Emotion Elicitation in a Socially Intelligent Service: The Typing Tutor

Andrej Košir [1,†,*] and Gregor Strle [2,†]

[1] Faculty of Electrical Engineering, University of Ljubljana, Tržaška cesta 25, 1000 Ljubljana, Slovenia
[2] Institute of Ethnomusicology, Research Centre of the Slovenian Academy of Sciences and Arts, Novi trg 2, 1000 Ljubljana, Slovenia; gregor.strle@zrc-sazu.si
[*] Correspondence: andrej.kosir@fe.uni-lj.si; Tel.: +386-1-4768-354
[†] These authors contributed equally to this work.

Abstract: This paper presents an experimental study on modeling machine emotion elicitation in a socially intelligent service, the typing tutor. The aim of the study is to evaluate the extent to which the machine emotion elicitation can influence the affective state (valence and arousal) of the learner during a tutoring session. The tutor provides continuous real-time emotion elicitation via graphically rendered emoticons, as an emotional feedback to learner's performance. Good performance is rewarded by the positive emoticon, based on the notion of positive reinforcement. Facial emotion recognition software is used to analyze the affective state of the learner for later evaluation. Experimental results show the correlation between the positive emoticon and the learner's affective state is significant for all 13 (100%) test participants on the arousal dimension and for 9 (69%) test participants on both affective dimensions. The results also confirm our hypothesis and show that the machine emotion elicitation is significant for 11 (85%) of 13 test participants. We conclude that the machine emotion elicitation with simple graphical emoticons has a promising potential for the future development of the tutor.

1. Introduction

Human communication requires a certain level of social and emotional intelligence, including the ability to detect and express non-verbal communication cues (e.g., facial expressions, gestures, body posture, movement), which in turn provide sustenance for verbal communication in given context. Individual's emotional involvement through expressions of affect and emotions is an essential indicator of their attention and overall engagement in communication and provides further clues about their affective state, intentions, and attitudes [1–4].

For humans, the ability to sustain communication in real-time, and recognize, respond to, and elicit affective states, seems almost effortless. For machines, on the other hand, even a minimal level of social intelligence is much harder, if not impossible, to replicate [5]. The computational modeling of affective processes is extremely difficult—not only because of the procedural aspects related that also need to run in real-time (e.g., recognition, detection, processing, modeling, and generation of affect), but primarily due to the complexity of human behavior and multimodality of human interaction [6]. The interplay between behavioral and psychological aspects of human behavior and the idiosyncratic nature and variable expressions of human affect and emotions pose further challenges for the affective computing [7]. Any advances in these areas can improve human-machine interaction (HCI). From the

applied perspective the core research challenges are to produce for humans meaningful output [8–10] and more generally to "enable computers to better serve people's needs" [7].

Being able to recognize, measure, model human affect, and generate appropriate affective feedback to establish a more naturalistic human-machine-communication (HMC) is key to the design of intelligent services. As in human communication, the use of affect and emotions in HMC is necessary for measuring and improving user experience and overall engagement [11–14]. Naturalistic HMC can also minimize user-service adaptation procedures and maximize the engagement and the intended use of a service [15–17]. For a socially intelligent service to be successful, it should be capable of expressing some level of social and emotional intelligence. Such service is socially intelligent when it is capable of reading, detecting, measuring and estimating user's verbal and/or non-verbal social signals, producing machine-generated reasoning and emotional feedback, and sustaining HMC in real-time [18].

The paper presents an experimental study on machine emotion elicitation in a socially intelligent service, the touch-typing tutor. The main hypothesis is that learner's affective state can be influenced by the machine-generated emotion expressions. The underlying assumption, not addressed in this paper, is that machine emotion elicitation can contribute to the learner's increased performance, attention, motivation, and overall engagement with a system [19]. In this paper, we evaluate the effect of machine emotion elicitation on the learner's affective state during a touch-typing session. For this purpose, two emotion models are integrated into the tutor. The machine emotion model of the tutor provides continuous real-time machine emotion elicitation via graphically rendered emoticons, as an emotional feedback to the learner's performance during the session. Simultaneously, the user emotion model of the tutor captures and maps learner's facial emotion expressions onto a 2-dimensional affective space of valence and arousal, for the evaluation of emotion elicitation effect.

The paper is structured as follows. After the general overview of related work in Section 2, Section 3 presents the architecture and the emotion models of the intelligent typing tutor. Section 4 presents the design of the experimental study. The experimental results are presented in Section 5, focusing on the evaluation of the machine emotion model, the correlation between the machine generated emoticon and learner's valence-arousal space, and the influence of the machine emotion elicitation on the learner's affective state. The paper ends with general conclusions and future work in Section 6.

2. Related Work

Research and development of a fully functioning socially intelligent service are still at a very early stage. However, various components that can ultimately lead to such a service have been under intensive development for several decades. We briefly present them below.

2.1. Social Intelligence, Social Signals, and Nonverbal Communication Cues

Various definitions of social intelligence exist, including skills in social interactions, ability to decode social information, exhibiting behavioral flexibility and adaptivity in different social contexts, among others. Thorndike coined the phrase "social intelligence" and defined it as the ability to "to act wisely in human relations" [20]. Gardner later argued for a theory of multiple intelligence, a more nuanced view of social intelligence which requires multiple cognitive abilities to adapt and respond to a particular social situation [21]. For Cantor and Kihlstrom, social intelligence is an "organizing principle", "a fund of knowledge about the social world", where differences in social knowledge cause differences in social behavior [22,23]. A broader definition, also used in this paper, is given by Vernon [24], who defines social intelligence as 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". Social intelligence is also viewed as an ability to express and recognize social cues and

behaviors [25,26], including various non-verbal cues, such as gestures, postures and face expressions, exchanged during social interactions [27].

Social signals (cues) are extensively being analyzed in the field of HCI [27–29], often under different terminology. For example, [30] 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. In this paper, we will use the term "social signal".

Behavioral Measurements

The development of wearable sensors has enabled acquisition of user data in near real-time. As a result, notable advances can today be found in research on multimodal expressions of human behavior, especially in the fields of social computing, affective computing, and HCI, with the development of several novel measurement related procedures and techniques. For example, behavioral measurements are used to extend communication bandwidth and develop smart technologies [32], conversational intelligence based on the environmental sensors [33], and estimation of user behavior (e.g., emotion estimation) [34,35]. Several research works use behavioral measurements to study human behavior and usability aspects in regard to technology use, analyzing stress estimation [36], workload estimation [37], cognitive load estimation [38–40], and classification [41], among many others. A good example where aforementioned measurements all come into play is in intelligent tutoring systems [42], which are also relevant to the topic of this paper.

2.2. Intelligent Tutoring Systems

Intelligent tutoring systems (ITS) use technology to enable meaningful and effective learning. Research has shown that ITSs are nearly as effective as human tutoring [43]. ITSs often use affective computing and emotion mediation to improve user experience and performance of learning [44–47]. Moreover, several ITSs 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 [47]. AutoTutor/Affective AutoTutor employs both affective and cognitive modeling to support learning and engagement, tailored to the individual user [47]. Some other examples include: Cognitive Tutor [48]—an instruction based system for mathematics and computer science, Help Tutor [49]—a meta-cognitive variation of AutoTutor that aims to develop better general help-seeking strategies for students, MetaTutor [50]—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 [51], 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 [52–54] and release of stress and frustration [55,56].

2.3. Computational Models of Emotion

One of the core requirements for a service to be socially intelligent 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 [10,57,58], which are mainly based on the appraisal theories of emotion [8].

The field of affective computing has developed several approaches to modeling, analysis, and interpretation of human emotions [5]. One of the most known and widely used emotion annotation and representation models is the valence-arousal emotion space, based on Russell's valence-arousal model of affect [59]. The valence-arousal space is used in many human to machine interaction settings [60–62], 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 [63], or more recently, models for the automatic and continuous analysis of human emotional behavior [5].

Several challenges remain, most notably in design, training, and evaluation of computational models of emotion [64], critical analysis and comparison, and their relevance for other research fields (e.g., cognitive science, human emotion psychology) [10]. Recent research on emotions argues that traditional emotion models might be overly simplistic [9], pointing out the notion of emotion is multi-componential and includes "appraisals, behavioral activation, action tendencies, and motor expressions" [65]. There is a need for the "multifaceted conceptualization" in the modeling of emotions that can be linked to "qualitatively different types of evaluations" used in the appraisal theories [66].

2.3.1. Emotion Recognition From Facial Expressions

Research on emotions has demonstrated the central importance of facial expressions in recognition of emotions [1]. As a result, various classification systems for mapping facial expressions and emotions had been developed [6], most of them based on emotions studied by [67]. For example, Ekman and Friesen had developed the Facial Action Coding System (FACS), where observable components of facial movements are encoded in 44 different Action Units (AUs)—various combinations of AUs represent various facial expressions of emotion [68]. Several other alternatives exist, such as the MPEG-4 Facial Animation Parameters (FAP) standard for virtually representing humans and humanoids [69] and the FACS+, which is based on computer vision system for observing facial motion [70].

2.3.2. Emotion Elicitation Using Graphical Elements

Research on emotion elicitation using simple graphical elements (such as emoticons and smilies) is far less common than research on emotion elicitation with affective agents such as facial expressions, humanoid faces, or emotional avatars [71–76]. Whereas several studies on emotion elicitation use different stimuli (e.g., pictures, movies, music) [77,78] and behavioral cues [72], only a few tackle the challenges of affective graphical emoticons for purpose of emotion elicitation [74]. In the intelligent typing tutor, the learner's emotions are elicited by the graphical emoticons (smilies) via tutor's dynamic graphical user interface. The choice of smilies was due to their semantic simplicity, unobtrusiveness, and ease of continuous adaptation (various degrees of a smile) and measurement. Using pictures as emotional stimuli would add additional cognitive load, possibly elicit multiple emotions, and distract the learner from a task at hand. The presented approach also builds upon the findings of the existing research in affective computing and emotion elicitation, which shows that the recognition of emotions represented by affective agents is intuitive for humans [79], that human face-like graphics increase user engagement [80,81], and that emotion elicitation with simple emoticons can be significant [56,76]. The latter is discussed in Section 5.

3. Emotion Elicitation in a Socially Intelligent Service: The Intelligent Typing Tutor

The following sections discuss general requirements of a socially intelligent service, the architecture of the intelligent typing tutor, and the models used for emotion elicitation in our experimental study.

3.1. General Requirements for Emotion Elicitation in a Socially Intelligent Service

A socially intelligent service should be capable of performing the following elements of social intelligence in regard to emotion elicitation:

1.　Conduct behavioral measurements and gather data for relevant user behavior. Human emotions are conveyed via behavior and non-verbal communication cues, such as face expression, gestures, body posture, voice color, among others. Selected behavioral measurements (pupil size, acceleration of the wrist, etc.) are believed to be correlated with user's affective state and other non-verbal communication cues.

2. Analyze, estimate, and model user emotions and non-verbal (social) communication cues via computational models. Selected behavioral measurements are used as an input to the computational model of user emotions and for modeling machine emotion feedback.
3. Use emotion elicitation to reward performance and improve user engagement. For example, the notion of positive reinforcement could be integrated into a service to reward and improve performance and user engagement, taking into account user's temporary affective state and other non-verbal communication cues.
4. Naturalistic affective interaction. A continuous feedback loop between a user and a service, based on non-verbal communication cues.
5. Context and task-dependent adaptation: adapt a service according to the design goals. For example, in the intelligent typing tutor case study, the intended goal is to reward learner's performance and progress. The touch-typing lessons are carefully designed and adapt in regard to the typing speed and difficulty to meet user's capabilities, temporary affective 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 a socially intelligent service that uses emotion elicitation to reward learner's performance is provided below.

3.2. The Architecture of the Intelligent Typing Tutor

The overall goal of the typing tutor is to reward learner's performance and thus improve the process of learning touch-typing. For this purpose, machine generated emotion elicitation is used together with the notion of positive reinforcement to respond to and reward learner's performance. In its current version, the tutor employs simple emoticon-like graphics for the machine emotion elicitation and state-of-the-art sensor technology to continuously measure and analyze learner's behavior.

Typing tutor's main building blocks consist of:

1. Web GUI: to support typing lessons and machine-generated graphical emoticons (see Figure 1). Data is stored on the server for later analysis and human annotation procedures. Such architecture allows for crowd-sourced testing and efficient remote maintenance.
2. Sensors: to gather behavioral data and monitor learner's status in real-time. The recorded data is later used to evaluate the effectiveness of the machine generated emotion elicitation. The list of sensors integrated into the typing tutor environment includes:

    -   Keyboard input recorder: to monitor cognitive and locomotor errors that occur while typing;
    -   Wrist accelerometer and gyroscope: to trace hand movement;
    -   Emotion recognition software: to extract learner's facial emotion expressions in real-time;
    -   Eye tracking software: to measure eye gaze and pupil size, and estimate learner's attention and possible correlates to typing performance.

3. Machine emotion model: for modeling machine-generated emotion expressions (through graphical emoticons; see Section 3.3.1).
4. User emotion model: for measuring learner's emotion expressions (and performance) during a tutoring session (see Section 3.3.2).
5. Typing content generator: which uses learner's performance to recommend the appropriate level of typing lectures.

The intelligent typing tutor is publicly available as a client-server service running in a web browser [82].
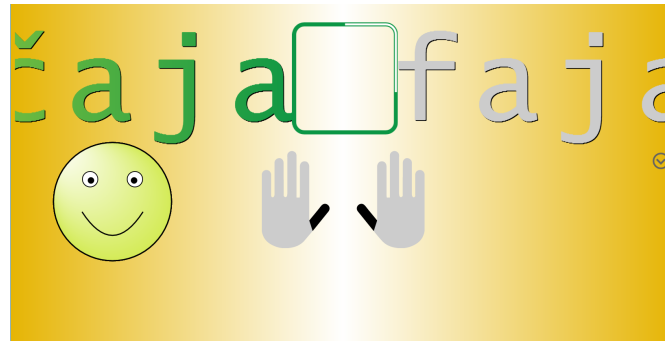
Figure 1. Socially intelligent typing tutor integrates touch-typing tutoring and machine-generated graphical emoticons (for emotion elicitation) via the graphical user interface.

3.3. Modeling Emotions in the Intelligent Typing Tutor

The typing tutor employs two emotion models: (a) the machine emotion elicitation model based on the graphical emoticons and (b) the user emotion model that maps learner's emotions onto a valence-arousal space. The role of emotion elicitation in the intelligent typing tutor is that of basic HMC and reward system. The underlying argument for the emotion elicitation is that by employing machine-generated emotion elicitation we can to some extent influence learner's affective state and thus reward their performance. The notion of positive reinforcement [83] was used in the design of the machine emotion model. According to the positive reinforcement theory, the rewarded behaviors will appear more frequently in future, and also improve learner's motivation and engagement [83–86]. In the tutor, the machine-generated positive emotion expressions (smilies) act as rewards for the learner's performance in the touch-typing practice. Learners are rewarded by the positive emotional response when they successfully complete a series of sub-tasks within a limited time (e.g., use correct fingering and keyboard input per letter). Negative reinforcement, where learners would be penalized for their mistakes, is not supported for two reasons. First, there is no clear indication that negative reinforcement would positively contribute to the learning experience and overall engagement, and second, such an approach would require additional dimension, adding additional complexity to the experimental study.

3.3.1. Machine Emotion Elicitation Model

The intelligent typing tutor uses the machine emotion elicitation model to reward learner's performance. It is based on the notion of positive reinforcement—the rewards come as positive emotional responses conveyed by the machine-generated graphical emoticon. The graphical emoticons were selected from the subset of the emoticons in Official Unicode Consortium code chart [87], via tutor's interface (as shown in Figure 1). These emotional responses range from neutral to positive (smiley) and act as stimuli to boost learner's performance.

The intensity of the emoticon Φ_m is defined by the interval $[0, 1]$, ranging from neutral to positive. The neutral emoticon is defined by the threshold eT that covers the lower 15% of the total emoticon intensity interval. This threshold eT represents the intensity of the neutral emoticon where the changes in its mouth curvature are visually still perceived as neutral. The threshold for the neutral emoticon was defined by the human observation of the Φ_m levels.

The machine emotion model was designed following the recommendations of the typing expert. The typing session starts with a neutral emoticon. As the session progresses, the emoticon intensity Φ_m updates gradually based on the performance values from the past time windows and the number of samples (see also Section 5). The performance is estimated as good or bad by the learner's typing errors and input delays (e.g., incorrect key input for the character presented on screen, the delay between the character presented and the key pressed, etc.). Note that the positive reinforcement (i.e. for the positive emoticon to come into effect and reward good performance) is not immediate and requires the learner

to show good performance over time. Such approach also avoids rapid switching between the different intensity levels. Consequently, there are short time intervals of good performances represented by the neutral emoticon, before the positive reinforcement starts.

To reward and improve learner's performance, as well as attention, motivation, and overall engagement, the machine-generated emotional feedback needs to function in real-time. To implement it successfully the designer needs to decide which behaviors need reinforcement and how often. However, in this early version, the positive reinforcement of the machine emotion model is linked exclusively to the learner's performance, with the aim to reward it. More realistic applications of the machine emotion elicitation model are planned to align the notion of positive reinforcement with other aspects, such as learner's affective state, attention, motivation, frustration level, and task-load, among others. Such an approach would also enable dynamic personalization, similar to the conversational RecSys [88].

3.3.2. User Emotion Model

The underlying argument for the emotion elicitation presented in this paper is that learner's affective state can be to some extent elicited by the machine-generated graphical emoticons.

The learner's affective state is measured with the valence-arousal model [59]. The user emotion model automatically measures learner's facial emotion expressions and maps them onto the valence-arousal space in real-time, using emotion recognition software Noldus FaceReader Version 6.1. [89].

Two independent linear regression models were used to analyze the emotion elicitation in the typing tutor, and to obtain the models' quality of fit and the proportion of the explained variance. The valence and arousal data gathered from learner's emotion expressions were fitted as dependent variables, whereas the machine-generated emoticon was fitted as an independent variable (Equation (1)). Analysis of the data showed the variation of the independent variable was sufficiently high to reliably estimate the goodness-of-fit measure R^2 (R^2 represents the coefficient of determination).

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

where Φ_m stands for the one dimensional parametrization of the intensity of the emoticon, for the interval $[0,1]$. Notations β_{1V} and β_{1A} are the emotion elicitation linear model coefficients for the dimensions of valence and arousal, β_{0V} and β_{0A} are the averaged effects of other influences on the dimensions of valence and arousal, and ε_V and ε_A are the independent variables of white noise.

There is no indication that emotion elicitation should be modeled as linear. Non-linear models might be more efficient at modeling emotion elicitation. However, the linear regression model was selected due to its efficient fitting procedure. Results show that the linear model can adequately capture the emotion elicitation process (see Section 5). Additionally, the residual plots (not reported here) showed that linear regression assumptions (homoscedasticity, the normality of residuals) are not violated.

4. Experimental Study

The experimental study consisted of 32 participants that completed the typing session, with the average duration approx. 17 min (1020 s) per session. The participants that did not complete the session, or the pre- and post- questionnaires, were excluded from the study. The participants were a homogeneous of the 4th year university female students (from the Faculty of Education, University of Ljubljana), with a similar level of typing expertise. The participants signed the experiment consent form and completed a short test with the pre-questionnaire inquiring about their typing technique and experience. The results showed none of the participants was skilled in touch-typing—they used

different typing patterns, required frequent visual (keyboard-to-screen) reference while typing, and engaged max. 3–4 fingers on each hand while typing. None of the participants had had previous experiences with a structured touch-typing practice.

For the typing session, the same set of expertly designed touch-typing lessons was given to all the participants. Additionally, the participants were required to complete four self-reports. Two self-reports were required from the participants before and after the experiment, inquiring about their attention, enthusiasm, tiredness, stress and task-load. Two identical in-session self-reports were given to the participants during the experiment, with three questions for rating their attention, task difficulty, and performance.

Additionally, the automatic interruptions (distractors) were integrated into the experiment to observe the participant's behavior during the touch-typing session in different attentional conditions (in line with the divided attention theory [90,91]). Three different experimental conditions were created: the primary task of touch-typing, the secondary task (given the distractor to the primary task), and switching from the primary to the secondary task and back. These experimental conditions were used to simulate the demanding modern working/learning environment where multitasking and interruptions (distractions) are very common.

The data was acquired in real-time using sensors and Noldus FaceReader emotion recognition software (as described in Section 3.2). For the analysis of results presented in the Section 5, 13 randomly selected participants were analyzed on the test segment of the overall experiment. The number of test participants is sufficient (comprising 40% of the participants) given the homogeneous study group. The main reason for not analyzing all 32 participants was due to the synchronization issues with our measuring system, which initially required manual synchronization of the data time stamps for each of the participants. Thus, the participants were first randomly selected and subsequently their data double-checked by manually. The duration of the test segment was 380 s, taken from the 6 to 12.2 min interval of the overall duration of the experiment (1020 s).

The test segment includes one in-session self-report and both interruptions, and is composed of the following steps:

1. Before the experiment: the instructions are given to the test users, who are informed by the personnel about the goal and the procedure of the experiment. Setting up sensory equipment;
2. At 00 s: the start of the experiment: the wrist accelerometer and the video camera are set on, and the experimental session starts;
3. At 30 s: the self-report with three questions, inquiring about the participant's attention, task difficulty, and performance. (The results of the self-reports are not presented in this paper.);
4. At 60 s: the machine-generated sound disruption of the primary task: "Name the first and the last letter of the word: mouse, letter, backpack, clock";
5. At 240 s: the machine-generated sound disruption of the primary task: "Name the color of the smallest circle" on the image provided in the tutor's graphical user interface. This cognitive task is expected to significantly disrupt the learner's attention away from the typing exercise;
6. The test segment ends at 380 s.

The following section presents the results of the experimental study and the evaluation of the emotion elicitation in the intelligent typing tutor.

5. Results of the Experimental Study

The analysis of the experimental data was conducted to measure the effect of emotion elicitation in the tutor. The estimation is based on the emotion elicitation model (1) fitting. The model (1) is fitted using linear regression on the measured data for the duration of the test segment. R^2s are computed for each discrete time instant of the test segment, using the window reaching 4 s (40 samples) in the past. Valence and arousal dimensions are treated as dependent variables, whereas the machine emotion elicitation expressed through the graphical emoticon is treated as the independent variable ranging from 0 (neutral elicitation) to 1 (maximum positive elicitation). R^2s for the two modes of

emotion elicitation are obtained—R^2 for the neutral elicitation with the emoticon threshold eT close to 0 (see Section 3.3.1), and R^2 for the positive elicitation in all other cases.

The data is 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 (for brevity, we skip resampling details). Mann-Whitney U test was used for the statistical hypothesis testing, for the analysis of valence and arousal, and for the comparison of mean R^2s. The significance level used in all tests was $\alpha = 0.05$.

The results of the experimental study are presented in the following sections. The results in Section 5.1 are presented in the context of the "positive reinforcement", the term we use throughout this paper to define the connection between the emoticon and the performance. We present the dynamics of the learner's affective state throughout the test segment. The supporting tables and figures show differences in R^2s and significance levels for the valence and arousal dimensions, and error rates and time interval durations, given the positive reinforcement. Section 5.2 tests our main hypothesis. It introduces additional criteria to conduct the measurement of the machine emotion elicitation and evaluate its influence on the affective state of the learner.

5.1. Measuring Learner's Affective State

The machine emotion elicitation model is based on the notion of positive reinforcement, defined by the learner's typing performance. The performance is estimated as good or bad based on the learner's typing errors and input delays and analyzed within a predefined time window (see Section 3.3.1).

To fit the regression models the 40 past samples from the time window of 4 s of real-time were used (see the previous section). These two values were selected as an optimum to the competitive arguments for more statistical power (requires more samples) and for enabling the detection of the time-dynamic changes in the effectiveness of emotion elicitation (requiring shorter time intervals, leading to fewer samples). Note that changing this interval from 3 to 5 s did not significantly affect the fitting results. Results are given in terms of R_V^2, R_A^2 representing the part of the explained variance for valence and arousal, and in terms of the p_V, p_A-values testing the null hypothesis regression models $H_{0V} = [R_V^2 = 0]$, $H_{0A} = [R_A^2 = 0]$, given the emoticon. To detect the time intervals when the affective state of the learner changes due to the positive reinforcement, we conducted the null hypothesis testing $H_0 = [R^2 = 0]$. The influence of the positive reinforcement is determined as the present where the null hypothesis is rejected.

Figure 2 and Table 1 show the valence-arousal ratings for five test participants. Note that, due to space constraints of this paper, the results presented in the figures are limited to five test participants. The X-axis times for all the figures are relative in seconds [s] and show the overall duration of the test segment. The figure shows the variations in the learner's affective state through time, with the higher values for the arousal dimension for the presented test participants.

Table 1 shows the average values of R^2s for the test participants—these values are part of the explained variance for the learner's affective state during the tutoring session. The average R^2s vary across test participants from 18.3% to 24.5% for valence and 19.7% to 31.4% for arousal, for all time intervals. If we average only over the time intervals when the positive reinforcement is active, the average value of R^2s varies across the participants from 32.5% to 39.3% for valence and 36.3% to 44.9% for arousal.
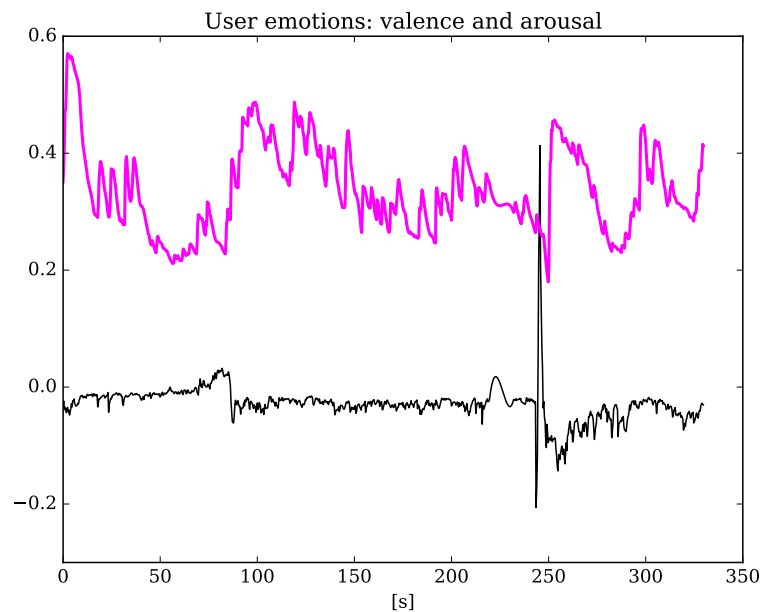
Figure 2. Learner's affective state: valence (black line) and arousal (magenta, light line) ratings for five test participants throughout the test segment.

Table 1. Average values of the explained variance for valence and arousal in % for all test participants. The table compares values for all the intervals vs. the time intervals when the positive reinforcement is active.

uID	Valence		Arousal	
	All int	Pos. reinf.	All int	Pos. reinf.
3	20.8	34.6	28.1	39.9
8	24.1	38.2	30.9	43.4
10	18.3	32.5	19.7	36.3
15	15.7	29.7	17.9	33.1
16	20.4	33.2	26.9	38.7
18	21.8	35.8	34.6	47.8
19	19.4	33.8	27.4	39.2
20	24.5	39.3	31.4	44.9
22	17.4	32.6	18.9	35.0
26	23.6	38.5	29.2	42.2
27	19.8	33.3	23.9	39.9
30	25.1	40.1	31.9	45.1
34	21.7	35.4	26.8	40.2

The learner's affective state given the positive reinforcement is represented by the *p*-values p_A and p_V in Figure 3.

The linear regression model (R^2) as a function of time for the valence and arousal dimensions is presented in Figure 4. The figure shows that the effect of positive reinforcement varies highly. Similar results were detected among all participants. The variations on the valence and arousal dimensions are probably due to general difficulties of capturing unstructured and noisy behavioral data, especially when using automatic emotion recognition from facial expressions in a real-world scenario. However, it is too early to draw any meaningful conclusion on the reasons for the high variations at this stage, as other potential factors influencing the measurement of learner's affective state (such as their emotion expressivity, task load, stress level, etc.) need to be evaluated.
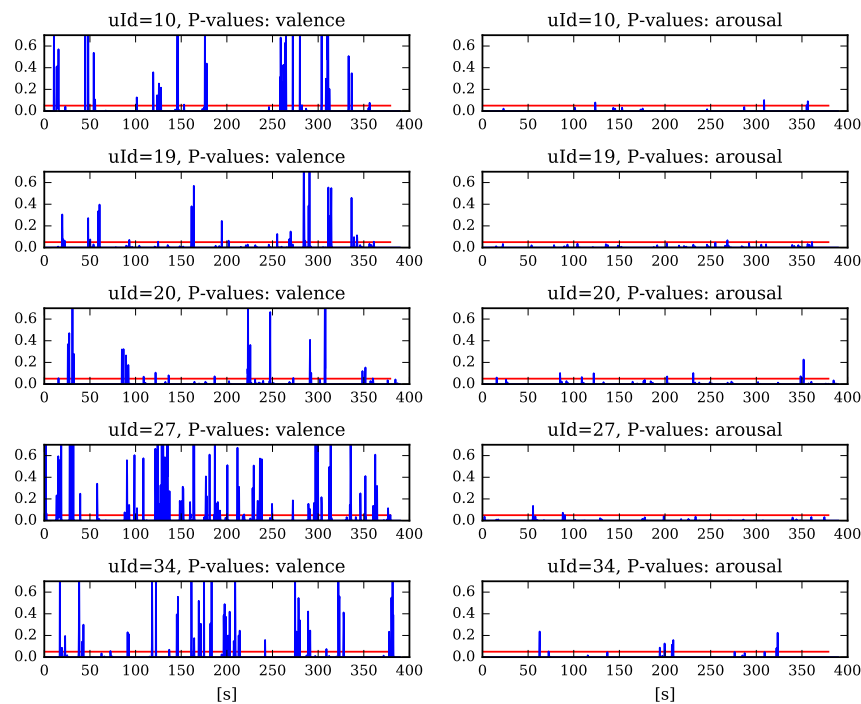
Figure 3. *P*-values for the null hypothesis testing $H_0 = [R^2 = 0]$ showing the effect of the positive reinforcement for five test participants, separately for valence (left) and arousal (right). The horizontal red line marks the significance level $\alpha = 0.05$, with *p*-values below the line indicating significant effect.
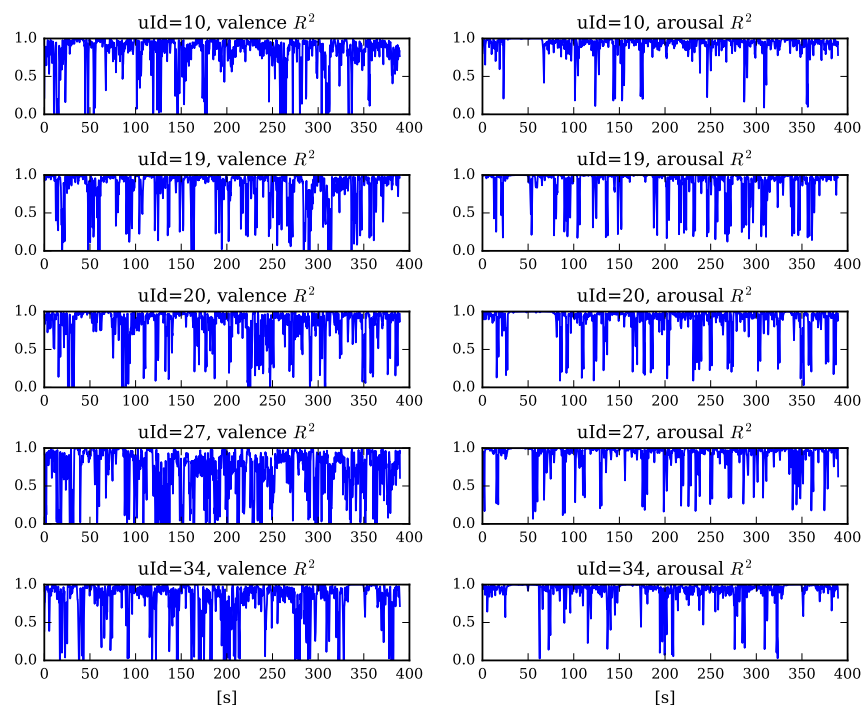

Figure 4. Linear regression model R^2 as a function of time for valence (left) and arousal (right), given the positive reinforcement, for five test participants.

Histograms (Figures 5 and 6) show the durations for the time intervals when the positive reinforcement is active vs. inactive, for the valence and arousal dimensions respectively. The results show longer time intervals (longer duration) for instances where the reinforcement is active: on average 18.34 s for valence and 25.64 s for arousal, compared to the average durations of 2.54 s for valence and 2.09 s for arousal when reinforcement is inactive. Such large differences in the average durations are due to the design of the positive reinforcement in the machine emotion model, which is based exclusively on the learner's performance.
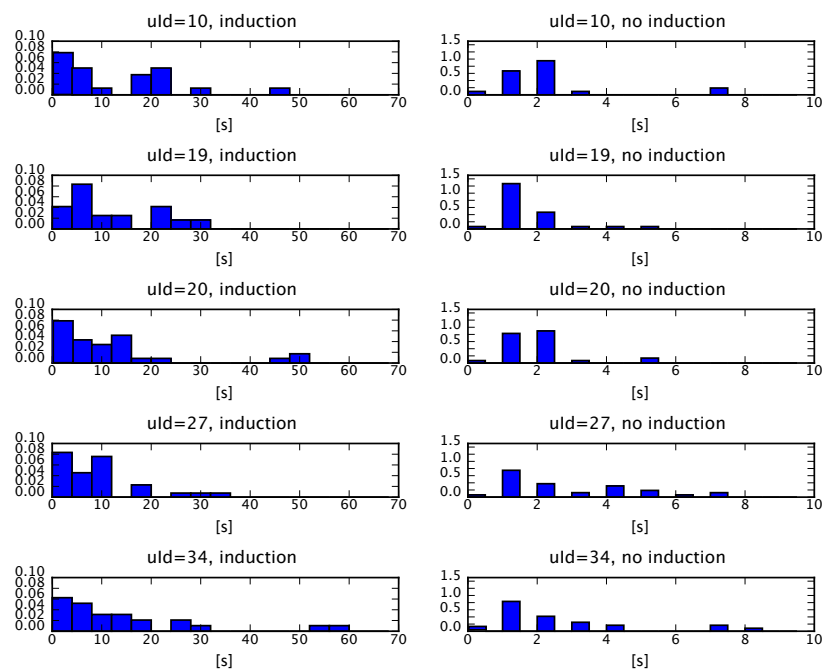

Figure 5. Histogram of the time interval lengths for the active vs. inactive reinforcement: valence.
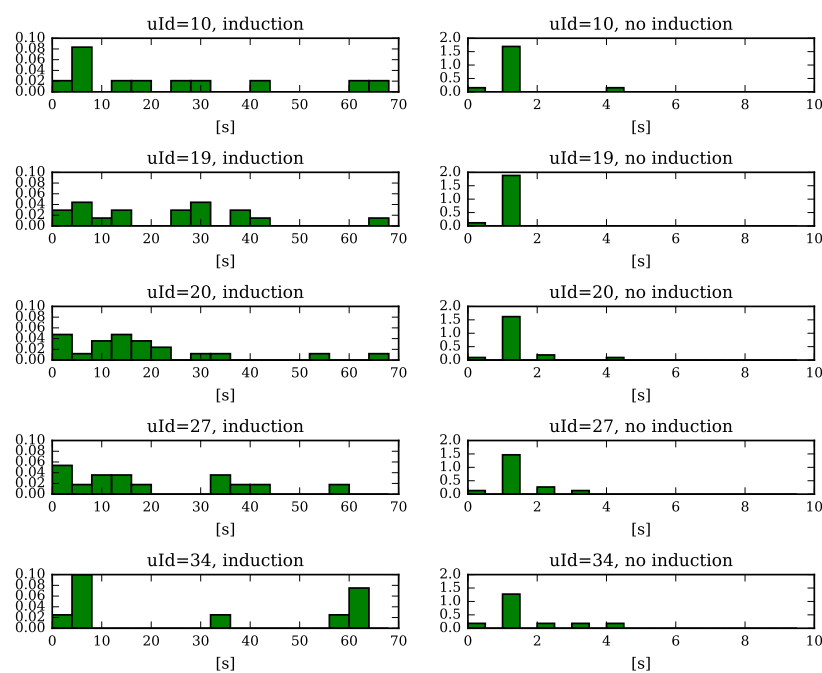

Figure 6. Histogram of the time interval lengths for the active vs. inactive reinforcement: arousal.

Additional insight is given by the analysis of the learners' error rates throughout the test segment, presented in Figure 7. Again, these results represent the underlying design of the machine emotion model, which only rewards learner's performance. As shown in Figure 7, the average number of errors (through time) is higher for the intervals with no reinforcement (defined by the neutral emoticon (red dot)), compared to the instances when reinforcement is active (positive emoticon (blue dot)). However, note that the occasional short time intervals of good performance are also represented by the neutral emoticon (red dot), due to the delay in the positive reinforcement (see Section 3.3.1). Missing values are due to the self-report (at 30 s) or the default interruption in the tutoring session (see Section 4), or when the typing was paused by the learner.
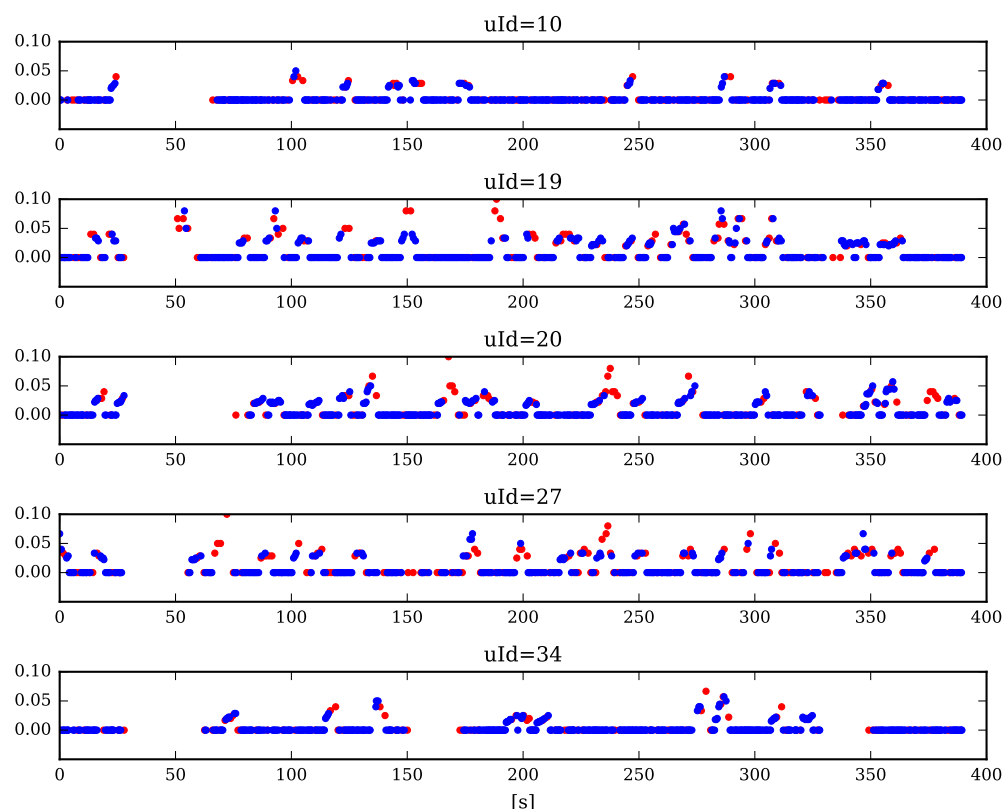

Figure 7. The average number of errors per second for the 5 s time interval: neutral (red) vs. positive (blue) emoticon state representing inactive vs. active positive reinforcement.

Presented results give some insight into the dynamics between the learner's affective state and the performance-based positive reinforcement of the machine emotion model. However, these results do not offer a reasonable proof for the influence of the machine emotion elicitation on the learner's affective state. It is impossible to infer what exactly influences learner's affective state—the emoticon or the performance—as they conflate the emoticon with the typing performance. We try to ameliorate this issue in the following section.

5.2. The Influence of the Machine Emotion Elicitation on the Learner's Affective State

The machine emotion elicitation model uses error-based performance evaluation as the underlying mechanism to generate the corresponding neutral and positive emoticon responses. This presents inherent limitation: the influence of machine emotion elicitation on learner's affective state is obfuscated by the typing performance. As the machine emotion elicitation is defined exclusively by the learner's typing performance, it is difficult to directly measure the effect that the machine emotion elicitation has on the learner's affective state. That is, it is difficult to argue that the learner's

affective state is influenced by the machine generated emoticons and not by the learner's good or bad performance. To test our hypothesis, additional measures are needed.

In this section, additional criteria are given that enable the measurement of the machine emotion elicitation effect on the learner's affective state. We do this by (to an extent) isolating the effect of the typing performance from the effect of the emoticons on the learner's affective state. The argument goes as follows. Emotion elicitation in the touch-typing tutor is defined by the learner's (good vs. bad) performance. Good performance is associated with the positive elicitation and represented by the positive emoticon, whereas the neutral elicitation is represented by the neutral emoticon and by both, bad and good performance. As noted in Section 3.3.1, the threshold for the positive reinforcement to come into effect and reward learner's good performance (with the positive emoticon) is not immediate. Consequently, there are time intervals of neutral elicitation, where the learner's performance is good, but not yet rewarded (the emoticon remains neutral), before the positive reinforcement comes in effect. Thus, we define the neutral elicitation by the instances of good performance where the emoticon is still in its neutral state. This way, we eliminate the effect of bad performance and simultaneously reduce the influence of good performance on the learner's affective state—the learner was already performing well even before the positive reinforcement came in effect.

The procedure goes as follows. First, we define the error-based threshold for the good performance, based on the expert advice and as designed in the machine emotion model. The regression model described in Section 5.1 is used for the fitting. The performance threshold is defined by the number of errors, given the time window and the number of samples. This step is crucial to avoid sampling of the isolated error-free instances, which would skew the results. Based on the predefined error threshold ($< 20\%$ error-ratio), the instances of bad performance are filtered out. Next, we define the emoticon threshold eT for the neutral and positive emoticons (see also Section 3.3.1). In the machine emotion model, the threshold for the neutral emoticon eT is set to proportionally represent the lower 15% of the total emoticon intensity Φ_m range $[0, 1]$. This range represents the intensity of the neutral emoticon where the changes in its mouth curvature are visually still perceived as neutral. These criteria isolate the effect of the typing performance from the effect of the emoticons by eliminating the instances of bad performance and only measuring the instances of positive and neutral emotion elicitation where the learner's performance was good.

Table 2 shows the p-values for the null hypothesis $H_0 = [R^2 = 0]$, confirming the correlation between the independent variable (positive emoticon) and the learner's affective state for valence (P_p-val.) and arousal (A_p-val.). The results show the correlation between the positive emoticon and the learner's affective state is significant for all 13 (100%) test participants on the arousal dimension and for 9 (69%) test participants on both dimensions. The hypothesis arguing the influence of the machine emotion elicitation on the learner's affective state is confirmed below.

Table 2. Correlation between the positive emoticon and the valence-arousal space: p-values for the null hypothesis $H_0 = [R^2 = 0]$.

uID	Val. p-val	Arou. p-val.
3	0.03	<0.01
8	<0.01	<0.01
10	<0.01	<0.01
15	0.069	<0.01
16	<0.01	<0.01
18	0.073	<0.01
19	<0.01	<0.01
20	<0.01	<0.01
22	0.054	<0.01
26	0.077	<0.01
27	0.033	<0.01
30	<0.01	<0.01
34	0.015	<0.01

Table 3 presents the mean valence and arousal values given the neutral (m(V)_neut., m(A)_neut) and the positive emoticon (m(V)_pos., m(A)_pos.), and the significance of the positive emoticon for valence and arousal (V_p-val, A_p-val).

Table 3. Machine elicitation for valence and arousal.

uID	m(V) neu	m(V) poz	p val	m(A) neu	m(A) poz	p val
3	0.094	0.012	<0.01	0.306	0.33	0.032
8	−0.097	−0.145	0.499	0.26	0.27	0.291
10	−0.031	−0.119	<0.01	0.308	0.289	<0.01
15	−0.01	0.01	0.085	0.37	0.31	<0.01
16	0.079	0.027	0.026	0.271	0.253	0.049
18	0.067	0.025	<0.01	0.274	0.269	0.199
19	−0.188	−0.241	0.045	0.319	0.342	<0.01
20	0.07	0.074	0.252	0.282	0.295	0.459
22	0.056	0.036	<0.01	0.277	0.285	0.353
26	0.114	0.015	<0.01	0.277	0.25	<0.01
27	−0.004	−0.026	<0.01	0.299	0.305	0.367
30	0.004	−0.012	0.01	0.262	0.259	0.199
34	−0.108	−0.1	0.28	0.297	0.365	0.04

The results presented in Table 3 show that the machine emotion elicitation is significant for 11 (85%) of 13 test participants: among these, for 9 (82%) test participants along the valence dimension and for 7 (64%) test participants along the arousal dimension. Moreover, 5 (45%) of the participants that exhibited significant elicitation effect (uID3, uID10, uID16, uID19, and uID26) were significantly influenced on both affective dimensions. Assuming the worst case scenario that none of the 19 remaining participants exhibit significant elicitation effect, the machine emotion elicitation is still significant with $p < 0.05$ for 34% of the participants that had taken part in the study.

6. Conclusions and Future Work

The paper evaluated the role of machine emotion elicitation in the intelligent typing tutor. The machine elicitation is based on the notion of positive reinforcement and rewards learner's good performance with emotional feedback via graphical emoticons. The main hypothesis of the paper argued that the machine emotion elicitation can influence learner's affective state during the tutoring session.

We conducted an experimental study with 32 participants and randomly selected 13 participants for the analysis of the experimental results. Given the machine emotion elicitation is performance-based, additional criteria had to be defined to enable the measurement of the emotion elicitation effect on the affective state of the learner. These criteria eliminate the effect of bad performance and simultaneously reduce the influence of good performance on the learner's affective state. Under these criteria, the linear regression model used for the evaluation only samples good performance instances for both modes of (neutral and positive) emotion elicitation.

The experimental results confirm the correlation between the positive emoticon and the learner's affective state with significant correlation for all 13 (100%) test participants on the arousal dimension and for 9 (69%) test participants on both dimensions. The results also confirm our hypothesis that the machine emotion elicitation can influence the affective state of a learner during a tutoring session. The machine emotion elicitation was significant for 11 (85%) test participants: for 9 (82%) on the valence dimension, for 7 (64%) on the arousal dimension, and for 5 (45%) test participants on both affective dimensions.

Capturing any type of behavioral data is difficult. The performance-based emotion elicitation varied strongly and frequently throughout the test segment for all test participants and on both affective dimensions. This is at least in part associated with the difficulty of capturing unstructured

and noisy behavioral data. Future work will focus on the reasons for high variation in the machine emotion elicitation. It is also necessary to consider alternative modes of machine emotion elicitation beyond simple graphic emoticons, as well as design more realistic applications to align the notion of positive reinforcement with other aspects, such as attention, motivation, frustration level, and task-load, among others.

Author Contributions: The authors contributed equally to the reported research and writing of the paper. Both authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ekman, P. Facial expression and emotion. *Am. Psychol.* **1993**, *48*, 384–392.
2. Horstmann, G. What do facial expressions convey: Feeling States, Behavioral Intentions, or Action Requests? *Emotion* **2003**, *3*, 150–166.
3. Tomasello, M. *Origins of Human Communication*; MIT Press: Cambridge, MA, USA, 2010.
4. Yiend, J. The effects of emotion on attention: A review of attentional processing of emotional information. *Cognit. Emot.* **2010**, *24*, 3–47.
5. Gunes, H.; Schuller, B.; Pantic, M.; Cowie, R. Emotion representation, analysis and synthesis in continuous space: A survey. In Proceedings of the 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011, Barbara, CA, USA, 21–25 March 2011; pp. 827–834.
6. Cowie, R.; Douglas-Cowie, E.; Tsapatsoulis, N.; Votsis, G.; Kollias, S.; Fellenz, W.; Taylor, J.G. Emotion recognition in human-computer interaction. *IEEE Signal Process. Mag.* **2001**, *18*, 32–80.
7. Picard, R.W. Affective computing: Challenges. *Int. J. Hum. Comput. Stud.* **2003**, *59*, 55–64.
8. Wehrle, T.; Scherer, G.R.; York, N. *Towards Computational Modeling of Appraisal Theories*; Oxford University Press: New York, NY, USA, 2001; pp. 350–365.
9. Broekens, J.; Bosse, T.; Marsella, S.C. Challenges in computational modeling of affective processes. *IEEE Trans. Affect. Comput.* **2013**, *4*, 242–245.
10. Marsella, S.; Gratch, J. Computationally modeling human emotion. *Commun. ACM* **2014**, *57*, 56–67.
11. Roda, C.; Thomas, J. Attention aware systems: Theories, applications, and research agenda. *Comput. Hum. Behav.* **2006**, *22*, 557–587.
12. Chua, R.; Weeks, D.J.; Goodman, D. Perceptual-Motor Interaction: Some Implications for Human-Computer Interaction. In *The Human-Computer Interaction Handbook*; Jacko, J.A., Sears, A., Eds.; CRC Press: Boca Raton, FL, USA, 2008; pp. 23–34.
13. Peters, C.; Castellano, G.; de Freitas, S. An exploration of user engagement in HCI. In Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots, Boston, MA, USA, 6 November 2009; pp. 9:1–9:3.
14. Marko, T.; Berardina, C.D.; Marco, G.; Ante, O.; Andrej, K. (Eds.) *Emotions and Personality in Personalized Services: Models, Evaluation and Applications*; Human-Computer Interaction Series; Springer: New York, NY, USA, 2016.
15. Biswas, P.; Robinson, P. A brief survey on user modelling in HCI. In *Intelligent Techniques for Speech Image and Language Processing*; Springer: New York, NY, USA, 2010.
16. Ahmed, E.B.; Nabli, A.; Gargouri, F. A Survey of User-Centric Data Warehouses: From Personalization to Recommendation. *Int. J. Database Manag. Syst.* **2011**, *3*, 59–71.
17. Bobadilla, J.; Ortega, F.; Hernando, A.; Gutiérrez, A. Recommender systems survey. *Knowl.-Based Syst.* **2013**, *46*, 109–132.
18. Vinciarelli, A.; Pantic, M.; Bourlard, H. Social signals, their function, and automatic analysis: A survey. In Proceedings of the 10th International Conference on Multimodal Interfaces (ICMI '08), Crete, Greece, 20–22 October 2008; pp. 61–68.
19. Harris, R.B.; Paradice, D. An Investigation of the Computer-mediated Communication of Emotions. *J. Appl. Sci. Res.* **2007**, *3*, 2081–2090.

20. Thorndike, E.L. Intelligence and its use. *Harper's Mag.* **1920**, *140*, 227–235.

21. Gardner, H. *Frames of Mind: The Theory of Multiple Intelligences*; Basic books Press: New York, NY, USA, 1984.

22. Cantor, N.; Kihlstrom, J. *Personality and Social Intelligence*; Century Psychology Series; Prentice Hall: Upper Saddle River, NJ, USA, 1987.

23. Kihlstrom, J.F.; Cantor, N. Social intelligence. In *The Cambridge Handbook of Intelligence*; Cambridge University Press: Cambridge, UK, 2011; pp. 564–581.

24. Vernon, P.E. Some Characteristics of the Good Judge of Personality. *J. Soc. Psychol.* **1933**, *4*, 42–57.

25. Albrecht, K. *Social Intelligence: The New Science of Success*, 1st ed.; John Wiley & Sons: Hoboken, NJ, USA, 2009.

26. Ambady, N.; Rosenthal, R. Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychol. Bull.* **1992**, *111*, 256–274.

27. Vinciarelli, A.; Valente, F. Social Signal Processing: Understanding Nonverbal Communication in Social Interactions. In Proceedings of the Measuring Behavior, Eindhoven, The Netherlands, 24–27 August 2010.

28. Pantic, M.; Sebe, N.; Cohn, J.F.; Huang, T. Affective multimodal human-computer interaction. In Proceedings of the 13th Annual ACM International Conference on Multimedia—MULTIMEDIA '05, Singapore, 6–11 November 2005; pp. 669–676.

29. Vinciarelli, A.; Pantic, M.; Heylen, D.; Pelachaud, C.; Poggi, I.; D'Errico, F.; Schroeder, M. Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Trans. Affect. Comput.* **2012**, *3*, 69–87.

30. Richmond, V.P.; McCroskey, J.C.; Hickson, M.L., III. *Nonverbal Behavior in Interpersonal Relations*, 7th ed.; Allyn & Bacon: Boston, MA, USA, 2011.

31. Pentland, A. *Honest Signals: How They Shape Our World*; MIT Press: Cambridge, MA, USA, 2008.

32. Fairclough, S.H. Fundamentals of physiological computing. *Interact. Comput.* **2009**, *21*, 133–145.

33. Derrick, D.C.; Jenkins, J.L.; Jay F. Nunamaker, J. Design Principles for Special Purpose, Embodied, Conversational Intelligence with Environmental Sensors (SPECIES) Agents. *AIS Trans. Hum.-Comput. Interact.* **2011**, *3*, 62–81.

34. Allanson, J.; Wilson, G. Physiological Computing. In Proceedings of the CHI '02 Extended Abstracts on Human Factors in Computing Systems, Minneapolis, MN, USA, 20–25 April 2002; pp. 21–42.

35. Allanson, J.; Fairclough, S.H. A research agenda for physiological computing. *Interact. Comput.* **2004**, *16*, 857–878.

36. Sandulescu, V.; Andrews, S.; David, E.; Bellotto, N.; Mozos, O.M. Stress Detection Using Wearable Physiological Sensors. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*; Springer: Heidelberg/Berlin, Germany, 2015; pp. 526–532.

37. Novak, D.; Beyeler, B.; Omlin, X.; Riener, R. Workload estimation in physical human—Robot interaction using physiological measurements. *Interact. Comput.* **2014**, *27*, 616–629.

38. McDuff, D.; Gontarek, S.; Picard, R. Remote Measurement of Cognitive Stress via Heart Rate Variability. In Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Chicago, IL, USA, 26–30 August 2014; pp. 2957–2960.

39. Ayzenberg, Y.; Hernandez, J.; Picard, R. FEEL: frequent EDA and event logging—A mobile social interaction stress monitoring system. In Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems Extended Abstracts CHI EA 12, Austin, TX, USA, 5–10 May 2012; p. 2357.

40. Jiang, X.; Zheng, B.; Bednarik, R.; Atkins, M.S. Pupil responses to continuous aiming movements. *Int. J. Hum.-Comput. Stud.* **2015**, *83*, 1–11.

41. Xia, V.; Jaques, N.; Taylor, S.; Fedor, S.; Picard, R. Active learning for electrodermal activity classification. In Proceedings of the 2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA, 12 December 2015; pp. 1–6.

42. Mello, S.K.D.; Craig, S.D.; Gholson, B.; Franklin, S.; Picard, R.; Graesser, A.C. Integrating Affect Sensors in an Intelligent Tutoring System. Available online: https://goo.gl/3EXu8R (accessed on 29 March 2017).

43. Vanlehn, K. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educ. Psychol.* **2011**, *46*, 197–221.

44. Partala, T.; Surakka, V. The effects of affective interventions in human-computer interaction. *Interact. Comput.* **2004**, *16*, 295–309.

45. Hascher, T. Learning and emotion: Perspectives for theory and research. *Eur. Educ. Res. J.* **2010**, *9*, 13–28.

46. Moridis, C.N.; Economides, A.A. Affective learning: Empathetic agents with emotional facial and tone of voice expressions. *IEEE Trans. Affect. Comput.* **2012**, *3*, 260–272.

47. D'mello, S.; Graesser, A. AutoTutor and Affective Autotutor: Learning by Talking with Cognitively and Emotionally Intelligent Computers That Talk Back. *ACM Trans. Interact. Intell. Syst.* **2013**, *2*, 23:1–23:39.

48. Anderson, J.R.; Corbett, A.T.; Koedinger, K.R.; Pelletier, R. Cognitive Tutors: Lessons Learned. *J. Learn. Sci.* **1995**, *4*, 167–207.

49. Aleven, V.; Mclaren, B.; Roll, I.; Koedinger, K. Toward meta-cognitive tutoring: A model of help seeking with a cognitive tutor. *Int. J. Artif. Intell. Educ.* **2006**, *16*, 101–128.

50. Azevedo, R.; Witherspoon, A.; Chauncey, A.; Burkett, C.; Fike, A. MetaTutor: A MetaCognitive Tool for Enhancing Self-Regulated Learning. Available online: https://goo.gl/dUMZo9 (accessed on 29 March 2017).

51. Mitrovic, A.; Martin, B.; Suraweera, P. Intelligent Tutors for All: The Constraint-Based Approach. *IEEE Intell. Syst.* **2007**, *22*, 38–45.

52. Shen, L.; Wang, M.; Shen, R. Affective e-Learning: Using emotional data to improve learning in pervasive learning environment related work and the pervasive e-learning platform. *Educ. Technol. Soc.* **2009**, *12*, 176–189.

53. Koedinger, K.R.; Corbett, A. Cognitive tutors: Technology bringing learning sciences to the classroom. In *The Cambridge Handbook of the Learning Sciences*; Sawyer, R.K., Ed.; Cambridge University Press: New York, NY, USA, 2006; pp. 61–78.

54. VanLehn, K.; Graesser, A.C.; Jackson, G.T.; Jordan, P.; Olney, A.; Rosé, C.P. When are tutorial dialogues more effective than reading? *Cogn. Sci.* **2006**, *30*, 1–60.

55. Hone, K. Empathic agents to reduce user frustration: The effects of varying agent characteristics. *Interact. Comput.* **2006**, *18*, 227–245.

56. Beale, R.; Creed, C. Affective interaction: How emotional agents affect users. *Int. J. Hum. Comput. Stud.* **2009**, *67*, 755–776.

57. Marsella, S.; Gratch, J.; Petta, P. Computational models of Emotion. *Blueprint Affect. Comput.* **2010**, *11*, 21–46.

58. Rodriguez, L.F.; Ramos, F. Development of Computational Models of Emotions for Autonomous Agents: A Review. *Cogn. Comput.* **2014**; pp. 351–375.

59. Russell, J.A. A circumplex model of affect. *J. Pers. Soc. Psychol.* **1980**, *39*, 1161–1178.

60. Yeh, Y.C.; Lai, S.C.; Lin, C.W. The dynamic influence of emotions on game-based creativity: An integrated analysis of emotional valence, activation strength, and regulation focus. *Comput. Hum. Behav.* **2016**, *55*, 817–825.

61. Tijs, T.; Brokken, D.; Ijsselsteijn, W. Creating an emotionally adaptive game. In *International Conference on Entertainment Computing*; Springer: Heidelberg/Berlin, Germany, 2008; pp. 122–133.

62. Plass, J.L.; Heidig, S.; Hayward, E.O.; Homer, B.D.; Um, E. Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learn. Instr.* **2014**, *29*, 128–140.

63. Dryer, D.C. Dominance and Valence: A Two-Factor Model for Emotion in HCI. Available online: https://goo.gl/uJvLhD (accessed on 29 March 2017).

64. Hudlicka, E. Guidelines for Designing Computational Models of Emotions. *Int. J. Synth. Emot.* **2011**, *2*, 26–79.

65. Shuman, V.; Clark-Polner, E.; Meuleman, B.; Sander, D.; Scherer, K.R. Emotion perception from a componential perspective. *Cogn. Emot.* **2015**, *9931*, 1–10.

66. Shuman, V.; Sander, D.; Scherer, K.R. Levels of valence. *Front. Psychol.* **2013**, *4*, 1–17.

67. Ekman, P.; Friesen, W.V. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Cues*; Prentice Hall: Englewood Cliffs, NJ, USA, 1975.

68. Ekman, P.; Friesen, W.V. *Facial Action Coding System: Investigatoris Guide*; Consulting Psychologists Press: Sunnyvale, CA, USA, 1978.

69. Ostermann, J. *Face animation in MPEG-4. MPEG-4 Facial Animation. The Standard, Implementation and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2002.

70. Essa, I.A.; Pentland, A.P. Coding, analysis, interpretation, and recognition of facial expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **1997**, *19*, 757–763.

71. Rivera, K.; Cooke, N.J.; Bauhs, J.A. The Effects of Emotional Icons on Remote Communication. In Proceedings of the Conference Companion on Human Factors in Computing Systems, Vancouver, BC, Canada, 13–18 April 1996; pp. 99–100.

72. Coan, J.A.; Allen, J.J. (Eds.) *Handbook of Emotion Elicitation and Assessment (Series in Affective Science)*, 1st ed.; Oxford University Press: Oxford, UK, 2007; p. 504.

73. Derks, D.; Bos, A.E.R.; von Grumbkow, J. Emoticons in computer-mediated communication: Social motives and social context. *CyberPsychol. Behav.* **2008**, *11*, 99–101.

74. Ganster, T.; Eimler, S.C.; Krämer, N.C. Same Same But Different!? The Differential Influence of Smilies and Emoticons on Person Perception. *Cyberpsychol. Behav. Soc. Netw.* **2012**, *15*, 226–230.

75. Lazzeri, N.; Mazzei, D.; Greco, A.; Rotesi, A.; Lanatà, A.; De Rossi, D. Can a Humanoid Face be Expressive? A Psychophysiological Investigation. *Front. Bioeng. Biotechnol.* **2015**, *3*, 64.

76. Dunlap, J.; Bose, D.; Lowenthal, P.R.; York, C.S.; Atkinson, M.; Murtagh, J. What sunshine is to flowers: A literature review on the use of emoticons to support online learning. *Emot. Des. Learn. Technol.* **2015**, *163*, 1–17.

77. Gross, J.J.; Levenson, R.W. Emotion elicitation using films. *Cogn. Emot.* **1995**, *9*, 87–108.

78. Uhrig, M.K.; Trautmann, N.; Baumgärtner, U.; Treede, R.D.; Henrich, F.; Hiller, W.; Marschall, S. Emotion Elicitation: A Comparison of Pictures and Films. *Front. Psychol.* **2016**, *7*, 180.

79. Churches, O.; Nicholls, M.; Thiessen, M.; Kohler, M.; Keage, H. Emoticons in mind: An event-related potential study. *Soc. Neurosci.* **2014**, *9*, 196–202.

80. Brave, S.; Nass, C. Emotion in Human-Computer Interaction. In *Human-Computer Interaction Fundamentals*, 2nd ed.; Jacko, J., Sears, A., Eds.; CRC Press: Boca Raton, FL, USA, 2003; Volume 29, pp. 81–96.

81. Jacko, J.A. (Ed.) *Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Third Edition (Human Factors and Ergonomics)*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2012.

82. Košir, A. The intelligent typing tutor, 2017. Available online: http://nacomnet.lucami.org/expApp/TypingTutorEng/ (accessed on 10 January 2017).

83. Skinner, B.F. Reinforcement today. *Am. Psychol.* **1958**, *13*, 94.

84. Amsel, A. Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychol. Rev.* **1962**, *69*, 306.

85. Bandura, A.; Walters, R.H. *Social Learning Theory*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1977.

86. Stipek, D.J. *Motivation to Learn: From Theory to Practice*, 2nd ed.; Allyn & Bacon: Boston, MA, USA, 1993; p. 292.

87. Official Unicode Consortium. Official Unicode Consortium code chart, 2017. Available online: http://www.unicode.org/ (accessed 3 January 2017).

88. Mahmood, T.; Mujtaba, G.; Venturini, A. Dynamic personalization in conversational recommender systems. *Inf. Syst. e-Bus. Manag.* **2013**, *12*, 213–238.

89. Noldus Information Technology. Noldus FaceReader, 2017. Available online: http://www.noldus.com/ (accessed on 10 January 2017).

90. Kahneman, D. *Attention and Effort*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1973; p. 246.

91. Hodgson, T.L.; Müller, H.J.; O'Leary, M.J. Attentional localization prior to simple and directed manual responses. *Percept. Psychophys.* **1999**, *61*, 308–321.