# MoodSwipe: A Soft Keyboard that Suggests Messages Based on User-Specified Emotions

**Chieh-Yang Huang**[1]    **Tristan Labetoulle**[2]    **Ting-Hao (Kenneth) Huang**[3]
**Yi-Pei Chen**[1]    **Hung-Chen Chen**[1]    **Vallari Srivastava**[4]    **Lun-Wei Ku**[1]

[1] Academia Sinica, Taiwan.
{`appleternity,ypc82,hankchen,lwku`}`@iis.sinica.edu.tw`
[2] IMT Atlantique, France. `contact@tristan-labetoulle.com`
[3] Carnegie Mellon University, USA. `tinghaoh@cs.cmu.edu`
[4] Institute of Engineering & Technology, DAVV, India. `vallari357@gmail.com`

## Abstract

We present MoodSwipe, a soft keyboard that suggests text messages given the user-specified emotions utilizing the real dialog data. The aim of MoodSwipe is to create a convenient user interface to enjoy the technology of emotion classification and text suggestion, and at the same time to collect labeled data automatically for developing more advanced technologies. While users select the MoodSwipe keyboard, they can type as usual but sense the emotion conveyed by their text and receive suggestions for their message as a benefit. In MoodSwipe, the detected emotions serve as the medium for suggested texts, where viewing the latter is the incentive to correcting the former. We conduct several experiments to show the superiority of the emotion classification models trained on the dialog data, and further to verify good emotion cues are important context for text suggestion.

## 1  Introduction

Knowing how and when to express emotion is a key component of emotional intelligence (Salovey and Mayer, 1990). Effective leaders are good at expressing emotions (Bachman, 1988); expressing positive emotions in group activities improves group cooperation, fairness, and overall group performance (Barsade and Gibson, 1998); and expressing negative emotions can promote relationships (Graham et al., 2008). However, in the mobile device era where text-based communication is part of life, technologies are rarely utilized to assist users to *express* their emotions properly via text. For instance, business people in heated disputes with their clients may need assistance to rephrase their angry messages into neutral descriptions before sending them. Similarly, people may have trouble finding the perfect words to show how much they appreciate a friend's help. Or people may want to deliberately express anger to extract concessions in negotiations, or to make a joke, such as with the "Obama's Anger Translator" skit, in which the comedian "translates" the U.S. President's calm statements into emotional tirades. While emotion classification has been used in helping users to better understand other people's emotions (Wang et al., 2016; Huang et al., 2017), these technologies have rarely been used to support user needs in expressing emotions. Most prior work focuses on interface design, for instance using kinetic typography or dynamic text (Bodine and Pignol, 2003; Forlizzi et al., 2003; Lee et al., 2006) , affective buttons (Broekens and Brinkman, 2009), or text color and emoticons (Sánchez et al., 2006) to enable emotion expression in instant messengers. Other work explores the relations between user typing patterns and their emotions (Zimmermann et al., 2003; Alepis et al., 2006; Tech, 2016). However, the text itself is still the essential part of text-based communication; visual cues and typing patterns are only minor factors. Other work even describes attempts to incorporate body signals such as fluctuating skin conductivity levels (DiMicco et al., 2002), thermal feedback (Wilson et al., 2016), or facial expressions (El Kaliouby and Robinson, 2004) to enrich emotion expression, but these require additional equipment and are less scalable.

In this paper we introduce ***MoodSwipe***[1], a soft keyboard that automatically suggests text according to user-specified emotions. As shown in Figure 1, MoodSwipe receives user input text from

---

[1]MoodSwipe is available at: `https://play.google.com/store/apps/details?id=sinica.moodswipe&hl=en`

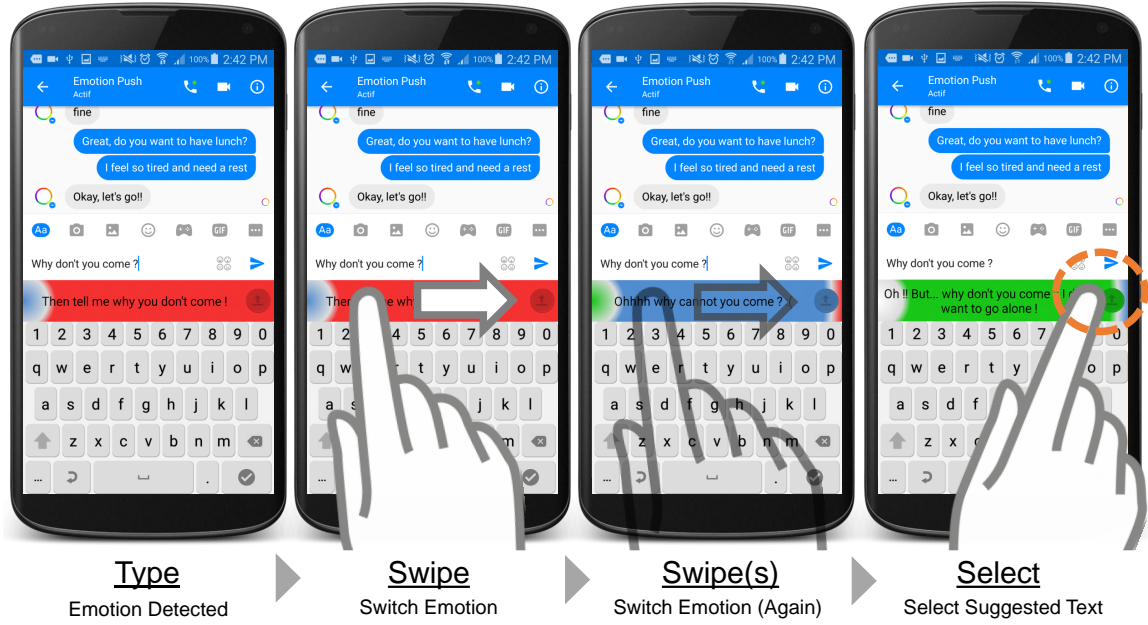| Type | Swipe | Swipe(s) | Select |
|------|-------|----------|--------|
| Emotion Detected | Switch Emotion | Switch Emotion (Again) | Select Suggested Text |

Figure 1: The user interface of MoodSwipe keyboard, which includes a standard soft keyboard, a color bar above the keyboard, and a circle button to the right of the color bar. Users swipe the color bar to specify their emotions and view the messages suggested for different emotions.

the keyboard, and immediately shows the detected emotion of the input message. Seven emotions (*Anger*, *Joy*, *Sadness*, *Fear*, *Anticipation*, *Tired* and *Neutral*) are detected and presented by their corresponding colors as defined by the research of psychologists and user studies (Wang et al., 2016). Users **swipe** the color bar to change the detected emotion according to their **mood** and view the messages suggested for different emotions. For example, if the user types "I disagree," MoodSwipe suggests a relevant message telling the user how one would say the same thing when he or she is happy, sad, or angry.

The contributions of this work are three-fold. First, we address the long-standing challenge of **collecting *self-reported* emotion labels for dialog messages**. Unlike posts on social media, where users often spontaneously tag their own emotions, self-reported emotion labels for dialog messages are expensive to collect. Users often feel disturbed when they are asked to annotate their own emotions on the fly. MoodSwipe provides a handy service which is entertaining and easy to use. It provides a natural incentive for users to label their own emotions on the spot. Second, MoodSwipe closes the loop of **bi-directional *interactive emotion sensing***. Most prior work powered by emotion detection focused on helping users when *receiving* messages (Wang et al., 2016) instead of when *sending* them. MoodSwipe enables auto-

| Emotion | Color | Suggestion |
|---------|-------|------------|
| Anger | Red (#F43131) | I'm in a pissy mood. I'm sorry. |
| Joy | Yellow (#E1D500) | I'm doing fine. |
| Sadness | Navy Blue (#518CCF) | I'm upset, but I'm fine. |
| Fear | Green (#17C617) | (no suggestion) |
| Anticipation | Orange (#E78300) | Oh, I'm fine, I'm widw awake. What's up? |
| Tired | Purple (#C350DF) | I'm working. |
| Neutral | White (#FFFFFF) | I'm fine. |

Figure 2: Mapping between colors and emotions, and example suggestion texts for "I am fine."

mated support, helping users express their emotions in text, and therefore supplies the missing piece for an emotion-sensitive text-based communication environment. Finally, MoodSwipe introduces a **new interaction paradigm**, in which users explicitly provide feedback to systems about *why* they select this suggested response. Classic response suggestion tasks such as dialog generation (Li et al., 2016) or automated email reply (Kannan et al., 2016) assume that the in-the-moment context of each user (e.g., the current emotion) is unknown. MoodSwipe opens up possibilities for users to explicitly and actively provide context on the fly, which the automated models can use to provide better suggestions.

## 2 The MoodSwipe Keyboard

The user interface and workflow of MoodSwipe are shown in Figure 1. The MoodSwipe keyboard interface contains three major parts: *(i)* a standard soft keyboard, *(ii)* a color bar above the keyboard, and *(iii)* a circle button to the right of the color

bar. When the user starts typing, MoodSwipe detects the emotion of the input text in real time, and the color bar background changes color on-the-fly to reflect the current emotion of the user input text. MoodSwipe's seven emotions and their colors are shown in Figure 2, which was developed based on psychological work and user studies (Wang et al., 2016). Based on the emotion MoodSwipe currently displays, the color bar also shows the user the suggested text. Those suggested messages for the input "I am fine" are also listed in Figure 2.

When the user *swipes* the color bar, one of MoodSwipe's seven emotion colors is brought up in descending order of the classification probability for the input message. The user swipes right to see the predicted emotion with a lower probability and its suggested text, or swipes left to see the previous one. In Figure 1, the user types "Why don't you come?" and MoodSwipe detects *Anger* (red) and suggests "Then tell me why you don't come!". The user swipes right to see the option "Ohhhh why cannot you come?" for emotion *Sadness* (blue). The user keeps swiping until he or she reaches a suitable message ("Oh!! But...why don't you come I don't want to go alone!") and then clicks the circle button at color bar's right side to replaces the user input with the suggested text. Then with this replacement, the user self-reports that the emotion label of the user's message "Why don't you come?" should be more *Fear* (green) than *Anger* (red) in the current dialog context.

MoodSwipe actively triggers emotion detection and updates color accordingly when the user presses the spacebar, which usually indicates the completion of a word, or has a 500ms pause after the last user input. To lighten server loads and reduce possible conflicts with the second condition, the minimum time interval between two triggers is set at 400ms. All users activities are recorded, including typed text, suggested texts, emotion labels selected/abandoned, and all the timestamps.

## 3 Use Cases

In this section we outline several possible use cases of MoodSwipe. First, MoodSwipe can help users to better **understand their own messages' emotions** perceived by other users. Our prior study (Huang et al., 2017) shows that not all users are clearly aware of what emotion their messages will convey. In this scenario, MoodSwipe is able to act as an early assistant or reminder when composing a message. Second, MoodSwipe can **assist users to better express themselves** when "words fail me." Sometimes users could experience strong emotions and have difficulty in finding good ways to express themselves via text, and they can type keywords into MoodSwipe to search for better messages from its dialog database. Third, users can **alternate the perceived emotions in their own texts** for various purposes. For instance, some people might need assistance to rephrase their angry messages into neutral descriptions, and some people may want to deliberately express anger to extract concessions in negotiations. Finally, MoodSwipe can be used as a tool to help new users quickly **adapt to the language style of a community.** MoodSwipe is powered by messages that were collected from young IM users. An elder new user who is not very familiar with the language styles of the young generation can use MoodSwipe to rephrase his/her sentence so that it can be better received by young users.

## 4 Back-end System, Experiments and Discussions

Two major functions of the MoodSwipe keyboard are to guess for the users the emotion of their current text message and to provide text suggestions based on that. In this section, we describe several models we developed and different settings used to evaluate their performance. Our advantage in conducting these experiments comes from our emotion-based chat app EmotionPush (Wang et al., 2016) and the social dialogs it has collected.

### 4.1 Experimental Materials

For the experiments, we adopted the Emotion-Push dataset (available soon). A total number of 162,031 message logs were collected for this dataset. To evaluate the performance of emotion classification, we had native speakers manually label the emotions of the randomly selected 8,818 messages. These manual emotion attribute were all chosen from the seven emotion labels defined for the keyboard. Among these 8,818 emotion labeled messages, 70% were for training, 10% for validation and 20% for testing. Two different emotion corpus, LJ40K (Yang and Liu, 2013) and the tweet data, are utilized for comparison. LJ40K contains 40 emotions and for each of the emotions, 1,000 blogs are collected. The 40 emotions are then mapped to the 7 emotions according to our

previous work (Wang et al., 2016). On the other hand, the tweet data is built by Twitter streaming API[2] with a filter of using the 40 emotions as hashtag. A total number of 19,480 tweets are collected and further categorized into 7 emotions.

The text message suggestion function provided by MoodSwipe recommends responses given the input text message and its corresponding emotion label from users. For the experiments, we selected messages from the labeled 8,818 messages according to two rules: 1) to ensure a proper turn, the previous message must be sent from another user instead of the same message owner, and 2) the emotion label must not be neutral. We drop a small number of short messages containing hindi or pure punctuation (e.g., "!!!") for which text suggestions cannot be found, as in these cases we are unable to evaluate the performance of different settings. A total of 1,366 messages were collected for the sentence suggestion experiment (707 Joy, 223 Anger, 189 Sadness, 124 Anticipation, 72 Fear and 51 Tired). For the evaluation, text suggestions are generated for these messages using MoodSwipe.

## 4.2 Emotion Classification

Two models are developed for emotion classification: the general CNN (Kim, 2014) with 125 filters, including 25 filters for each filter length ranging from 1 to 5, and the LSTM (Hochreiter and Schmidhuber, 1997). These two models are trained on blog data, tweet data, and our dialog data, respectively, and then tested on the dialog data. In Table 1 we report the results of three major emotions with these two models, as the other emotions are minor and the training data insufficient to build a reliable model (anticipation 1.77%, tired 0.8%, fear 1.19%, total 3.77%). Only accuracies trained on the dialog data for three major emotions are all over 0.9 (see CNN[3] and LSTM[3]), which supports the use of dialogs in MoodSwipe. Considering time-consuming issue, we adopted CNN as the final model for MoodSwipe.

## 4.3 Text Suggestion & Results

The purpose of the experiments for test message suggestion is to determine whether the system generates better suggested texts given the user-specified emotion. We designed a retrieval-based model utilizing Lucene (McCandless et al., 2010)

| Model | Joy | Anger | Sadness | Neutral |
|---|---|---|---|---|
| (Wang et al., 2016) | .779 | .771 | .853 | .323 |
| CNN[1] | .832 | .960 | .750 | .513 |
| CNN[2] | .645 | .942 | .503 | .222 |
| CNN[3] | .905 | .962 | .973 | .820 |
| LSTM[1] | .230 | .967 | .963 | .222 |
| LSTM[2] | .596 | .959 | .516 | .222 |
| LSTM[3] | .906 | .965 | .964 | .816 |

Table 1: Accuracy of the emotion classification task tested on dialog data while trained on blog[1], tweet[2] and dialog[3] data.

and then applied it on the EmotionPush dataset which contains 162,031 social dialog messages. When searching for similar messages, Lucene first applies term matching on the dataset using query message. Messages containing at least one same word would be candidates which is then ranked by BM25 (Robertson et al., 2009). When the user received a message and is composing a response, the user manually specifies an emotion (*e.g.*, *Anger*) that he/she wants to convey in the message, and the following two settings for generating responses.

1. **[Baseline]:** Given the message that the user received, MoodSwipe first retrieves its most similar message (by Lucene) from the database, and then returns the response of that retrieved message as the suggestion.

2. **[+Emotion]:** The procedure is identical as [Baseline], just that the suggested text must convey the user-specified emotion. For instance, if the user specifies *Anger*, MoodSwipe takes the message that the user received and from database finds its most similar message whose response's emotion is labelled as *Anger* as the suggestion.

Note that the emotions in our database are annotated by automatic algorithms instead of humans.

To assess the quality of suggested messages in each setting, we used the 1,366 messages collected in sentence suggestion experiment to conduct human evaluations with crowd workers recruited via Amazon Mechanical Turk. For each message, we first show the crowd workers its 10 previous messages in the original chat log. We then show the following three messages, in a random order, as the *follow-up line candidates* of the displayed chat log: 1) the actual user input response, 2) the suggested texts in [Baseline], and 3) the suggested texts in [+Emotion]. Workers are asked to rank these three candidate messages based on their

| Setting | Clarity | Comfort | Responsiveness |
|---|---|---|---|
| Rank of Messages and Suggested Texts | | | |
| Input | 1.522 | 1.570 | 1.531 |
| **Baseline** | 2.245 | 2.220 | 2.244 |
| **+Emotion** | 2.233 | 2.210 | 2.225 |
| Good Suggestion Rate (%) | | | |
| **Baseline** | 26.12 | 28.38 | 26.44 |
| **+Emotion** | 26.09 | 28.65 | 26.70 |

Table 2: Human evaluation results. In Baseline the user-specified emotions are not available.

| Good Suggestion Rate (%) | | |
|---|---|---|
| Setting | Anger | Anticipation | Fear |
| **Baseline** | 40.36 | 21.29 | 31.39 |
| **+Emotion** | 37.49 | 20.32 | 25.28 |
| Setting | Joy | Sadness | Tired |
| **Baseline** | 25.35 | 29.31 | 27.45 |
| **+Emotion** | 28.18 | 26.56 | 29.41 |

Table 3: Good suggestion rates of *comfort* for messages of different emotions.

*clarity*, *comfort*, and *responsiveness* of being the follow-up line of the given chat log (Liu et al., 2010) ($rank = 1, 2, or 3$. Lower is better.) For each message, we recruit 5 workers and average their results. Table 2 shows the average ranking and the "Good Suggestion Rate" of each setting, which is the proportion of the suggested messages that have a better (lower) ranking than the original user input response. While the original input responses still have a better (lower) average ranking, Table 2 shows that 26% to 28% of the suggested texts are considered good by crowd workers and thus could be useful to users. Results also show that the [+Emotion] setting on average generates slightly better suggestions than [Baseline] setting in all three aspects.

With the consideration that in the three evaluation aspects *comfort* is most relevant to emotions, we further analyze the *comfort* result of each emotion (Table 3.) Among all emotions, the Good Suggestion Rates for *Anger* messages are the highest (40.36% and 37.49%), which are even much higher than the average rates (about 28% as shown in Table 2). This result suggests that our method is particularly useful for expressing Anger emotions.

These results show the potential benefits of including emotion signals in a response suggestion application. While the simple retrieval-based model (where emotions act only as "filters") may be of limited use, MoodSwipe is still able to suggest responses that are better than the user responses around 25% of time. We believe that when MoodSwipe is deployed, more data can be collected and a more sophisticated models can be developed to boost the benefit of emotion context.

## 4.4 Collecting User-reported Labels

One merit of MoodSwipe is the capability to collect user-reported labels. MoodSwipe can obtain labels from two major user actions, select and swipe, respectively.

1. **[Select]** When the user first types a response, browses all suggestions, and finally selects one suggested text, MoodSwipe can record the user's original typed text and label its emotion as that of the selected text.

2. **[Swipe]** When the user first types a response, swipes directly to a specific emotion (*e.g.*, Joy) and stops there, even without selecting the suggested text, it often still indicates that the user wants to express this emotion (*e.g.*, Joy.) Therefore, MoodSwipe can record the user's current typed text and label it as the same emotion, even the user does not select the suggested text eventually.

## 5 Conclusion

We have developed the sender side MoodSwipe to cooperate with the receiver side applications and complete the emotion sensitive communication framework. MoodSwipe provided a convenient interface which facilitating users on using the modern emotion classification and text suggestion techniques. In MoodSwipe, data are labeled automatically according to frond-end cues in the background. We show that the user-specified emotion can benefit text suggestion, though the performance can still be improved by increasing the size of the dialog database. MoodSwipe is available at Google Play, and a demo video is provided at: https://www.youtube.com/watch?v=SZ1biWoiq3Y

## References

Efthymios Alepis, Maria Virvou, and Katerina Kabassi. 2006. Affective student modeling based on microphone and keyboard user actions. In *Advanced*

*Learning Technologies, 2006. Sixth International Conference on*, pages 139–141. IEEE.

Wallace Bachman. 1988. Nice guys finish first: A symlog analysis of us naval commands. *The SYMLOG practitioner: Applications of small group research*, pages 133–154.

Sigal G Barsade and Donald E Gibson. 1998. Group emotion: A view from top and bottom. *Research on managing groups and teams*, 1(4):81–102.

Kerry Bodine and Mathilde Pignol. 2003. Kinetic typography-based instant messaging. In *CHI'03 Extended Abstracts on Human Factors in Computing Systems*, pages 914–915. ACM.

Joost Broekens and Willem-Paul Brinkman. 2009. Affectbutton: Towards a standard for dynamic affective user feedback. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pages 1–8. IEEE.

Joan Morris DiMicco, Vidya Lakshmipathy, and Andrew Tresolini Fiore. 2002. Conductive chat: Instant messaging with a skin conductivity channel. In *Proceedings of Conference on Computer Supported Cooperative Work*.

Rana El Kaliouby and Peter Robinson. 2004. Faim: integrating automated facial affect analysis in instant messaging. In *Proceedings of the 9th international conference on Intelligent user interfaces*, pages 244–246. ACM.

Jodi Forlizzi, Johnny Lee, and Scott Hudson. 2003. The kinedit system: affective messages using dynamic texts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 377–384. ACM.

Steven M Graham, Julie Y Huang, Margaret S Clark, and Vicki S Helgeson. 2008. The positives of negative emotions: Willingness to express negative emotions promotes relationships. *Personality and Social Psychology Bulletin*, 34(3):394–406.

Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.

Chieh-Yang Huang, Lun-Wei Ku, et al. 2017. Challenges in providing automatic affective feedback in instant messaging applications. *arXiv preprint arXiv:1702.02736*.

Anjuli Kannan, Karol Kurach, Sujith Ravi, Tobias Kaufmann, Andrew Tomkins, Balint Miklos, Greg Corrado, László Lukács, Marina Ganea, Peter Young, et al. 2016. Smart reply: Automated response suggestion for email. *arXiv preprint arXiv:1606.04870*.

Yoon Kim. 2014. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*.

Joonhwan Lee, Soojin Jun, Jodi Forlizzi, and Scott E Hudson. 2006. Using kinetic typography to convey emotion in text-based interpersonal communication. In *Proceedings of the 6th conference on Designing Interactive systems*, pages 41–49. ACM.

Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky. 2016. Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541*.

Leigh Anne Liu, Chei Hwee Chua, and Günter K Stahl. 2010. Quality of communication experience: definition, measurement, and implications for intercultural negotiations. *Journal of Applied Psychology*, 95(3):469.

Michael McCandless, Erik Hatcher, and Otis Gospodnetic. 2010. *Lucene in Action, Second Edition: Covers Apache Lucene 3.0*. Manning Publications Co., Greenwich, CT, USA.

Stephen Robertson, Hugo Zaragoza, et al. 2009. The probabilistic relevance framework: Bm25 and beyond. *Foundations and Trends® in Information Retrieval*, 3(4):333–389.

Peter Salovey and John D Mayer. 1990. Emotional intelligence. *Imagination, cognition and personality*, 9(3):185–211.

J Alfredo Sánchez, Norma P Hernández, Julio C Penagos, and Yulia Ostróvskaya. 2006. Conveying mood and emotion in instant messaging by using a two-dimensional model for affective states. In *Proceedings of VII Brazilian symposium on Human factors in computing systems*, pages 66–72. ACM.

Cornell Tech. 2016. This smartphone keyboard app can read your emotions.

Shih-Ming Wang, Chun-Hui Li, Yu-Chun Lo, Ting-Hao K Huang, and Lun-Wei Ku. 2016. Sensing emotions in text messages: An application and deployment study of emotionpush. *arXiv preprint arXiv:1610.04758*.

Graham Wilson, Dobromir Dobrev, and Stephen A Brewster. 2016. Hot under the collar: Mapping thermal feedback to dimensional models of emotion. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4838–4849. ACM.

Yi-Hsuan Yang and Jen-Yu Liu. 2013. Quantitative study of music listening behavior in a social and affective context. *Multimedia, IEEE Transactions on*, 15(6):1304–1315.

Philippe Zimmermann, Sissel Guttormsen, Brigitta Danuser, and Patrick Gomez. 2003. Affective computing—a rationale for measuring mood with mouse and keyboard. *International journal of occupational safety and ergonomics*, 9(4):539–551.