

Analiza a trei instrumente de evaluare a utilizabilității, a satisfacției și a experienței utilizatorului în context românesc

Daniela-Maria Andrei, Adriana-Mihaela Guran

Universitatea Babeș-Bolyai, Facultatea de Psihologie și Științe ale Educației
Str. Republicii, Nr. 37, Cluj-Napoca

E-mail: danielaandrei@psychology.ro, adriana@cs.ubbcluj.ro

Rezumat. Prezenta lucrare își propune să investigheze principalele constructe utilizate în cercetarea percepției utilizatorilor cu privire la calitatea produselor interactive în domeniul interacțiunii om-calculator, și anume, utilizabilitate, satisfacție și experiența utilizatorilor. Pentru atingerea acestui obiectiv a fost proiectată o cercetare vizând verificarea calităților psihometrice ale instrumentelor de măsurare ale acestor constructe. Trei instrumente consacrate în literatura de specialitate au fost traduse și adaptate în limba română – AttrakDiff 2.0., Questionnaire for User Interaction Satisfaction (QUIS) și Software Usability Measurement Inventory (SUMI). Acestea au fost completate de 152 utilizatori ai site-ului web www.psychology.ro. Rezultatele obținute în cadrul studiului pun în evidență indici de fidelitate buni pentru toate cele trei instrumente, însă și deficiențe la nivelul validității de construct a scalelor, deficiențe mai accentuate în cazul utilizabilității sistemelor interactive. Astfel, configurația factorială originală se păstrează cu mici modificări în cazul Attrakdiff 2.0, este ușor modificată în cazul QUIS și nu poate fi pusă în evidență în cazul SUMI. Datele obținute sugerează necesitatea considerării atente a conceptului central măsurat atunci când folosim astfel de instrumente de măsurare a percepțiilor utilizatorilor cu privire la calitatea produselor interactive.

Cuvinte cheie: Utilizabilitate, Satisfacție, Experiența utilizatorului, Validitate de construct

1. Introducere

Ca și disciplină relativ nouă, interacțiunea om-calculator (HCI) are în centrul preocupărilor sale studiul modului în care oamenii proiectează, implementează și utilizează sisteme interactive și a fenomenelor asociate acestora, precum și a modului în care tehnologia interactivă afectează indivizii, organizațiile și societatea (Myers, Hollan, & Cruz, 1996, Hewett, Baecker, Card, Carey, Gasen, Mantei, Perlman, Strong, & Verplank, 1992). Elementul central al acestei discipline este desemnat de conceptul de

interacțiune, cel responsabil de a trasa o punte de legătură între lumea tehnologică (cea a calculatorului) și cea umană (utilizatori). Interacțiunea este astfel procesul care justifică alăturarea disciplinelor care vin dinspre științele informaticii celor care vin dinspre psihologie aplicată (Carroll, 1997; Preece, Rogers, Sharp, Benyon, Holland & Carey, 1994).

În acest cadru de studiu al proceselor de interacțiune, un rol deosebit de important îl ocupă percepțiile oamenilor, sau utilizatorilor, cu privire la sistemele interactive. În mod clasic, aceste percepții au fost cel mai intens studiate în cadrul a două paradigme: cea a utilizabilității (sau satisfacției) și cea a acceptării tehnologiei (Wixom & Todd, 2005). Cele două paradigme sunt relaționate de fapt între ele, de cele mai multe ori modelele clasice de acceptare a tehnologiei postulând o relație liniară între satisfacția utilizatorilor și acceptarea tehnologiei (Davis, 1986). Recent, o altă linie de cercetare puternică a început să se dezvolte în acest domeniu, vizând în principal aspecte legate de întreaga experiență a utilizatorilor cu tehnologia (Hassenzahl, 2008; Hassenzahl, 2004; Hassenzahl, 2002; Draper, 1999; Gaver & Martin, 2000; Jordan, 2000; Monk & Frohlich, 1999) și punând accentul pe aspecte ale experienței care merg mai departe decât simpla rezolvare de sarcini cu ajutorul tehnologiei interactive.

2. Fundamentare teoretică

2.1 Utilizabilitatea sistemelor interactive

Utilizabilitatea sistemelor interactive reprezintă unul dintre subiectele care au provocat un interes susținut în comunitatea științifică specializată în interacțiunea om-calculator. Acest interes major al comunității s-a datorat în principal legăturii puternice postulate între utilizabilitatea tehnologiilor și succesul general al produsului, fie el un obiect, un produs soft sau un site web (Kuniavsky, 2003; Nielsen, 1993; Norman, 2002). Cu toate acestea, istoria acestui concept nu este una liniară în domeniul studiului tehnologiilor interactive. Această istorie debutează cu o perioadă incipientă, în care centrarea principală era pe tehnologie și soluțiile tehnice, pe funcționalitate, mai degrabă decât pe ușurința utilizării (Bannon, 1991). În această primă perioadă termenul de utilizabilitate nu este recunoscut de comunitatea științifică ca având relevanță pentru domeniu și, în consecință, aspectele relaționate cu acesta sunt foarte puțin sau deloc studiate, iar în bugetele de

dezvoltare ale tehnologiilor nu erau prevăzute cheltuieli legate de studiul utilizatorilor (Kuniavsky, 2003; Norman, 2002).

Lucrurile s-au modificat substanțial pe măsură ce soluțiile tehnologice au evoluat și sistemele interactive au început să devină din ce în ce mai accesibile publicului general. Această evoluție a comutat interesul de pe soluțiile tehnice pe problemele legate de ușurința de a utiliza aceste soluții (Bannon, 1991) și pe plusul de valoare asociat acestei ușurințe, care a început să fie conceptualizată ca un avantaj competitiv (Baven, 2008; Kuniavsky, 2003). Aceste condiții au dus la structurarea cercetării în acest domeniu și contribuțiile științifice pe tema utilizabilității au început să devină vizibile. Caracteristica dezvoltării acestui domeniu a constat într-o orientare mai degrabă dinspre tehnologie decât dinspre partea umană (Pitariu, Andrei, & Guran, 2009), cu focalizare pe acele atribute ale tehnologiei care au rolul de a facilita interacțiunea (Bevan, 2008) și devenind mai degrabă un indicator al calității sistemelor (Hartson, Andre, & Williges, 2000) decât al calității interacțiunii.

Această observație este susținută și de analiza principalelor definiții ale conceptului de utilizabilitate, care, în marea lor majoritate, fac referire la caracteristici ale tehnologiei. De exemplu, Nielsen (1993) definește utilizabilitatea în termeni de atribute ale sistemelor (ușor de învățat, eficient, ușor de memorat, cu puține erori, satisfăcător). Dix. et al. (2004) definesc utilizabilitatea prin ușurința de a învăța sistemul, flexibilitate și robustețe, toate aceste trăsături fiind identificate în produs. O privire critică asupra acestor definiții ne arată că, chiar și în această perioadă de maxim interes pentru utilizabilitate, punctul central era reprezentat de tehnologie și nu de utilizatori, oameni (Pitariu et al., 2009). Această caracteristică a abordării conceptului trebuie înțeleasă și în contextul eforturilor de dezvoltare a unor instrumente adecvate pentru măsurarea utilizabilității, deoarece ceea ce înțelegem prin utilizabilitate este în mare parte determinat de modul în care o măsurăm (Hornbaek, 2006). Am arătat deja (Andrei, 2009) faptul că abordarea cel mai des întâlnită în construirea acestor instrumente este una factorială, pornind mai degrabă de la observații ale utilizatorilor și experților, și mai puțin de la eforturi de conceptualizare și rafinare teoretică. Aceste caracteristici particulare au două urmări majore. Prima este mai degrabă de ordin metodologic și afectează validitatea conceptului (Andrei, 2009; Hornbaek, 2006) iar a doua ține de aplicabilitatea datelor obținute în urma evaluărilor, în condițiile în care, într-un mediu foarte dinamic și

complex, succesul unui anumit produs depinde de modul în care acesta reușește să echilibreze o serie de constrângeri multiple venite din diverse surse, nu doar cele dinspre tehnologie sau utilizatori (Kuniavsky, 2003; Norman, 2002).

2.2 De la utilizabilitate la satisfacția utilizatorului

Perspectiva tradiționalistă asupra utilizabilității a început să se modifice odată cu propunerea unei definiții care a devenit extrem de populară. ISO (1998) definește utilizabilitatea ca măsura în care anumiți utilizatori pot utiliza tehnologia cu eficiență, eficacitate și satisfacție, pentru a atinge scopuri specifice. Schimbarea perspectivei este evidentă prin faptul că centrul de interes este comutat dinspre caracteristici ale tehnologiei înspre ceea ce rezultă din utilizarea acesteia. De asemenea, această definiție a însemnat totodată o comutare a interesului dinspre caracteristicile tehnologiei înspre reacția subiectivă rezultată din utilizare, și anume satisfacția utilizatorului.

Prin satisfacția utilizatorului se înțelege, de cele mai multe ori, atitudinea sau reacția emoțională pe care utilizatorul o are față de un sistem interactiv, lipsa disconfortului (ISO, 1998; Wixom & Todd, 2005). Consolidarea liniei de cercetare pe satisfacția utilizatorului nu înseamnă o părăsire totală a domeniului utilizabilității, și acest lucru este vizibil din examinarea modalităților de măsurare a satisfacției. Cele mai multe instrumente de evaluare a satisfacției utilizatorului încă păstrează o orientare pe caracteristicile tehnologiei, (DeLone & McLean, 1992, Wixom & Todd, 2005). În analiza critică realizată asupra literaturii de specialitate, Wixom & Todd (2005) ajung la concluzia că instrumentele existente pentru măsurarea satisfacției utilizatorului reprezintă din punct de vedere conceptual doar câteva dimensiuni supraordonate, și oferă o soluție eficientă pentru identificarea și examinarea structurii caracteristicilor sistemului sau informațiilor. Aceeași concluzie este susținută și de analiza modului de dezvoltare a instrumentelor de satisfacție comparativ cu cele de utilizabilitate (Andrei, 2009), dar și de gradul foarte ridicat de dezacord în ceea ce privește modalitățile de măsurare a conceptului inventariate în literatura de specialitate (Hornbaek, 2006). Aceste argumente ne sugerează că este posibil ca, atunci când vorbim de utilizabilitate și de satisfacție, să nu

vorbim de fapt de două concepte diferite, ci de două perspective diferite ale aceluiași concept.

2.3 De la utilizabilitate și satisfacție, la calitate în utilizare și experiența utilizatorului

Una din soluțiile identificate în literatura de specialitate cu privire la impasul în care măsurarea conceptului de satisfacție a ajuns a constatat în reorientarea interesului dinspre atitudini pozitive și lipsa disconfortului înspre experiența utilizatorului (Hornbaek, 2006). Această recomandare coincide cu consolidarea tot mai puternică a unei noi perspective, și anume cea a calității în utilizare impusă de noul standard ISO (2008) care a contribuit la reorientarea interesului de pe calitățile tehnologiei înspre capacitatea de a satisface nevoile și scopurile utilizatorilor (Bevan, 2008). De asemenea se observă o presiune din partea comunității științifice de a renunța la evaluarea calității sistemelor după ce acestea au fost finalizate, și o preocupare pentru înglobarea cerințelor de calitate în procesul de proiectare. Aceste preocupări au pus în evidență nevoia de a măsura calitatea utilizării luând în considerare și atribute nonfuncționale ale tehnologiei (Law et al., 2008). Prezenta conceptualizare a calității în utilizare susține reorientarea interesului pe procesul de interacțiune dintre oameni, produsul interactiv și munca lor (van Welie, 2001), în alte cuvinte, spre o mai bună înțelegere a experienței utilizatorului (Kuniavsky, 2003).

Acest termen a devenit un termen intens utilizat pentru a desemna o perspectivă holistică asupra calității produselor interactive și o completare a modelelor de calitate tradiționale cu concepte care nu țin direct de utilitatea produselor (Hassenzahl, 2004). El țintește să cuprindă toate aspectele relevante din interacțiunea cu un anumit produs. Contribuția majoră a acestei perspective se concretizează însă în sublinierea nevoii de a trece dincolo de viziunea mult prea simplificată, axată pe funcționalitate și pe sarcini pe care o propune abordarea clasică a utilizabilității, spre o viziune care să înglobeze rolul emoțiilor, afectelor și a modului în care se construiesc experiențele de utilizare (Hassenzahl, 2004; Hassenzahl & Tractinsky, 2006).

Deși nu ignoră rolul atributelor și caracteristicilor tehnologice, această perspectivă reușește să mute accentul dinspre aceste caracteristici înspre nevoile diferite ale utilizatorului în funcție de care aceste caracteristici sunt percepute și înspre consecințe ale utilizării cum ar fi satisfacția, plăcerea,

amuzamentul (Batra & Ahtola, 1990; Huang, 2003, Hassenzahl, 2003). Centrul de interes major în această perspectivă este reprezentat de utilizator, prin recunoașterea deschisă a caracterului subiectiv al experienței utilizatorului. Abordările de până acum tratau caracteristicile tehnologiei ca pe ceva extern utilizatorului, acesta din urmă fiind un receptor al acestor caracteristici și un evaluator al lor prin prisma anumitor criterii. Cercetătorii din experiența utilizatorilor specifică deschis faptul că nu există caracteristici ale tehnologiei, există doar caracteristici percepute. Cu alte cuvinte, utilizatorii construiesc caracterul unui produs, iar la această construcție nu contribuie doar caracteristicile tehnologiei, ci în primul rând variabilele interne, personale, ce țin de individ și variabilele care țin de situație și care pot activa variabilele individuale (Hassenzahl, 2004).

Chiar dacă comunitatea științifică pare să fi răspuns pozitiv la această perspectivă, persistă încă multiple neînțelegeri asupra semnificației conceptului. Puținele modele propuse (Logan, 1994; Jordan, 2000; Hassenzahl, 2003) sunt mai mult sau mai puțin reduționiste și reprezintă doar începutul eforturilor de a cristaliza relațiile conceptuale complexe pe care le presupune.

2.4. Necesitatea realizării studiilor de adaptare a instrumentelor

Deși interesul pentru domeniul interacțiunii om-calculator a crescut considerabil în România, o analiză a cercetării românești în acest domeniu a arătat faptul că preocuparea pentru conceptul de utilizabilitate, dar mai ales pentru cel de satisfacție și experiența utilizatorului este mai puțin evidentă, tematicile legate de design și arhitectura interfețelor fiind prevalente (Andrei, 2009). Mai mult, chiar și în cazul în care utilizabilitatea este tematica de interes a cercetării, foarte rar sunt utilizate instrumente consacrate în literatura de specialitate pentru măsurarea acesteia (ex. SUS). Cel mai adesea se recurge la utilizarea unor scale construite ad-hoc, fără o investigație serioasă a modului de construcție și a calităților psihometrice ale acestora, deși există studii riguroase de elaborare de instrumente de evaluare (Balog & Pribeanu, 2009).

Un alt motiv care susține nevoia validării acestor instrumente este relaționat cu aplicabilitatea interculturală a instrumentelor propuse cel mai adesea într-o cultură anglo-saxonă. Deși interacțiunea cu și accesul la tehnologie a devenit aproape universal, există indicii cu privire la faptul că

experiențele de interacțiune sunt modelate cultural (Orlikowski, 2000), ca urmare o atenție deosebită trebuia acordată calităților instrumentelor folosite într-o cultură diferită în care acestea au fost foarte puțin studiate.

În lumina acestor argumente în studiul de față am adaptat trei instrumente – AttrakDiff 2.0 (Hassenzahl, Burmester, & Koller, 2003), QUIS (Questionnaire for User Interaction Satisfaction) (Chin, Diehl, & Norman, 1988) și SUMI (Software Usability Measurement Inventory) (Porteous, Kirakowski, & Corbett, 1993) – pentru măsurarea experienței utilizatorului, a satisfacției utilizatorului, respectiv a utilizabilității, în limba română, și am testat indicii de fidelitate și validitate de construct ai acestora.

3. Metodă

3.1. Participanți

Cele 3 instrumente, în forma lor finală, au fost reunite într-o broșură unică de evaluare care a fost aplicată unui eșantion de 213 participanți. Dintre aceștia, 152 au completat în întregime protocoalele care au fost ulterior introduse în analiză. Participanții au fost în totalitate studenți la nivel de master sau licență ai Facultății de Psihologie și Științe ale Educației, care au participat în mod voluntar la studiu și care au avut interacțiuni repetate cu produsul interactiv evaluat – site-ul web al catedrei de psihologie (www.psychology.ro). Dintre participanți doar 4,69% au fost studenți la nivel de master iar restul de 95,31% au fost studenți la nivel licență. Vârsta participanților a variat de la 20 la 25 ani cu o medie de vârstă de 21 ani.

3.2. Instrumente

În prezentul studiu s-a realizat procesul de traducere și adaptare pentru trei instrumente raportate în literatura de specialitate, destinate măsurării conceptelor investigate – experiența utilizatorului, satisfacția utilizatorului și utilizabilitate.

3.2.1. Experiența utilizatorului

Pentru măsurarea experienței utilizatorului am ales spre validare instrumentul AttrakDiff 2.0 (Hassenzahl, Burmester, & Koller, 2003) fiind unul dintre puținele instrumente susținut de un model teoretic (Hassenzahl,

2003; Hassenzahl, 2004; Hassenzahl, 2008) și pentru care au fost raportate procedura de construcție precum și date legate de indicii psihometrici și de evoluția diferitelor forme ale instrumentului (Hassenzahl, 2004; Hassenzahl, 2002; Hassenzahl et al., 2000; Kunze, 2001).

Instrumentul este construit pe baza unui diferențiator semantic și evaluează 4 dimensiuni principale. Primele 3 dimensiuni vizează caracterul perceput al tehnologiei – caracterul pragmatic perceput (PQ), caracterul hedonic perceput – stimulare (HQS) și caracterul hedonic perceput – identificare (HQI). Autorii indică faptul că scorul variabilei PQ indică de fapt utilizabilitatea percepută a produsului (Hassenzahl, 2004). Scorul HQI indică percepții cu privire la capacitatea tehnologiei de a comunica o anumită identitate celorlalți, iar scorul HQS indică nivelul perceput de noutate, provocare și stimulare. Atributele HQI sunt în primul rând de natură socială, iar cele ale HQS sunt în principal relaționate cu dezvoltarea personală. Ultima variabilă, cea de-a patra, măsoară evaluările generale realizate asupra produsului tehnologic (pe două subcomponente: bun și atractiv). Instrumentul totalizează 28 de itemi, câte 7 pentru fiecare dimensiune măsurată, construiți pe modelul diferențiatorului semantic și folosind o scală likert cu 7 trepte între cele două adjective ale fiecărui item.

Autorii instrumentului raportează indici de consistență internă alpha Cronbach buni pentru toate variabilele instrumentului (HQI, $\alpha=0,85$; HQS, $\alpha=0,95$; PQ, $\alpha=0,90$) și de asemenea, raportează validarea modelului conceptual care stă la baza instrumentului cu ajutorul măsurătorilor efectuate (Hassenzahl, 2004; Hassenzahl, 2008).

3.2.2. Utilizabilitate

Pentru măsurarea utilizabilității s-a recurs la traducerea și adaptarea instrumentului SUMI (HFRG 1994, MUSiC project). Motivul principal al alegerii acestui instrument l-a constituit faptul că la baza construcției sale s-a aflat tocmai un efort de a separa aspectele care țin de experiența emoțională relaționată cu utilizarea de celelalte aspecte care erau văzute a fi relaționate cu competența (Kirakowski, 1994). Analiza instrumentelor existente până în acel moment în domeniu arătase faptul că, dincolo de operaționalizările date conceptului de utilizabilitate, toate păreau să convergă spre un factor general care ținea de afect, de reactivitatea emoțională la utilizarea tehnologiei (satisfacție). Centrați pe mutarea interesului dinspre satisfacție cu utilizarea spre calitatea softului percepută

de utilizatori, autorii au început printr-o analiză atentă a instrumentelor dezvoltate până atunci.

Una din conceptualizările inițiale ale acestei dimensiuni a calității softului percepută de utilizatori (Dzida, Herda & Itzfeldt, 1978) avea la bază presupunerea că fiecare dimensiune este definită printr-un set de proprietăți ale sistemului, fiecare din aceste caracteristici fiind asociate în principal cu o anumită dimensiune. Ulterior, istoricul construcției instrumentelor de evaluare a încercat să se îndepărteze puțin de această perspectivă centrată pe tehnologie și a pus accent pe satisfacția utilizatorilor cu privire la anumite elemente ale tehnologiei.

Considerând ambele abordări insuficiente, prima deoarece punea un accent prea puternic pe caracteristicile sistemului, iar a doua deoarece punea accent doar pe reacția utilizatorilor, fără să ofere informații valoroase asupra modalităților de intervenție în tehnologie, cercetătorii din domeniu au încercat să separe elementul atitudinal (satisfacția) de elementul funcțional al tehnologiei (competența) (Kirakowski, 1987, Kirakowski & Corbett, 1988). Dezvoltarea ulterioară a SUMI a venit să completeze aceste demersuri, focalizându-se pe elementul de competență a tehnologiei, propunându-și să o dezvolte și să examineze posibilă structură a acestuia. Ca urmare, demersurile inițiale ale acestuia au fost centrate din nou pe aspectele care țin de gradul în care utilizatorii se simt susținuți de către tehnologie (competența).

Construirea instrumentului a fost realizată prin metoda factorială, pornind de la un număr mai ridicat de itemi colectați din discuții cu utilizatori, din instrumentele anterioare sau din recomandările experților HCI. Prin analize factoriale repetate ale acestor itemi s-au obținut 5 factori: *Eficiență* (utilizatorii simt că sistemul îi asista în sarcinile pe care le au de îndeplinit), *Afect* (reacția emoțională generală la tehnologie), *Ajutor* (măsura în care sistemul se poate auto-explica, și adecvarea soluțiilor de ajutor), *Control* (măsura în care utilizatorul se simte în control atunci când utilizează tehnologia și nu se simte controlat) și *Învățare* (viteza și ușurința cu care utilizatorul simte că a ajuns să stăpânească sistemul).

Ulterior, scala rezultată a fost testată, însă majoritatea testărilor au implicat evaluarea validității relative la criteriu. Altfel spus, majoritatea studiilor de validare au fost centrate pe demonstrarea utilității scalei în a discrimina între două produse tehnologice diferite precum și pe abilitatea subscalelor de a se comporta diferit în relație unele cu altele, astfel încât să

poată identifica puncte tari și puncte slabe ale scalei. Deși nu sunt raportate studii asupra relaționării teoretice a celor 5 factori, se poate extrage ideea conform căreia aceste dimensiuni se comportă relativ independent unele de altele. Mai mult, interesul autorilor scalei nu a evoluat înspre rafinarea conceptuală a utilizabilității, ci tocmai înspre o analiză individualizată pe itemi a chestionarului care să susțină o aplicabilitate practică mult mai ridicată a scorurilor obținute și să poată susține cât mai eficient intervenția în re-proiectare (Kirakowski, 1994).

3.2.3. Satisfacția utilizatorului

Pentru evaluarea satisfacției utilizatorului am recurs la traducerea și adaptarea instrumentului QUIS (Questionnaire for User Interaction Satisfaction) dezvoltat de o echipă multidisciplinară de la Universitatea Maryland, College Park (Chin, Diehl, & Norman, 1988). Acest instrument a fost dezvoltat pentru a măsura satisfacția subiectivă a utilizatorilor față de aspecte specifice ale interfeței. Chestionarul a fost proiectat în așa fel încât să poată fi configurat în funcție de produsul tehnologic evaluat, prin includerea doar a acelor secțiuni care se aplică tehnologiei și oferă informații relevante utilizatorilor.

La baza alegerii acestui instrument au stat eforturile echipei de cercetare de a depăși deficiențele semnalate în literatura de specialitate cu privire la fidelitatea, validitatea și standardizarea instrumentelor de evaluare a interfețelor om-calculator (Ives, Olson, & Baroudi, 1983; Hornbaek, 2006). În literatura de specialitate sunt raportate date care susțin o fidelitate bună a chestionarului în condițiile evaluării unei varietăți ridicate de interfețe (Harper & Norman, 1993).

În studiul nostru am recurs la traducerea și adaptarea formei scurte a QUIS 5., care conține o evaluare a satisfacției generale cu tehnologia pe 6 dimensiuni, urmată de măsurători organizate ierarhic pentru 4 factori specifici ai interfeței (ecran, terminologie și informația oferită de sistem, învățare și capacitățile sistemului). Itemii aparținând fiecărei dintre aceste dimensiuni sunt construiți în forma unui diferențiator semantic, cele două adjective (aspect negativ – aspect pozitiv) fiind măsurate printr-o scală likert cu 7 trepte. Pentru studiul de față, forma oficială de 27 itemi a fost redusă la o formă cu 23 itemi prin eliminarea a 4 itemi care nu erau adecvat tipului de tehnologie evaluată – site web (Tullis & Stetson, 2004)

Procesul de traducere și adaptare

În acord cu recomandările și standardele în domeniu, fiecare instrument a fost tradus independent de către 3 persoane, dintre care două persoane cu experiență în domeniul evaluării psihometrice și al interacțiunii om calculator și o persoană specializată în traduceri. Ulterior, s-a realizat o ședință comună în cadrul căreia variantele de traducere au fost revăzute, discutate, adaptate cultural și s-au luat deciziile privind formele finale ale itemilor. S-au realizat câteva reformulări ale itemilor atât pentru a obține o mai bună adaptare a acestora la populația românească, cât și pentru a adecva instrumentul la tipul de tehnologie interactivă ce urma să fie evaluată. Forma finală a fost supusă unei retroversiuni în limba engleză pentru a putea evalua gradul de apropiere a celor două forme în limba engleză.

Această formă a instrumentelor a fost pretestată pe un eșantion de 10 participanți, completarea fiind însoțită de interviu. În cadrul acestui interviu s-au urmărit reacțiile generale ale participanților față de instrument, identificarea itemilor problematici, care generează dificultăți de înțelegere și răspuns.

În urma acestei pretestări s-au operat mici modificări de exprimare a itemilor și s-a ajuns la forma finală a celor 3 instrumente.

3.3. Procedură

Cele 3 scale au fost completate în formă creion-hârtie. Instructajul oferit la începutul aplicării a subliniat scopul de cercetare a datelor colectate, asigurarea anonimatului participanților și importanța rezultatelor evaluării pentru îmbunătățirea unui serviciu destinat în primul rând studenților. Participanții au completat chestionarele fără limită de timp, iar la final au depus chestionarele completate într-o cutie destinată colectării acestora poziționată la ieșirea din clădirea facultății.

4. Rezultate

4.1. Evaluarea experienței utilizatorului

Statistici descriptive

O primă fază în analiza instrumentului a constituit-o analiza statisticilor

descriptive. Pe baza acestora putem observa faptul că cele patru dimensiuni ale chestionarului reprezintă valori apropiate de media scalelor, cu abateri standard apropiate de 1, iar indicii de oblicitate și boltire sugerează existența unor abateri moderate de la normalitate.

Tabelul 1. Statistici descriptive AttrakDiff 2.0.

	Minimum	Maximum	Medie	Abatere standard	Skewness	Kurtosis
PQ	1.29	6.43	4.41	1.05	-0.67	0.14
HQI	1.00	6.86	4.46	1.08	-0.55	0.72
HQS	1.00	6.14	3.64	1.12	-0.18	-0.48
ATT	1.00	6.86	4.49	1.10	-0.48	0.55

Din evaluarea statisticilor descriptive se poate observa faptul că site-ul www.psychology.ro este perceput a fi mai degrabă un site cu caracteristici hedonice, de identificare, atributele hedonice – identificare prezentând media cea mai ridicată. Acest rezultat era oarecum de așteptat, având în vedere faptul că este vorba despre un site de prezentare al Catedrei de psihologie, cu care studenții se identifică puternic pe perioada în care studiază în cadrul facultății (Boroș, 2007)

Fidelitatea Attrakdiff 2.0

Un al doilea indicator psihometric evaluat a fost reprezentat de fidelitatea scalelor. Pentru aceasta, am folosit indicatorul de consistență internă a scalelor, alpha Cronbach, indicatorul folosit și de autorii scalei originale. Rezultatele obținute sunt prezentate în Tabelul 2., alături de rezultatele obținute de autorii scalei originale.

Din analiza rezultatelor legate de fidelitate se poate observa faptul că instrumentul tradus și adaptat prezintă indici de validitate buni, apropiați de valorile obținute pentru alte versiuni ale scalelor în limba germană.

Analiza factorială exploratorie

Pentru a verifica configurația factorilor pentru versiunea tradusă și adaptată a scalei, am recurs la o analiză factorială exploratorie, având în vedere că modelul teoretic care stă la baza instrumentului nu a fost încă validat pe populație românească, dar și faptul că în această primă fază am fost interesați de verificarea și rafinarea instrumentului în cazul culturii noastre.

Byrne (2010) susține faptul că procedurile de analiză factorială confirmatorie (CFA) pentru validarea unui instrument sunt indicate atunci când instrumentele investigate au atins un anumit grad de maturitate, și când există date suficiente cu privire la structura lor factorială. Instrumentele cu care lucrăm noi nu se regăsesc în aceste categorii. AttrackDiff 2.0 este un instrument foarte recent propus și există relativ puține date legate de validitatea acestuia. QUIS și SUMI sunt instrumente deja consacrate pe piața domeniului, însă pentru acestea există extrem de puține date publicate în literatură cu privire la validitatea de construct și configurația lor factorială. Deoarece în cazul acestor situații, literatura avertizează asupra utilizării inadecvate a CFA (Byrne, 2010), am ales să investigăm în primul rând datele analizei factoriale exploratorii pentru toate cele 3 instrumente adaptate.

Tabelul 2. Indici de fidelitate pentru scala AttrackDiff.

Dimensiunile Attrackdiff 2.0	Alpha Cronbach eșantion Rom (N=152)	Alpha Cronbach (Hassenzahl, Burmester & Koller, 2001)	Alpha Cronbach (Hassenzahl, Kekez & Burmester, 2002) (N=46)	Alpha Cronbach (Hassenzahl, 2004) (N=33)
PQ	0,79	0,83-0,85	0,89	0,90
HQI	0,83	0,73-0,83	0,88	0,85
HQs	0,85	0,76-0,90	0,88	0,95
ATT	0,88	-	0,91	-

În cazul AttrackDiff 2.0, metoda de extragere a factorilor a fost cea a componentelor principale. Înainte de inspecția rezultatelor analizei factoriale, am analizat câțiva indici legați de adecvarea datelor la această analiză. Astfel, eșantionul nostru de date s-a dovedit a avea un indice excelent al gradului de adecvare la analiza factorială ($KMO= 0,91$) (Field, 2000). De asemenea, determinantul matricei de corelații obținut (0,0018) este mai mare decât valoarea prag de 0,00001, ceea ce indică un risc scăzut pentru existența unei situații de multicolaritate, iar testul de sfericitate Bartlett este semnificativ, indicând existența unor corelații între variabilele măsurate. Toate aceste teste indică faptul că datele obținute sunt adecvate analizei factoriale.

În pasul următor au fost analizate datele obținute în urma analizei factoriale. În urma analizei s-au extras 4 factori ($Eigenvalue > 1$) care explică

împreună 60,54% din varianța totală. Deoarece autorii inițiali ai instrumentului afirmă faptul că cele 3 calități reprezintă variabile independente, necorelate (Hassenzahl, 2004; Hassenzahl et al., 2003), am optat pentru o rotire Varimax cu normalizare Kaiser. Pentru a testa această decizie am realizat și o analiză factorială exploratorie limitând factorii la 3 și utilizând o rotație Direct Oblimin. Analizând corelațiile interfactoriale, nu am identificat corelații mai mari de 0,32 (Fiddell, 2007) ceea ce susține decizia alegerii unei rotații ortogonale. Rezultatele obținute au diferit de cele obținute pentru scala originală, și pot fi analizate în Tabelul 3.

Tabelul 3. Rezultatele analizei factoriale exploratorii intermediare pentru scala AttrakDiff.

Item	Factor 1	Factor 2	Factor 3	Factor 4	Comunalitate
1				0.907	0.837
2		0.618			0.581
4		0.477	0.334		0.367
5	0.723				0.679
6	0.437	0.305	0.350		0.412
8	0.634				0.596
10	0.712				0.617
11			0.720		0.613
12	0.740				0.604
13			0.759		0.666
14		0.435	0.435		0.538
15		0.516			0.547
16			0.727		0.720
18		0.559	0.580		0.724
20	0.692				0.642
22		0.630			0.581
23		0.678			0.558
24			0.634		0.694
25		0.550	0.534		0.601
27		0.757			0.601
28	0.738				0.635
Eigenvalue	8,27	2,25	1,17	1,00	
%Varianță	39,4	10,73	5,60	4,80	

Ceea ce se poate observa este faptul că există câțiva itemi care prezintă probleme datorită comunalității scăzute și/sau datorită unui grad de saturație

similar în doi sau mai mulți factori. Pentru a putea interpreta factorii obținuți, acești itemi au fost înlăturați din analiză.

Astfel, primul factor apare saturat în itemi care evaluează calitățile pragmatice ale tehnologiei (PQ) și este și factorul care explică cea mai mare parte din varianța totală (39,40%). Cel de-al doilea factor este saturat în itemi care evaluează calitatea hedonică stimulare (HQS) și explică 10,73 % din varianța totală. Al treilea factor este saturat în itemi care evaluează identificarea ca și calitate hedonică (HQI) și explică 5,60 % din varianța totală. Spre deosebire de scala originală, în cazul varianței în limba română itemul nr 1 nu a încărcat scala de calitate pragmatică (PQ) și a constituit un factor specific care explică singur un procent din varianța aproape similar cu cel al scalei de identificare (4,80%). Acest item ține de calitatea percepută Uman vs. Tehnic, care se pare că în eșantionul nostru nu este relaționată cu calitățile pragmatice, dar nici cu cele hedonice de stimulare sau identificare. Problematika acestui item va trebui verificată ulterior, ea putând duce la construirea unei alte dimensiuni ale caracterului perceput al tehnologiei în construirea experienței utilizatorului. De asemenea, datele noastre arată că, pentru eșantionul nostru, și pentru forma în limba română a scalei, există probleme cu itemii 4, 6, 14, 18 și 25 care sunt aproximativ la fel de saturați în doi sau 3 factori. La o analiză atentă putem vedea că aceștia țin de câteva aspecte. Astfel, aceștia țin de caracterul inventiv, creativ, al site-ului, și par a încărcă atât pe factorul de identificare cât și pe factorul de stimulare, deși în cadrul modelului aceștia sunt destinați să măsoare factorul de identificare. Același lucru se întâmplă și cu itemul 14 care vizează aspecte de relaționare cu ceilalți prin intermediul tehnologiei, însă în cazul nostru este saturat și în factorul de stimulare. Itemul 6, deși destinat să măsoare aspecte legate de identificare, în cazul nostru este saturat în toți factorii, și cu precădere în cel care ține de evaluarea caracterului pragmatic. Se pare că pentru eșantionul nostru, un produs profesionist înseamnă în primul rând un produs care funcționează bine, și abia apoi un produs cu care doresc să mă identific sau care mă stimulează tocmai datorită calității sale.

În concluzie, pentru acuratețea analizelor în stadiul 2, am ales ca pentru studiul de față să renunțăm la acei itemi care s-au dovedit problematici și să păstrăm doar itemii bine saturați în factorii obținuți. Într-o fază ulterioară de perfecționare a instrumentului în limba română va fi necesară reconsiderare acestor itemi și conducerea unor analize calitative de adâncime pentru a extrage acele atribute care pot cel mai acurat măsura fiecare dintre caractere.

O nouă analiză factorială, desfășurată doar pe itemii păstrați, a confirmat menținerea factorilor și a saturației itemilor prezentați în acești factori (Tabelul 4).

Tabelul 4. Rezultatele analizei factoriale exploratorii intermediare pentru scala AttrakDiff.

Itemi	Factor 1	Factor 2	Factor 3	Comunalitate
2		0.631		0.563
5	0.748			0.586
8	0.580			0.572
10	0.709			0.628
11			0.735	0.622
12	0.734			0.599
13			0.785	0.693
15		0.557		0.521
16			0.756	0.720
20	0.659			0.622
22		0.657		0.514
23		0.707		0.603
24			0.622	0.659
27		0.768		0.633
28	0.744			0.648
Eigenvalue	6.12	1.99	1,06	
%Varianță	40.83	13.28	7,11	

Scalele inventarului QUIS

Ca și în cazul precedent, am pornit analiza de la investigarea statisticilor descriptive pentru toate scalele QUIS.

Din analiza acestor date se poate observa că cele 5 dimensiuni ale chestionarului oscilează în jurul mediei, iar abaterea standard sunt apropiate de 1. Indicii de oblicitate și boltire sugerează existența unor distribuții normale. Analiza mediilor ne indică faptul că evaluările generale pentru site-ul www.psychology.ro se situează în jurul mediei, aspectele percepute mai pozitiv fiind cele relaționate de ușurința învățării de a opera cu site-ul ($m=5,06$, $as=1,18$) iar cele percepute a fi mai negative fiind cele legate de structura paginii ($m=3,72$, $as=0,91$). Observăm că această dimensiune este și cea care prezintă abaterea standard cea mai mică, ceea ce indică un grad ridicat de acord cu privire la aceste aspecte în eșantionul investigat.

Tabelul 5. Analize descriptive pentru scala de satisfacție QUIS

	N	Medie	Abatere Standard	Skewness	Kurtosis
General	152	4.33	1.22	-.350	-.145
Ecran	152	3.72	.91	.720	.057
Terminologie	152	4.48	1.12	-.361	.021
Învățare	152	5.06	1.18	-.409	-.510
Capacități	152	4.82	1.18	-.483	.145

Al doilea tip de indici investigați au fost indicii de consistență internă alpha Cronbach.

Tabelul 6. Indicii de consistență internă alpha Cronbach.

Dimensiunile QUIS 5	Alpha Cronbach eșantion Rom (N=152)	Alpha Cronbach - Chin et al., 1988 (N=150)
General	0,87	0,93
Ecran	0,78	0,933-0,939
Terminologie	0,78	
Învățare	0,81	
Capacități	0,80	

Rezultatele obținute indică o bună consistență internă pentru toate subscalele instrumentului analizat. Valorile obținute de noi sunt asemănătoare valorilor obținute în studiile de dezvoltare a instrumentului.

Analiza factorială exploratorie

Pentru a verifica configurația factorilor pentru versiunea tradusă și adaptată a scalei, am recurs la o analiză factorială exploratorie. Acest tip de analiză se impune mai ales datorită faptului că autorii scalei oficiale în limba engleză nu indică un model conceptual clar care a stat la baza construcției chestionarului și raportează date foarte puține legate de configurația factorială a scalei (Chin et al., 1988). Mai mult, în această primă fază a studiului eram interesați de configurația factorilor pe cultura noastră și rafinarea adaptării instrumentului.

Pentru extragerea factorilor am folosit metoda componentelor principale. Înainte de a trece la interpretarea rezultatelor obținute, am

analizat câțiva indici legați de adecvarea datelor la această analiză. Astfel, eșantionul nostru de date s-a dovedit a avea un indice excelent al gradului de adecvare la analiza factorială ($KMO= 0,90$) (Field, 2000). De asemenea, determinantul matricei de corelații obținut ($0,0006$) este mai mare decât valoarea prag de $0,00001$, ceea ce indică un risc scăzut pentru existența unei situații de multicoliniaritate, iar testul de sfericitate Bartlett este semnificativ, indicând existența unor corelații între variabilele măsurate. Toate aceste teste indică faptul că datele obținute sunt adecvate analizei factoriale.

Tabelul 7. Rezultatele analizei factoriale exploratorii pentru instrumentul QUIS

itemi	Factor 1	Factor 2	Factor 3	Factor 4	Comunalitate
quis6		0.601			0.572
quis7	0.749				0.700
quis8	0.746				0.689
quis9	0.434		0.434		0.550
quis10			0.738		0.601
quis11	0.698				0.654
quis12	0.580			0.423	0.656
quis13	0.498			0.572	0.623
quis14				0.798	0.702
quis15			0.614		0.638
quis16			0.777		0.719
quis17	0.537		0.534		0.610
quis18	0.620				0.604
quis19	0.775				0.701
quis20		0.792			0.718
quis21		0.744			0.731
quis22		0.627			0.589
quis23	0.500	0.460			0.562
Eigenvalue	7.91	1.43	1.21	1.05	
%Varianță	43.97	7.98	6.76	5.83	

În pasul următor au fost analizate datele obținute în urma analizei factoriale. În urma analizei s-au extras 4 factori ($Eigenvalue > 1$) care explică împreună 64,55% din varianța totală. Autorii instrumentului original nu specifică relația dintre subscalele instrumentului, însă având în vedere faptul că ele doresc să evalueze calități diferite ale tehnologiei, am pornit de la

presupunerea că acestea nu vor fi intercorelate. Ca urmare am optat pentru o rotire Varimax cu normalizare Kaiser. Rezultatele obținute diferă de prescripțiile oficiale ale scalei și pot fi analizate în tabelul 3. Din analiza factorială au fost eliminați primii 5 itemi deoarece aceștia vizau evaluări cu un grad diferit de generalitate asupra site-ului. Deși studiile asupra construcției scalei nu fac referire la statutul acestor variabile, am optat pentru o excludere a acestora, similar cu cazul variabilei care țin de judecarea atractivității site-ului în ceea ce privește AttrakDiff 2.0 (Hassenzahl, 2004; Hassenzahl et al., 2003).

Ceea ce se poate observa este faptul că există câțiva itemi care prezintă probleme datorită unui grad de saturație similar în doi factori. Pentru a putea interpreta factorii obținuți, acești itemi au fost înlăturați din analiză.

Astfel, primul factor apare saturat în itemi care țin de **structura și conținutul paginii** web (succesiunea paginilor, organizarea informațiilor, calitatea informațiilor suplimentare, poziția mesajelor pe pagină, mesaje de ajutor) și este și factorul care explică cea mai mare parte din varianța totală (43,97%). Cel de-al doilea factor este saturat în itemi care evaluează **capacitățile/caracteristicile siteului**, însă aici vedem că intră și itemul legat de ușurința de citire a literelor, probabil relaționată cu celelalte calități ale site-ului. Acest factor explică 7,98% din varianța totală a datelor. Al treilea factor adună itemi care măsoară **ușurința cu care putem învăța** operarea site-ului și include și un item legat de terminologia intuitivă. Este foarte probabil ca o terminologie intuitivă să faciliteze învățarea operării, ceea ce justifică plasarea itemului aici. Acest al treilea factor explică 6,76% din varianța totală. Ultimul factor prezintă un singur item clar, și anume un item legat de mesajele de eroare. Este posibil ca acest item să fi constituit un factor separat datorită faptului că mesajele de eroare caracteristice doar site-ului să fie foarte rare. Cel mai frecvent întâlnite mesaje de eroare sunt mesajele de eroare standard. Acest lucru ar putea explica de ce acest item nu este influențat de ceilalți factori caracteristici percepției site-ului. Rezultatele obținute de noi sunt similare cu ceea ce a fost raportat de către creatorii ai scalei care observă și ei că itemii legați de mesajele de eroare nu încarcă factorul pe care erau desemnați să încarce (Chin et al., 1988). Acest ultim factor explică 5,83% din varianță dar, datorită suspiciunii legate de faptul că ar putea măsura aspecte necaracteristice site-ului în sine vom renunța la el pentru analizele următoare.

Observăm astfel că dintre cele 4 dimensiuni de evaluare a satisfacției doar 3 se pot menține în urma analizei factoriale exploratorii. Itemii care țineau de terminologia și informația dată de site prezintă încărcături pe mai mulți factori, iar alții încarcă complet unul dintre cei 3 factori exemplificați. Posibile argumente pentru această situație sunt relaționate de faptul că aspectele vizate de anumiți itemi nu au acoperire în structura și funcționalitățile site-ului, ca urmare studenții i-ar fi putut interpreta foarte diferit. Din itemii acestui grupaj care au acoperire în funcționalitățile site-ului, cum este cel referitor la terminologia folosită, par a fi mai degrabă relaționați cu ușurința învățării de a opera cu site-ul. Ca urmare, nu se justifică menținerea acestei dimensiuni.

Itemii păstrați în urma analizei factoriale exploratorii inițiale au fost introduși din nou într-o analiză factorială exploratorie, iar configurația factorilor și încărcăturile observate s-au păstrat.

Tabelul 8. Rezultatele analizei factoriale exploratorii finale pentru instrumentul QUIS.

itemi	Factor 1	Factor 2	Factor 3	Comunalitate
quis6		0.597		0.497
quis7	0.729			0.687
quis8	0.734			0.688
quis10			0.734	0.581
quis11	0.719			0.660
quis15			0.634	0.628
quis16			0.816	0.754
quis18	0.725			0.601
quis19	0.824			0.724
quis20		0.832		0.701
quis21		0.777		0.736
quis22		0.615		0.580
Eigenvalue	5.52	1.26	1.04	
%Varianță	46.00	10.57	8.72	

Scalele inventarului SUMI

Analiza datelor descriptive pentru SUMI

Tabelul 9 ne indică faptul că toate cele 5 dimensiuni ale scalei se află situate în jurul mediilor raportate de autorii inițiali (Porteous, Kirakowski, & Corbett, 1993). De asemenea, indicii de oblicitate și boltire indică abateri

moderate de la distribuția normală. Din analiza acestor date nu putem discrimina între cele cinci dimensiuni ale utilizabilității care stau la baza chestionarului, toate fiind situate în jurul mediei din acest punct de vedere putem considera că, în cazul acestei evaluări acestea nu se comportă diferențiat, lucru care îngreunează utilizarea lor în practică.

Tabelul 9. Statistici descriptive pentru instrumentul SUMI.

	N	Min	Max	Medie	Abatere Standard	Skewness	Kurtosis
eficacitate	152	10	30	21.22	4.87	-0.27	-0.61
ajutor	151	10	29	22.03	4.25	-0.68	-0.07
afect	152	10	30	21.95	5.17	-0.59	-0.28
control	152	10	29	21.53	4.03	-0.18	-0.34
învățare	152	11	30	22.91	4.15	-0.43	-0.32
general	152	26	72	53.77	10.86	-0.39	-0.53

Rezultatele obținute indică o bună consistență internă pentru toate subscalele instrumentului analizat. Valorile obținute de noi sunt asemănătoare valorilor obținute în studiile de dezvoltare a instrumentului.

Tabel 10. Fidelitatea scalelor SUMI

Dimensiunile SUMI	Alpha Cronbach eșantion Rom (N=152)	Alpha Cronbach Kirakowski, 1994 (N=143)	Alpha Cronbach Kirakowski, 1994 (N=1100)
eficacitate	0,81	0,77	0,81
ajutor	0,87	0,80	0,83
afect	0,77	0,80	0,85
control	0,70	0,65	0,71
învățare	0,83	0,77	0,82
general	0,91	0,90	0,92

Analiza factorială

Pentru chestionarul de utilizabilitate SUMI, indicii de adecvare a datelor pentru analiza factorială au arătat că aceste date, în structura lor actuală, nu sunt adecvate pentru analiza factorială datorită multicoliniarității variabilelor manifeste. Chiar și în urma identificării celor mai mari indici de corelație între anumiți itemi și eliminarea acestora din analiza factorială, determinantul încă se menține la o valoare cu mult mai mică față de valoarea de reper. Realizarea analizei factoriale exploratorii chiar și în aceste condiții a dus la extragerea unei structuri factoriale imposibil de

interpretat, cu un factor general care explică majoritatea varianței datelor, și alți 11 factori pe care încarcă un singur item sau maxim doi. Aceasta structură factorială nu se apropie de cea sugerată în contribuțiile empirice care susțin dezvoltarea testului (Kelly, 1994; Kirakowski, 1994). Această situație ridică semne de întrebare cu privire la acuratețea analizelor factoriale invocate în aceste contribuții, însă niciodată prezentate în detaliu, cu toate datele statistice aferente. În forma actuală, instrumentul SUMI pare să prezinte serioase deficiențe în ceea ce privește validitatea de construct. Prin comparație cu celelalte două instrumente, acesta din urmă prezintă cele mai serioase probleme în acest sens, subscalele raportate nefiind regăsite nici măcar parțial în structura datelor. De asemenea, s-a încercat și o analiză factorială confirmatorie, forțând datele să încarce factorii pe care erau destinați să încarce, însă indicatorii obținuți au pus în evidență un grad foarte redus de potrivire a datelor cu soluția în 5 factori ($\chi^2=1120,2$, $df=623$ ($p=,000$); $CFI=0,782$; $GFI=0,702$; $RMSEA=0,073$), chiar și atunci când sunt incluse în model unele corelații la nivelul erorilor. Mai mult, analiza indicilor de modificare nu pun în evidență soluții care să fie susținute și teoretic pentru îmbunătățirea modelului factorial.

Corelații inter-scale

Un ultim aspect verificat a fost cel al corelațiilor inter-scale pentru a observa modelul în care se structurează relațiile între diferitele aspecte ale percepțiilor utilizatorilor.

Rezultatele obținute pun în evidență încă o dată deficiențele la nivelul validității de construct prin coeficienți reduși de validitate discriminativă. Din nou, deși în literatură se raportează corelații reduse între scale, datele noastre arată corelații pozitive medii constante între majoritatea subscalelor acestor instrumente. Acești coeficienți ridică, la rândul lor, semne de întrebare asupra modului în care structura datelor se suprapune peste structura teoretică invocată a sta la baza construirii măsurătorilor.

5. Concluzii

În studiul de față am fost interesați în primul rând de dezvoltarea și testarea versiunii în limba română a unora dintre cele mai vehiculate scale în literatura de specialitate pentru măsurarea a trei concepte esențiale în studiul percepțiilor utilizatorilor cu privire la produsele interactive.

În ceea ce privește măsurarea experienței utilizatorului, versiunea tradusă a AttrakDiff 2.0 se dovedește a fi consistentă intern, iar în privința validității de construct, cei trei factori principali ai atributelor tehnologiei au putut fi puși în evidență. Se solicită însă o atenție sporită acordată anumitor itemi care, în forma actuală a scalei par să încarce pe mai mulți factori. De asemenea, trebuie considerat aspectul uman vs. tehnic în cultura noastră deoarece rezultatele obținute pun în evidență în mod sistematic un alt factor latent, în afara celor trei specificați în instrument, care ar putea să stea la baza acestui atribut perceput. De asemenea, contrar datelor raportate în literatură care susțin corelații reduse mai ales între atributele pragmatice și cele hedonice, în cazul nostru au existat corelații semnificative între scale. Ca urmare trebuie investigată mai departe structura relațiilor între aceste scale pentru a obține o imagine adecvată a comportamentului diferitelor tipuri de atribute. Deși autorii inițiali susțin că anumiți coeficienți de corelație situații în jurul valorii de 0,3 și 0,4 sunt de așteptat și nu constituie motive pentru a pune la îndoială independența celor trei concepte (Hassenzahl, 2008), aceste aspecte merită investigații ulterioare.

Tabloul 11. Intercorelații între scalele celor 3 instrumente (toate corelațiile sunt semnificative la $p < 0.01$).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
PQ	1.00													
HQS	0.37	1.00												
HQI	0.45	0.55	1.00											
quies_gen	0.60	0.58	0.60	1.00										
ecran	-0.13	-0.33	-0.18	-0.35	1.00									
Quis_ter	0.46	0.53	0.56	0.66	-0.21	1.00								
quis_inv	0.55	0.46	0.52	0.74	-0.28	0.71	1.00							
Quis_cap	0.50	0.43	0.54	0.63	-0.17	0.67	0.60	1.00						
eficacit	0.49	0.53	0.51	0.69	-0.39	0.58	0.67	0.61	1.00					
afect	0.49	0.60	0.59	0.76	-0.37	0.62	0.66	0.57	0.81	1.00				
ajutor	0.45	0.55	0.52	0.66	-0.41	0.65	0.67	0.51	0.80	0.85	1.00			
control	0.42	0.43	0.43	0.58	-0.23	0.57	0.56	0.62	0.80	0.71	0.71	1.00		
invatare	0.41	0.25	0.35	0.53	-0.21	0.36	0.52	0.51	0.69	0.61	0.55	0.67	1.00	
general	0.49	0.58	0.55	0.70	-0.37	0.63	0.66	0.57	0.91	0.91	0.91	0.82	0.68	1.00

Rezultatele obținute pentru instrumentul de măsurare a satisfacției utilizatorilor pun în evidență o bună consistență internă a scalelor, însă probleme semnificative de validitate de construct a acestora. Structura factorială pusă în evidență nu se potrivește cu cea descrisă de propunătorii

scalei. Mai mult, chiar și aceștia raportează probleme legate de configurația factorială a scalei, factorii presupuși a constitui scala nefiind regăsiți în încărcăturile factoriale ale scalei (Chin et al., 1988). Mai mult, o analiză calitativă a acestor factori pun în evidență faptul că ceea ce stă la baza construcției itemilor meniți să evalueze satisfacția utilizatorilor nu este o reacție emoțională la utilizarea tehnologiei (satisfacție), ci mai degrabă caracteristici structurale ale sistemului, în acord cu unele observații critice ale literaturii (Bevan, 2008; Wixom & Todd, 2005).

Aceeași observație este susținută și de rezultatele obținute în ceea ce privește măsurarea utilizabilității cu ajutorul SUMI. Aici, chiar dacă din nou am obținut indici de fidelitate buni, structura pe dimensiuni a scalei nu poate fi nici pe departe aproximată în configurația factorială a scalei. Deși autorii instrumentului (Kirakowski, 1994) pun la baza dezvoltării instrumentului tocmai eforturile de separare a factorului afectiv (satisfacție) de cel de competență (utilizabilitate), datele obținute de noi nu par să confirme această structură. Ceea ce pun în evidență datele este convergența majorității itemilor către o scală generală, care cuprinde atât itemi care măsoară aspecte emoționale cât și itemi care măsoară competența. Acest rezultat pare să confirme încă o dată observațiile din literatura de specialitate cu privire la faptul că nu există o diferență conceptuală între satisfacție și utilizabilitate (Bevan, 2008; Wixom & Todd, 2005).

O ultimă categorie de rezultate care vin și ele să pună în discuție aceeași validitate de construct a instrumentelor de evaluare a utilizabilității sunt cele relaționate cu corelațiile între diversele scale ale instrumentelor. Aceste corelații, majoritatea moderat pozitive, pun în evidență o validitate discriminativă redusă care trebuie adresată, nu doar pentru clarificarea conceptuală a domeniului, ci și pentru utilitatea practică a acestor scale.

În ansamblu, datele noastre par să se alăture studiilor existente în literatura de specialitate care argumentează persistența unor deficiențe de validitate a instrumentelor și a modalităților de măsurare a utilizabilității și a satisfacției utilizatorilor (Hornbaek, 2006; Gray & Salzman, 1998a, Gray & Salzman, 1998b, Hartson et al., 2000; Melone, 1990). Literatura propune căutarea soluțiilor la aceste probleme în diverse locuri: definirea clară a utilizabilității și dezvoltarea unor instrumente de măsură aplicabile în mai multe contexte precum și a unor designuri experimentale valide (Gray & Salzman, 1998a); distincția între și compararea măsurătorilor obiective și subiective ale utilizabilității, dezvoltarea de măsurători ale memorării și

învățării, validarea și standardizarea măsurătorilor satisfacției, și măsurarea la nivel macro și micro (Hornbaek, 2006); completarea cu date calitative, clasificarea problemelor de utilizabilitate și derivarea unor criterii clare de evaluare a instrumentelor de utilizabilitate (Hartson et al., 2006). Cu toate acestea, pentru o mai bună fundamentare teoretică și de cercetare a acestor aspecte, deosebit de necesară domeniului HCI (Gray & Salzman, 1998a), soluția susținută de noi este comutarea accentului de pe caracteristici intrinseci tehnologiei, înspre mecanismele psihologice care pot explica modul în care anumite percepții și anumite consecințe ale acestor percepții (în termeni de emoții, cogniții evaluative și comportamente) se formează. Numai apelând la un astfel de cadru explicativ putem integra observația lui McClelland (1998 apud Olson & Moran 1998) cu privire la caracterul evolutiv, emergent al utilizabilității. Am arătat că menținând și conceptul măsurat în aceasta formă de emergență, sau continuă remodelare, nu putem obține măsurători valide. Dacă însă vom pune în evidență modul în care funcționează mecanismele prin care se explica apariția sau inhibiția anumitor emoții, evaluări, comportamente, vom putea ajunge la o anumită putere de predicție a acestora indiferent de volatilitatea soluțiilor tehnologice.

Referințe

- Andrei, D.M. (2009). Particularități ale procesului de implementare a tehnologiei intranet în contextul culturii organizaționale românești – date preliminare. Referat doctorat nepublicat, Cluj-Napoca: Universitatea Babeș-Bolyai.
- Balog, A., Pribeanu, C. (2009). Developing a measurement scale for the evaluation of AR-based educational systems. *Studies in Informatics and Control*, 18, 137-148.
- Bannon, L. (1991). From Human Factors to Human Actors. The Role of Psychology and Human Computer Interaction Studies in System Design. J. Greenbaum & M. Kyng (Eds.) *Design at Work. Cooperative design of computer system*. 25-44. New-Jersey: LEA.
- Batra, R., & Ahtola, O.T. (1990). Measuring the hedonic and utilitarian sources of consumer attitudes. *Marketing Letters*, 2, 159-170.
- Bevan, N. (2008). A framework for selecting the most appropriate usability measures. In *COST 294-MAUSE Workshop: Critiquing Automated Usability Evaluation Methods*.
- Bevan, N. (2008). UX, usability and ISO standards. Now let's do it in practice: *User experience evaluation methods in product development. Proceedings of the workshop on CHI 2008* Accesat Online în data de 20.05.2010 la:
www.cs.tut.fi/ihte/CHI08_workshop/papers/Bevan_UXEM_CHI08_06April08.pdf

- Byrne, B.M. (2010). *Structural Equation Modelling with AMOS: basic concepts, applications, and programming* (2nd Edition). New York: Taylor & Francis Group.
- Carroll, J.M. (1997). Human computer interaction: psychology as a science of design. *Annual Review of Psychology*, 48, 61-83.
- Chin, J. P., Diehl, V. A., & Norman, K. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of CHI '88 Conference on Human Factors in Computing Systems*, 213-218.
- Davis, F.D., Bagozzi, R.P., & Warshaw, P.R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35, 982-1003.
- Delone, W., & McLean, E. (1992). Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research* 3, 60-95.
- Dix, A., Finlay, J. E., Abowd, G. D., & Beale, R. (2004). *Human-Computer Interaction*. Harlow: Pearson Education Limited.
- Draper, S.W. (1999). Analysing fun as a candidate software requirement. *Personal Technology*, 3, 1-6.
- Dzida, W., S. Herda, and W. D. Itzfeldt. 1978. User perceived quality of interactive systems. *IEEE Transactions on Software Engineering* 4, 270-276.
- Field, A. (2000). *Discovering statistics using SPSS for Windows*. London: Sage.
- Gaver, W. W., & Martin, H. (2000). Alternatives. Exploring information appliances through conceptual design proposals. In *Proceedings of the CHI 2000 Conference on Human Factors in Computing*. New York: ACM.
- Gray, W. D. & Salzman, M. C. (1998a). Damaged merchandise? A review of experiments that compare usability evaluation methods. *Human-Computer Interaction* 13, 203-261.
- Gray, W. D. & Salzman, M. C. (1998b) "Repairing damaged merchandise: A rejoinder," *Human-Computer Interaction* 13, 325-335.
- Harper, B. D. & Norman, K. L. (1993). Improving User Satisfaction: The Questionnaire for User Interaction Satisfaction Version 5.5. In *Proceedings of the 1st Annual Mid-Atlantic Human Factors Conference*, Virginia Beach, VA, 224-228.
- Hartson, H. R., Andre, T. S. & Williges, R. C., (2001). Criteria for evaluating usability evaluation methods. *International Journal of Human Computer Interaction*, 13, 373-410.
- Hassenzahl, M. & Tractinsky, N. (2006). User Experience - a research agenda [Editorial]. *Behavior & Information Technology*, 25, 91-97.
- Hassenzahl, M. (2002). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, 13, 479-497.
- Hassenzahl, M. (2003). The thing and I: understanding the relationship between user and product. In M. Blythe, C. Overbeeke, A. F. Monk, & P. C. Wright (Eds.), *Funology: From usability to enjoyment* (pp. 31-42). Dordrecht: Kluwer.
- Hassenzahl, M. (2004). The Interplay of Beauty, Goodness, and Usability in Interactive

- Products. *Human Computer Interaction*, 19, 319-349.
- Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th French-speaking Conference on Human Computer Interaction IHM '08*, 11-15.
- Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttracDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität [AttracDiff: A questionnaire to measure perceived hedonic and pragmatic quality]. In J. Ziegler & G. Szwillus (Eds.), *Mensch&Computer 2003. Interaktion in Bewegung* (pp. 187–196). Stuttgart, Leipzig: B. G. Teubner.
- Hassenzahl, M., Kekez, R., & Burmester, M. (2002). The importance of a software's pragmatic quality depends on usage modes. In H. Luczak, A. E. Cakir, & G. Cakir (Eds.), *Proc. International Conference on Work With Display Units (WWDU 2002)*; pp. 275–276). Berlin: ERGONOMIC Institut für Arbeits- und Sozialforschung.
- Hassenzahl, M., Platz, A., Burmester, M., & Lehner, K. (2000). Hedonic and ergonomic quality aspects determine a software's appeal. *Proceedings of the CHI 2000 Conference on Human Factors in Computing*. New York: CM.
- Hewett, T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., & Verplank, W. (1992). *ACM SIHCHI curricula for human-computer interaction*. New York, NY: ACM Press.
- Huang, M-H. (2003). Designing Website Attributes to Induce Experiential Encounters. *Computers in Human Behavior*, 19, 425-442.
- ISO 9241-11 (1998). *Ergonomic requirements for office work with visual display terminals (VDTs)*, (Part 11: Guidance on usability). New York: ISO
- ISO FDIS 9241-210 (2008). *Human-centred design process for interactive systems*. New-York: ISO.
- Ives, B., Olson, M. H., & Baroudi J. J. (1983). The Measurement of User Information Satisfaction. *Communications of the ACM*, 26, 785-793.
- Jordan, P. (2000). *Designing pleasurable products: An introduction to the new human factors*. London: Taylor & Francis Group.
- Kelly, M. (1994). *MUSiC Final Report Parts 1 and 2: the MUSiC Project*. Hampshire, UK: Brameur Ltd,
- Kirakowski, J. (1987). *The Computer User Satisfaction Inventory. IEE Colloquium on Evaluation Techniques for Interactive System Design, II*, London.
- Kirakowski, J. (1994). The use of questionnaire methods for usability assessment. Accesat electronic în data de 15.03.2009 la: <http://sumi.ucc.ie/sumipapp.html>
- Kirakowski, J. (1998). *SUMI user handbook*. York University College: Human Factors Research Group.
- Kirakowski, J., & Corbett, M. (1988). Measuring User Satisfaction. In D. M. Jones & R. Winder (Eds.), *People and Computers IV*. Cambridge: Cambridge University Press.
- Kuniavsky, M. (2003). *Observing the user experience. A practitioner's guide to user research*. San Francisco: Elsevier.

- Kunze, E. -N. (2001). How to get rid of boredom in waiting-time-gaps of terminal-systems. In M. G. Helander, H. M. Khalid, & T. Ming Po (Eds.), *Proceedings of The International Conference on Affective Human Factors Design*. London: Asean Academic Press.
- Law, E. L-C., Hvannberg, E.T., & Cockton, G. (2008). A Green Paper on Usability Maturation In E. L-C., Law, E.T., Hvannberg, & G., Cockton (Eds). *Maturing Usability. Quality in interaction, software and value*. London: Springer-Verlag Limited.
- Logan, R.J. (1994). Behavioral and emotional usability: Thomson Consumer Electronics. In M. Wiklund (Ed.), *Usability in Practice*. Cambridge, MA: Academic Press.;
- Melone, N.P. (1990). A theoretical assessment of the user-satisfaction construct in information system research. *Management Science*, 36, 76-91.
- Monk, A. F., & Frohlich, D. (1999). Computers and fun. *Personal Technology*, 3, 91
- Myers, B., Hollan, J., Cruz, I., Bryson, S., Bulterman, D., Catarci, T., Citrin, W., Glinert, E., Grudin, J., & Ioannidis, Y. (1996). Strategic directions in human-computer interaction. *ACM Computing Surveys*, 28, 794-809.
- Nielsen, J. (1993). *Usability Engineering*. San Francisco: Morgan Kaufmann Elsevier.
- Norman, D.A. (2002). *The design of every day things*. New York: Basic Books.
- Olson, G.M. & Moran, T.P. (1998). Commentary on „Damaged Merchandise?”. *Human-Computer Interaction*, 13, 263-323.
- Orlikowski, W.J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science*, 11, 404-428
- Pitariu, HD., Andrei, D., Guran, M. A. (2009). Social research methods used in moving the traditional usability approach towards a user-centered design approach. *Journal of information technology and web engineering*, 4, 36-53.
- Porteous M., Kirakowski, J., & Corbett, M. (1993). SUMI User Handbook. *Human Factors*
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-computer interaction*. Addison-Wesley.
- Research Group*, University College Cork, Ireland.
- Tullis, T.S., & Stetson, J.N. (2004). A comparison of questionnaires for assessing website usability. In *Usability Professionals' Association Conference*, Minneapolis, Minnesota (June 2004).
- van Welie, M. (2001). *Task-Based user interface design*. Amsterdam, NL: SIKS.
- Wixom, B.H. and Todd, P.A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, 16, 85-102.