities in real contexts with students there emerged the need for including scaffolding mechanisms *before and after the activity* too. For this reason, a new version of the system (Melero et al., *in press*) was developed. This version provides to the learner scaffolding mechanisms before, during and after the activity. The learner can select an 'introduction guidance' section *before starting the activity* (to understand better the context and the goals of the activity), *during the activity* hints can be manually selected by the learner if his/her has difficulties finding the correct way for solving a question. In parallel, during the activity the system monitors and collects information related to the learner's position, time and answering attempts. This information is shown to the learner as graphics *after the activity*. The goal of this graphics is to help the student to understand and reflect about their performance during the activity, for instance: Where she had more interaction difficulties to solve a question, or where the student interacts adequately.



Context: Formative assessment to learn local culture using PDAs and WIFI

Figure 4. m-AssIST model applied to the FAML scenario



Context: Assessment in situ using smartphones and QuesTInSitu

Figure 5. m-AssIST model applied to QuesTInSitu

By applying the m-AssISt model, we have realized that the QuesTInSitu system could be improved by including more facilities to support the *interactions among learners*. All the activities performed with this tool were carried out by groups of students (2 or 3 using the same smartphone) and other groups doing the same activity in parallel, for this reason it is important to consider the interactions among the members of the same group and with other groups. In addition, the interaction with indoors environments could be enriched by integrating indoors positioning sensors to the system. As Sichitiu, & Ramadurai (2004) show in their study, wireless sensors networks have the potential to become the pervasive sensing technology of the near-future. In particular with assessment purposes, sensors will allow us to collect specific interaction information in a space but also manipulating an object (i.e., plants inside a greenhouse of the botanical garden), this information will be especially useful to support specific HOTS.

Conclusions

This paper has discussed the benefits of using mobile devices to support authentic assessment. The paper highlights the importance of using scaffolding mechanisms to assist learners when they interact with the environment, physical/digital objects, other learners and the activity itself. The selected scaffolding mechanisms have to be applied to facilitate the performance of meaningful interactions which is the based to support authentic tasks and HOTS (Woo & Reeves, 2007).

In this context, the m-AssIST model is presented as a model that captures the emergent elements of an authentic m-assessment activity. The model highlights the importance of considering the different interactions possibilities that learners have during a test-based activity. Monitoring those interactions through the mobile device allows the provision of specific scaffolding mechanisms before, during and after the activity. We have shown how the inclusion of scaffolding strategies transforms the m-assessment system in a 'more capable peer' assisting the learner during the activity to enhance the performance of meaningful interactions and as a consequence the correct achievement of HOTS. In this sense, the identification of the specific HOTS related to the Activity have to be considered in order to understand which specific features of mobile devices are going to be used. For instance, in order to support the identification of a bird, a layer of augmented reality in the natural park can be used to guide the student; or the mobile phone can be used in a group of students to *explore* different positions of the botanical garden and *debate* in real time the correct location of a plant.

We have used the m-AssIST model to analyse the characteristics and limitations of three existing m-assessment activities designed to support meaningful interactions in real environment and with physical/virtual objects. We would like to clarify, on the one hand, that other m-based test cases (different subject matters or learning levels) could be also analysed through the m-AssIST model. The model can be used to analyse existing scenarios (and systems), but also to design new ones and help practitioners and technicians to reflect which are the scaffolding mechanisms that can be used to support the particular interactions of an m-assessment context. On the other hand, although the examples selected were focused on the use of PDAs and mobile phones, other mobile devices can be used to assess the learners' authentic tasks in real settings. For instance, Clark et al., (2012) provide a good example that shows how Microsoft Kinect can be used for assessment of postural control, recording the actions done by the users in a space giving the opportunity of using the entire body (and not only the hands) to make interactions that are computationally assessed. This last works links with our future research interest which deals on analysing the usefulness of mobile technology to support affective computing and intelligent interaction with assessment purposes.

Acknowledgements

This work has been partially funded by the Learning Layers project. Learning Layers is a 7th Framework Large-scale integrating project co-funded by the European Commission and the EEE (TIN2011-28308-C03-03) project.

References

Bennett, R. E. (1990). Towards intelligent assessment: An integration of constructed-response testing, artificial intelligence and model-based measurement. *ETS Research Report Series, 1990*(1), i-41.

Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment. *Optimising new modes of assessment: In search of qualities and standards* (pp. 13-36). doi: 10.1007/0-306-48125-1_2

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, *31*(4), 399-413.

Boud, D. (2013). *Enhancing learning through self-assessment*. New York, NY: Routledge.

Boyle, A., & Hutchison, D. (2009). Sophisticated tasks in e-assessment: what are they and what are their benefits? *Assessment & Evaluation in Higher Education*, *34*(3), 305-319.

Bloom, B., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners*. Handbook 1: Cognitive domain. New York, NY: Longman.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, *18*(1), 32-42.

Brown, E. (Ed.) (2010). Education in the wild: contextual and location-based mobile learning in action. A report from the STELLAR Alpine Rendez-Vous workshop series. Nottingham, UK: University of Nottingham: Learning Sciences Research Institute (LSRI).

Bull, J., & McKenna, C. (2003). *A blueprint for computer-assisted assessment*. London, UK: Routledge.

Chen, Y. S., Kao, T. C., & Sheu, J. P. (2003). A mobile learning system for scaffolding bird watching learning. *Journal of Computer Assisted Learning*, *19*(3), 347-359.

Choi, J. I., & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. *Educational Technology Research and Development*, *43*(2), 53-69.

Clark, R. A., Pua, Y. H., Fortin, K., Ritchie, C., Webster, K. E., Denehy, L., & Bryant, A. L. (2012). Validity of the microsoft kinect for assessment of postural control. *Gait & posture*, *36*(3), 372-377.

Cook, J. (2010). Mobile phones as mediating tools within augmented contexts for development. *International journal of mobile and blended learning*, *2*(3), 1-12.

Conole, G., & Warburton, B. (2005). A review of computer-assisted assessment. *Research in learning technology*, *13*(1).

Crespo, R. M., Najjar, J., Derntl, M., Leony, D., Neumann, S., Oberhuemer, P., ... & Kloos, C. D. (2010, April). Aligning assessment with learning outcomes in outcome-based education. In *Education Engineering (EDUCON)* (pp. 14-16). Madrid, Spain: IEEE.

Davies, S. (2010). *Effective Assessment in a Digital Age*. UK. Retrieved October 13, 2014 from http://www.jisc.ac.uk/digiassess

de-Marcos, L., Hilera, J. R., Barchino, R., Jiménez, L., Martínez, J. J., Gutiérrez, J. A., ... & Otón, S. (2010). An experiment for improving students performance in secondary and tertiary education by means of m-learning auto-assessment. *Computers & Education*, *55*(3), 1069-1079.

Dourish, P. (2004). What we talk about when we talk about context. *Personal and Ubiquitous Computing*, *8*(1), 19-30.

Engeström, Y., Miettinen, R., & Punamäki, R. L. (Eds.). (1999). *Perspectives on activity theory*. Cambridge, UK: Cambridge University Press.

Eraut, M. (2007). Learning from other people in the workplace. *Oxford Review of Education*, *33*(4), 403-422.

Facer, K., Joiner, R., Stanton, D., Reid, J., Hull, R., & Kirk, D. (2004). Savannah: Mobile gaming and learning? *Journal of Computer Assisted Learning*, *20*(6), 399-409.

Gulikers, J. T., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, *52*(3), 67-86.

Hillman, D. C., Willis, D. J., & Gunawardena, C. N. (1994). Learner interface interaction in distance education. An extension of contemporary models and strategies for practitioners. *The American Journal of Distance Education*, *8*(2), 30−42.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, *7*(2), 174–196.

Hwang, G. J., & Chang, H. F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education*, *56*(4), 1023-1031.

IMS. (2006). *IMS question and test interoperability assessment test, section, and item information model*. Retrieved October 13, 2014 from http://www.imsglobal.org/question/qtiv2p1pd2/imsqti_infov2p1pd2.html

Kaminsky, S., & Podell, D. M. (1997). *Educational psychology: Windows on teaching*. New York, NY: McGraw-Hill College.

Lieberman, D. A., & Linn, M. C. (1991). Learning to learn revisited: Computers and the development of self-directed learning skills. *Journal of research on computing in education*, *23*(3), 373-95.

Louys, A., Hernández-Leo, D., Schoonenboom, J., Lemmers, R.,Pérez-Sanagustín, M. (2009). Self-development of competences for social inclusion using the TENCompetence infrastructure. *Educational Technology & Society*, *12*(3), 70-81.

Mayotte, S. (2010). Online assessment of problem solving skills. *Computers in Human Behavior, 26*(6), 1253-1258.

Melero, J., Sun, J., Hernández-Leo, D., Santos, P., & Blat, J. (2015). How was the activity? A visualization support for a case of situated m-learning design. *British Journal of Educational Technology, 46*(2), 317-329.

Muirhead, B., & Juwah, C. (2004). Interactivity in computer-mediated college and university education: A recent review of the literature. *Educational Technology & Society*, *7*(1), 12−20.

Santos, P., Pérez-Sanagustín, M., Hernández-Leo, D., & Blat, J. (2011). QuesTInSitu: From tests to routes for assessment in situ activities. *Computers & Education*, *57*(4), 2517-2534.

Santos, P., Hernández-Leo, D., Pérez-Sanagustín, M., & Blat, J. (2012). Modeling the computing based testing domain extending IMS QTI: Framework, models and exemplary implementations. *Computers in Human Behavior*, *28*(5), 1648-1662.

Santos, P., Hernández-Leo, D., & Blat, J. (2014). To be or not to be in situ outdoors, and other implications for design and implementation, in geolocated mobile learning. *Pervasive and Mobile Computing*, *14*, 17-30

Sharples, M., Taylor, J., & Vavoula, G. (2005). Towards a theory of mobile learning. *Proceedings of mLearn 2005*, *1*(1), 1-9.

Sichitiu, M. L., & Ramadurai, V. (2004). Localization of wireless sensor networks with a mobile beacon. *Proceedings of Mobile Ad-hoc and Sensor Systems* (pp. 174-183). doi: 10.1109/MAHSS.2004.1392104

Susono, H., & Shimomura, T. (2006). Using mobile phones and QR codes for formative class assessment. *Current developments in technology-assisted education*, *2*, 1006-1010.

Tarasewich, P. (2003). Towards a comprehensive model of context for mobile and wireless computing. *Proceedings* of the *AMCIS Conference* (pp.114-124). Tampa, FL: AISeL.

Thompson, B. W. (2005). The transforming effect of handheld computers on nursing practice. *Nursing administration quarterly*, *29*(4), 308-314.

Vygotsky, L. (1978). *Mind in society*: *The development of higher psychological processes* (Cole, M., Eds.). Cambridge, MA: Harvard University Press.

Wegerif, R. (2002). *Literature review in thinking skills, technology and learning.* Retrieved from https://telearn.archives-ouvertes.fr/hal-00190219/

Woo, Y., & Reeves, T. C. (2007). Meaningful interaction in web-based learning: A social constructivist interpretation. *The Internet and Higher Education*, *10*(1), 15-25.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, *17*(2), 89-100.

Zhou, F., Duh, H. B. L., & Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality* (pp. 193-202). Cambridge, UK: IEEE Computer Society. doi: 10.1109/ISMAR.2008.4637362