Emotion Elicitation in Socially Intelligent Services: the Intelligent Typing Tutor Study Case

Andrej Košir
University of Ljubljana, Faculty of Electrical Engineering
Tržaška cesta 25
Ljubljana, Slovenia
andrej.kosir@fe.uni-lj.si

Marko Meža
University of Ljubljana, Faculty of Electrical Engineering
Tržaška cesta 25
Ljubljana, Slovenia
marko.meza@fe.uni-lj.si

Janja Košir
University of Ljubljana, Faculty of Education
Krdeljeva ploščad
Ljubljana, Slovenia
janja.kosir@pef.uni-lj.si

Matija Svetina
University of Ljubljana, Faculty of Fine Arts
Tržaška cesta 2
Ljubljana, Slovenia
matija.svetina@ff.uni-lj.si

Gregor Strle
Scientific Research Centre SAZU
Novi trg 2
Ljubljana, Slovenia
gregor.strle@zrc-sazu.si

ABSTRACT

The paper discusses the challenges of user emotion elicitation in socially intelligent services, based on the experimental design and results of the intelligent typing tutor. Human-machine communication (HMC) of the typing tutor is supported by the continuous real-time emotion elicitation of user's expressed emotions and the emotional feedback of the service, through the graphically rendered emoticons. It is argued that emotion elicitation is an important part of successful HMC, as it improves the communication loop and increases user engagement. Experimental results show that user's valence and arousal are elicited during the typing practice, on average 18% to 25% of the time for valence and 20% to 31% of the time for arousal. However, the efficiency of emotion elicitation varies greatly throughout the use of the service, and also moderately among users. Overall, the results show that emotion elicitation, even via simple graphical emoticons, has significant potential in socially intelligent services.

Keywords

affective computing, emotion elicitation, social intelligence, human-machine communication, intelligent tutoring systems

1. INTRODUCTION

Bridging the gap between modern digital services and the increasing demands and (often insufficient) capabilities of a wide range of users is a challenging task. In recent years, much focus has been given to user adaptation procedures in socially intelligent services, including user modeling, recommender systems, human-machine communication (HMC), among many others [10], [1], [11]. While there have been substantial advances in many of these areas, the state-of-the art technology still lacks satisfactory means to efficiently meet various user needs and/or tailor to their capabilities. As the potential for new users of technology supported services is growing (e.g. groups of elderly users), so is the digital divide [42]. This gap may manifest itself in many forms. It may deprive a particular user group of efficient use of a service (e.g., due to the lack of technological proficiency), it may be limited in scope and only partially attend to users needs (e.g., the use of multiple services for a series of common, integrated tasks), or for some user groups offer no accessibility to a service altogether (e.g., e-banking for the elderly users). In general, it results in frustration and increased cognitive load, requiring significant effort to use a service (e.g. interaction, navigation, finding information, etc.), instead of a service adapting to user needs and capabilities.

One way to address these issues is to establish and sustain efficient (close-to-human) communication level between a user and a service, with HMC at the core of contextualization and adaptation procedures. Whereas natural (human-to-human) communication is innate and in general requires minimal effort for the actors involved to sustain it, HMC is void of both innateness and context, as well as of nonverbal (auditory, visual, olfactory) cues. Thus, for a modern digital service to be successful, it should be capable of expressing minimal social intelligence [45]. Another important and inherent property of natural communication is its continuity in real-time. HMC should be able to exhibit some level of social intelligence by generating and processing social signals in near-real-time.[1] To sustain the feedback loop the user should be at least minimally engaged, with nonverbal (social) signals (such as emotions) elicited at a continuous (minimal delay) rate. Ideally, effective HMC should minimize the user-service adaptation procedures and maximize the engagement and the intended use of a service. In other words, a service is socially intelligent when it is ca-

EMPIRE 2016, September 16, 2016, Boston, MA, USA.

[1]The maximal tolerated delay is about 0.5 seconds.

pable of reading (measuring and estimating) user's social signals (verbal and/or non-verbal communication signals), producing machine generated feedback on these signals, and sustaining and adapting according such HMC.

In general, we believe it is possible to alleviate some of the main obstacles towards more effective user-service adaptation procedures by addressing the following:

- *Non-intrusive user data acquisition.* Some types of user data (e.g., user's emotion state) should be tracked in near-real-time. The problem is users do not like obtrusive data gathering methods (e.g., to repeatedly fill in questionnaires or use wearable sensors in everyday situations). The state-of-the-art techniques for non-intrusive user data acquisition are limited and can not provide sufficient high quality user data for the efficient user-service adaptation procedures;

- *Contextualization.* Contextualization refers to the definition of circumstances relevant for specific user-service adaptation. Effective user adaptation is highly context-sensitive as user involvement, attention and motivation, as well as preferences, are to a large extent context dependent. The emergent technologies of Internet of things (IoT), wearable computing, ubiquitous computing, and others, offer various building blocks to model specific contextualization tasks, however, user interaction data is typically not taken into account;

- *Service functionality and content adaptation for the user.* Ideally, user adaptation procedure is successful when the service is able to adapt to (and improve upon) the user needs and preferences in near real-time. As a result, the adaptation mechanisms of the service need to go beyond generally applicable adaptation procedures to address the specific task-dependent and user-interaction scenarios.

The aim of the paper is to analyze the efficiency of emotion elicitation in a socially intelligent service. The underlying assumption is that emotion elicitation should be an integral part of HMC, as it can greatly improve user-service adaptation procedure. For this purpose, the experiment was conducted using the socially intelligent typing tutor. The tutor is a web-based learning service designed to elicit emotions and thus improve learner's attention and overall engagement in the touch-typing training. Emotion elicitation is utilized together with the notion of positive reinforcement, where the learner is being rewarded for her efforts through the emotional feedback of the service. Moreover, the tutor is able to model and analyze learner's expressed emotions and measure the efficiency of emotion elicitation in the tutoring process.

The paper is structured as follows. Section 2 presents related work, while Section 3 discusses general aspects of emotion elicitation in socially intelligent services and then presents the socially intelligent typing tutor. Section 4 presents the experimental results on emotion elicitation in the intelligent typing tutor. The paper ends with a general conclusion and future work.

2. RELATED WORK

The research and development of a fully functioning socially intelligent service is still at a very early stage. However, various components that will ultimately enable such services are under intensive development for several decades. We briefly present them grouped according to the following subsections.

2.1 Social intelligence, social signals and non-verbal communication cues

There are many definitions of social intelligence applicable in this context [23]. The wider definition used here is by Vernon [44], who defines social intelligence as the person's "ability to get along with people in general, social technique or ease in society, knowledge of social matters, susceptibility to stimuli from other members of a group, as well as insight into the temporary moods or underlying personality traits of strangers". Furthermore, social intelligence is demonstrated as the ability to express and recognize social cues and behaviors [2], [6], including various non-verbal cues (such as gestures, postures and face expressions) exchanged during social interaction [47].

Social signals are extensively being analyzed in the field of human to computer interaction [47], [46], often under different terminology. For example, [33] use the term 'social signals' to define a continuously available information required to estimate emotions, mood, personality, and other traits that are used in human communication. Others [31] define such information as 'honest signals' as they allow to accurately predict the non-verbal cues and, on the other hand, one is not able to control the non-verbal cues to the extent one can control the verbal form. Here, we will use the term social signal.

2.2 Socially intelligent learning services

Several services exist that support some level of social intelligence, ranging from emotion-aware to meta-cognitive. One of the more relevant examples is the intelligent tutoring system AutoTutor/Affective AutoTutor [15]. AutoTutor/Affective AutoTutor employs both affective and cognitive modelling to support learning and engagement, tailored to the individual user [15]. Some other examples include: Cognitive Tutor [7] – an instruction based system for mathematics and computer science, Help Tutor [3] – a meta-cognitive variation of AutoTutor that aims to develop better general help-seeking strategies for students, MetaTutor [9] – which aims to model the complex nature of self-regulated learning, and various constraint-based intelligent tutoring systems that model instructional domains at an abstract level [28], among many others. Studies on affective learning indicate the superiority of emotion-aware over non-emotion-aware services, with the former offering significant performance increase in learning [37], [22], [43].

2.3 Computational models of emotion

One of the core requirements for socially intelligent service is the ability to detect and recognize emotions, and exhibit the capacity for expressing and eliciting basic affective (emotional) states. Most of the literature in this area is dedicated to the affective computing and computational models of emotion [26], [25], [34], which are mainly based on the appraisal theory of emotions [48]. Several challenges remain, most notably the design, training and evaluation of computational models of emotion [20], their critical analysis and comparison, and their relevancy for other research fields (e.g., cognitive science, human emotion psychology), as most computational models of emotion are overly simplistic [12].

2.4 Physiological sensors

The development of wearable sensors enabled the acquisition of user data in near-real-time, as well as the research and estimation of user's internal states (such as emotion and stress level estimation) that started more than a decade ago [5], [4]. Notable advances can also be found in the fields of psychological computing and HCI, with the development of several novel measurement related procedures and techniques. For example, psychophysiological measurements are being employed to extend the communication bandwidth and develop smart technologies [18], along with the design guidelines for conversational intelligence based on the environmental sensors [14]. Several studies deal with human stress estimation [36], workload estimation [30], cognitive load estimation [27], [8], among others, and specific learning tasks related to physiological measurements [49], [21].

2.5 Human emotion elicitation

The field of affective computing has developed several approaches to modeling, analysis and interpretation of human emotions [19]. The most known and widely used emotion annotation and representation model is the Valence-Arousal-Dominance (VAD) emotion space, an extension of Russell's valence-arousal model of affect [35]. The VAD space is used in many human to machine interaction settings [50], [40], [32], and was also adopted in the socially intelligent typing tutor (see section 3.3.2). There are other attempts to define models of human emotions, such as specific emotion spaces for human computer interaction [16], or more recently, models for the automatic and continuous analysis of human emotional behaviour [19]. Recent research on emotion perception argues that traditional emotion models might be overly simplistic, pointing out the notion of emotion is multi-componential, and includes "appraisals, psychophysiological activation, action tendencies, and motor expressions" [38]. Consequently, and relevant to the interpretations of valence in the existing models, some researchers argue there is a need for the "multifaceted conceptualization of valence" that can be linked to "qualitatively different types of evaluations" used in the appraisal theories [39].

Research of emotion elicitation via graphical user interface is far less common. Whereas several studies on emotion elicitation use different stimuli (e.g., pictures, movies, music) [41] and behavior cues [13], none to our knowledge tackle the challenges of graphical user interface design for the purpose of emotion elicitation.

In the intelligent typing tutor, user emotions are elicited by the graphical emoticons (smileys) via the dynamic graphical user interface of the service. The choice of emoticons was due to their semantic simplicity, unobtrusiveness, and ease of continuous measurement – using pictures as a stimuli would add additional cognitive load and likely evoke multiple emotions. This approach also builds upon the results of previous research, which showed that human face-like graphics increase user engagement, that the recognition of emotions represented by emoticons is intuitive for humans, and that emotion elicitation based on emoticons is strong enough to be applicable [17]. The latter assumption is verified in this paper.

3. EMOTION ELICITATION IN SOCIALLY INTELLIGENT SERVICES: THE TYPING TUTOR STUDY CASE

The following sections discuss the role of emotion elicitation in socially intelligent services and its importance for efficient HMC. General requirements and the role of emotion elicitation are discussed in the context of our study case – the intelligent typing tutor. Later sections present the design of the intelligent typing tutor and its emotion elicitation model.

3.1 General requirements for a socially intelligent service

A given service is socially intelligent if it is capable of performing the following elements of social intelligence:

1. Read relevant user behavior cues: human emotions are conveyed via behaviour and non-verbal communication cues such as face expression, gestures, body posture, color of the voice, etc.

2. Analyze, estimate and model user emotions and non-verbal (social) communication cues via computational model: behavior cues are used to estimate user's temporary emotion state. Selected physiological measurements (pupil size, acceleration of the wrist, etc.) are believed to be correlated with user's emotion state and other non-verbal communication cues. These are used as an input to the computational model of user emotions and other non-verbal communication cues.

3. Integrate and model machine generated emotion expressions and other non-verbal communication cues: for example, the notion of positive reinforcement could be integrated into a service to improve user engagement, taking into account user's temporary emotion state and other non-verbal communication cues.

4. Generate emotion elicitation to improve user engagement: continuous feedback loop between user emotion state and machine generated emotion expressions for purpose of emotion elicitation.

5. Context and task-dependent adaptation: adapt the service according to the design goals. For example, in the intelligent typing tutor case study, the intended goal is to improve learner's engagement and progress. The touch-typing lessons are carefully designed and adapt in terms of typing speed and difficulty to meet individual's capabilities, temporary emotion state and other non-verbal communication cues.

Such service is capable of sustaining efficient, continuous and engaging HMC. It also minimizes user-service adaptation procedures. An early-stage example of socially intelligent service is provided below.

3.2 Typing tutor as a socially intelligent service

The overall goal of the socially intelligent typing tutor is to improve the process of learning touch-typing. For this purpose, emotion elicitation is integrated into HMC together with the notion of positive reinforcement, to amplify the attention, motivation, and engagement of the individual learner. In its current form, the rudimentary model of emotion elicitation utilizes emoticon-like graphics via the graphical user interface of the service, presented to the learner in real-time (see section 3.3). The tutor uses state-of-the-art technology (3.2.1) and is able to model, measure and analyze emotion elicitation throughout the tutoring process.

3.2.1 Architecture and design

Typing tutor's main building blocks consist of:

1. Web GUI: to support typing lessons and machine generated emotion expressions via emoticons (see Fig. 1);

2. Sensors: to conduct physiological measurements and monitor user status (wrist accelerometer, camera, emotion-recognition software to estimate user emotions, eye gaze, pupil size, etc.);

3. Computational model: for measuring user emotions and attention in the tutoring process;

4. Recommender system: for modelling machine generated emotion expressions;

5. Typing content generator: which follows typing lectures designed by the expert.

Real-time sensors are integrated into the service to gather physiological data about the learner. The recorded data is later used to establish the weak ground truth of learner's attention and the efficiency of emotion elicitation. Both are further estimated through the human annotation procedure, based on the carefully designed operational definition and verified using psychometric characteristics. The list of sensors integrated in the tutor includes:

- Keyboard: to monitor cognitive and locomotor errors that occur while typing;

- Video recorder: to extract learner's facial emotion expressions in real-time;

- Wrist accelerometer and gyroscope: to trace the hand movement;

- Eye tracking: to measure pupil size and estimate learner's attention and possible correlates to typing performance.

The intelligent typing tutor is publicly available as a client-server service running in a web browser (http://nacomnet.lucami.org/test/desetprstno_tipkanje). Data is stored on the server for later analyses and human annotation procedures. Such architecture allows for crowd-sourced testing and efficient remote maintenance.

3.3 Emotion elicitation in the intelligent typing tutor

The role of emotion elicitation in the intelligent typing tutor is that of efficient HMC and reward system. The positive reinforcement assumption [29] is used in the design of the emotion elicitation model. Positive reinforcement argues that learning is best motivated by a positive emotional responses from the service when learners ratio of attention over fatigue goes up, and vice versa. Here, machine generated positive emotion expressions act as rewards, with the aim to improve learner's attention, motivation and engagement during the touch-typing practice. The learner is rewarded by a positive emotional response from the service when she invest more effort into practice (the service does not support negative reinforcement). According to the positive reinforcement assumption, the rewarded behaviors will appear more frequently in the future. Negative reinforcement is not used for two reasons: there is no clear indication how negative reinforcement would contribute to the learning experience, and it would require an introduction of additional dimension, making the research topic of the experiment even more complex.

3.3.1 Machine emotion model

The intelligent typing tutor uses emotion elicitation to reward any behavior leading to the improvement of learner's engagement with the service. The rewards come as positive emotional responses conveyed by the emoticon via graphical user interface. The machine generated emotion responses range from neutral to positive (smiley) and act as stimuli for user (learner) emotion elicitation. For this purpose, a subset of emoticons from Official Unicode Consortium code chart (see http://www.unicode.org/) was selected and emoticon-like graphical elements were integrated into the newly designed user interface of the service shown in Fig. 1.

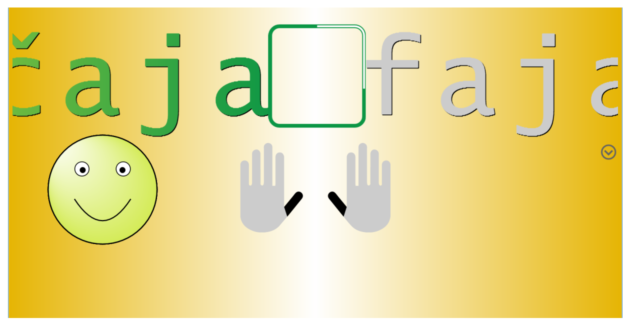

Figure 1: Socially intelligent typing tutor integrates touch-typing tutoring and machine generated emoticons (for emotion elicitation) via its graphical user interface.

Emotional responses are computed according to the learning goals of the tutor. To improve learner's attention and overall engagement in the touch-typing practice, the emotional feedback of the service needs to function in real-time. As mentioned above, the positive reinforcement assumption acts as the core underlying mechanism for modelling machine generated emotions. At the same time such mechanism is suitable for dynamic personalization, similar to the conversational RecSys [24]. In order to implement it successfully, the designer needs to decide on 1. which behaviors need to be reinforced to appear more frequently, and 2. which rewards, relevant for the learner, need reinforcement.

3.3.2 User emotion model

User (learner) emotions are elicited via tutor's graphical user interface, based on the machine generated emotion expressions from (3.3.1). The VAD emotion model is used for representation and measurement of learner elicited emotions, similar to [16]. The VAD dimensions are then measured in real-time by emotion recognition software (see section 4.1).[2]

[2]Here, we only discuss valence Φ_{uV} and arousal Φ_{uA}, the

Two independent linear regression models are used to model user emotion elicitation as a response to the machine generated emoticons. The models are fitted as follows: the measured values of user emotion elicitation for valence and arousal are fitted as dependent variables, whereas the machine generated emotion expression is fitted as an independent variable (Eq.1). The aim is to obtain the models' quality of fit and the proportion of the explained variance in emotion elicitation.

$$\Phi_{uV} = \beta_{1V}\Phi_m + \beta_{0V} + \varepsilon_V, \quad \Phi_{uA} = \beta_{1A}\Phi_m + \beta_{0A} + \varepsilon_A, \tag{1}$$

where Φ_m stands for one dimensional parametrization of the machine emoticon graphics, ranging from 0 (neutral emoticon) to 1 (maximal positive emotion expression). Notations β_{1V} and β_{1A} are user emotion elicitation linear model coefficients, β_{0V} and β_{0A} are the averaged effects of other influences on user emotion elicitation, and ε_V and ε_A are independent variables of white noise.

The linear regression model was selected due to the good statistical power of its goodness of fit estimation R^2. There is no indication that emotion elicitation is linear, but we nevertheless believe the choice of the linear model is justified. The linear model is able to capture the emotion elicitation process, detect emotion elicitation, and provide valid results (see section 4.2). Residual plots (not reported here) show that linear regression assumptions (homoscedasticity, normality of residuals) are not violated.

To further support our argument for emotion elicitation in the intelligent typing tutor, we statistically tested our hypothesis that a significant part of learner's emotions is indeed elicited by the machine generated emoticons. We did this with the null hypothesis testing $H_0 = [R^2 = 0]$ (see section 4.2), which demonstrated good power compared to the statistical tests by some of the known non-linear models.

4. USER EXPERIMENT: THE ESTIMATION OF USER EMOTION ELICITATION

The following sections give an overview of the user experiment and results on emotion elicitation in the intelligent typing tutor.

4.1 User experiment

The experiment consisted of 32 subjects invited to practice touch-typing in the intelligent typing tutor (see 3.2), with the average duration of the typing session approx. 17 minutes (1020 seconds). The same set of carefully designed touch-typing lessons was given to all test subjects. User data was acquired in real-time using sensors (as described in section 3.2), and used as an input to the computational model of machine generated emotion expressions, and recorded for later analysis. For the preliminary analysis presented here, five randomly selected subjects were analysed on the segment of the overall duration of the experiment.[3] The test segment spans from 6 to 11.5 mins (330 seconds) of the experiment.

The test segment used for the analysis is composed of the following steps:[4]

1. Instructions are given to the test users: users are personally informed about the goal and the procedure of the experiment (by the experiment personnel);

2. Setting up sensory equipment, start of the experiment: a wrist accelerometer is put on, the video camera is set on, and the experimental session time recording is started (at 00 seconds);

3. At 60 seconds: machine generated sound disruption of the primary task: "Name the first and the last letter of the word: mouse, letter, backpack, clock";

4. At 240 seconds: machine generated sound disruption of the primary task, "Name the color of the smallest circle", in the figure (Fig 2). This cognitive task is expected to significantly disrupt learner's attention away from the typing exercise;

5. The test segment ends at 330 seconds.

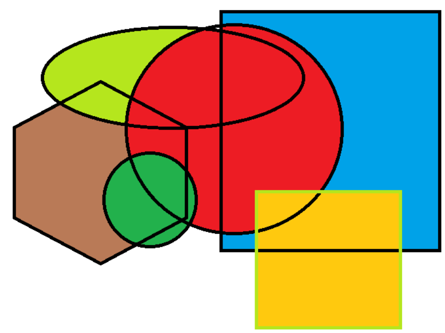

Figure 2: Graphics shown during the second disruption (Step 4) at 240 seconds of the test segment

During the experiment, users' emotion expressions are analyzed using Noldus Observer video analysis software http://www.noldus.com. The recordings are in sync with the machine generated emoticons, readily available for analysis (see next section 4.2).

4.2 Experimental results

The analysis of the experimental data was conducted to measure the effectiveness of emotion elicitation. The x-axis times for all graphs presented below are relative in seconds [s], for the whole duration of the test segment (330 seconds). The estimation is based on the emotion elicitation model (1) fitting. To detect the time when the emotion elicitation is present, we conducted the null hypothesis testing $H_0 = [R^2 = 0]$ at risk level $\alpha = 0.05$. The emotion elicitation is determined as present where the null hypotheses is rejected, and not present otherwise.

An example of valence and arousal ratings for a randomly selected subject is shown in Fig. 3.

The model (1) is fitted using linear regression on the measured data for the duration of the test segment. The data is

two primary dimensions for measuring emotion elicitation.

[3]To simplify the presentation of the experiment results. Note that similar results were found for the remaining subjects.

[4]Due to limited space, the two disruption parts of the experiment (Steps 3. and 4.) are not further discussed.

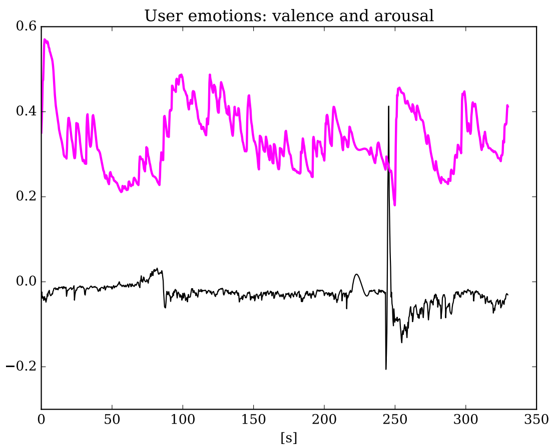

Figure 3: Valence (black line) and arousal (magenta, light line) ratings of learner's emotional state throughout the test segment.

sampled in a non-uniform manner due to the technical properties of the sensors (internal clocks of sensors are not sufficiently accurate, etc.). The data is approximated by continuous smooth B-splines of order 3, according to the upper frequency limit of measured phenomena, and uniformly sampled to time-align data (we skip re-sampling details here).

To fit the regression models the 40 past samples from the current (evaluation) time representing 4 seconds of real-time were used. These two value were selected as an optimum according to competitive arguments for more statistical power (requires more samples) and for enabling to detect time-dynamic changes in the effectiveness of emotion elicitation (requiring shorter time interval leading to less samples). Note that changing this interval from 3 to 5 seconds did not significantly affect the fitting results. Results are given in terms of R_V^2, R_A^2 representing the part of explained variance of valence and arousal when the elicitation is known, and in terms of a p_V, p_A-values testing the null hypothesis regression models $H_{0V} = [R_V^2 = 0]$, $H_{0A} = [R_A^2 = 0]$, respectively. The time dynamics of emotion elicitation is represented by p-values p_A and p_V on Fig. 4.

In order to estimate the effect of emotion elicitation, the percentages were computed on the number of times the elicitation was significant. The analyzed time intervals were uniformly sampled every 2 seconds. The results are shown in Table 1. It turned out that the test interval sampling had no significant impact on the results.

Table 1: Proportion q of the time when the measured emotion elicitation is significant. Notation red. q stands for the reduced efficiency, which is 5% lower than the measured one. Measured for the five selected test subjects.

	Valence		Arousal	
User Id	q %	red. q %	q %	red. q %
1	47.7	45.3	43.2	41.1
2	68.3	65.0	72.2	68.6
3	60.0	57.0	61.3	58.2
4	51.6	49.1	60.6	57.6
5	62.3	59.4	61.9	58.8

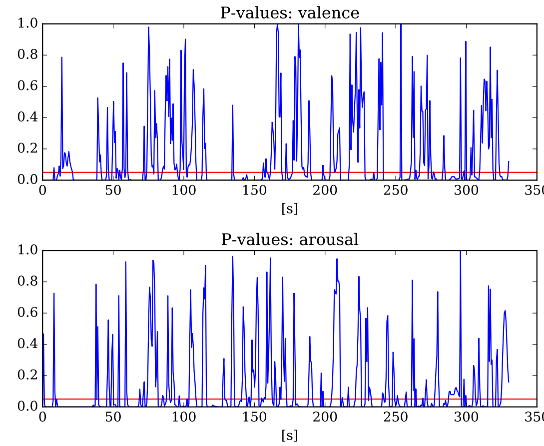

Figure 4: P-values for the null hypothesis testing $H_0 = [R^2 = 0]$ of emotion elicitation for a randomly selected subject, separately for valence (top) and arousal (bottom). The horizontal red line marks the risk level $\alpha = 0.05$, with p-values below the line indicating significant emotion elicitation effect.

We also analyzed the reduced percentages. These are 5% lower than the measured ones, since the significance testing was performed at a risk level $\alpha = 0.05$ and approximately 5% detections are false (type I. errors). Note that Bonferroni correction does not apply here. However, we nevertheless computed the above given percentages using Bonferroni correction and it turned out the percentages drop approximately to one half of the reported values.

The strength of emotion elicitation is shown in the linear regression model R^2 as a function of time (Fig. 5).

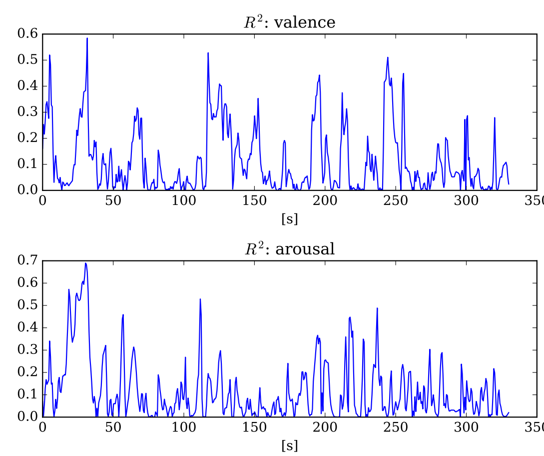

Figure 5: Linear regression model R^2 of emotion elicitation for a randomly selected test subject, separately for valence (top) and arousal (bottom).

The strength of emotion elicitation effect is significant, but also varies highly (Fig. 5). Similar results were detected among all test subjects. However, it is too early to draw any meaningful conclusions on the reasons for high variability at this stage, as many of the potential factors influencing emotion elicitation need further analysis.

To estimate the average strength of emotion elicitation, the average values of R^2 were computed for the five selected subjects (as in Table 1) – these values are part of the explained variance for learner emotions when the machine generated emotion is known. The average value of R^2 varies across test subjects from 18.3% to 24.5% for valence and 19.7% to 31.4% for arousal, for all time intervals (when significant or non-significant elicitation is present). If we average only over the time intervals when the elicitation is significant, the average value of R^2 varies across test subjects from 32.5% to 39.3% for valence and 36.3% to 44.9% for arousal (see Table 2).

Table 2: Average values for the explained variance for valence and arousal in %: for all time intervals and for the time intervals when emotion elicitation is significant. Measured for the five selected test subjects.

	Valence		Arousal	
User Id	All int.	Signif. int.	All int.	Signif. int.
1	18.3	32.5	19.7	36.3
2	19.4	33.8	27.4	39.2
3	24.5	39.3	31.4	44.9
4	19.8	33.3	23.9	39.9
5	21.7	35.4	26.8	40.2

Observe that there is considerably less variability among the subjects in terms of elicitation strength (average R^2), compared to the proportions of time the elicitation is significant (see Table 1).

5. CONCLUSION AND FUTURE WORK

The paper discussed the efficiency of emotion elicitation in socially intelligent services. The experiment was conducted using the socially intelligent typing tutor. The overall aim of the intelligent typing tutor is to elicit emotions and thus improve learning and engagement in the touch-typing training. Emotion elicitation is utilized together with the notion of positive reinforcement. The tutor is able to model and analyze learner's expressed emotions and measure the efficiency of emotion elicitation in the process. Experimental results show that the efficiency of emotion elicitation is significant, but at times also varies highly for the individual learner and moderately among learners.

Future work will focus on reasons for variations in emotion elicitation by analyzing potential factors, such as the effects of machine generated emotion expressions on emotion elicitation, learner's emotional state, cognitive load, attention, and engagement, among others.

6. REFERENCES

[1] E. B. Ahmed, A. Nabli, and F. Gargouri. A Survey of User-Centric Data Warehouses: From Personalization to Recommendation. *International Journal of Database Management Systems*, 3(2):59–71, 2011.

[2] K. Albrecht. *Social Intelligence: The New Science of Success*. Pfeiffer, 1 edition, February 2009.

[3] V. Aleven, B. Mclaren, I. Roll, and K. Koedinger. Toward meta-cognitive tutoring: A model of help seeking with a cognitive tutor. *International Journal of Artificial Intelligence in Education*, 16(2):101–128, 2006.

[4] J. Allanson and S. H. Fairclough. A research agenda for physiological computing. *Interacting with Computers*, 16(5):857–878, 2004.

[5] J. Allanson and G. Wilson. Physiological Computing. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, pages 21–42, 2002.

[6] N. Ambady and R. Rosenthal. Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 111:256–274, 1992.

[7] J. R. Anderson, A. T. Corbett, K. R. Koedinger, and R. Pelletier. Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, 4(2):167–207, 1995.

[8] Y. Ayzenberg, J. Hernandez, and R. Picard. FEEL: frequent EDA and event logging – a mobile social interaction stress monitoring system. *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts CHI EA 12*, page 2357, 2012.

[9] R. Azevedo, A. Witherspoon, A. Chauncey, C. Burkett, and A. Fike. MetaTutor: A MetaCognitive Tool for Enhancing Self-Regulated Learning. *Annual Meeting of the American Association for Artificial Intelligence Symposium on Metacognitive and Cognitive Educational Systems*, pages 14–19, 2009.

[10] P. Biswas and P. Robinson. A brief survey on user modelling in HCI. *Intelligent Techniques for Speech Image and Language Processing SpringerVerlag*, 2010.

[11] J. Bobadilla, F. Ortega, a. Hernando, and a. Gutiérrez. Recommender systems survey. *Knowledge-Based Systems*, 46:109–132, 2013.

[12] J. Broekens, T. Bosse, and S. C. Marsella. Challenges in computational modeling of affective processes. *IEEE Transactions on Affective Computing*, 4(3):242–245, 2013.

[13] J. A. Coan and J. J. Allen, editors. *Handbook of Emotion Elicitation and Assessment (Series in Affective Science)*. Oxford University Press, 1 edition, 4 2007.

[14] D. C. Derrick, J. L. Jenkins, and J. Jay F. Nunamaker. Design Principles for Special Purpose, Embodied, Conversational Intelligence with Environmental Sensors (SPECIES) Agents. *AIS Transactions on Human-Computer Interaction*, 3(2):62–81, 2011.

[15] S. D'mello and A. Graesser. Autotutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Trans. Interact. Intell. Syst.*, 2(4):23:1–23:39, Jan. 2013.

[16] D. C. Dryer. Dominance and valence: a two-factor model for emotion in HCI. In *Proceedings of AAAI Fall Symposium Series on Affect and Cognition*, pages 111 – 117. AAAI Press, 1998.

[17] J. Dunlap, D. Bose, P. R. Lowenthal, C. S. York, M. Atkinson, and J. Murtagh. What sunshine is to flowers: A literature review on the use of emoticons to support online learning. *Emotions, Design, Learning and Technology*, pages 1–17, 2015.

[18] S. H. Fairclough. Fundamentals of physiological computing. *Interacting with Computers*, 21(1-2):133–145, 2009.

[19] H. Gunes, B. Schuller, M. Pantic, and R. Cowie.

Emotion representation, analysis and synthesis in continuous space: A survey. *2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011*, pages 827–834, 2011.

[20] E. Hudlicka. Guidelines for Designing Computational Models of Emotions. *International Journal of Synthetic Emotions*, 2(1):26–79, 2011.

[21] X. Jiang, B. Zheng, R. Bednarik, and M. S. Atkins. Pupil responses to continuous aiming movements. *International Journal of Human-Computer Studies*, 83:1–11, 2015.

[22] A. C. K. Koedinger. *The Cambridge Handbook of the Learning Sciences*, chapter Cognitive tutors: Technology bringing learning sciences to the classroom, pages 61–78. Cambridge University Press, New York, 2006.

[23] J. F. Kihlstrom; and N. Cantor. Social Intelligence. In R. Sternberg, editor, *Handbook of intelligence*, pages 359–379. Cambridge, U.K.: Cambridge University, 2nd edition, 2000.

[24] T. Mahmood, G. Mujtaba, and A. Venturini. Dynamic personalization in conversational recommender systems. *Information Systems and e-Business Management*, pages 1–26, 2013.

[25] S. Marsella and J. Gratch. Computationally modeling human emotion. *Communications of the ACM*, 57(12):56–67, 2014.

[26] S. Marsella, J. Gratch, and P. Petta. Computational models of emotion. In *Blueprint for Affective Computing (Series in Affective Science)*. Oxford University Press, 2010.

[27] D. McDuff, S. Gontarek, and R. Picard. Remote Measurement of Cognitive Stress via Heart Rate Variability. *Proceedings of 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2957–2960, 2014.

[28] A. Mitrovic, B. Martin, and P. Suraweera. Intelligent tutors for all: The constraint-based approach. *IEEE Intelligent Systems*, 22(4):38–45, 2007.

[29] A. Neuringer. Operant variability: Evidence, functions, and theory. *Psychonomic Bulletin & Review*, 9(4):672–705, 2002.

[30] D. Novak, B. Beyeler, X. Omlin, and R. Riener. Workload estimation in physical human–robot interaction using physiological measurements. *Interacting with Computers*, page iwu021, 2014.

[31] A. Pentland. *Honest Signals: How They Shape Our World*. The MIT Press, 2008.

[32] J. L. Plass, S. Heidig, E. O. Hayward, B. D. Homer, and E. Um. Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction*, 29:128–140, 2014.

[33] V. P. Richmond, J. C. McCroskey, and M. L. Hickson III. *Nonverbal Behavior in Interpersonal Relations (7th Edition)*. Pearson, 7 edition, 4 2011.

[34] L.-F. Rodriguez and F. Ramos. Development of Computational Models of Emotions for Autonomous Agents : A Review. *Cognitive Computing*, pages 351–375, 2014.

[35] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[36] V. Sandulescu, S. Andrews, E. David, N. Bellotto, and O. M. Mozos. Stress Detection Using Wearable Physiological Sensors. *Artificial Computation in Biology and Medicine Lecture Notes in Computer Science*, pages 526–532, 2015.

[37] L. Shen, M. Wang, and R. Shen. Affective e-Learning: Using emotional data to improve learning in pervasive learning environment related work and the pervasive e-learning platform. *Educational Technology & Society*, 12:176–189, 2009.

[38] V. Shuman, E. Clark-Polner, B. Meuleman, D. Sander, and K. R. Scherer. Emotion perception from a componential perspective. *Cognition and Emotion*, 9931(November):1–10, 2015.

[39] V. Shuman, D. Sander, and K. R. Scherer. Levels of valence. *Frontiers in Psychology*, 4(MAY):1–17, 2013.

[40] T. Tijs, D. Brokken, and W. Ijsselsteijn. Creating an emotionally adaptive game. *Lecture Notes in Computer Science*, 5309 LNCS:122–133, 2008.

[41] M. K. Uhrig, N. Trautmann, U. Baumgärtner, R.-D. Treede, F. Henrich, W. Hiller, and S. Marschall. Emotion Elicitation: A Comparison of Pictures and Films. *Frontiers in psychology*, 7(February):180, 2016.

[42] J. van Dijk. Digital divide research, achievements and shortcomings. *Poetics*, 34(4-5):221–235, 2006.

[43] K. VanLehn, A. C. Graesser, G. T. Jackson, P. Jordan, A. Olney, and C. P. Rosé. When are tutorial dialogues more effective than reading? *Cognitive science*, 30:1–60, 2006.

[44] P. E. Vernon. Some characteristics of the good judge of personality. *The Journal of Social Psychology*, 4(1):42–57, 1933.

[45] A. Vinciarelli, M. Pantic, H. Bourlard, and A. Pentland. Social signals, their function, and automatic analysis: a survey. *Proceedings of the 10th international conference on Multimodal interfaces*, pages 61–68, 2008.

[46] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder. Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Transactions on Affective Computing*, 3(1):69–87, 2012.

[47] A. Vinciarelli and F. Valente. Social Signal Processing: Understanding Nonverbal Communication in Social Interactions. In *Proceedings of Measuring Behavior*, 2010.

[48] T. Wehrle, G. R. Scherer, and N. York. Towards Computational Modeling of Appraisal Theories. *Appraisal*, pages 350–365, 2001.

[49] V. Xia, N. Jaques, S. Taylor, S. Fedor, and R. Picard. Active learning for electrodermal activity classification. In *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–6. IEEE, 2015.

[50] Y.-c. Yeh, S. C. Lai, and C.-W. Lin. The dynamic influence of emotions on game-based creativity: An integrated analysis of emotional valence, activation strength, and regulation focus. *Computers in Human Behavior*, 55:817–825, 2016.