

# Influența acurateții perceptuale asupra experienței utilizatorului unei aplicații de învățare a biologiei

Costin Pribeanu

Institutul Național de Cercetare-Dezvoltare în Informatică – ICI București

Bd. Mareșal Averescu nr.8-10, București

pribeanu@ici.ro

## REZUMAT

Interesul pentru aspectele hedonice ale interacțiunii cu sistemele de învățare a crescut spectaculos în ultimii ani. Tehnicile de interacțiune bazate pe realitate îmbogățită (AR – Augmented Reality) oferă noi oportunități pentru creșterea valorii motivaționale a sistemelor de e-learning. În acest articol se prezintă un index formativ care măsoară acuratețea percepției vizuale și auditive a unei aplicații de realitate îmbogățită pentru învățarea biologiei. Analiza relației dintre indicatorii măsurați formativ și un construct reflectiv, care măsoară plăcerea percepută, permite evidențierea influenței pe care o are utilizabilitatea asupra experienței utilizatorului.

## Cuvinte cheie

Utilizabilitate, realitate îmbogățită, model de măsurare, index formativ, plăcere percepută, e-learning.

## Clasificare ACM

D.2.2: Design tools and techniques. H5.2 User interfaces.

## INTRODUCERE

Realitatea îmbogățită oferă un nou tip de experiență pentru utilizatorii sistemelor de e-learning. Integrarea unor obiecte reale specifice disciplinei țintă și interacțiunea directă cu acestea crează un sentiment de control asupra procesului educațional și mărește motivația de a învăța. O problemă specifică acestor sisteme este utilizabilitatea, care se manifestă mai ales la nivelul tehnicilor de interacțiune [7].

Platforma ARTP (Augmented Reality Teaching Platform), a fost dezvoltată în cadrul unui proiect european și cuprinde trei aplicații care implementează trei scenarii de învățare distincte. Prima aplicație, care implementează un scenariu de învățare a biologiei, a fost evaluată din punctul de vedere al utilizabilității pe baza metodelor de evaluare formativă din HCI [12]. De asemenea, intenția de utilizare a ARTP a fost estimată cu ajutorul unui model structural [2]. Deși modelul structural a permis verificarea unor ipoteze relevante, varianța explicată a fost destul de mică, ceea ce sugerează unele limite ale modelului de măsurare.

O cauză este caracterul heterogen al constructului care măsoară ușurința în utilizare, în care au fost incluși o serie de itemi specifici tehnologiei AR. Cerințele de consistență internă și unidimensionalitate a scalei au impus eliminarea unor itemi, reducând astfel puterea explicativă a modelului structural. O altă cauză este dezvoltarea modelului de măsurare la nivelul ARTP, pe baza observațiilor colectate după utilizarea a două scenarii de învățare, relațiile

cauzale fiind mai degrabă generale pentru ARTP decât specifice unei aplicații.

În acest articol se prezintă un index măsurat cu indicatori formativi, care măsoară acuratețea percepției vizuale și auditive a unei aplicații de realitate îmbogățită pentru învățarea biologiei. Integrarea acestui index într-un model structural care conține un construct măsurat reflectiv permite analiza influenței acurateții percepției vizuale și auditive asupra plăcerii de a învăța cu această aplicație.

Restul acestui articol este structurat după cum urmează. În secțiunea următoare se prezintă succint câteva aspecte metodologice. În continuare, se descrie indexul formativ, se analizează validitatea modelului de măsurare și relațiile cauzale pe baza a două modele structurale. Articolul se încheie cu concluzii și direcții de continuare a cercetărilor.

## ASPECTE METODOLOGICE

### Evaluarea utilizabilității

Standardul ISO 9126:2001 definește utilizabilitatea prin capabilitatea produsului software de a fi ușor de înțeles, învățat, utilizat și considerat atractiv de către utilizator, atunci când este folosit în condiții specificate [10].

Există o mare varietate de metode utilizate în evaluarea utilizabilității, fiecare având avantaje și dezavantaje. Cele mai frecvent utilizate metode se pot grupa în două categorii: inspecția de utilizabilitate (efectuată de regulă prin evaluare euristică, evaluare bazată pe recomandări) și testarea cu utilizatori. Testarea cu utilizatori permite aplicarea unor instrumente de evaluare (chestionare) cu itemi închiși și deschiși, care permit studierea utilizabilității dintr-o perspectivă mai largă, care să și alți factori de interes, cum sunt utilitatea și experiența utilizatorului.

În majoritatea studiilor privind modelele de măsurare se utilizează doar două constructe relevante pentru HCI: ușurința în utilizare percepută (PEOU – Perceived Ease of Use) și plăcerea percepută (PE – Perceived Enjoyment). Cel mai adesea, acestea au fost utilizate în relație cu utilitatea percepută (PU – Perceived Usefulness) în contextul studierii impactului pe care îl au asupra intenției în utilizare.

### Măsurarea și estimarea unui index formativ

Un model de măsurare descrie relația dintre un construct și măsurile acestuia (itemi sau indicatori) în timp ce un model structural descrie relația între diferite constructe [5]. Specificarea corectă a modelului de măsurare este o

condiție necesară înainte de a analiza relațiile cauzale pe baza modelului structural [1]. Relația dintre construct și măsuri poate fi de la construct la măsuri (model reflectiv) sau de la măsuri la construct (model formativ), așa cum se arată în Figura 1.

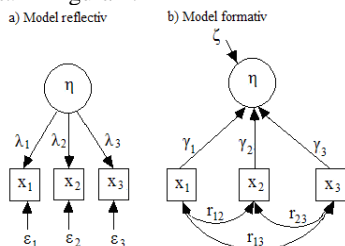


Figura 1. Modele de măsurare

În modelul de măsurare reflectiv relația cauzală este de la construct la indicatori, care sunt manifestări ale variabilei latente. Caracteristic modelelor reflective este faptul că o modificare în variabila latentă determină modificări simultane în toate variabilele manifest. Toate măsurile în acest model trebuie să fi corelate pozitiv iar modelul de măsurare trebuie să demonstreze validitate convergentă și discriminantă.

O alternativă este modelul de măsurare formativ, în care relația cauzală este de la măsuri la construct. Acest model are câteva caracteristici distincte față de modelul reflectiv: indicatorii nu sunt interschimbabili (fiecare capturează o cauză distinctă), nu există ipoteze specifice asupra intercorelațiilor, indicatorii nu au asociat un termen de eroare [3]. Un model formativ izolat este neidentificat și ca atare nu poate fi estimat. În multe din lucrările care publică modele de măsurare formativă, autorii folosesc termenul de index în locul celui de scală, pentru setul de indicatori.

Diamantopoulos et al. [4] menționează trei căi de identificare a modelelor formative, bazate pe regula 2+ (specificarea a două efecte ale constructului formativ asupra altor variabile), și anume: adăugarea de doi indicatori reflectivi, adăugarea a două constructe măsurate reflectiv și adăugarea unui construct reflectiv și a unui indicator reflectiv.

Prima cale conduce la un model MIMIC (*Multiple Indicators Multiple Causes*). A doua cale este adăugarea a două constructe măsurate reflectiv, dacă acestea pot fi incluse în rețeaua nomologică, pe baza unor considerații teoretice. Alegerea variabilelor efect afectează coeficienții  $\gamma$  atât în ceea ce privește semnificația cât și magnitudinea influenței [6]. Ultima modalitate de specificare este o combinație a primelor două. În acest caz, adăugarea unui indicator reflectiv ca o măsură globală (care capturează esența constructului formativ) permite identificarea și serveste, totodată, validării.

În ceea ce privește încrederea și validitatea, cerințele de consistență internă a scalei și unidimensionalitate nu sunt aplicabile modelelor formative. Literatura de specialitate recomandă validitatea de conținut, validitatea indicatorilor și validitatea externă [4, 12]. În privința conținutului, este necesară acoperirea domeniului, având în vedere că în acest caz măsurile definesc constructul. Itemii trebuie să aibă coeficienți  $\gamma$  semnificativi. Un alt mod de apreciere a

validității indicatorilor este corelația semnificativă cu un indicator extern (o măsură globală).

Estimarea modelului trebuie să demonstreze, prin indicii de calitate (adecvare) faptul că variabila latentă măsurată formativ mediază efectele indicatorilor asupra variabilelor măsurate reflectiv (coeficienți  $\beta$  semnificativi).

Așa cum subliniază mai mulți autori, estimarea aceluiași index cu alte variabile efect produce alte rezultate, atât în ceea ce privește relațiile cauzale cât și varianța explicată de către model [4, 6]. Din acest motiv se recomandă ca specificarea variabilelor efect să fie parte din specificarea constructului formativ și nu o etapă ulterioară. Cadrul de lucru teoretic avut în vedere și variabilele efect alese determină setul de indicatori ales (domeniul indexului) iar realizarea empirică variază de la un studiu la altul.

## DEZVOLTAREA ȘI ESTIMAREA INDEXULUI FORMATIV

### Eșantionul utilizat

Acest studiu utilizează un eșantion colectat în 2008 în cadrul evaluării unei aplicații care implementează un scenariu de învățare a biologiei. ARTP este un mediu AR de tip desktop: utilizatorii au în față un ecran „see-through”, pe care sunt suprapuse imagini ale obiectelor virtuale (imagini generate de calculator) peste imaginea observată a unui obiect real [14].

Chestionarul de utilizabilitate a cuprins 28 de itemi care măsoară ergonomia platformei (ERG), ușurința în utilizare (PEOU), utilitatea percepută (PU), plăcerea percepută (PE) și intenția de utilizare (INT). Mai multe detalii privind experimentul și instrumentul de evaluare pot fi găsite în lucrările anterioare [2, 12, 13].

Eșantionul inițial de 139 de observații a fost analizat din punctul de vedere al valorilor marginale (*outliers*) și normalității cu ajutorul programelor SPSS 17.0 for Windows și AMOS 17.0. Pe baza analizei datelor, au fost eliminate pe rând un număr de 9 observații astfel încât eșantionul de lucru de 130 de observații prezintă abateri moderate de la normalitate fiind adecvat pentru prelucrarea cu metode SEM (Structural Equation Modeling).

### Specificarea și estimarea indexului formativ

Lista inițială de indicatori, care măsoară aspecte specifice utilizabilității ARTP cuprinde 10 itemi care pot fi grupați în două categorii, care se referă la confortul și ușurința în interacțiune / colaborare, respectiv acuratețea percepției vizuale și auditive.

Relația dintre prima grupă de itemi și plăcerea percepută a fost analizată într-o lucrare anterioară [13]. Din a doua grupă, care face obiectul acestui studiu, fac parte patru itemi: ERG5 (Observarea prin ecran este clară), PEOU2 (Suprapunerea dintre proiecție și obiectul real este clară), PEOU5 (Înțelegerea explicațiilor vocale este ușoară) și PEOU6 (Citirea textului pe ecran este ușoară).

În vederea estimării constructului formativ, au fost adăugați doi itemi reflectivi, care măsoară plăcerea de a învăța cu acest sistem (PE5) și ușurința în învățarea modului de operare (PEOU3). În acest fel, s-a obținut un model MIMIC, care permite formularea de ipoteze privind relația cauzală dintre indicatorii măsurați formativ și

variabila latentă, precum și două ipoteze privind efectele indexului formativ asupra ușurinței în utilizare și caracterului plăcut al învățării cu ARTP. Modelul structural a fost estimat în AMOS 17.0.

Modelul MIMIC care cuprinde patru indicatori formativi este prezentat în Figura 2. Indexul este un construct compozit, asociat cu percepția vizuală și auditivă a ARTP și definit de patru indicatori care măsoară: claritatea observării prin ecran (obiectul real și dispozitivul de interacțiune), claritatea suprapunerii dintre proiecție (cu rol de îmbogățire / evidențiere a unui obiect selectat) și obiectul real, ușurința înțelegerii explicațiilor vocale (determinată de calitatea sunetului, a exprimării și dicției) și ușurința de a citi informația pe ecran.

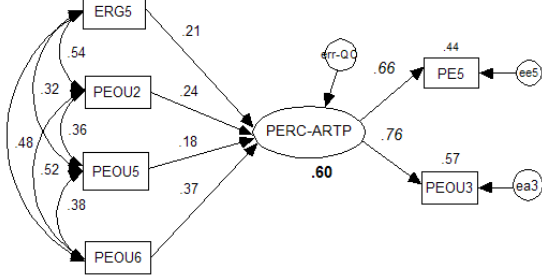


Figura 2. Modelul MIMIC pentru indexul PERC-ARTP

Toți coeficienții  $\gamma$  sunt semnificativi justificând includerea în index, astfel:  $\gamma_{ERG5} = 0.21$  ( $p=0.037$ ),  $\gamma_{PEOU2} = 0.24$  ( $p=0.023$ ),  $\gamma_{PEOU5} = 0.18$  ( $p=0.045$ ) și  $\gamma_{PEOU6} = 0.37$  ( $p<0.001$ ). Așa cum se observă, efectul direct al itemului ERG5 asupra indexului este mai mult decât dublu față de efectele celorlalți doi itemi. Colinearitatea indicatorilor, analizată cu statistica VIF (variation inflation factor), a fost 1.618, mult sub pragul recomandat de 10 [3].

Coeficienții  $\beta$  sunt semnificativi ( $p<0.001$ ), fapt care denotă suportarea a două ipoteze privind medierea de către variabila latentă a influenței pe care o au indicatorii formativi asupra plăcerii de a învăța cu ARTP și ușurinței generale în înțelegere a modului de utilizare a ARTP. Variabila latentă definită de indexul formativ are o influență mai mare asupra ușurinței în înțelegere decât asupra plăcerii de a învăța.

Itemul PEOU3 are un caracter general, servind astfel și pentru validarea de tip criteriu a unui set de indicatori care măsoară ergonomia și ușurința utilizării ARTP. Analiza pe baza coeficientului de corelație Pearson denotă relații lineare semnificative ( $p<0.001$ ) între PEOU1 și itemii care formează indexul, astfel:  $R_{ERG5} = 0.540$ ,  $R_{PEOU2} = 0.462$ ,  $R_{PEOU5} = 0.366$  și  $R_{PEOU6} = 0.404$ .

Indicii de adecvare ai modelului au valori foarte bune, peste pragurile minime recomandate [8]:  $\chi^2=6.244$ ,  $DF=3$ ,  $CFI=.985$ ,  $GFI=.984$ ,  $SRMR=0.023$ . Varianța totală explicată de model pentru PERC-ARTP este 60%. Analiza rezultatelor denotă proporționalitatea efectelor structurale (indexul mediază complet efectele) [6].

O altă modalitate de identificare a indexului formativ în vederea estimării este includerea unui construct reflectiv (PE) și a unui indicator măsurat reflectiv.

Având în vedere că modelul de măsurare include un construct reflectiv, este necesară verificarea consistenței

interne și unidimensionalității acestuia. Consistența subscalei PE a fost .856 (Cronbach's alpha) ceea ce este satisfăcător ( $\text{prag}>0.7$ ). Analiza dimensionalității a fost făcută pe baza analizei comunalităților și gradului de saturație al fiecărei variabile (*factor loading*), corespunzător prgurilor date de Hair et al. (2006). Itemul PE2 a fost eliminat întrucât a avut comunalitățile sub pragul de 0.4.

Analiza validității convergente s-a făcut pe baza comparării fiabilității compuse (CR) și varianței medii extrase (AVE) cu valorile prag. Întrucât  $CR=0.861$  ( $>0.7$ ) iar  $AVE=0.554$  ( $>0.5$ ), constructul PE compus din cinci itemi are validitate convergentă.

În Figura 3 este prezentat modelul structural care include constructul PE și itemul PEOU1. Modelul a fost estimat în AMOS 17.0 și rezultatele au confirmat validitatea la nivel de indicator, după cum urmează:  $\gamma_{ERG5} = 0.24$  ( $p=0.010$ ),  $\gamma_{PEOU2} = 0.21$  ( $p=0.026$ ),  $\gamma_{PEOU5} = 0.18$  ( $p=0.034$ ) și  $\gamma_{PEOU6} = 0.35$  ( $p<0.001$ ). O comparație cu modelul MIMIC arată diferențe mici în ceea ce privește coeficienții  $\gamma$ .

Coeficienții  $\beta$  sunt semnificativi (la un prag de semnificație  $p<0.001$ ) și au valori ridicate. Varianța explicată de către model este de  $R^2=0.56$  pentru indexul formativ și  $R^2=0.79$  pentru constructul PE. În ceea ce privește itemii asociați cu plăcerea percepută, influența cea mai mare a indexului formativ se reflectă în caracterul captivant al exercițiilor (PE4) și caracterul plăcut al învățării (PE5).

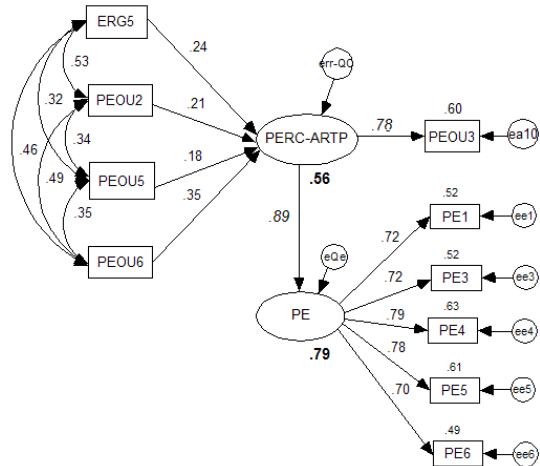


Figura 3. Model structural cu indexul PERC-ARTP și un construct reflectiv

Indicii de adecvare a modelului au valori foarte bune, peste pragurile recomandate [8]:  $\chi^2=33.170$ ,  $DF=28$ ,  $CFI=0.990$ ,  $GFI=.952$ ,  $SRMR=0.038$ . Analiza efectelor directe și indirecte arată că nu există efecte directe ale indicatorilor formativi asupra variabilelor efect PE și PEOU1 (indexul mediază complet efectele) [6]. Estimarea cu al doilea model structural confirmă validitatea indexului formativ conform cu recomandările din literatura de specialitate [3, 4, 6,10].

Analiza modelului structural confirmă ipoteza unui efect pozitiv semnificativ PEOU→PE, verificată într-un studiu anterior [2], efect care denotă o legătură între aspectele pragmatice și cele hedonice ale utilizabilității. De asemenea, studiul a permis identificarea unui index

formativ asociat cu acuratețea percepției vizuale și auditive, relevant pentru platforma ARTP și a demonstrat influența pe care o are claritatea observației prin ecranul “see-through” asupra experienței utilizatorului.

Așa cum s-a arătat în [9], cele mai scăzute medii s-au înregistrat pentru itemii PEOU2 ( $M=3.62$ ,  $SD=1.02$ ), ERG5 ( $M=3.76$ ,  $SD=1.07$ ). De asemenea, se confirmă, printr-o metodă diferită, faptul că acuratețea percepției vizuale este o problemă critică pentru sistemele bazate pe AR. În acest sens, analiza unui model structural bazat pe itemi formativi este o metodă complementară.

Există unele limite inerente acestui studiu de caz. În primul rând, construirea indexului a fost făcută pe un eșantion existent, având ca scop studierea unui model de măsurare alternativ, pentru analiza influenței unor măsuri care au fost eliminate din rațiuni de unidimensionalitate și validitate.

O altă limitare inerentă este faptul că instrumentul de evaluare elaborat inițial a fost primul de acest gen pentru tehnologii AR de tip desktop. În acest sens, studiul literaturii de specialitate nu a permis identificare de scale de măsurare existente care să permită adaptarea unor constructe cu itemi validați într-o rețea nomologică.

În fine, chiar specificul ARTP a creat probleme în interpretarea relațiilor cauzale dintre variabilele latente și măsuri. Pentru o discuție mai detaliată și referințe privind problematica definirii unor constructe ca reflectiv sau formative, vezi [4, 11].

#### CONCLUZII ȘI DIRECȚII DE CONTINUARE

Specificarea unui construct cu măsuri formative a fost utilă pentru o mai bună înțelegere a unor aspecte specifice utilizabilității ARTP și a direcțiilor în care trebuie rafinat modelul de măsurare. În acest sens, acuratețea percepției vizuale și auditive vor fi avute în vedere la elaborarea unui nou chestionar de utilizabilitate. Concluzii asemănătoare s-au desprins și din rezultatele obținute în studiile anterioare [2, 9, 12, 13].

Studiul relațiilor cauzale pe baza unor modele de măsurare alternative, prezentate în [13] și în această lucrare, a permis definirea a două indexuri formative care măsoară aspecte specifice utilizabilității ARTP. Integrarea acestor constructe în modele structurale a furnizat astfel elemente suplimentare privind utilizabilitatea aplicației și influența pe care o are asupra experienței utilizatorului.

Se intenționează repetarea studiului pentru aplicația care implementează scenariul de învățare a chimiei, pe baza eșantionului din 2008. Rezultatele din ambele studii, împreună cu analiza rezultatelor din studiile anterioare vor fi utilizate în elaborarea unui nou model de măsurare, pentru o aplicație îmbunătățită care va implementa un scenariu de învățare a chimiei pe ARTP. Acest model va cuprinde atât constructe reflectiv cât și indexuri reflectiv. Noul model de măsurare va fi testat în 2011.

#### Confirmare

Această lucrare a fost finanțată din Programul Nucleu TEHSIN 503/2009.

#### REFERINȚE

1. Anderson, J.C., Gerbing, D.W. (1988). Structural Equation Modelling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin* 103 (3), 411-423.
2. Balog, A., Pribeanu, C. (2010). The Role of Perceived Enjoyment in the Students' Acceptance of an Augmented Reality Teaching Platform: a Structural Equation Modelling Approach. *Studies in Informatics and Control*, 19 (3), 319-330.
3. Diamantopoulos, A., Winklhofer, H. (2001) Index construction with formative indicators : an alternative to scale development. *Journal of Marketing Research* 28, 269-277.
4. Diamantopoulos, A., Riefler, P., Roth, K. (2008) Advancing formative measurement models. *Journal of Business Research* 61, 1203-1218.
5. Edwards, J., Bagozzi, R. (2000) On the nature and direction of relationship between constructs and measures. *Psychological Methods* 5(2), 155-174.
6. Franke, G., Preacher, K., Rigdon, E. (2008) Proportional structural effects of formative indicators. *Journal of Business Research* 61, 1229-1237.
7. Gabbard, J., Swann, E. (2008) Usability engineering for augmented reality: Employing user-based studies to inform design. *IEEE Transactions on Visualization and Computer Graphics* 14(3), 513-525.
8. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L. (2006). *Multivariate Data Analysis*. 6<sup>th</sup> Ed., Prentice Hall, 2006.
9. Iordache, D.D., Pribeanu, C. (2011) Impactul problemelor de utilizabilitate asupra utilității percepute și experienței utilizatorului unui scenariu de învățare bazat pe realitate îmbogățită. *Revista Română de Interacțiune Om-Calculator* 4(Număr special – RoCHI 2011), în curs de publicare.
10. ISO 9126-1:2001 Software Engineering - Software product quality. Part 1: Quality Model.
11. Jarvis, C.B., Mackenzie, S., Podsakoff, M. (2003) A critical review of construct indicators and measurement models misspecification in marketing and consumer research. *Journal of Consumer Research* 30, 199-218.
12. Pribeanu, C., Iordache, D.D., Balog, A. (2008) Evaluarea utilizabilității unui scenariu de învățare a biologiei implementat pe o platformă de realitate îmbogățită. *Revista Română de Interacțiune Om-Calculator* 1(1), 39-56.
13. Pribeanu, C. (2011) Studiu de caz în specificarea unui index formativ pentru analiza utilizabilității unei aplicații de e-learning. *Revista Română de Interacțiune Om-Calculator* 4(1), în curs de apariție.
14. Wind, J., Riege, K., Bogen M., 2007. Spinnstube@: A Seated Augmented Reality Display System, *Virtual Environments: Proceedings of IPT-EGVE – EG/ACM Symposium*, 17-23.