A Comparison of Emotion Elicitation Methods for Affective Driving Studies

Michael Braun
BMW Group Research, New Technologies, Innovations
Garching, Germany
LMU Munich
Munich, Germany
michael.bf.braun@bmw.de

Bastian Pfleging
LMU Munich
Munich, Germany
bastian.pfleging@ifi.lmu.de

Simon Weiser
LMU Munich
Munich, Germany
simon.weiser@campus.lmu.de

Florian Alt
Bundeswehr University
Munich, Germany
LMU Munich
Munich, Germany
florian.alt@unibw.de

AutomotiveUI '18 Adjunct,, September 23–25, 2018, Toronto, ON, Canada
ACM 978-1-4503-5947-4/18/09.
https://doi.org/10.1145/3239092.3265945

Abstract
Advances in sensing technology enable the emotional state of car drivers to be captured and interfaces to be built that respond to these emotions. To evaluate such emotion-aware interfaces, researchers need to evoke certain emotional states within participants. Emotion elicitation in driving studies poses a challenge as the driving task can interfere with the elicitation task. Induced emotions also lose intensity with time and through secondary tasks. This is why we have analyzed different emotion elicitation techniques for their suitability in automotive research and compared the most promising approaches in a user study. We recommend using autobiographical recollection to induce emotions in driving studies, and suggest a way to prolong emotional states with music playback. We discuss experiences from a a driving simulator study, including solutions for addressing potential privacy issues.

Author Keywords
Emotion Elicitation; Affective Computing; Driving Studies.

CCS Concepts
•Human-centered computing → HCI design and evaluation methods;

Figure 1: Russell's circumplex model of affect can be used to classify emotions in a two-dimensional space. An optimal driving state would be medium arousal in combination with high valence.

Introduction

Research in affective computing is based on the assumption that machines will be able to assess users' emotions in the near future. In order to teach them how to understand these human emotions, as well as to test systems that react to them, we need ways of putting study participants into distinct emotional states. This way, the ground truths for machine learning algorithms can be specified, and user interfaces can be evaluated in controlled settings [12].

When testing affective systems, the driving task might be in direct competition with the elicitation task and both tasks require mental capacity from the the participant. For this reason, emotion elicitation is mostly done before the actual driving task, which unfortunately allows the induced emotional state to diminish as time passes. Our goal was to find an approach that could either be used while driving or prior to driving but with prolonged effects.

Related Work

Affective computing is a research domain with various applications in the automotive domain, such as the adaptation of user interfaces by affect sensitive technology [2] or the detection and improvement of driver states to avoid dangerous driver behavior, e.g. road rage [4]. Emotional states which are thought to influence driving performance are fear, happiness, anger, depression, and boredom, among others [6]. These can be classified in Russell's circumplex model of affect, consisting of the dimensions arousal and valence [13]. Especially extreme states can lead to a divergence from safe driving performance, exemplary in situations of high or low mental load. A safe driving state would be positive valence and medium arousal (see Figure 1), which can also be explained by the Yerkes-Dodson law: performance is best in a medium level of arousal [9].

Rosalind Picard classifies affective studies based on the way emotions are elicited, the study environment, whether emotions are required to be expressed or felt, and participants' knowledge of the study's goals and their participation [11]. A perfect setting would be a real-world scenario with inherent, event-elicited emotions and a hidden apparatus and purpose. This is however ethically and practically impossible. Studies on affective systems mostly use active elicitation, also because research suggests that emotions are felt more strongly when participants know about the desired inducement [10].

Common elicitation methods include showing images and videos [3, 8] (both of which have been shown to be equally effective [14]), autobiographical recollection [1, 5], the Velten mood induction method [7], and in the automotive context, inducement through driving scenarios. The latter two approaches are limited in scope since none of the methods induce emotions related to both valence and arousal.The Velten task is only intended to set a positive or negative mood and driving scenarios can only induce different levels of cognitive load.

We chose to use both music videos as well as autobiographical recollection due to their high versatility. In contrast to other methods, they might be capable of inducing every combination of arousal and valence on the spectrum.

Study

The study was conducted using a static driving simulator on an easy highway track. The simulation was displayed on a wide screen display and controlled by a consumer grade steering wheel and pedal combination. Participants experienced 4 inducement techniques for 4 different emotional states (see Figure 2) which we combined in a balanced latin square distribution to avoid carry-over-effects. 8 participants

(3 female, 5 male) between the ages of 22 to 27 completed 4 rides each. Each ride began with an 8 minute elicitation session and ended with an interview. We also sampled the intensity of the induced emotions at 1 minute intervals during the ride.

We compared the inducement of emotional states with music videos taken from the DEAP dataset [8] and the approach of autobiographical recollection. Both methods were also enhanced by introducing an additional playback of music during the driving task, with the intension of extending the emotional effect. The songs were also taken from the DEAP dataset (see Table 1).

We induced four combinations in the extreme areas of low and high arousal and valence (see Figure 2). Before each ride, test users experienced one of these four emotion inducement methods:

MV Participants watch a music video which has been rated to induce the desired emotion in the DEAP dataset.

MV+ Same approach as MV, with another instance of music playback while driving (the same song).

AR Autobiographic recollection requires the participant to think of and write down an experience from their past in order to recall a specific emotion. When driving, they had to recount this story to themselves without the experimenter in the room.

AR+ Same approach as AR, with the according song from the DEAP dataset playing along the ride.

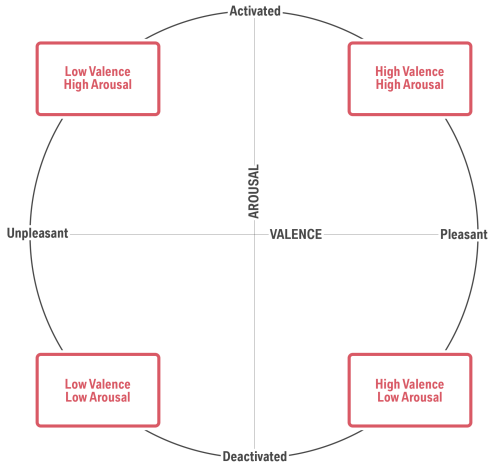

Figure 2: Valence-arousal model showing the four extremes we induced in the study.

Emotion	*Artist*	*Title*
Happiness	M. Franti & Spearhead	Say Hey (I Love You)
Contentment	Grand Archives	Miniature birds
Anger	Dead To Fall	Bastard Set of Dreams
Sadness	James Blunt	Goodbye My Lover

Table 1: Music titles form the DEAP database for extreme values of arousal and valence, named after the closest emotion categories.

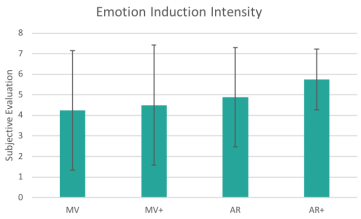

Figure 4: Subjectively rated intensity of emotions right after the inducement task. Autobiographical recollection shows higher tendencies and a lower standard deviation than music videos.

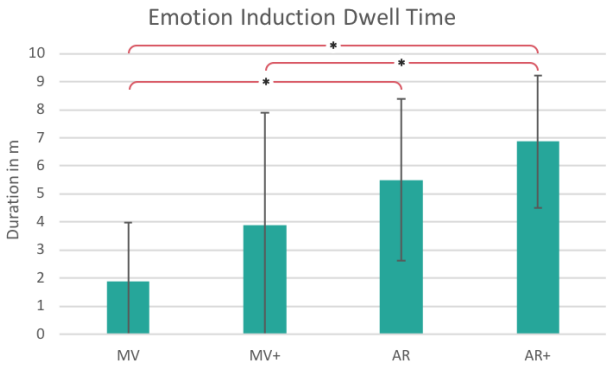

Figure 3: The dwell time of induced emotions was found to be significantly longer for autobiographial recollection than for watching music videos. Additional music playing after the inducement task could also prolong the felt emotions.

Results

Participants rated the effectiveness of the emotion induction during the ride, as well as in an interview afterward. The approaches using music videos were each assessed as effective by 5/8 respondents, while 7/8 said autobiographical recollection worked for them. Feedback during the ride features the same tendencies, with AR+ performing best in initial emotion inducement, as seen in Figure 4. Participants experiencing the music video method for low valence and high arousal stated they felt rather stressed than angry. Over time, emotions induced with music videos faded significantly faster than emotions induced with autobiographical recollection $(p < 0.005)$. Music playback during the driving task proved to prolong the induced emotional state, as shown in Figure 3. In the end participants ranked which emotion was most intense: while the states were equally distributed, 7/8 rated first were induced with AR or AR+.

Conclusion

The goal of this study was to decide on which emotion elicitation method is suitable to use in driving studies. Our findings suggest that autobiographical recollection works really well in inducing emotions and is versatile to use. A significant advantage lies in the fact that the stimulus is generated by the user themselves, thus giving little leeway for misinterpretation. The active recounting can also be done while driving, allowing a less sharp transition from elicitation to the driving task.

We have since applied the method in a driving simulator study where we have learned that participants might be hesitant to talk aloud, which should however be enforced due to a better outcome. Measures to ensure privacy and thereby motivate participants to talk openly are encouraged. For instance, turning off the microphone connection between simulator and control room and telling participants they can only interact through showing of hands for the following elicitation period. In our experiment this helped build trust towards the experimenter and helped to achieve very satisfactory inducement success.

In summary, autobiographical recollection is the first choice to elicit emotions in driving studies and music playback can be used to prolong the effect of induced emotions. We are looking forward to testing this approach in combination with novel emotion-aware interfaces as so far we have only looked at the elicitation of emotions. We hope to stimulate a discussion whether this approach can be used to evaluate affective user interfaces in automotive environments.

REFERENCES

1. Richard C. Baker and Daniel O. Guttfreund. 1993. The effects of written autobiographical recollection induction procedures on mood. *Journal of Clinical Psychology*

49, 4 (1993), 563–568. DOI: http://dx.doi.org/10.1002/1097-4679(199307)49:4<563::AID-JCLP2270490414>3.0.CO;2-W

2. Florian Eyben, Martin Wöllmer, Tony Poitschke, Björn Schuller, Christoph Blaschke, Berthold Färber, and Nhu Nguyen-Thien. 2010. Emotion on the Road: Necessity, Acceptance, and Feasibility of Affective Computing in the Car. *Advances in Humam-Computer Interaction* 2010, Article 5 (jan 2010), 17 pages. DOI: http://dx.doi.org/10.1155/2010/263593

3. James J. Gross and Robert W. Levenson. 1995. Emotion elicitation using films. *Cognition and Emotion* 9, 1 (1995), 87–108. DOI: http://dx.doi.org/10.1080/02699939508408966

4. Leon James. 2000. *Road rage and aggressive driving: Steering clear of highway warfare*. Prometheus Books.

5. Joris H. Janssen, Paul Tacken, J.J.G. (Gert-Jan) de Vries, Egon L. van den Broek, Joyce H.D.M. Westerink, Pim Haselager, and Wijnand A. IJsselsteijn. 2013. Machines Outperform Laypersons in Recognizing Emotions Elicited by Autobiographical Recollection. *Human-Computer Interaction* 28, 6 (2013), 479–517. DOI: http://dx.doi.org/10.1080/07370024.2012.755421

6. Myounghoon Jeon and Bruce N Walker. 2011. What to detect? Analyzing factor structures of affect in driving contexts for an emotion detection and regulation system. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 55. Sage Publications Sage CA: Los Angeles, CA, 1889–1893.

7. Pamela M. Kenealy. 1986. The velten mood induction procedure: A methodological review. *Motivation and Emotion* 10, 4 (01 Dec 1986), 315–335. DOI: http://dx.doi.org/10.1007/BF00992107

8. Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. 2012. Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing* 3, 1 (2012), 18–31. DOI: http://dx.doi.org/10.1109/T-AFFC.2011.15

9. Yerkes Robert M. and Dodson John D. 1908. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology* 18, 5 (1908), 459–482. DOI: http://dx.doi.org/10.1002/cne.920180503

10. Jürgen H. Otto, Harald A. Euler, and Heinz Mand. 2000. *Emotionspsychologie: ein Handbuch*. Beltz, Psychologie-Verlag-Union.

11. Rosalind W. Picard. 1995. Affective Computing. (1995).

12. Rosalind W. Picard, Elias Vyzas, and Jennifer Healey. 2001. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE transactions on pattern analysis and machine intelligence* 23, 10 (2001), 1175–1191.

13. James Russell. 1980. A Circumplex Model of Affect. 39 (12 1980), 1161–1178.

14. Meike K. Uhrig, Nadine Trautmann, Ulf Baumgärtner, Rolf-Detlef Treede, Florian Henrich, Wolfgang Hiller, and Susanne Marschall. 2016. Emotion Elicitation: A Comparison of Pictures and Films. *Frontiers in Psychology* 7 (2016), 180. DOI: http://dx.doi.org/10.3389/fpsyg.2016.00180