Affective Metacognitive Scaffolding and Enriched User Modelling for Experiential Training Simulators: A Follow-Up Study

Gudrun Wesiak[1,2], Adam Moore[3], Christina M. Steiner[1], Claudia Hauff[4], Conor Gaffney[5], Declan Dagger[5], Dietrich Albert[1,2], Fionn Kelly[6], Gary Donohoe[6], Gordon Power[5], and Owen Conlan[3]

[1] Knowledge Technologies Institute, Graz University of Technology, Austria
[2] Department of Psychology, University of Graz, Austria
{gudrun.wesiak,christina.steiner,dietrich.albert}@tugraz.at
[3] KDEG, School of Computer Science and Statistics, Trinity College, Dublin, Ireland
{mooread,owen.conlan}@scss.tcd.ie
[4] Delft University of Technology, the Netherlands
c.hauff@tudelft.nl
[5] EmpowerTheUser, Trinity Technology & Enterprise Campus, The Tower, Dublin, Ireland
{conor.gaffney,declan.dagger,gordon.power}@empowertheuser.com
[6] Department of Psychiatry, School of Medicine, Trinity College, Dublin, Ireland
fkelly@stpatsmail.com, donoghug@tcd.ie

Abstract. The ImREAL project is researching how to meaningfully augment and extend existing experiential training simulators. The services developed support self-regulated, goal-, and application-oriented learning in adult training. We present results from a study evaluating a medical interview training simulator that has been augmented by an affective metacognitive scaffolding service and by user modelling exploiting social digital traces. Data from 152 medical students participating in this user trial were compared to the results of a prior trial on an earlier technology version. Findings show that students perceived the learning simulator positively and that the enhanced simulator led to increased feelings of success, less frustration, higher technical flow, and more reflection on learning. Interestingly, this cohort of users proved reluctant to provide their open social IDs to enrich their user models.

Keywords: training simulator, self-regulated learning, affective metacognitive scaffolding, user modelling, social digital traces, evaluation.

1 Introduction

Today's technology-enhanced learning market offers a variety of adaptive cognitive systems that automatically recognize individual learners' needs and are able to create engaging and motivating learning experiences. In this context, experiential training simulators are gaining increasing popularity and importance as learning technologies for adult training. Adult learning is characteristically self-directed, experienced-based,

D. Hernández -Leo et al. (Eds.): EC-TEL 2013, LNCS 8095, pp. 396–409, 2013.

goal-oriented, intrinsically motivating, and relevant to the real world application context [1]. Immersive simulated environments for experiential learning need to correspond to these principles of andragogy by creating a situational context in which diverse kinds of skills relevant for job practice can be acquired and practiced [2]. Additionally, training environments should be tailored to the individual, taking into account each learner's goals, prior experiences, and metacognitive competence [3]. This makes self-regulation and metacognition, which have a long research tradition in pedagogy and psychology (e.g. [4],[5]), key topics in the definition of sound conceptual underpinnings of learning technologies, in general, and training simulators, in particular [3],[6].

The EU-funded ImREAL project[1] is researching how to augment existing simulators to increase their meaningfulness and their relevance by real-world activity modelling, enriched user modelling, and affective metacognitive scaffolding.

This paper is a follow-up to earlier work on developing, integrating, and testing a metacognitive scaffolding service (MSS) in a medical interview training simulator [7],[8]. After the initial evaluation of the training simulator (section 2) with integrated ImREAL services showed positive effects on learning motivation and perceived performance, the simulator augmentation has been further expanded: the metacognitive scaffolding service has been refined and also extended by incorporating affective aspects and the user modelling has been enriched by exploiting social digital traces as a basis for adaptation (section 2.1). Herein, we present the related ImREAL services and how their integration in the medical training simulation is shown to add value to simulation development and learning experiences. A user trial has been conducted to investigate the effects of the updated and extended augmentation. The results are compared to the previous user trial with metacognitive scaffolding only. Potential improvements to the prior version regarding self-regulated learning support, learning experience, learning performance, motivation, and service integration are analysed (section 3). Obtained outcomes and their implications for future work are discussed in section 4.

2 Experiential Training Simulator and Augmentation

EmpowerTheUser (ETU), a Dublin-based SME, has developed the innovative RolePlay Simulation Platform, empowering training experts with easy to use, rapid tools for developing immersive simulations. The platform provides a test-bed for the simulations in the ImREAL project and consists of three core tools: a Simulation Development Tool, an adaptive RolePlay Simulator, and an Analytics Dashboard. The Platform is actively used in training areas such as sales, customer services training, leadership, management and clinical interviewing (Fig. 1). Each RolePlay Simulation supports two running modes: assess mode (where all learner decisions are scored and a detailed performance report is presented) and practice mode (where learners can explore the scenario with all the learning interventions like coaching activated). With its assess mode the platform also serves as a psychometric profiling, behavioural measurement and skills assessment tool.

---

[1] `http://www.imreal-project.eu`

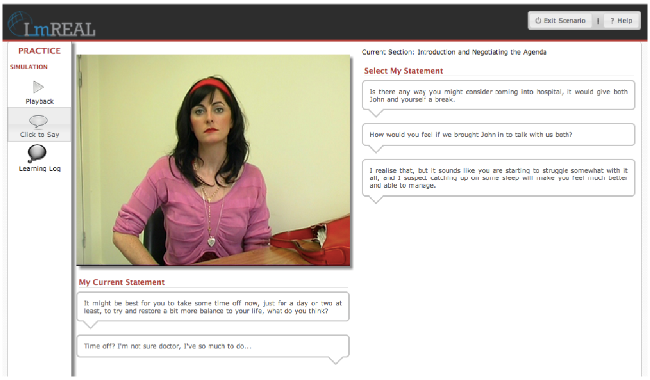

Fig. 1. ETU medical simulations: Interviewing a manic bipolar patient

2.1 Augmenting the ETU Simulator

During the ImREAL project a subset of the platform's learning triggers were extended to facilitate augmentations by the ImREAL suite of learning services, namely: coaching trigger (provides in-context feedback on the decision path) and reflection trigger (supports critical thinking and the transfer of learning). These learning triggers are modelled and configured *before* running a simulation by the subject expert/simulation developer. The focus in this paper is the impact of the ImREAL augmentations on ETU's RolePlay Simulation Platform in the context of medical training, particularly teaching student doctors about effective doctor-patient communication. ETU works closely with psychiatry and psychology experts at Trinity College Dublin's Medical School to create simulation scenarios for doctor-patient communication ranging from medical health to general medicine. Two enhancement opportunities are possible with the ImREAL suite of learning services: 1) augmenting reflection triggers with dynamic metacognitive scaffolding and 2) augmenting coaching triggers with more culturally oriented content.

To investigate the educational relevance of the simulator after implementation of these services, the following evaluation questions were addressed:

- Referring to the requirement of learning simulations providing authentic situational contexts [2],[3]: Does the simulation provide learning experiences that are *relevant* for users?
- Relating to the simulator augmentation: Are the *services* well *integrated* in the simulator and learning experience?

From Metacognitive Scaffolding to Affective Metacognitive Scaffolding. It has previously been described [8] how scaffolding is an important part of the educational process, supporting learners in knowledge acquisition and learning skill development.

As with the previously reported trial, metacognitive scaffolding was provided using calls to a RESTful service developed as part of the ImREAL project. The service utilises technology initially developed for the ETTHOS model [9] and presents items from the Metacognitive Awareness Inventory [10]. In an evolution from last year's scaffolding, however, prompts are displayed according to an underlying map created by the simulation's instructional designers.

It is commonly acknowledged that emotions play an important role in learning, exerting effects on information processing and performance [11]. In order to improve the provision of scaffolding to the learner, the existing MSS has been extended by an additional, affective element, thus becoming the Affective Metacognitive Scaffolding Service (AMSS). To obtain information of users' current affective state, an explicit Smiley-Based Affect Indicator (SBAI) [12] was inserted into the simulation feedback interface (*cf.* Fig. 2) and the enriched analysis and user modelling services, as described in the next section, were also utilised to provide additional information, where available, about the learner and their input.

With affect now included in the user model, and with a richer understanding of the learners, also the selection of the most salient metacognitive prompts to be displayed was changed. More priority was given to the weightings provided by the instructional designers. If high confidence values for learner affect attitudes were available, these were used to prioritise appropriate prompts (e.g. a prompt considered as encouraging might have a more prominent weighting for learners with negative affect indication).

The augmentation services (as presented below) are the basis for providing additional information to the learner model. If a learner provides a Twitter ID and textual input into the system, an affect stereotype is derived from sentiment analysis of their tweet stream or entry to update the model in real time. Affect stereotypes are positive (i.e. optimistic), neutral or negative (i.e. pessimistic). The augmentation service also provides a unique ID for each learner, which allows AMSS to send back the user model values that it derived and updated.

In addition, the presentation of the scaffolding service was slightly changed in this version of the simulation, with the open text box for collecting reflection consistently prefaced with a short text: "Reflect now on your learning: Was this last part of the simulation useful for you?", as shown in Fig. 2. Also, as the scaffolding was now provided as an XML bundle, styling was delegated to the simulation presentation engine, resulting in a more completely integrated interface and styling.

To integrate AMSS, ETU's Platform was extended in a number of ways: (a) The Simulation Development Tool was extended to accommodate mapping between simulation sections and the thinking prompts of AMSS. It also allows a reflection trigger to be tagged as dynamically "metacognitive". (b) The adaptive RolePlay Simulator can then interpret these new metacognitive reflection triggers, access the AMSS service and present the metacognitive prompts *in-situ*. (c) As part of the simulation self report dashboard, learners also have access to the AMSS SBAI affect reporting tool.

In the evaluation of AMSS the following questions were addressed in our study:

- Is *self-regulated learning* supported?
- Does the AMSS lead to better *learning experience?*
- Does the AMSS lead to a better *learning performance?*
- Does the AMSS enhance *motivation/affect*?

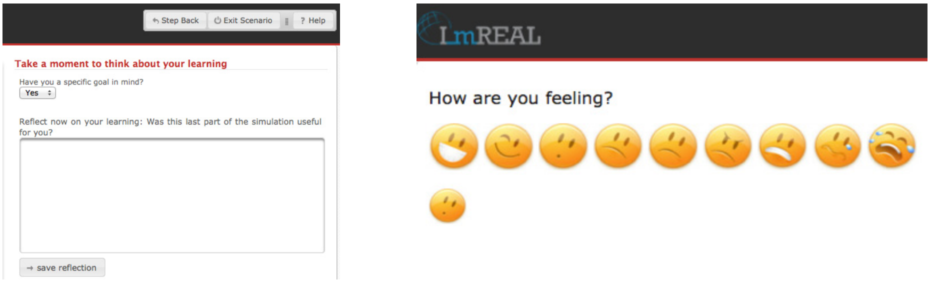

Fig. 2. Reflection trigger augmentation: Metacognitive scaffolding and capturing affective state

Enriched User Models Based on Digital Traces. The adaptation and personalization of e-learning systems requires user modelling, which involves collecting user information either explicitly, by asking the users a series of questions about their knowledge levels and skills or implicitly, by deriving a user profile from already existing data. Regarding the latter, digital traces on social Web platforms may give insights about relevant learner characteristics. With Facebook counting more than a billion monthly active users, Twitter and LinkedIn over 200 million users each[2], it is likely that today, most users of an e-learning system are regularly active on one (or several) social Web platforms. Exploiting the increasing amount of these digital traces learners create and share on social media is a potential resource for addressing the fact that adult learners arrive at a learning situation with highly diverse motivations and aims, real-world experiences and interests, knowledge and competence levels. Since digital traces are publicly available on the Web, we hypothesised that there will be few trust and privacy concerns for the users.

In our work on user model augmentation, we have investigated the utility of extracting different types of information nuggets about a learner from his social Web profiles, such as for example, the learner's geographic history (where has the learner been) [13], the learner's academic knowledge (which topics does the learner know something about) [14], the learner's high-level interests [15], and sentiment expressed in social media [16].

The U-Sem framework [17] developed in the ImREAL project is a service infrastructure for enriching and mining user data. Work on U-Sem focuses on how to derive user profiles from the social Web, how to enrich existing user profiles with that information, and how to verify the correctness of certain aspects of the automatically derived user models. U-Sem allows researchers and developers to create and design Web services for enriching and analysing user-related data, which are then made available to client applications, for instance a location detection service [13] used as a basis for providing adaptive, culturally-oriented coaching triggers.

To integrate U-Sem into the ETU simulator, its platform was extended in a number of ways: (1) The Simulation Development Tool was extended to support a

[2] See `http://newsroom.fb.com/Key-Facts`, `http://blog.twitter.com/2013/03/celebrating-twitter7.html`, and `http://press.linkedin.com/about` for Facebook, Twitter, and LinkedIn.

range of adaptive dimensions supported by the ImREAL services (e.g. cultural awareness). It also allows a coaching trigger to be tagged based on a specified adaptive dimension; (2) The adaptive RolePlay Simulator provides a mechanism for capturing a user's social tags (e.g. Twitter name or Flicker ID), registering them with ImREAL services and then adapting the coaching triggers to the user.

To evaluate this conceptual and technical approach for enriching user models the following questions were of interest:

- Do learners actively *use* one or more *social network(s)*?
- Do learners *feel comfortable* to freely provide their open social IDs?
- Does the augmented user model *correctly reflect* user characteristics?

3 Empirical Study

An empirical study was conducted with the goal of investigating the above outlined evaluation questions on the quality and impact of the service augmentation to the simulator. Overall 152 students from Trinity College, Dublin took part in this second user trial (UT2) and used the medical interview training on the ETU simulator with integrated AMSS and U-Sem services. Data collection was done during (log data) and after (questionnaire data) the training with the simulator. In addition, a cohort characterisation survey (on demography, metacognition, social network use, etc.) was conducted ahead of the trial.

3.1 Method

Instruments. Aligned with the evaluation questions outlined in Section 2 the following instruments were used:

Real-world relevance of the simulation was evaluated by two survey items.

Service integration of the ImREAL services in the simulator was measured by specific questions regarding the integration of AMSS and the flow of the simulation and learning experience (from a technical perspective).

Self-regulated learning (SRL) was assessed via an SRL questionnaire [18] with nine subscales measuring the general use of cognitive, metacognitive, and resource management strategies. In addition, interaction data and text entries tracked by the simulator served to investigate SRL behaviour during the training.

Learning experience covered workload and flow experience as aspects referring to how users perceive and experience the training with the simulator. These were measured with the NASA-TLX [19] and the Flow Short Scale [20], respectively. In addition, it was assessed how users experienced the thinking prompts provided by the AMSS, via a set of items provided in the post-simulation survey.

Learning performance in the simulator was evaluated via the objective assessment procedures built into the ETU simulator. Thereby, dialogue scores relating to the individual steps of the medical interview scenario and competence scores on several interview skills were used. The scoring constructs are derived from the skills defined in the Calgary-Cambridge model for medical interviewing. The constructs are then loaded in the simulation model and each decision path loads a particular weighting onto the respective skill constructs that are represented by a particular decision. All weightings across all paths taken by the learner during an attempt are then computed to produce the learners' performance score. The subscale 'performance' of the NASA-TLX served as additional subjective measure.

Motivation and affect. State motivation was explicitly queried by four specific questions; for affect assessment the Total Affective State Scales [21] were used as a subjective self-report. In addition, the SBAI of the AMSS provided the opportunity of capturing affect self-reports during the learning activity.

Questions on *social network use* and attitudes relating to *user privacy and trust* in the context of providing open social IDs for personalisation of the learning experience were part of the cohort characterisation study. Social IDs were queried in the cohort survey, as well as directly in the simulator for user model enrichment.

Participants and Procedure. UT2 participants were on average 22.81 years old (*SD* = 3.79) ranging from 19 to 45 years (note: *N*=95 for the cohort survey) and with equal gender ratio. Students participated from February until March 2013 as part of their curriculum. Before starting their training with the simulator participants completed the cohort characterization survey. During the training students could use the simulator as long and as often as they wished and deal with two psychiatric scenarios: mania and depression. Both scenarios could be entered in assess as well as practice mode. ImREAL services are integrated in the latter, whereas learning performance scores were assigned in assess mode. When entering the assess mode students were asked to predict their own scores regarding their performance in different steps of the clinical interview (introduction, eliciting information, outlining a management plan, closing the interview) and for interviewing skills (e.g. empathy, communication). The average overall duration of interacting with the simulation was 27.24 minutes (*SD*=11.03). After finishing their training, students provided their feedback on the simulator via an online survey covering the instruments listed above. Completion time of this survey was on average 10.75 minutes (*SD*=4.59).

3.2 Results

From the 152 participants using the simulator in the assess mode, different subsamples answered the cohort characterization and the post-simulation survey. Thus, sample sizes vary throughout this section. Where applicable results of this user trial (UT2) are compared to those of the prior user trial (UT1; 143 students using the simulator with MSS only in 2011/2012) reported in [8]. Comparisons are therefore based on independent samples, however participants in both studies were medical students from Trinity College, Dublin, with no significant differences regarding their sex or age. Log data was collected separately for the two scenarios. A scenario-specific analysis had shown that the scenarios were used to a different extent in the two user trials, but closer analysis yielded the same trends with respect to data distribution across the interview phases and the prediction scores and ETU scores per trial. Therefore, the data was aggregated across the two scenarios.

Real World Relevance. In both trials, participants were asked to (1) rate how relevant they think the experience with the ETU simulator was as preparation for real clinical interviews and (2) how well prepared (in percent 0-100) they felt for the interview based on their experience with the ETU learning environment. With a median of 3 on a 4pt. scale, the 40 participants from UT2 viewed the experience as rather relevant for their future interviews with patients, which does not differ significantly from UT1 ratings. With an average score of 59 (*SD*=16.16; *N*=37) on the second question UT2 students evaluated the relevance of the simulator higher than those of UT1 (*M*=51.3, *SD*=17.8; *N*=33), however the difference did not reach statistical significance ($T_{(68)}$=1.9, *p*=.062).

Service Integration. Integration of the metacognitive scaffolding, i.e. the thinking prompts is seen as neutral to positive. The perception of technical flow increased from UT1 to UT2 (M_{UT1}=3.75, *SD*=1.16; M_{UT2}=4.69, *SD*=1.15; $T_{(75)}$=-3.551, *p*=.001) indicating an appropriate integration of the service in the simulator.

Self-Regulated Learning. Data (*N*=25) from the nine SRL subscales range between means of 56.42 (*SD*=11.9) for ‘memorising strategies’ and 71.74 (*SD*=10.1) for ‘elaboration strategies’. A MANOVA for independent samples revealed no difference between the two user trials, neither for the nine univariate nor for the multivariate analyses (for the latter $F_{(9,36)}$=.623, *p*=.77). Since for both trials all subscale means are above 50%, students from all samples show an average to good use of cognitive and metacognitive strategies (higher values indicate better strategies).

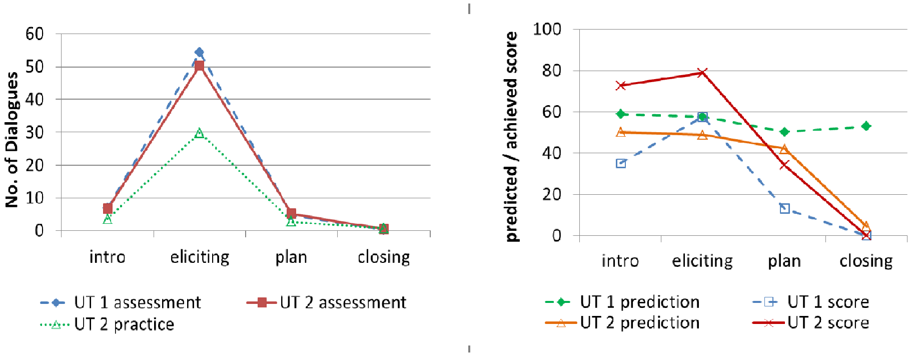

Fig. 3. Number of dialogues (left), predicted and achieved scores (right) in User Trials 1 and 2

Data on effort in terms of time spent with the simulation and number of dialogues were logged by the simulator. Because both measures are correlated (*r*=.487 for assess and .8 for practice mode, *p*<.001), and only the number of dialogues is reported here. Fig. 3 (left hand) compares the number of dialogues in different steps of the clinical interview during the assess modes of UT1 and UT2 and the practice mode of UT2. Mann-Whitney U-tests revealed no difference between the two user trials, irrespective of the step within the interview. For UT2, the assess mode was used more often and also more intensively than the practice mode. From the 152 students using the assess mode, 71 entered the practice mode and 58 processed dialogues as well. On the right hand side of Fig. 3, scores predicted by the students before entering the interviews and those actually obtained during the interview are shown. Whereas students from UT1 overestimated their performance for the introduction and planning phases, UT2 students underestimated their performance for the introduction and eliciting phases and overestimated it for the planning phase (*T*-tests for UT1: all $T_{(17)} \geq 6.5$, $p<.001$; Wilcoxon-tests for UT2: all $z \geq 3.2$, $p \leq .001$).

Another indicator for SRL strategies is the kind of information students recorded during the learning process. Overall 1092 thinking prompts were provided during the second user trial (*M*=15 per student, *SD*=16.9). With each prompt a text entry field popped up for collecting users' reflections (AMSS text entries) and students were asked about the usefulness of the last part of the simulation. From 69 yes/no responses 88.4% were positive. Additionally, the ETU simulator's notepad could be used at any time to reflect or take notes. Table 1 compares the number and types of UT2 notes (taken via AMSS or ETU) to those of UT1, coded by one researcher assessing the content of each entry and assigning its content to at least one of four types – Position, Technical, Patient Notes and Reflection [22]. Where the content was of more than one type, each was coded against that entry. The proportion of entries that were actually reflective increased in UT2, while the portion of all other entry types (positional, technical, and patient notes) decreased. Thus the type of notes taken by the students supports the assumption that AMSS fosters metacognition and reflection.

Table 1. Rounded percentages of content types for entries from the note-taking tools.

	Users	Text entries	Position	Technical	Notes	Reflection
UT 1	50	107	17	57	16	66
UT 2	35	86	1	13	7	93

Note: Entries can be coded to more than one type, thus percentages may exceed 100.

Learning Experience. 21 students of UT2 answered 10 questions on the perception of how helpful the *thinking prompts* were. On 5-pt. rating scales (from '1 not at all' to '5 very much') the overall score reached 3.55 (*SD*=.72), single scores ranged between 3.10 and 3.81. This indicates that the prompts are helpful with regard to content, timing, support in planning, monitoring, improving, and analysing one's learning performance. Comparisons with UT1 showed no significant differences between the overall or single item scores (all $p >.2$ for unrelated samples *T*-tests).

Two further measures used for learning experience are workload and flow. Average NASA-TX *workload* scores for the two user trials are depicted in Fig. 4. Multivariate MANOVA results revealed significantly lower overall workload for UT2 (multivariate $F_{(6,52)}=7.6$, $p<.001$). Results from single subscales indicate that the training in the simulator leads especially to load for effort and mental demand. Univariate analyses yield effects on performance and frustration ($F_{(1,52)}=48.4$, $p<.001$ for performance and $F_{(1,52)}=7.3$, $p=.009$ for frustration), indicating that students in UT2 felt less frustrated and had a stronger feeling to be successful than UT1 students.

Ratings for flow (on 7-pt. scales, with higher ratings indicating higher fluency and smoothness of the learning process and higher involvement in the task) from 40 UT2 users were compared to UT1 (*N*=37). The average UT2 rating of 4.75 (*SD*=.79) for overall flow was significantly higher than in UT1 with *M*=4.35 (*SD*=.91; $T_{(75)}=-2.12$, $p=.038$).

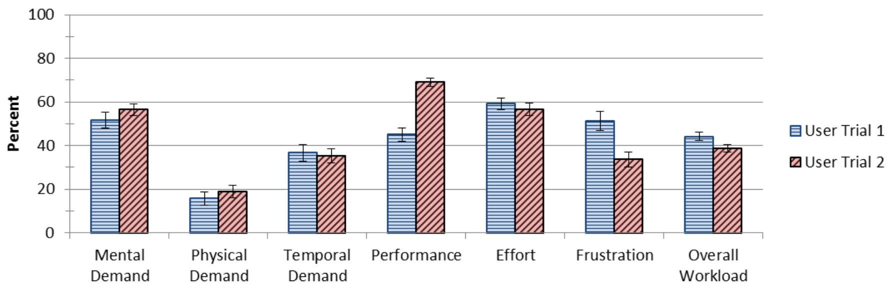

Fig. 4. Mean scores (and SD) from the NASA-TX workload scales for the two user trials

Learning Performance. ETU-scores assigned for each interview step increased significantly from UT1 to UT2 for the introduction, eliciting, and planning phases (*U*-tests: all $z \geq 7.9$, all $p<.001$, see also Fig. 3, right hand side). For UT2, also scores for the interview skills empathy, communication, eliciting information, summarising, and transition were available for both scenarios. Averaged student scores (*N*=152) for the different skills ranged between 61.4 (*SD*=18.5) for transition and 89.7 (*SD*=15.5) for eliciting information with a mean of 77.0 (*SD*=11.4). Considering a possible score maximum of 100, students' objective performance was rather high, which is also reflected in the section scores and subjective feelings of success (NASA-TLX performance) reported above.

Motivation and Affect. Motivation, measured by four questions regarding the motivation to learn about clinical interviewing, to improve one's interview skills, and to apply them to real situation, remained at a high level (on 4-pt. rating scales all $Md \geq 3$). To account for the added affective scaffolding service in UT2 students have explicitly been asked about their *affective states*. The overall mean score derived from seven subscales of the TASS (*M*=64.4, *SD*=16.4) shows that students were in a positive affective state regarding different dimensions like mood, motivation or thinking activity after they had finished their training in the simulator.
The *Smiley Based Affect Indicator* (SBAI) was displayed 352 times to 158 unique users, i.e. between 1 and 8 times per user (*M*=2.23, *SD*=1.36, *Md*=2). Only one learner, however, actually made use of the affect report to indicate her current emotional state, while all other learners did not. It is not clear, though, what the reasons for not using the SBAI were – this will need further investigation by consulting students. As learners did not provide any Twitter IDs in the simulation, sentiment analysis of tweet streams to derive affect information was not possible.

Social Network Use, User Privacy and Trust. From 95 participants completing the cohort characterization survey, 81% use Facebook, 20% Twitter, and less than 5% LinkedIn, Flickr or MySpace. Asked for their social network IDs, only 13 students (13.7%) provided their usernames on different social media (mainly Facebook). On 4pt.-rating scales (from totally fine to will not provide) students also indicated that they are rather 'nervous' about providing their IDs or that they would not do it (*Md*=4, *M*=3.12, *SD*=1.12). Open responses from 60 students on reasons why they would/would not provide their IDs concern mainly privacy issues, such as use for friends and family only (25 entries), followed by another group of comments (13 entries) indicating insecurity and a lack of trust about what happens with their information. These results were underlined by participants' behaviour during their training with the simulator, where nobody provided a Twitter or Flickr ID, such that user model enrichment via the U-Sem service was not enabled and students could not experience additional adaptive coaching based on their enriched user model. Of course, with only 19 Twitter and one Flickr user (as per cohort survey), the number of users who could have potentially provided their IDs was very low in the first place. To investigate the correctness of user model augmentation, for the social IDs collected in the cohort survey a comparison of information derived via U-Sem services with explicitly queried or available learner data was attempted. The Twitter-based interest profile [15] could be derived for two users, confirming their medical background and studies: for one user a quite narrow set of interest topics could be identified, with one third clearly associated with health-related aspects; for the second user a slightly wider interest profile resulted, including but not limited to health and education related topics. A comparison of information from location detection [13] with cohort survey data was not possible due to the unavailability of Flickr user IDs.

4 Discussion and Conclusion

We presented an evaluation study of the improved ETU simulator with enriched user modelling and affective metacognitive scaffolding. The study was a follow up to a first user trial with metacognitive scaffolding only. This allowed a more specific evaluation of the services, because differences could be directly attributed to the two stages of development. Each type of advancement led to different research questions, which we want to take up again now.

Simulator Augmentation. The RolePlay Simulation Platform used for clinical interview training was enhanced with ImREAL services via U-Sem and AMSS. The extended simulator was evaluated with respect to service integration and perceived relevance of the learning experience. Relevance ratings did not increase significantly between the two trials, but users consider the training definitely as a relevant preparation for interviews with real patients. The simulation embeds an authentic scenario relevant to the real-world application context, as called for by adult learning theory [1],[2]. Technical flow ratings increased from the last user trial, which indicates a good integration of the services in the simulator and a smooth interplay of software components, thus allowing a fluent interaction with the simulation.

Affective Metacognitive Scaffolding Services. The already existing MSS was enhanced by adding an affective element to prompts. Also, the selection of prompts was modified by giving higher priorities to instructional designer weightings and taking into account possible affective impacts of prompts. That is, each of the scaffolding prompts was examined and rated as to whether it would have a neutral, negative or positive impact. These ratings were added to the rulebase of the scaffolding intervention selector. The simulation designer now also creates the mapping and alignment of metacognitive prompts to simulation steps. Furthermore, the SBAI was integrated into the feedback interface of the simulation. Research questions concerned the aspects SRL, learning experience, learning performance, motivation and affect. Effort and performance scores as well as students' reflection notes show that students are supported in their SRL activities. Nevertheless, students still find it difficult to estimate how well they will perform, whereby under- as well as over-estimations occurred.

With the improved service, the notes students take are increasingly of reflective nature, which points to an additional benefit that users get out of the enhanced simulation. However, the number of unique users who actually worked with the thinking prompts is rather low: 30 students left entries via AMSS and/or the ETU notepad. One important reason is that the ImREAL services are only available in practice mode, but almost 60% of the students did the training only in the assess mode (which is prerequisite for the practice mode). Thus, for future applications, it is necessary to find ways to attract more students into the practice mode and to actually use the prompts. Similar to this, the reluctance to use the SBAI needs to be further investigated. Potential reasons might be unawareness of this self-report tool, but also that students did not see any added value of using it. In the second case, more detailed information about the benefit for their learning experience (e.g. selection of prompts adapted to a user's current affective state) could be a possible solution.

With regard to learning experience and performance, results indicate that the enhancement of the simulator led to an improved learning experience for the students. In UT2 prompts were generally perceived as helpful. Furthermore, students from UT2 rated the overall feeling of flow higher than in UT1. Thus, the improved simulator was able to convey a stronger feeling of task involvement and fluency. With regard to workload, frustration went down and the feeling to be successful at a given task (performance) increased from UT1 to UT2. High performance is also confirmed by the objective scores assigned by the ETU simulator, where students obtained an average of 77 (out of 100). Their performance had increased from UT1 in the introduction, eliciting, and planning phase of the interviews. Motivation and affect measures showed that students had a high motivation regarding their learning task and that they were in positive affective states after the training.

Enriching User Models based on Digital Traces. The main idea behind U-Sem is to extract and structure relevant information from the social Web services users are active in and to use this information for enriching user models from client applications. These applications can consequently update their user profiles according to the newly received information and provide adaptive services without having the users to take initial assessments or fill out long surveys on their preferences. The Web services utilised by U-Sem for mining users' digital traces are Twitter and Flickr. Natural requirements for this procedure are users who actively use social platforms and the willingness of users to provide their social network IDs. There are two main points we can derive from this study. First, the use of social networks other than Facebook is not as widely spread as we initially assumed. Especially for Flickr, which is a main source for the digital traces used by U-Sem, only one participant indicated use of this platform. Second, the type of data used by U-Sem is publicly available on the Web and participants indicated that they believe that social networks are rather open (*Md* = 3.5 on a 4-pt. scale from 1-totally private to 4-totally open). Nevertheless, they were very reluctant to provide their social network IDs. Main reasons given were to maintain privacy and that the people behind the service are not known. Thus, for future research, these two aspects need to be taken into account by: (a) considering the type of sample (which might influence the type and extent of social networks used) and (b) by fostering trust in the service in order to prompt users to provide their user ID (e.g. by giving detailed explanation on the type of information that is used and on the way the information is retrieved, or by providing more information about the research group behind the service). By increasing users' willingness to share their user IDs, the correctness of user information derived from digital traces can be investigated in more detail in a next step.

Conclusion. This paper reports on further augmentations to the ETU RolePlay Simulation Platform with ImREAL services. Whilst further work needs to be performed to investigate outstanding issues of unwillingness to engage in affective reporting and supplying social IDs, we have shown that these augmentations further enhance the flow and fidelity of the simulation, leaving learners more motivated, engaged and competent. We finish with a quote from a user: 'I really like the simulation: it is very interesting and is directly associated to a real world problem.'

Acknowledgments. The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement no 257831 (ImREAL project) and could not be realised without the close collaboration between all ImREAL partners.

References

1. Knowles, M.: Andragogy in action. Jossey-Bass, San Francisco (1984)
2. Swartout, W.: Lessons learned from virtual humans. AI Magazine 31, 9–20 (2010)
3. Hetzner, S., Steiner, C., Dimitrova, V., Brna, P., Conlan, O.: Adult self-regulated learning through linking experience in simulated and real world: A holistic approach. In: Kloos, C.D., Gillet, D., Garcia, R.M.C., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 166–180. Springer, Heidelberg (2011)
4. Zimmerman, B.J.: A social cognitive view of self-regulated academic learning. Journal of Educational Psychology 81, 329–339 (1989)
5. Flavell, J.H.: Metacognition and cognitive monitoring. A new area of cognitive-developmental inquiry. American Psychologist 34, 906–911 (1979)
6. Azevedo, R., Hadwin, A.F.: Scaffolding self-regulated learning and metacognition – Implications for the design of computer-based scaffolds. Instructional Science 33, 367–379 (2005)
7. Moore, A., Hetzner, S., Pannese, L., Steiner, C., Brna, P., Conlan, O.: Affective metacognitive scaffolding for the enhancement of experiential training for adult learners. In: Workshop on Awareness and Reflection in Learning Networks (ARNets11), EC-TEL 2011, Palermo, Italy (2011)
8. Berthold, M., Moore, A., Steiner, C.M., Gaffney, C., Dagger, D., Albert, D., Kelly, F., Donohoe, G., Power, G., Conlan, O.: An initial evaluation of metacognitive scaffolding for experiential training simulators. In: Ravenscroft, A., Lindstaedt, S., Kloos, C.D.,

Hernández-Leo, D. (eds.) EC-TEL 2012. LNCS, vol. 7563, pp. 23–36. Springer, Heidelberg (2012)
9. Macarthur, V., Moore, A., Mulwa, C., Conlan, O.: Towards a cognitive model to support self-reflection: Emulating traits and tasks in higher order schemata. In: Workshop on Augmenting the Learning Experience with Collaborative Reflection, EC-TEL 2011, Palermo, Italy (2011)
10. Schraw, G., Sperling Dennison, R.: Assessing metacognitive awareness. Contemporary Educational Psychology 19, 460–475 (1994)
11. Pekrun, R.: The impact of emotions on learning and achievement: towards a theory of cognitive/motivational mediators. Applied Psychology 41, 359–376 (1992)
12. Moore, A., Ascolese, A.: Explicit tools for affective indicators. In: Workshop on Augmenting the Learning Experience with Collaborative Reflection, EC-TEL 2011, Palermo, Italy (2011)
13. Hauff, C., Houben, G.-J.: Placing images on the world map: A microblog-based enrichment approach. In: SIGIR 2012, pp. 691–700 (2012)
14. Hauff, C., Houben, G.-J.: Deriving knowledge profiles from twitter. In: Kloos, C.D., Gillet, D., Garcia, R.M.C., Wild, F., Wolpers, M. (eds.) EC-TEL 2011. LNCS, vol. 6964, pp. 139–152. Springer, Heidelberg (2011)
15. Tao, K., Abel, F., Gao, Q., Houben, G.-J.: TUMS: Twitter-based user modeling service. In: García-Castro, R., Fensel, D., Antoniou, G. (eds.) ESWC 2011. LNCS, vol. 7117, pp. 269–283. Springer, Heidelberg (2012)
16. Gao, Q., Abel, F., Houben, G.-J., Yu, Y.: A Comparative Study of Users' Microblogging Behavior on Sina Weibo and Twitter. In: Masthoff, J., Mobasher, B., Desmarais, M.C., Nkambou, R. (eds.) UMAP 2012. LNCS, vol. 7379, pp. 88–101. Springer, Heidelberg (2012)
17. Abel, A., Celik, I., Hauff, C., Hollink, L., Houben, G.-J.: U-Sem: Semantic enrichment, user modeling and mining of usage data on the social web. In: International Workshop on Usage Analysis and the Web of Data, USEWOD 2011 (2011)
18. Fill Giordano, R., Litzenberger, M., Berthold, M.: On the Assessment of strategies in self-regulated learning (SRL) – Differences in adolescents of different age group and school type. Poster Presented at the 9. Tagung der Österreichischen Gesellschaft für Psychologie, Salzburg, April 8-10 (2010)
19. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, N.M.P.A. (ed.) Human Mental Workload, pp. 239–250. North Holland Press, Amsterdam (1988)
20. Rheinberg, F., Vollmeyer, R., Engeser, S.: Die Erfassung des Flow-Erlebens [The assessment of flow]. In: Stiensmeier-Pelster, J., Rheinberg, F. (eds.) Diagnostik von Motivation und Selbstkonzept, pp. 261–279. Hogrefe, Göttingen (2003)
21. Nicolescu, A.B., Lulow, L.L., Ogden, C.A., Le-Niculescu, H., Salomon, D.R., Schork, N.J., Caligiuri, M.P., Lohr, J.B.: PhenoChipping of psychotic disorders: A novel approach for deconstructing and quantitating psychiatric phenotypes. American Journal of Medical Genetics Part B (Neuropsychiatric Genetics), 141B, 653–662 (2006)
22. Moore, A., Gaffney, C., Dagger, D., Kelly, F., Donohoe, G., Power, G., Conlan, O.: Preliminary content analysis of reflections from experiential training simulator support. In: Workshop on Self-Regulated Learning in Educational Technologies (SRL@ET), ITS 2012. Chania, Greece (2012)