

Studiu de caz în specificarea unui index formativ pentru analiza utilizabilității într-o aplicație de e-learning

Costin Pribeanu

Institutul Național de Cercetare-Dezvoltare în Informatică – ICI București
Bd. Mareșal Averescu, Nr. 8-10, 011455, București
E-mail: pribeanu@ici.ro

Rezumat. În prezent există puține abordări în evaluarea utilizabilității care utilizează modele structurale pentru analiza relațiilor cauzale între factorii care influențează calitatea în utilizare și / sau intenția de utilizare. De asemenea, există puține modele de măsurare care includ constructe compozite cu scopul de a ținti mai multe dimensiuni. Un subiect de interes în literatura de specialitate din ultimii ani este adecvarea modelelor de măsurare la natura constructului și la direcția relațiilor cauzale dintre construct și măsurile acestuia. În acest articol se prezintă o sinteză a preocupărilor recente privind metodologia de specificare și validare a modelelor de măsurare formative. Metodologia a fost aplicată la dezvoltarea unui model structural având ca variabilă centrală un construct compozit, măsurat cu indicatori formativi. Scopul acestui studiu este analiza mai fină a unor aspecte specifice ale utilizabilității unei aplicații educaționale, prin intermediul a două constructe (unul formativ și unul reflectiv) și analiza relației dintre utilizabilitate și plăcerea percepută.

Cuvinte cheie: model de măsurare, index formativ, utilizabilitate, plăcere percepută.

1. Introducere

Evaluarea utilizabilității aplicațiilor informatice are o importanță deosebită într-o economie modernă, având în vedere impactul pe care îl au problemele de utilizabilitate asupra eficacității, eficienței și satisfacției cu care un utilizator își îndeplinește sarcinile de lucru. În ultima decadă se observă o expansiune a preocupărilor în direcția experienței utilizatorului (UX – User eXperience), ca urmare a creșterii interesului față de aspectele hedonice ale interacțiunii cu calculatorul.

O disciplină care s-a dezvoltat independent de interacțiunea om-calculator (HCI) este modelarea factorilor care influențează intenția de utilizare și utilizarea propriu-zisă a unei aplicații, precum și relațiile cauzale dintre acești factori. În mod tradițional, modelele structurale au la bază scale

de măsurare reflectivă, în care o variabilă latentă este măsurată prin intermediul unor indicatori observabili (Churchill, 1979; Edwards & Bagozzi, 2000). Adeseori, estimarea modelelor se face cu metode bazate pe ecuații structurale (SEM – Structural Equation Modelling).

În prezent există o dezbatere continuă privind caracterul formativ sau reflectiv al unor constructe, modalitatea de includere a indicatorilor formativi în modelele de măsurare și gradul de adecvare a scalelor de măsurare care sunt folosite frecvent (Diamantopoulos, 2008).

Deși există o zonă de intersecție care s-a extins de la ușurința în utilizare la experiența utilizatorului, în literatura din HCI există relativ puține studii care să integreze specificarea unor modele structurale în metodologiile de evaluare a utilizabilității. De asemenea, au fost dezvoltate puține scale de măsurare care țintesc utilizabilitatea (ușurința în utilizare fiind numai un aspect al acesteia). Un avantaj important al modelelor structurale este posibilitatea de a analiza relația dintre diferitele aspecte ale utilizabilității, cum ar fi ușurința în învățare sau ușurința în operare, utilitatea percepută a aplicației și experiența utilizatorului. Studiul acestor relații are o importanță crescândă în sistemele de e-learning, care țintesc introducerea unor metode moderne în educație.

O categorie de sisteme de e-learning, care implementează concepte bazate pe teoriile constructiviste ale învățării în scopul creșterii eficacității pedagogice și motivației de a învăța, sunt sistemele bazate pe realitate îmbogățită (AR – Augmented Reality). Un exemplu este platforma ARTP (Augmented Reality Teaching Platform), dezvoltată în cadrul unui proiect european. ARTP cuprinde trei aplicații care implementează trei scenarii de învățare distincte.

Prima aplicație, destinată învățării biologiei, a fost evaluată din punctul de vedere al utilizabilității și eficacității pedagogice, atât pe baza metodelor de evaluare formativă¹ din HCI (Pribeanu et al, 2008; Iordache & Pribeanu, 2009) cât și cu ajutorul unui model de măsurare reflectiv (Balog & Pribeanu, 2010). Deși modelul structural a permis verificarea unor ipoteze relevante pentru tehnologiile AR cu scop educațional, varianța explicată la

¹ Pentru a evita orice confuzie de termeni, se cuvine menționat faptul că nu există nici o legătură între evaluarea formativă și modelele de măsurare formative.

nivelul diferitelor constructe a fost destul de mică, ceea ce sugerează unele limite ale modelului de măsurare.

O cauză este caracterul eterogen 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 ale scalei au impus eliminarea unor itemi, reducând astfel puterea explicativă a modelului structural.

O altă cauză este numărul relativ mic de observații raportat la numărul de itemi, la limită față de cerințele metodelor de modelare bazate pe ecuații structurale. Din acest motiv, dezvoltarea modelului de măsurare a fost făcută la nivelul ARTP, pe baza observațiilor colectate după utilizarea a două scenarii de învățare (biologie și chimie). În acest sens, relațiile cauzale sunt mai degrabă generale pentru o platformă educațională AR de tip desktop decât specifice unei aplicații care implementează un anumit scenariu de învățare.

În acest articol se prezintă sintetic unele considerații metodologice privind dezvoltarea constructelor formative și integrarea acestora în modele structurale, care pot fi apoi estimate cu ajutorul SEM. Metodologia este aplicată la specificarea a două modele structurale, care integrează un construct formativ. Spre deosebire de modelul dezvoltat anterior, aceste modele sunt mai restrânse și au ca scop studiul influenței pe care o au trei caracteristici specifice aplicației care implementează scenariul de biologie asupra ușurinței în utilizare și experienței utilizatorului.

2. Considerații metodologice

2.1 Evaluarea utilizabilității

Standardul ISO 9241-11 (1994) a definit utilizabilitatea ca măsură în care un anumit produs poate fi utilizat de către utilizatori specificați, pentru a îndeplini obiective specificate, într-un context de lucru specificat, cu eficacitate, eficiență și satisfacție. 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. Același standard definește calitatea în utilizare prin intermediul a patru sub-caracteristici: eficacitate, productivitate,

siguranță și satisfacție. Această definiție sugerează o perspectivă mai largă asupra evaluării utilizabilității, care să includă și alte aspecte.

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 prin evaluare euristică, evaluare bazată pe recomandări) și testarea cu utilizatori. Testarea cu utilizatori are avantajul că permite aplicarea unor instrumente de evaluare (chestionare) cu itemi închiși și deschiși, precum și colectarea unor măsuri ale eficacității și eficienței. Chestionarele de utilizabilitate permit analiza utilizabilității dintr-o perspectivă mai largă, care include și alți factori de interes, cum sunt utilitatea și experiența utilizatorului.

Există numeroase studii care au ca scop modelarea factorilor care influențează intenția de utilizare și utilizarea propriu-zisă a unei aplicații, precum și relațiile cauzale dintre acești factori. În majoritatea cazurilor se utilizează două constructe relevante pentru HCI: ușurința în utilizare percepută (PEOU – Perceived Ease of Use) și plăcerea percepută (PE – Perceived Enjoyment). Aceste constructe sunt analizate din perspectiva relației cu utilitatea percepută (PU – Perceived Usefulness) în contextul studierii impactului pe care îl au asupra intenției de utilizare a sistemului (Davis, 1989).

Modelele structurale au fost rareori folosite în evaluarea utilizabilității sistemelor interactive, probabil datorită caracterului laborios al aplicării metodologiei de validare a modelului de măsurare, cerințelor privind mărimea eșantionului și expertizei necesare în utilizarea unor instrumente software specializate pentru estimarea modelului structural. De asemenea costul ridicat nu încurajează utilizarea în etapele de evaluare formativă, în care sunt mai eficiente inspecția de utilizabilitate și / sau testarea cu un număr relativ mic de utilizatori.

2.2 Modele de măsurare formative și reflectivă

Teoria clasică a testării și analiza factorială au influențat abordările asupra măsurării variabilelor latente, atât în psihologie cât și în științele sociale (Bollen & Lennox, 1991). 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 (Edwards & Bagozzi, 2000). Specificarea corectă a modelului de măsurare este o condiție

necesară înainte de a fi analizat modelul structural (Anderson & Gerbing, 1988). Conform cu Jarvis et al. (2003), în literatura de specialitate există multe cazuri de specificări greșite ale modelelor de măsurare.

Modelul de măsurare specifică relația dintre construct și măsuri, care poate fi de la construct la măsuri (model reflectiv) sau de la măsuri la construct (model formativ), așa cum se arată în Figura 1.

În modelul de măsurare reflectiv relația cauzală este de la construct la indicatori, care sunt manifestări ale variabilei latente. Fiecare item (x_i) este o funcție liniară a variabilei latente (η) plus eroarea de măsurare (ε_i). Caracteristic modelelor reflective este faptul că o modificare în variabila latentă determină modificări simultane în toate variabilele manifest. De asemenea, toate măsurile în acest model trebuie să fie corelate pozitiv. Pentru a permite o analiză corectă a modelului structural, modelul de măsurare trebuie să aibă validitate convergentă și discriminantă.

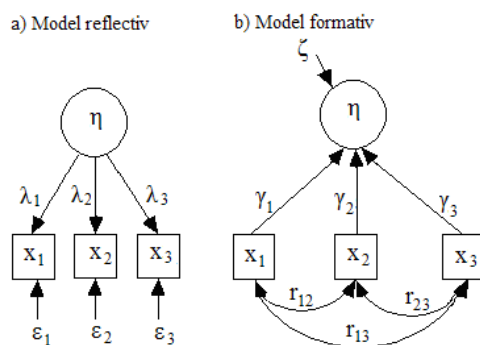


Figura 1. Modele de măsurare

O alternativă este modelul de măsurare formativ (Figura 1b) în care relația cauzală este de la măsuri la construct. În acest caz, o variabilă latentă este suma produselor dintre măsuri (x_i) și efectul pe care îl au asupra variabilei latente (γ_i), plus un termen de eroare (ζ). Acesta din urmă, nu reprezintă o eroare de măsurare ci o perturbație (*disturbance*), care capturează efectul cauzelor care nu sunt cuprinse de setul de itemi (Diamantopoulos, 2006).

Acest model are câteva caracteristici distincte față de modelul reflectiv (Diamantopoulos & Winklhofer, 2001): indicatorii nu sunt interschimbabili (fiecare capturează o cauză distinctă), nu există ipoteze specifice asupra

intercorelațiilor, indicatorii nu au asociat un termen de eroare. Se recomandă intercorelarea indicatorilor aceluiași construct (Bollen & Lennox, 1991).

2.3 Alegerea modelului de măsurare

Adecvarea modelului de măsurare nu este o problemă simplă. Așa cum arată Jarvis et al (2003), în literatura de specialitate au fost publicate multe modele de măsurare inadecvat specificate. În analiza făcută de autori, pe baza analizei a 178 de articole publicate în 4 reviste de top din marketing, se constată că gradul de inadecvare a modelelor de măsurare este destul de ridicat: 29% raportat la un total de 1192 de constructe. Variația de la o revistă la alta este de asemenea mare, de la 17% la 34% cazuri de specificare inadecvată.

Mai mult decât atât, se menționează dificultățile pe care le-au avut chiar autorii studiului în clasificarea unor constructe (14% din total), care aveau atât caracteristici reflectiv cât și formative. În cadrul acestui studiu, ei au propus un set de criterii pentru determinarea naturii reflectiv sau formative a unui construct. Aceste criterii sunt sintetizate în Tabelul 1 (majoritatea sunt formulate ca întrebări).

Tabelul 1. Criterii pentru determinarea modelului de măsurare (adaptat după Jarvis et al, 2008)

Criteriu	Model formativ	Model reflectiv
Direcția cauzalității	itemi → construct	construct → itemi
Itemii definesc caracteristici sau sunt manifestări ale constructului	caracteristici	manifestări
Modificările în itemi produc modificări în construct	Da	Nu
Modificările în construct produc modificări în itemi	Nu	Da
Indicatori sunt interschimbabili	Nu	Da
Conținut similar al itemilor	Nu	Da
Eliminarea unui item alterează domeniul constructului	Da	Nu
Covarianța între itemi	nu este necesară	este necesară
Rețeaua nomologică a itemilor constructului	poate să difere	este aceeași

Dificultățile de stabilire a modelului de măsurare cel mai adecvat sunt menționate și de Wilcox et al (2008), care arată că un construct nu este în mod inerent reflectiv sau formativ, astfel încât în multe cazuri cercetătorul are posibilitatea de a opta pentru un model sau altul. În acest sens, se poate considera că studierea unor modele de măsurare alternative este utilă și poate permite o analiză mai fină a unor aspecte specifice.

2.4 Identificarea modelelor formative

Un model formativ izolat este neidentificat și ca atare nu poate fi estimat. De menționat că 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 (2008) menționează trei modalități 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 este recomandată și de Jarvis et al (2003), întrucât parametrii măsurării sunt mai puțini senzitivi la schimbări în parametrii structurali. Constelația rezultată este referită ca model MIMIC (Multiple Indicators Multiple Causes). Mai multe exemple de modele MIMIC sunt prezentate de Cadogan et al (2008).

A doua cale este adăugarea a două constructe măsurate reflectiv, care este justificată dacă acestea pot fi incluse în rețeaua nomologică, pe baza unor considerații teoretice. Alegerea constructelor care specifică efectele constructului formativ afectează interpretarea prin influențarea coeficienților γ în ceea ce privește completitudinea medierii efectelor, semnificația și magnitudinea influenței (Franke et al, 2008).

Ultima modalitate de specificare este o combinație a primelor două, aplicabilă atunci când există o singură relație structurală cu un construct măsurat reflectiv. În acest caz, adăugarea unui indicator reflectiv ca o măsură globală (care capturează esența constructului formativ) permite identificarea și servește, totodată, validării.

2.5 Evaluarea modelelor formative

În ceea ce privește încrederea și validitatea unui index formativ, cerințele de consistență internă a scalei și unidimensionalitate nu sunt aplicabile modelelor formative. Diamantopoulos & Winklhofer (2001) menționează următoarele criterii de apreciere a validității unui index formativ: validitatea de conținut, validitatea indicatorilor și validitatea externă.

În ceea ce privește validitatea de conținut, este necesară acoperirea domeniului, având în vedere că în acest caz măsurile definesc constructul.

În ceea ce privește validitatea indicatorilor, coeficienții γ trebuie să fie semnificativi. Un alt mod de apreciere a validității indicatorilor este corelația semnificativă cu un indicator extern (o măsură globală). Multicolinearitatea între indicatori nu este permisă.

Estimarea modelului structural trebuie să demonstreze, prin indicii de calitate (adecvare), faptul că variabila latentă măsurată formativ mediază complet efectele indicatorilor asupra variabilelor măsurate reflectiv (coeficienți β semnificativi). Sunt importante semnul și magnitudinea coeficienților β , care măsoară influența constructului formativ asupra constructelor reflective incluse în model (validare la nivel nomologic).

Așa cum subliniază Franke et al (2008), estimarea aceluiași index cu alte variabile efect produce alte rezultate, atât în ceea ce privește relațiile cauzale cât și varianța explicată de către model. Este posibil ca relațiile cauzale dintre indicatorii formativi și index să fie semnificative într-un model și ne semnificative în altul. Din acest motiv, ei recomandă ca specificarea variabilelor efect să fie parte din specificarea constructului formativ și nu o etapă ulterioară. De asemenea, nu trebuie să existe influențe directe ale indicatorilor formativi asupra variabilelor efect.

Așa cum arată și Diamantopoulos (2006), alegerea variabilelor efect este la fel de importantă ca alegerea indicatorilor care formează indexul. Cadrul de lucru teoretic avut în vedere și variabilele efect alese determină setul de indicatori ales (domeniul indexului). De asemenea, Wilcox et al (2008) arată că natura constructului măsurat formativ se modifică de la un model la altul, în funcție de predicție. În acest sens, chiar dacă numele constructului se păstrează, realizarea empirică variază de la un model la altul.

3. Studiu de caz

3.1 Participanți, echipament și metodă

Studiul de caz utilizează un eșantion colectat în 2008 în cadrul evaluării utilizabilității unei aplicații care implementează un scenariu de învățare a biologiei. În cele ce urmează vom prezenta succint contextul evaluării, având în vedere că mai multe detalii pot fi găsite în lucrările anterioare (Pribeanu et al, 2008; Balog & Pribeanu, 2010).

Experimentul a fost organizat în ICI București. La testare au participat 139 de elevi din clasa a 8-a, provenind de la 3 școli generale din București

(câte 2 clase de la fiecare școală). Elevii au venit în grupuri de 6-8, acompaniați de un profesor.

ARTP este un mediu AR de tip „seated” (Wind, Riege & Bogen, 2007): 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. În cazul aplicației care implementează scenariul de învățare a biologiei, obiectul real este un mulaj aplatizat al sistemului digestiv uman, care poate fi partajat de doi elevi care stau față în față.

Ca instrument de interacțiune a fost utilizat un dispozitiv de poziționare și selectare alcătuit dintr-o telecomandă Wii Nintendo pe care este montată o baghetă metalică la capătul căreia este o bilă colorată din plastic. Telecomanda este utilizată și pentru selectarea unui item din meniu.

În Figura 1 se pot observa două eleve care au testat scenariul de biologie în 2008. Configurația din ICI cuprinde 4 module iar cele două mulaje sunt partajate de către doi utilizatori care stau față în față.



Figura. 1. Elevi testând scenariul pentru biologie în 2008

Chestionarul de utilizabilitate prezentat în Anexa 1 cuprinde 28 de itemi. Primii 24 itemi ținesc mai multe dimensiuni: ergonomia platformei (ERG), ușurința în utilizare (PEOU), utilitatea percepută (PU), plăcerea percepută (PE) și intenția de utilizare (INT). Ultimii 4 itemi sunt generali și măsoară ușurința în utilizare, utilitatea pentru învățare, caracterul atractiv al învățării și caracterul incitant al utilizării.

Participanții au fost solicitați să evalueze itemii pe o scală Likert cu 5 trepte (de la 1 – dezacord total, la 5 – acord total). Gradul de fidelitate a

scalei (Cronbach's Alpha) a fost 0.942, ceea ce este acceptabil. Media și abaterea standard pentru fiecare item sunt prezentate în Anexa 1.

Eșantionul inițial de 139 de observații a fost analizat din punctul de vedere al valorilor marginale (*outliers*) și normalității cu ajutorul programului SPSS 17.0 for Windows. În acest scop au fost calculate valorile marginale univariate (pe baza scorurilor standardizate z) și multivariate (pe baza distanței Mahalanobis) și au fost analizați indicatorii de asimetrie și exces (pe baza scorurilor standardizate z). Pe baza analizei, 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.

3.3 Specificarea și estimarea unui index formativ

Există două motivații ale acestui studiu. Prima este legată de elaborarea unui model structural pentru evaluarea fiecărui scenariu de învățare în parte. A doua este legată de utilizarea unui model de măsurare alternativ, care să includă itemi specifici unei platforme de învățare bazată pe AR (itemi care au fost eliminați în studiul menționat anterior, din rațiuni de consistență internă a scalei și unidimensionalitate).

Din punct de vedere teoretic, scopul studiului este de a analiza în detaliu aspecte specifice relației dintre utilizabilitate și experiența utilizatorului. În acest sens, interesează itemii cuprinși în constructele ERG, PEOU și PE din chestionarul de utilizabilitate.

Lista inițială de indicatori, care măsoară aspecte specifice utilizabilității ARTP având astfel potențial de influențare a experienței utilizatorului și ușurinței în utilizare, cuprinde 10 itemi: ERG1-ERG5, PEOU2, PEOU5-PEOU7 și PEOU9. Din punct de vedere conceptual, acești itemi pot fi grupați în două categorii, care se referă la confortul și ușurința în operare / interacțiune, respectiv acuratețea percepției vizuale și auditive.

Din prima categorie, care face obiectul acestui studiu, fac parte șase itemi: ERG1, ERG2, ERG3, ERG4, PEOU7 și PEOU9. În vederea estimării constructului formativ, au fost adăugați doi itemi măsurați reflectiv, care țintesc plăcerea de a învăța cu acest sistem (PE5) și ușurința generală în utilizare (PEOU10). 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 cu programul AMOS 17.0 (Arbuckle, 2007).

Analiza relațiilor cauzale dintre indicatori și variabila latentă a condus la eliminarea succesivă a trei dintre indicatori, care aveau coeficienți γ ne semnificativi. Modelul MIMIC revizuit, care cuprinde trei indicatori formativi este prezentat în Figura 3. Indexul este un construct compozit, asociat cu confortul general al utilizării ARTP și definit de trei indicatori care măsoară: confortul postului de lucru (constrâns de sistemul de ecrane), ușurința selectării unui articol din meniu (cu ajutorul telecomenzii) și ușurința colaborării între elevi (care partajează același obiect real).

Toți coeficienții γ sunt semnificativi (la un prag de semnificație $p < 0.05$) și au valori care argumentează pentru includerea în index, astfel: $\gamma_{ERG4} = 0.43$ ($p < 0.001$), $\gamma_{PEOU7} = 0.16$ ($p = 0.049$), $\gamma_{PEOU9} = 0.31$ ($p < 0.001$). Așa cum se observă, itemul ERG4 are cea mai mare influență asupra indexului, urmat de itemul PEOU9. Întrucât colinearitatea indicatorilor nu este de dorit, aceasta a fost analizată pe baza statisticii VIF (*variation inflation factor*), care a avut valoare 1.618, cu mult sub pragul recomandat de 10.0.

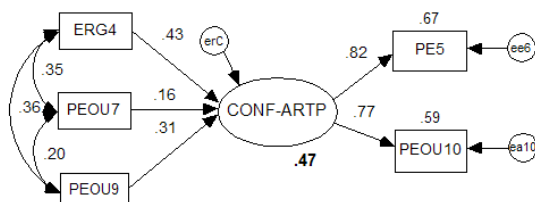


Figura 3. Model MIMIC pentru indexul CONF-ARTP

De asemenea, coeficienții β sunt semnificativi (la un prag de semnificație < 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 utilizare. Variabila latentă definită de indexul formativ are o influență mai mare asupra plăcerii de a învăța decât asupra ușurinței generale în utilizare.

Itemul PEOU10 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 PEOU10 și itemii care formează indexul, astfel: $R_{ERG4} = 0.444$, $R_{PEOU5} = 0.346$ și $R_{PEOU9} = 0.412$.

Indicii de adecvare ai modelului au valori foarte bune, peste pragurile recomandate (Hair et al, 2006): $\chi^2=1.506$, $DF=2$, $\chi^2/df=0.753$, $TLI=1.017$, $CFI=1.000$, $RMSEA=0.000$, $SRMR=0.015$. Varianța totală explicată de model pentru CONF-ARTP este 47%. Analiza efectelor directe și indirecte a arătat că nu există efecte directe ale indicatorilor formativi asupra itemilor efect, astfel încât indexul mediază complet efectele itemilor săi (Franke et al., 2008).

O altă modalitate de identificare a indexului formativ în vederea estimării este includerea unui construct reflectiv (PE) și a unui indicator măsurat reflectiv. În acest caz, este necesară verificarea prealabilă a consistenței interne și unidimensionalității constructului. Consistența sub-scalei PE a fost .856 (Cronbach's alpha) ceea ce este satisfăcător (prag>0.7). Analiza dimensionalității a fost făcută pe baza analizei comunalităților și gradului de saturație (*factor loading*) al fiecărei variabile, corespunzător pragurilor date de Hair et al. (2006). Itemul PE2 a fost eliminat întrucât a avut comunalitățile sub pragul de 0.4. De asemenea, itemul PE3 a fost eliminat întrucât gradul de saturație a fost sub pragul de 0.55.

Analiza validității convergente s-a făcut pe baza comparării fiabilității compuse (CR) și varianței medii extrase (AVE) cu valorile prag. Întrucât $CR=0.855$ (>0.7) iar $AVE=0.596$ (>0.5), constructul PE compus din patru itemi are validitate convergentă.

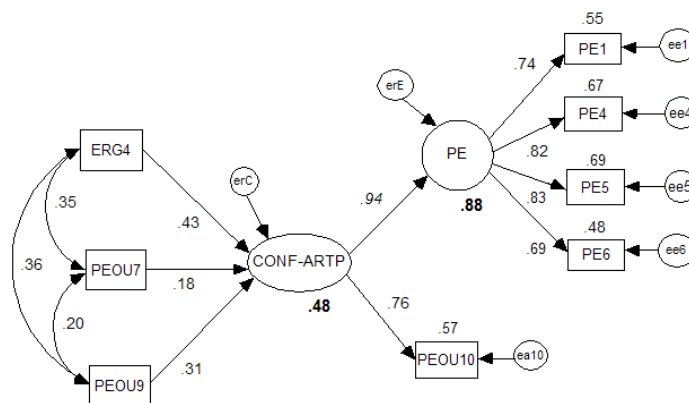


Figura 4. Model structural cu indexul CONF-ARTP cu un construct reflectiv

În Figura 4 este prezentat modelul structural care include constructul PE și itemul PEOU10. Modelul a fost estimat în AMOS 17.0 și rezultatele au

confirmat validitatea la nivel de indicator, după cum urmează : $\gamma_{\text{ERG4}} = 0.43$ ($p < 0.001$), $\gamma_{\text{PEOU5}} = 0.18$ ($p = 0.032$), $\gamma_{\text{PEOU9}} = 0.31$ ($p < 0.001$). O comparație cu modelul MIMIC arată diferențe foarte mici în ceea ce privește varianța explicată și coeficienții γ .

Coeficienții β sunt semnificativi (la un prag de semnificație $p < 0.001$) și au valori ridicate, de 0.94 și 0.76. Varianța explicată de către model este de $R^2 = 0.48$ pentru indexul formativ și $R^2 = 0.88$ 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 și caracterul plăcut al învățării.

Indicii de adecvare a modelului au valori foarte bune, peste valorile minime recomandate (Hair et al, 2006): $\chi^2 = 22.226$, $DF = 16$, $\chi^2/df = 1.389$, $CFI = 0.984$, $RMSEA = 0.055$, $SRMR = 0.038$.

Rezultatele estimării cu cel de al doilea model confirmă validitatea indexului formativ, conform cu recomandările din literatura de specialitate (Diamantopoulos & Winklhofer, 2001; Jarvis et al, 2003; Diamantopoulos 2008, Franke et al, 2008).

Analiza modelului structural confirmă ipoteza unui efect pozitiv semnificativ $\text{PEOU} \rightarrow \text{PE}$, verificată într-un studiu anterior (Balog & Pribeanu, 2010), efect care denotă o legătură între aspectele pragmatice și cele hedonice ale utilizabilității. De asemenea, studiul a permis identificarea unui index formativ, relevant pentru platforma ARTP. În cazul scenariului de învățare a biologiei acest index este definit de trei măsuri : caracterul confortabil al postului de lucru, ușurința utilizării telecomenzii și ușurința colaborării cu colegii.

Există unele limite inerente acestui studiu, care este primul studiu de caz privind dezvoltarea unui index formativ realizat în țară.

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. Rezultatele obținute confirmă concluziile studiilor anterioare, care sugerau rafinarea modelului de măsurare.

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 identificarea unor 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, problemă care este dificilă în sine așa cum au arătat studiile lui Jarvis et al. (2003), Wilcox et al (2008) și Diamantopoulos (2008).

4. Concluzii și intenții de viitor

Specificarea unui index formativ a fost utilă pentru analiza unor aspecte specifice utilizabilității aplicației care implementează scenariul de biologie pe ARTP. Rezultatele acestui studiu sunt complementare rezultatelor obținute anterior cu o scală de măsurare reflectivă (Balog & Pribeanu, 2010) aplicată la nivelul platformei ARTP. În acest sens, se poate considera că rezultatele sunt încurajatoare pentru specificarea unor modele de măsurare alternative.

Concluzii asemănătoare se desprind și din rezultatele obținute pe cele două eșantioane din 2007, în care au fost utilizate metode cantitative și calitative. În ambele studii, analiza comentariilor studenților a permis atât evidențierea unor probleme de utilizabilitate specifice cât și a unor valențe specifice ARTP și / sau a fiecărei aplicații în parte, care nu pot fi analizate numai cu ajutorul chestionarului de utilizabilitate.

În general, acest studiu de caz confirmă faptul că o abordare bazată pe mai multe metode și triangulare mărește încrederea în rezultate, și permite atât o validare convergentă cât și identificarea unor aspecte care nu pot fi acoperite de o singură metodă sau un singur model, oferind astfel o perspectivă mai largă asupra sistemului studiat.

Se intenționează repetarea studiului pentru scenariul de chimie, pe baza eșantionului din 2008. Rezultatele din ambele studii, coroborate cu analiza rezultatelor din studiile anterioare vor fi utilizate în elaborarea unui nou chestionar de utilizabilitate, pentru evaluarea unei versiuni îmbunătățite a aplicației care implementează scenariul de învățare a chimiei pe ARTP. Acest chestionar va cuprinde atât scale reflective cât și indexuri formative. Noul model de măsurare va fi testat în 2011.

Confirmare

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

Referințe

- 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.
- Arbuckle, J.L. (2007). *AMOS 16.0 User's Guide*. Amos Development Corporation.
- 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.
- Bollen, K., Lennox, R. (1991) Conventional wisdom on measurement : a structural perspective. *Psychological Bulletin* 110 (2), 305-314.
- Cadogan, J., Souchon, A., Procter, D. (2008) The quality of market-oriented behaviors : Formative index construction. *Journal of Business Research* 61, 1263-1277.
- Churchill, G.A. (1979). A Paradigm for Developing better Measures of Marketing Constructs. *Journal of Marketing Research* 16, 64-73.
- Davis, F.D (1989). Perceived usefulness, perceived easy of use, and user acceptance of information technology. *MIS Quaterly* 13, 319-340.
- Diamantopoulos, A., Winklhofer, H. (2001) Index construction with formative indicators : an alternative to scale development. *Journal of Marketing Research* 28, 269-277.
- Diamantopoulos, A. (2006) The error term in formative measurement models : interpretation and modeling implications. *Journal of Modeling in Management* 1(1), 7-17.
- Diamantopoulos, A., Riefler, P., Roth, K. (2008) Advancing formative measurement models. *Journal of Business Research* 61, 1203-1218.
- Edwards, J., Bagozzi, R. (2000) On the nature and direction of of relationship between constructs and measures. *Psychological Methods* 5(2), 155-174.
- Fornell, C., Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18 (1), 39-50.
- Franke, G., Preacher, K., Rigdon, E. (2008) Proportional structural effects of formative indicators. *Journal of Business Research* 61, 1229-1237.
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L. (2006). *Multivariate Data Analysis*. 6th Ed., Prentice Hall, 2006.
- ISO/DIS 9241-11:1994 Information Technology – Ergonomic requirements for office work with visual display terminal (VDTs)- *Guidance on usability*.
- ISO 9126-1:2001 Software Engineering - Software product quality. Part 1: Quality Model.
- Iordache, D.D., Pribeanu, C. (2009) Comparison of Quantitative and Qualitative Data from a Formative Usability Evaluation of an Augmented Reality Learning Scenario. *Informatica Economică Journal*, 13 (3), 67-74.
- 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.
- Law, E. L-C., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J. (2009). Understanding, scoping and defining user experience: A survey approach. *Proceedings of CHI 2009 – User Experience*, 719-728.
- 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.
- Wilcox, J., Howell, R., Breivik, E. (2008) Questions about formative measurement. *Journal of Business Research 61*, 1219-1228.
- Wind, J., Riege, K., Bogen M., 2007. Spinnstube®: A Seated Augmented Reality Display System, *Virtual Environments: Proc. IPT-EGVE – EG/ACM Symposium*, 17-23.

Anexa 1. Chestionarul de utilizabilitate și rezultatele obținute

Cod	Item	M	SD
ERG1	Ajustarea ecranului "see-through" este ușoară	4.07	0.80
ERG2	Ajustarea ochelarilor stereo este ușoară	4.26	0.79
ERG3	Ajustarea căștilor este ușoară	4.24	0.91
ERG4	Postul de lucru este confortabil	4.22	0.91
ERG5	Observarea obiectului real prin ecran este clară	3.76	1.07
PEOU1	Înțelegerea modului de operare cu aplicația realitate îmbogățită este ușoară	4.02	0.86
PEOU2	Suprapunerea dintre proiecție și obiectul real este clară	3.62	1.02
PEOU3	Învățarea modului de lucru cu aplicația de realitate îmbogățită este ușoară	4.04	0.94
PEOU4	Reamintirea modului de lucru cu aplicația de realitate îmbogățită este ușoară	3.95	0.89
PEOU5	Înțelegerea explicațiilor vocale este ușoară	4.10	0.93
PEOU6	Citirea informației pe ecran este ușoară	4.06	1.02
PEOU7	Selectarea unui item din meniu este ușoară	3.95	1.10
PEOU8	Corectarea erorilor este ușoară	3.79	1.04
PEOU9	Colaborarea cu colegii este ușoară	4.00	1.07
PU1	Utilizarea aplicației mă ajută să înțeleg mai rapid lecția	4.14	0.98
PU2	Utilizând aplicația voi obține rezultate mai bune la teste	3.84	1.04
PU3	După utilizarea aplicației voi ști mai multe despre acest subiect	4.05	0.93
PE1	Sistemul face învățarea mai interesantă	4.35	0.87
PE2	Lucrul în grup cu colegii este stimulant	4.06	0.95
PE3	Îmi place să interacționez cu obiecte reale	3.91	1.05
PE4	Efectuarea exercițiilor este captivantă	4.15	1.01
INT1	Aș dori să dispun de acest sistem în școală	4.43	0.97
INT2	Intenționez să utilizez acest sistem pentru învățare	4.05	1.02
INT3	Voi recomanda altor colegi să utilizeze acest sistem	4.19	0.97
PEOU10	În general, apreciez că sistemul este ușor de utilizat	4.14	0.89
PU4	În general, consider că sistemul este util pentru învățare	4.10	0.90
PE5	În general, îmi place să învăț cu acest sistem	4.09	0.97
PE6	În general, apreciez că sistemul este incitant	4.13	0.89