

O ABORDARE METODOLOGICĂ INTEGRATĂ ÎN EVALUAREA UNEI PLATFORME DE REALITATE ÎMBOGĂȚITĂ DESTINATĂ ÎNVĂȚĂRII

Dragoș Daniel Iordache

Institutul Național de Cercetare - Dezvoltare în Informatică - ICI București

iordache@ici.ro

Rezumat: În cadrul acestei lucrări se prezintă metodologia integrată de evaluare a unei platforme de predare-învățare bazată pe tehnologia de realitate îmbogățită (augmented reality) dezvoltată în cadrul proiectului european ARiSE (Augmented reality in school environment). În acest sens, prima parte a lucrării este rezervată unei scurte introduceri în problematica tehnologiei de realitate îmbogățită. Următoarele secțiuni abordează evaluarea aplicațiilor educaționale și prezentarea în detaliu a metodologiei integrate de evaluare a platformei de învățare. Lucrarea se încheie cu concluzii și direcții de lucru viitoare.

Cuvinte cheie: realitate îmbogățită, metodologie integrată evaluare.

Abstract: This paper presents an integrated methodology for assessing a teaching-learning platform based on technology augmented reality developed in the European project ARiSE (Augmented reality in the school environment). The first part of the paper presents a brief introduction to the issue of augmented reality technology. The following sections address the evaluation of educational applications and detailed presentation of the integrated methodology for assessing the learning platform. The paper ends with conclusions and directions for future work.

Key words: augmented reality, mixed method evaluation .

1. Introducere

Între aplicațiile educaționale de mare actualitate, se numără și cele bazate pe tehnologia de realitate îmbogățită (AR). Realitatea îmbogățită reprezintă o variație a mediilor virtuale sau a realității virtuale [1]. În timp ce realitatea virtuală presupune imersia totală a utilizatorului în mediul artificial, fără ca acesta să aibă posibilitatea de a vedea lumea reală din jurul lui, realitatea îmbogățită permite utilizatorului să vadă lumea reală, cu obiecte reale pe care se suprapun obiecte virtuale. Astfel, realitatea îmbogățită mai degrabă suplimentează realitatea (o îmbogățește) decât să o înlocuiască complet. La nivel mental, obiectele reale și cele virtuale îi vor apărea utilizatorului ca fiind coexistente în același spațiu. Noutatea oferită de realitatea îmbogățită se referă astfel la posibilitatea de a utiliza obiecte obișnuite pentru a interacționa cu tehnologia, mai familiară decât dispozitivele de intrare ale calculatorului. În plus, pot fi implementate diferite modalități de vizualizare a obiectelor. Obiectul real poate fi rotit și manipulat, fiind însoțit de proiecția virtuală (augmentare) însoțindu-l și antrenând astfel mai multe tipuri de analizatori (vizual, auditiv, tactil) în actul de învățare.

În cadrul proiectului ARiSE a fost dezvoltată platforma ARTP (Augmented Reality Teaching Platform), pe care au fost implementate trei aplicații de învățare: una pentru învățarea biologiei, una pentru învățarea chimiei și una care vizează cooperarea la distanță (remote collaboration) și crearea de conținut interdisciplinar.

ARTP este o platformă de tip desktop [16]: utilizatorii au în față un ecran „see-through” la nivelul căruia sunt suprapuse imagini ale obiectelor virtuale peste imaginea observată a unui obiect real. Platforma a fost realizată de către Fraunhofer IAIS și a fost înregistrată sub marca Spinnstube®. Unul dintre obiectivele dezvoltării ARTP, a fost dezvoltarea unei tehnologii AR pentru mediul școlar care să răspundă exigențelor învățării de tip constructivist și proiectării centrate pe utilizator și sarcină [8]. Dezvoltarea ARTP oferă de asemenea suport și pentru modelele predării bazate pe simulare, care au ca principii generale libertatea de acțiune, interacțiunea prin implicare și angajarea în sarcină.

Scopul acestei lucrări este de a prezenta principalele componente ale unei metodologii integrate care a fost utilizată pentru evaluarea ARTP. Metodologia include o varietate de metode, tehnici și instrumente. Printre avantajele acestei abordări se pot enumera: posibilitatea de a analiza o

diversitate de aspecte; cercetătorul poate elabora și testa o teorie întemeiată pe metode mixte [10]; posibilitatea de a oferi răspunsuri la o arie largă de întrebări de cercetare întrucât cercetătorul nu este limitat la o singură metodă sau abordare de cercetare; posibilitatea de a oferi explicații cuprinzătoare și complete la una și aceeași întrebare de cercetare; cercetătorul se poate folosi de avantajele oferite de o metodă pentru a depăși limitele altei metode prin utilizarea ambelor în cadrul aceluiași studiu; posibilitatea de a desprinde concluzii semnificative datorită convergenței și triangulării rezultatelor.

2. Evaluarea aplicațiilor educaționale

Eforturile de integrare a aplicațiilor de realitate îmbogățită în câmpul educației sunt de dată relativ recentă. Prin urmare, preocupările specialiștilor de a elabora metodologii de evaluare a acestor aplicații nu au un trecut îndelungat. În prezent există o diversitate de ghiduri metodologice specifice evaluării la nivelul diferitelor discipline școlare. Acestea sunt focalizate în special pe caracteristicile metodologice ale testelor docimologice și mai puțin pe evaluarea unor programe și aplicații inovatoare destinate învățării respectivelor discipline școlare.

Nu există o metodă unică, validă pentru evaluarea pedagogică a sistemelor de realitate îmbogățită destinate mediului școlar. Acest câmp de cercetare se află încă într-un stadiu incipient de dezvoltare. Cadrul adecvat pentru evaluarea pedagogică este în curs de constituire. Cu toate acestea, există abordări și modele destinate evaluării sistemelor de e-learning și domeniului sistemelor de învățare în mediu virtual (virtual learning environment - VLE). Oliver [13] a identificat patru tipuri de evaluare:

- evaluarea de tip analiză de cerințe (requirements analysis) - presupune o evaluare a situației curente, a punctelor tari și a punctelor slabe și a măsurii în care resursele luate în calcul sunt de ajutor. La acest nivel sunt relevante informații referitoare la adecvarea echipamentelor, la atitudinile și abilitățile utilizatorilor potențiali. Analiza necesităților mai puțin frecventă în cadrul procesului de dezvoltare, acest fapt având deseori consecințe nefavorabile.
- evaluarea formativă - implică colectarea de informații cu privire la acele resurse care sprijină procesul de dezvoltare și asigură funcționalitatea unui prototip. Una dintre problemele analizei cerințelor constă în dificultatea utilizatorilor potențiali de a imagina beneficiile unui produs care nu există încă. În cazul evaluării formative există un produs real (sau o componentă a acestuia) care poate fi perceput și analizat.
- evaluarea sumativă - intervine la sfârșitul ciclului de dezvoltare având drept scop verificarea gradului de atingere a obiectivelor propuse. În cazul resurselor de predare și de învățare, acest lucru poate implica demonstrarea existenței unei diferențe reale între alternativele supuse evaluării.
- evaluarea integrativă - ia în calcul dificultatea demonstrării că o schimbare la nivelul practicii sau al cunoștințelor ar putea să fie atribuită unei singure variabile. Prin urmare, abordarea integrativă încearcă să ia în considerare o multitudine de factori care pot influența variabila supusă cercetării. În acest sens este importantă corelarea rezultatelor obținute prin intermediul unei diversități de metode de cercetare.

Modelul procesului de evaluare a calității produselor software [3] presupune, într-o primă fază, stabilirea cerințelor de evaluare și apoi specificarea, proiectarea și execuția evaluării. Printre cele mai importante caracteristici ale procesului de evaluare se pot menționa: repetabilitatea, reproductibilitatea, imparțialitatea și obiectivitatea.

Există o distincție metodologică între evaluarea de sistem și evaluarea de proces [4]. Evaluarea de sistem urmărește gradul de realizare a finalităților macro-structurale în anumite limite de spațiu și timp, care permit: măsurarea și aprecierea unor aspecte centrate asupra raporturilor dintre învățământ și viața socio-economică și culturală; corelarea calității învățământului cu contribuția sa la dezvoltarea socială. Evaluarea de proces urmărește gradul de realizare a obiectivelor micro-structurale, acțiune complexă determinată de: finalitățile macro-structurale; corelațiile profesor-

elev, rezultate școlare – metodologie folosită; corespondențele pedagogice dintre elementele activității didactice (obiective – conținuturi - metodologie); operațiile de măsurare și apreciere și instrumentele oficiale instituționalizate pentru consemnarea rezultatelor.

Totodată, sunt relativ puține abordări sistematice în ceea ce privește evaluarea sistemelor informatice pentru educație, deși numărul acestor sisteme este în continuă creștere. Pe plan internațional s-au propus numeroase abordări în evaluare, dar există puține metodologii dedicate sistemelor de e-learning. În acest context se impune necesitatea elaborării unei metodologii mixte de evaluare a aplicațiilor educaționale bazate pe tehnologia de AR care să propună o abordare sistematică privind evaluarea acestor sisteme prin integrarea de metode de cercetare adecvate.

3. Metodologia integrată de evaluare a ARTP

3.1 Integrarea dintre cantitativ și calitativ

Evaluarea tehnologiei de realitate îmbogățită a presupus adoptarea unei metodologii integrate de cercetare. În mod tradițional, demersurile de evaluare în științele sociale sunt legate fie de paradigma cantitativă, fie de cea calitativă. Din această perspectivă, metodele mixte de cercetare reprezintă o nouă paradigmă, care a dobândit o atenție considerabilă în diferite domenii cum sunt: științele sociale, informatica și managementul [10].

Cele mai importante metode calitative utilizate pentru evaluarea ARTP au fost: protocolul „gândire cu voce tare”, metoda învățării în perechi și metoda focus grup. Protocolul „gândire cu voce tare” (TAP) este o metodă folosită frecvent în ultima perioadă în cercetările de tip educațional datorită bogăției de date care pot fi derivate din metodologie. În cadrul evaluării ARTP, protocolul gândirii cu voce tare a fost folosit în evaluare cu un număr restrâns de subiecți în vederea realizării unei triangulări a rezultatelor prin compararea acestora cu cele obținute la nivelul celorlalte metode de evaluare utilizate.

Metoda învățării în perechi reprezintă o abordare specifică evaluării aplicațiilor informatice destinate copiilor. Folosirea învățării în perechi a ajutat la evaluarea aplicației, dar și la obținerea unor informații importante despre modul în care copiii folosesc sistemul sau comunică despre acesta (ce fel de indicații și termeni folosesc și unde anume diferă limbajul lor de cel folosit de proiectanții aplicației sau de adulți în general).

Sesiunile de evaluare ARTP folosind „protocolul gândire cu voce tare” și metoda învățării în perechi s-au realizat în perioada 4-16 martie 2010 în prezența unui evaluator cu rol de facilitator. În prima etapă, procedura de lucru pentru TAP a constat, pentru fiecare utilizator, în patru etape: introducere, exercițiul de „încălzire”, sarcinile de „gândire cu voce tare” și o discuție concluzivă. Sesiunile de testare au fost înregistrate audio și video. Apoi, fiecare elev a testat ARTP în prima fază, a jucat rolul de instructor pentru unul dintre elevii care nu au lucrat cu platforma. Ei au avut sarcina să îi instruiască pe elevii care nu au participat la sesiunile TAP cu privire la modul de utilizare a ARTP și cu privire la realizarea sarcinilor de lucru. Au rezultat astfel mai multe perechi de elevi care au interacționat activ în vederea realizării sarcinilor prezentate. După parcurgerea sesiunilor de testare, toți elevii au completat un scurt chestionar în care au menționat principalele puncte tari și limite ale tehnologiei utilizate [7].

Focus grupul a fost utilizat cu scopul investigării opiniilor cadrelor didactice din învățământul gimnazial și superior cu privire la oportunitatea introducerii ARTP în școlile din România. Această metodă a permis și culegerea unor recomandări valoroase cu privire la dezvoltarea de noi sisteme educaționale bazate pe tehnologia de realitate îmbogățită. Focus grupul [9] a fost realizat în sediul ICI București, după testarea de către cadrele didactice a ARTP în perioada 11 – 18 mai 2009. La evaluare au participat în total 16 cadre didactice din învățământul gimnazial (6 de la Școala Nr.172, 4 de la Școala Nr. 56 și 6 de la Școala Nr. 193 din București). În prima fază, fiecare profesor a testat în detaliu aplicația corespunzătoare materiei pe care o predă. Apoi, au fost organizate 2 grupuri: un grup format din profesori de biologie și altul format din 8 profesori de chimie.

Metodele cantitative utilizate pentru evaluarea ARTP s-au concentrat pe testările cu utilizatori

din cadrul școlilor de vară care au avut la bază chestionarul și metode și tehnici statistice de prelucrare și interpretare a datelor.

În cadrul cercetării au fost elaborate următoarele chestionare:

- chestionare de evaluare a cunoștințelor elevilor;
- chestionarul de evaluare a utilizabilității.

Înainte și după fiecare sesiune de testare elevii au completat două teste de cunoștințe: unul pentru biologie și altul pentru chimie. Pre-testele și post-testele au fost identice sub raportul conținutului, diferența constând în variația momentului în care au fost aplicate.

Chestionarul de utilizabilitate are la bază un model conceptual inspirat din teoria acceptării tehnologiei (TAM). Modelul acceptării ARTP implică cinci constructe: ergonomia platformei, ușurința în utilizarea percepută a aplicației, utilitatea percepută, satisfacția (enjoyment) percepută și intenția de utilizare. Prin utilizarea de constructe precum plăcerea (enjoyment) percepută și utilitatea percepută, rezultatele evaluării utilizabilității pot fi integrate cu rezultatele evaluării pedagogice.

Chestionarul a fost mai întâi administrat elevilor pe parcursul școlilor de vară [2, 14] cu scopul îmbunătățirii utilizabilității prototipurilor dezvoltate în cadrul proiectului ARiSE. După aceste sesiuni a fost administrat din nou elevilor de la școlile din România care au participat la testarea cu utilizatori. Rezultatele au fost utilizate pentru a valida modelul conceptual [6]. Rezultatul final a fost scala validată cu 28 de itemi care poate fi utilizată pentru evaluarea sumativă a viitoarelor scenarii de învățare care pot fi dezvoltate pentru ARTP.

3.2 Integrarea dintre formativ și sumativ

Metodologia integrată de evaluare a ARTP a presupus atât evaluarea formativă, cât și evaluarea sumativă a platformei de învățare. În funcție de momentul și scopul evaluării, evaluarea sistemelor interactive poate fi formativă sau sumativă. Evaluarea formativă este efectuată iterativ, pe parcursul ciclului de dezvoltare a unei aplicații informatice, cu scopul identificării și remedierii problemelor de utilizabilitate cât mai devreme [15].

Sistemele interactive sunt proiectate în mod iterativ. În acest sens, metodele de evaluare formativă sunt utilizate în cadrul unei bucle de proiectare - evaluare. De regulă, în evaluarea formativă a utilizabilității se colectează măsuri cantitative (numărul de probleme de utilizabilitate pe niveluri de severitate: major, moderat și minor) și măsuri calitative (descrierea detaliată a fiecărei probleme și a contextului în care apare).

Evaluarea formativă a ARTP a fost realizată prin intermediul sesiunilor de evaluare euristică, realizate pe parcursul ciclului de dezvoltare a ARTP. Tot în această categorie intră și testarea cu utilizatori a fiecărui scenariu de învățare pe parcursul școlilor de vară din cadrul proiectului ARiSE.

Evaluarea sumativă a presupus parcurgerea unor sesiuni de lucru la care au participat atât elevii cât și cadre didactice de specialitate. Elevii au parcurs mai multe scenarii didactice sub forma unor programe demonstrative, lecții și exerciții prin intermediul ARTP. Atât pe parcursul desfășurării sesiunilor de testare, cât și la final au fost utilizate instrumentele de colectare a datelor.

Un aspect important al oricărei abordări în evaluarea formativă este utilitatea pe care o au rezultatele evaluării asupra procesului de dezvoltare. În esență, întrebarea este în ce măsură rezultatele sunt valorificate ulterior în procesul de proiectare și care este gradul de recuperare al costurilor activității de evaluare. Raportul de evaluare trebuie să fie atât util cât și utilizabil [11]. Din acest motiv este necesară includerea, în descrierea problemelor de utilizabilitate, a recomandărilor de remediere.

3.3 Integrarea dintre evaluarea euristică și testarea cu utilizatori

Cele mai uzuale metode de evaluare formativă a sistemelor interactive sunt: evaluarea euristică, inspecția cognitivă, evaluarea bazată pe recomandări și testarea cu utilizatori. În cadrul

metodologiei propuse s-a realizat o triangulare a rezultatelor evaluării euristice cu rezultatele testării cu utilizatori.

Evaluarea euristică este o metodă efectuată de un număr redus de evaluatori, care examinează o interfață cu utilizatorul, judecă respectarea unui set de principii de utilizabilitate (euristici) și elaborează o listă de probleme de utilizabilitate clasificate pe categorii de severitate, corespunzător impactului estimat asupra performanțelor utilizatorului sau acceptării [12].

Testarea cu utilizatori este definită ca o metodă empirică de evaluare implicând participanți având caracteristici apropiate de cele ale utilizatorilor reali ai produsului care va fi evaluat. Prin testarea cu utilizatori se înregistrează comportamentul utilizatorului cu ajutorul unor tehnici specifice cum sunt observarea, chestionarul sau protocoalele verbale. O tehnică frecvent utilizată în evaluarea formativă a utilizabilității este protocolul de «gândire cu voce tare» (TAP – Think Aloud Protocol).

Evaluarea euristică produce un set de probleme de utilizabilitate care sunt considerate anticipate, în timp ce testarea cu utilizatori identifică probleme de utilizabilitate reale. Diferența între cele două seturi o reprezintă problemele fals pozitive (alarme false, identificate de evaluarea euristică, dar neconfirmate de testarea cu utilizatori) și fals negative (problemele de utilizabilitate reale care nu au fost anticipate de evaluarea euristică).

Pentru măsurarea validității, eficacității și eficienței evaluării euristice, au fost propuși mai mulți indicatori [5]:

- validitate: probleme anticipate-confirmate / probleme anticipate.
- completitudine (thoroughness): probleme anticipate-confirmate / probleme reale
- eficacitate generală (overall effectiveness): validitate * completitudine.

Calcularea acestor indicatori presupune evaluarea cu ambele metode, pe baza aceluiași scenariu de utilizare, fapt care mărește costurile evaluării.

Pe de altă parte, în cazul testării cu utilizatori pot fi colectate și alte tipuri de măsuri: măsuri ale eficacității (rata de îndeplinire a sarcinii, numărul de erori), ale eficienței (timpul de execuție) și satisfacției (prin aplicarea de chestionare, cu grad variabil de complexitate). În acest fel, se obțin date cantitative / calitative suplimentare față de cele obținute de evaluarea euristică. Repetarea evaluării cu o altă metodă, dar utilizând același set de sarcini relevante, astfel încât rezultatele să fie comparabile, mărește încrederea în rezultatele obținute.

Fiecare dintre cele două metode necesită competențe specifice. Compararea rezultatelor obținute permite analiza procedurii, identificarea greșelilor de evaluare și îmbunătățirea practicii de evaluare. Validarea rezultatelor evaluării euristice prin testare cu utilizatori permite și o comparație corectă a rezultatelor obținute de către fiecare expert în parte, fapt care îmbunătățește expertiza individuală.

Evaluarea euristică a avut loc pe parcursul procesului de dezvoltare a ARTP și a furnizat dezvoltatorilor, în timp util, recomandări de remediere a problemelor identificate.

4. Concluzii

În această lucrare a fost prezentată o metodologie integrată de evaluare a unei tehnologii de realitate îmbogățită destinată mediului școlar. Integrarea mai multor abordări metodologice a oferit posibilitatea triangulării rezultatelor care a permis obținerea unei imagini mai cuprinzătoare asupra obiectului evaluării și o înțelegere mai profundă a aspectelor critice.

Studiile efectuate cu ajutorul metodologiei integrate au permis analize comparative utile ale calității ergonomice, valorii educaționale și motivaționale ale ARTP. O concluzie importantă a evaluării cu metode cantitative este valoarea educațională și motivațională ridicată pe care o au capacitățile specifice realității îmbogățite. Analiza rezultatelor evaluării cu metode calitative a arătat că ambele aplicații (de învățare a biologiei și chimiei) dezvoltate pe ARTP prezintă atât

avantaje și oportunități în plan didactic, cât și limite.

Rezultatele obținute în studiile anterioare [2, 6, 7, 8, 9] au arătat că valoarea motivațională este mai mare în cazul aplicației de chimie decât a celei de biologie. Întrucât tehnologia AR este relativ scumpă, este de dorit ca utilizarea acesteia în mediul școlar să fie făcută diferențiat, pentru fiecare disciplină, în funcție de efectele estimate și de valoarea adăugată față de tehnologia de e-learning tradițională.

BIBLIOGRAFIE

1. **AZUMA, R.:** A Survey of Augmented Reality. *PRESENCE: Teleoperators and Virtual Environments*, Vol. 6, No. 4, 1997, pp. 355-385.
2. **BALOG, A.; PRIBEANU, C.; IORDACHE, D.:** Augmented Reality in Schools: Preliminary Evaluation Results from a Summer School. *International Journal of Social Sciences* Volume 2 Number 3. ISSN 2070-3872, 2007, pp.163-166.
3. **BALOG, A. (ed.):** Calitatea sistemelor interactive. Editura Matrix Rom, Bucuresti, 2004.
4. **CRISTEA, S.:** Dicționar de pedagogie, Grupul Editorial Litera Educațional, Chișinău, 2002.
5. **HARTSON, H. R.; ANDRE, T. S.; WILLIGES, R. C.:** Criteria for evaluating usability evaluation methods. *Internat. Journal of Human-Computer Interaction* 13, 2001, pp.373–410.
6. **IORDACHE, D. D.; PRIBEANU, C.:** Cum ajută realitatea îmbogățită înțelegerea chimiei – un model bazat pe indicatori cauzali. (Ed.) *Interacțiune Om-Calculator* 2013. ISSN, Editura MatrixROM București..
7. **IORDACHE, D. D.:** Evaluarea unei aplicații de învățare bazată pe realitate îmbogățită utilizând protocolul « gândire cu voce tare » și metoda învățării în perechi. *Revista Română de Interacțiune Om-Calculator*. Vol. 5 (1) 2012, pp. 185-198.
8. **IORDACHE, D. D.; PRIBEANU, C.; VILKONIS, R.:** A user-centred design approach to the development of AR-based educational systems”. *Proceedings of ICTNSE International Conference - 2007, Siauliai 3-4 December 2007*, pp. 66-70.
9. **IORDACHE, D. D.; NEACȘU, I.:** Evaluarea a două aplicații de realitate îmbogățită destinate mediului școlar – focus grup cu cadre didactice. *Revista Romana de Interacțiune Om-Calculator* 5 (2), 2012, pp. 19-24.
10. **JOHNSON, B.; CHRISTENSEN, L.:** Educational research, Quantitative, Qualitative, and Mixed Approaches (3rd ed.), Thousand Oaks, CA: Sage Publications, 2008.
11. **MOLICH, R.; JEFFRIES, R.; DUMAS, J.:** Making usability recommendations usable. *Journal of Usability Studies*, 2(1), 2007, pp. 162-179.
12. **MOLICH, R.; NIELSEN, J.:** Improving a human-computer dialogue, *Communications of the ACM* 33(3), 1990, pp. 338-348.
13. **OLIVER, M.:** The ELT Toolkit. University of North London, 1998.
<http://www.unl.ac.uk/tltc/elt/toolkit.pdf>
14. **PRIBEANU, C.; IORDACHE, D. D.:** Evaluating the motivational value of an augmented reality system for learning chemistry. Holzinger, A. (Ed.) *Proceedings of USAB 2008*. LNCS 5298 Springer, 2008, pp. 31-42.
15. **THEOFANOS, M.; QUESENBERRY, W.:** Towards the Design of Effective Formative Test Reports, *Journal of Usability Studies*, Issue 1, Vol.1, 2005, pp. 27-45.
16. **WIND, J.; RIEGE, K.; BOGEN M.:** Spinnstube®: A Seated Augmented Reality Display System, *Proceedings of IPT-EGVE – EG/ACM Symposium, 2007*, pp. 17-23.