Is Reading Mirrored in the Face? A Comparison of Linguistic Parameters and Emotional Facial Expressions

Egon Werlen
egon.werlen@ffhs.ch

Ivan Moser
ivan.moser@ffhs.ch

Christof Imhof
christof.imhof@ffhs.ch

Per Bergamin
per.bergamin@ffhs.ch

Institute for Research in Open-, Distance- and eLearning (IFeL)
Swiss Distance University of Applied Sciences (FFHS)

Abstract

The ongoing digitalisation facilitates measuring emotional characteristics of texts (e.g. lexical emotional valence), and emotional face expressions (e.g. facial emotional valence). In this context, a text was lexically analysed with the revised Berlin Affective Word List (BAWL-R), and videos of 91 subjects reading this text were analysed with a facial emotion recognition software. We hypothesized that lexical emotional valence predicts readers' facial emotional valence. The result was significant but explained nearly no variance (0.3%). Detecting emotional face expressions is a well established method, which means that the mostly neutral face expressions of our participants may be a result of the non-social reading situation.

1 Introduction

Digital reading is becoming increasingly frequent and important. This trend is not only affecting business correspondence (e-mail, e-collaboration) or online shopping, but also reading fictional and non-fictional texts. Digital reading can be defined in contrast to "traditional" reading by the texts' own characteristics. Digital texts can be structured as hypertexts permitting to navigate through multiple different documents, but they can be linear as well. They may contain multimedia such as sound and dynamic visualisations. They are presented on a computer screen, a tablet, an e-reader, a smart phone or via virtual reality goggles. Technology permits digital reading to become more social and interactive by sharing and commenting on information (Kaakinen et al., 2018). The digitalisation of knowledge management and learning is also increasing. Regardless of the digital application, reading and understanding of words, terms and entire sentences or texts, as well as reactions to them, are fundamental processes. Since the publication of "Affective Computing" by Picard (1995) measuring and analysing of emotions and emotional processes has been increasing. Digitalisation enables new forms of text capture, display and processing, including the estimation of the emotional content of a text.

In this study we investigated the emotional characteristics of texts in conjunction with the emotional reaction of readers. We examined if the intensity of positive or negative expressions on the reader's face corresponds with the emotional potential of texts. This fundamental research is at the basis of the development of sensors in adaptive learning systems. Being able to estimate the reactions of readers when reading texts with known characteristics allows preparation of learning material in such a way that it can be presented to learners according to their needs.

2 Theoretical background

Humans evaluate every kind of event concerning their actual emotional state, for instance a text or its individual words, with respect to valence (negative / positive), as well as novelty and relevance for an individual's goals or needs (Ellsworth and Scherer, 2003). In many emotion theories this kind of evaluation is referred to as appraisal. Appraisal is a central compo-

In: Mark Cieliebak, Don Tuggener and Fernando Benites (eds.): Proceedings of the 3rd Swiss Text Analytics Conference (SwissText 2018), Winterthur, Switzerland, June 2018

nent of emotion (Ellsworth and Scherer, 2003; Frijda, 1993). We can distinguish primary and secondary appraisal. Between the primary and secondary appraisal cognitive processes take place. In a paper, focusing on emotions and language Koelsch et al. (2015) emphasize furthermore that “affective processes can be observed prior to, and independent from, ‘higher’ cognitive appraisal processes” (p. 13). The primary (unconscious) appraisal on low-level neural circuits begins about 200-300 ms after a stimulus has been perceived. The brain reacts to emotionally valenced words, representing a distinction between positive and negative affective words, (e.g. studies about emotional reactions: Citron 2012; Ponz et al. 2014). An example for a primary appraisal is the unconscious emotional reaction to a negative word like “murder” that may influence the following cognitive processing. The cognitive processing can begin between 500 and 600 ms (Citron, 2012; Ponz et al., 2014) after the initial perception of a stimulus and ends depending on task characteristics. The secondary appraisal succeeds after this time span has passed and has - in contrast to the primary appraisal - conscious characteristics (e.g. the reader remembers an impressive film about a murder).

2.1 Emotional valence

Emotional valence, a result of the primary appraisal, is the experience of one’s own actual positive or negative state. It is a first dimension of the circumplex model as proposed by Russell and Barrett Feldman (1999). The second dimension of the circumplex model is emotional arousal, i.e. the subjective amount of activation or energy. Together, these two dimensions form the core affect, “the most elementary consciously accessible affective feelings (and their neurophysiological counterparts) that need not be directed at anything” (S. 806). In this paper, we analyze the emotional valence 1) with a textual analysis of a fictional text read by the participants - the **lexical emotional valence** - and 2) by the emotional reaction of the readers expressed on their faces while reading the text - the **facial emotional valence**.

2.1.1 Lexical emotional valence

A guide for textual structure analysis is the 4x4 matrix for text analysis by Jacobs (2015). He proposes a classification with four text characteristics and four hierarchical levels. The four text characteristics are based on a list of Jakobson (1979). The metric characteristics concern the structuring of a language in units (e.g. line, verse) by means of rhythm and articulation (e.g. rime, assonance, alliteration). The phonological characteristics affect the function and sound of phonemes. The syntactic characteristics pertain to the combination of words and word groups in larger units like sentences. The semantic characteristics concern the meanings of sentences, parts of sentences, words, components of words, or characters per se. The four hierarchical levels are the sub-lexical level (i.e. the characteristic of components of words like phonemes, articulation), the lexical level (i.e. single words without consideration of the textual content), the interlexical relations between words, phrases, or sections, and the supra-lexical level concerning whole sentences and stories. In a structural text analysis, there are several ways to measure different characteristics of a text, for instance by rating the lexical emotional valence of a text, i.e. estimating how a reader may emotionally experience the text. In this sense, lexical emotional valence is part of the emotional potential of a text, a combination of “information structure, coherence, and implicit information combined with expressive language” (Schwarz-Friesel 2015, p. 167). A common method is to rate the whole text with the help of a few questions, a questionnaire or with the Self-Assessment Manikins scale (SAM; Bradley and Lang 1994). Alternatively, the lexical emotional valence of a text can be assessed by consulting data bases that contain valence ratings of thousands of words (e.g. Affective Norms for English Words - ANEW; Bradley and Lang 1999; Berlin Affective Word List - BAWL-R; Võ et al. 2009) and calculating the average valence of all words of a text, a section, or a sentence. A similar method uses lexical data bases that contain categorized words. The amount of emotional words within a text offers an estimation of its emotional content (e.g. Linguistic Inquiry and Word Count - LIWC; Pennebaker et al. 2015; Regressive imagery dictionary - RID; Martindale 2008; Coh-Metrix; McNamara and Graesser 2012).

2.1.2 Facial emotinal valence

The most commonly used tool for systematically measuring emotions including emotional valence visible in the face is the Facial Action Coding System (FACS) by Ekman and Friesen (1978). FACS “is a comprehensive, anatomically based system for measuring all visually discernible facial movements" (Rosenberg 2005; p. 13). It measures the activity of 44 unique action units (AUs) and several positions and movements of the head and the eyes. FACS is based on facial anatomy, but because facial muscles can act in different ways and contract in different regions there is no 1:1 correspondence between an AU and a facial muscle. "FACS coding procedures also allow for coding of the intensity of each facial action on a five-point intensity scale, for the timing of facial actions, and for the coding of facial expressions in terms of 'events'. An event is the AU-based description of each facial expression, which may consist of a single AU or many AUs contracted as a single expression" (Rosenberg 2005; p. 13).

2.2 Hypothesis

We assumed that the emotional valence of words, the lexical emotional valence, influences the reading experience (e.g. Altmann et al. 2012). A theoretical framework that supports this notion is the Quartet Theory of Human Emotions of Koelsch et al. (2015). In this framework, four distinct interacting brain regions form the 'affect system' whose activity interacts with the activity of the 'effector system'. The 'affect system' defines four classes of emotions originating from four distinct cerebral regions: brainstem (e.g. ascending activation), diencephalon (e.g. pain, pleasure), hippocampus (e.g. attachment related affects), orbitofrontal cortex (e.g. moral affects). The 'effector systems' include action tendencies behaviour (e.g. approaching to or moving away from a stimuli), modulation of physiological arousal (e.g. heart activity or breathing), attention and memory (e.g. selection of information for memory processing or storage), and motor expression (e.g. facial expressions or vocalizations). The 'affect and effector systems' 'create' together the 'emotion percept', a pre-verbal subjective feeling, that can be expressed in symbolic code (e.g. spoken language). Important for our purpose, the motor expression part of the 'effector systems', for instance via facial expressions, manifests the emotions of the affect system (e.g. pain/pleasure, i.e. positive/negative valence).

We hypothesized that the lexical emotional valence of a text, respectively its sections, predicts the readers' facial emotional valence. We suppose that the course of the emotional valence expressed on the reader's face follows the course of the valence of words in a text and its sections.

In spite of a large amount of research in emotion and in text analysis, we are aware of one other research group (Wegener et al., 2017) that ventures into similar territory by combining linguistic and literary analyses of texts with readers' emotional response data (eye-tracking, facial gestures, comprehension, etc.) with the goal of constructing a database for emotion detection and mapping textual triggers for readers' emotion during literary text reading. The EmoLiTe database aims to synchronize information about the reading process (e.g. interview data, reader annotations, likability scores), the research context (e.g. experiment design, experimentor), and contextual information (e.g. crowed-soureced data about the stimulus text and their authors, linguistic and literary analysis of the texts).

3 Methods

3.1 Experimental design

The data for this analysis stem from a reading experiment investigating the influence of cognitive load and emotional reactions on reading performance. The participants filled in several questionnaires and were instructed to read the first part of a fictional text with neutral content. When finished, they were asked how they felt just in that moment (valence, arousal) and were instructed to indicate the difficulty of the text. Then the participants had to retell the story and answered five multiple choice questions to assess what they memorized about the text. This procedure was repeated with the second and the third part of the text, one of which had negative undertones while the other had positive ones. The texts were presented on three different screen sizes simulating different reading devices (smartphone 5", tablet 10", and laptop 15") and with two different levels of readability (easy vs difficult). Screen size and readability were presented by chance. The experiment was followed by a series of pictures of persons displaying different emotions and nature images (landscapes, snakes, spiders) to measure emotional reactions.

3.2 Stimulus material: Story of the text

The text presented in the experiment consisted of three parts telling an overall story. The first (neutral) part describes a park in a town with a pond surrounded by a path with three people sitting on different benches (an old man, a young man and a young woman). The second (negative) part tells the story of an apparent break-up between the young couple. The text contains scenes with disgusting fantasies and describes outbursts of violence. The third (positive) part reveals the whole ordeal as a prank orchestrated by the couple's friends and ends by describing how the two reconcile and dream of holidays by the sea.

3.3 Sample

103 students attending secondary school participated in the study. Most of the participants were women (87%) with an average age of 17.8 (SD: 1.2; range 16-21). For this analysis, the data of 91 students were usable. Data were lost due to missing videos of the respective participant's face. There were no differences between the complete sample and the subsample in age ($F[1,101]=0.56$, $p=.813$) and gender ($F[1,101]=0.19$, $p=.657$).

3.4 Measurements

The semantic lexical analysis of the text was executed with the revised form of the Berlin Affective Word List BAWL-R (Võ et al., 2009). The BAWL-R is a list containing about 2900 words (nouns, verbs, and adjectives) from the CELEX database (Baayen et al., 1993), rated on valence, arousal and imageability. The list also includes psycholinguistic factors (e.g. number of letters, phonemes, word frequency, accent). Valence was rated on a 7-point Likert scale (-3 very negative through 0 neutral to +3 very positive). Based on the story's content, the three text parts were divided in five (neutral), six (negative), respectively seven sections (positive). Each of the three text parts existed in an easily and a difficultly readable version that were randomly assigned to the subjects. Readability was scored for with the Flesch Index (Flesch, 1948; Amstad, 1978), using the web based Flesch-Index calculator of Peter Schoell (`http://fleschindex.de/berechnen`). The easy versions of the neutral and negative text part had a Flesch score of 85, respectively 81 (easy to read), the positive text part had a score of 77 (fairly easy to read). The difficult versions of the neutral and negative text part had a Flesch score of 54, respectively 52 (fairly difficult to read), the positive text part had a score of 30 (difficult to read). The lexical emotional valence was calculated for each section and for the easy and difficult version separately.

The facial emotional valence expressed on the participants' faces was measured objectively using the FaceReader® (version 7) software by Noldus® which is based on Ekman's FACS and utilizes 21 of its action units. The FaceReader® rated the videos of all participants whose faces were filmed. It automatically calculates emotional valence by subtracting the intensity of the most intense negative expression (sadness, disgust, anger or fear) from the intensity of positive expression (i.e. happiness) in a specific timeframe. The values are between –1 and 1. To combine the two data sets, the reading time of each section of the three text parts was estimated with the number of characters in each section in relation to the number of characters of the whole text part (e.g. neutral). For example: the neutral text part (easy to read) contained 2618 characters (100%), its first section 654 characters, i.e. 24.98%. Therefore, each subject's reading time for the first section of the neutral easy to read text was estimated to be 24.98% of the whole reading time of the neutral easy to read text.

Besides the above-mentioned variables we measured subjective feelings, perceived text difficulty, and knowledge about the texts read. In addition, the participants filled in questionnaires assessing sociographic issues, reading habits, visual memory span, and actual mood (German adaptation of the Positive Activation Negative Activation Scale PANAS; Watson et al. 1988; Schallberger 2005). At the end of the experiment session, the participants responded to questions about text difficulty, emotional reactions and cognitive load. The mood questionnaire was filled in a second time, followed by questions about the experiment itself (e.g. atmosphere, conditions). Results of these variables were not reported here.

3.5 Statistics

Most statistical analyses were conducted with R (3.3.4; R-Core-Team 2017). For the generalized linear mixed-model, we used the lme4 package of Bates et al. (2014), the p-values were calculated by

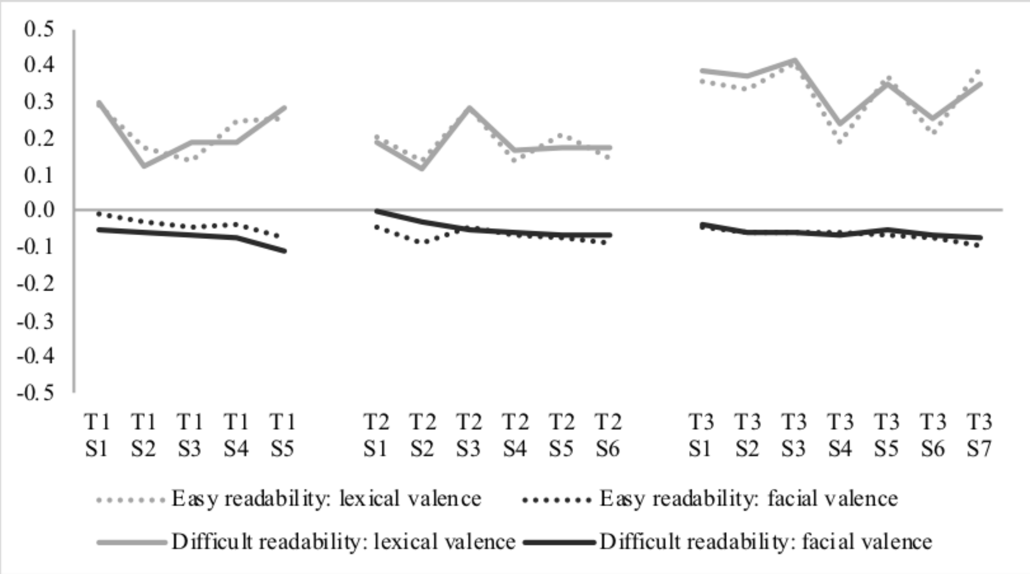

Figure 1: Time course of lexical and facial emotional valence over the sections (S1 to S5, resp. S6 and S7) of the three parts of the fictional text (T1 =neutral; T2=negative; T3=positive); The BAWL-R-values (lexical emotional valence) are standardized from -1 to 1.

means of Satterthwaite's approximation with lmerTest (Kuznetsova et al., 2017), and the pseudo-R^2 of the fixed effects with MuMln by Nakagawa and Schielzeth (2013). Other analyses like the comparison of the included and excluded subjects, and the comparisons of the lexical and facial emotional valence between the three text parts and between the readability levels (easy, difficult) were conducted with JASP (JASP-Team, 2018).

4 Results

The differences of the lexical emotional valence between the three text parts (neutral, negative, positive) estimated with the BAWL-R was statistically significant (F[2,1128]=22.80, p<.001). A post hoc Tukey test revealed that the lexical emotional valence of the positive text part differs significantly from the two other text parts. The lexical emotional valence did not differ between the two readability levels (easy, difficult; F[1,1129]=0.41, p=.523). The neutral text part had a lexical emotional valence of 0.66 (SD: 0.87), the negative part of 0.52 (SD: 1.18), and the positive part of 1.12 (SD: 1.12). The average facial emotional valence measured with FaceReader® did not differ between the three text parts (F[2.1595]=0.31, p=.733), nor between the two readability levels (F[1,1596]=0.02, p=.890).

The time course of the lexical emotional valence resulting from the semantic lexical analysis shows similar values for the easy and difficult texts (the upper two lines in Figure 1).
Overall, the course of the facial emotional valence on the lower two lines of Figure 1 shows values that are all in a narrow range.

We tried to predict the course of the facial emotional valence with a generalized linear mixed model (GLMM) based on the lexical emotional valence of the texts. We included fixed effects for readability (easy vs. difficult), type of text (neutral vs. negative vs. positive), and lexical emotional valence (BAWL-R), including all higher order interaction terms. Furthermore, we included random intercepts and random slopes for the effect of time of each participant. Text difficulty and type of text were effect-coded in order to interpret the regression weights at the grand mean (instead of a reference category).

The GLMM yielded a significant prediction of the facial emotional valence by the lexical emotional valence (β=0.02, CI=[0.02; 0.02], p<.001). All other effects (influences of type of text and readability, and all interactions) were significant as well (all $p < .001$, see Table 1). However, it is important to note that all fixed effects combined only explained 0.3% of variance in the facial emotional valence (pseudo-R^2 of the fixed

	Facial emotional valence		
	β	**CI (95%)**	**p**
Fixed Parts			
(Intercept)	-0.05	-0.06 – -0.04	<.001
Lexical emotional valence	0.02	0.02 – 0.02	<.001
Negative text	-0.02	-0.02 – -0.02	<.001
Positive text	0.01	0.01 – 0.01	<.001
Readability difficult	0.02	0.02 – 0.02	<.001
Lex.emo.valence: difficult	0.03	0.03 – 0.03	<.001
Lex.emo.valence: negative text	-0.02	-0.02 – -0.02	<.001
Lex.emo.valence: positive text	-0.02	-0.02 – -0.02	<.001
Negative text: difficult	0.01	0.01 – 0.01	<.001
Positive text: difficult	-0.02	-0.02 – -0.02	<.001
Lex.emo.valence: neg. text: difficult	-0.03	-0.03 – -0.03	<.001
Lex.emo.valence: pos. text: difficult	0.04	0.03 – 0.04	<.001
Random Parts			
σ^2	0.015		
$\tau_{00,vp}$	0.016		
ρ_{01}	-0.218		
N_{vp}	91		
ICC_{vp}	0.516		
Observations	1328120		
R^2/Ω_0^2	.540/.540		

Table 1: Full Model of the prediction of the facial emotional valence by lexical emotional valence, readability and type of text.

effects).

The interaction effect between lexical valence and readability on facial valence was also significant (β=0.02, CI=[0.02; 0.02], p<.001). This effect refers to all three text parts. Visual inspection of Figure 2 revealed distinct differences between the slopes of the easy and difficult text in all three parts. Therefore, we calculated the model for all three text parts separately. All three analyses yielded significant results. The largest effect of lexical emotional valence on facial emotional valence can be found in the negative text part (β=0.09, CI=[0.08; 0.09], p<.001). But, the explained variance remained very low (1.4%; pseudo-R^2 of the fixed effects). The pseudo-R^2 in the neutral and the positive text part were even smaller (0.8% / 0.03%).

5 Discussion

Statistically, facial emotional valence is significantly predicted by the lexical emotional valence. The explained variance of the fixed effects was very low at 0.3%, challenging the possible conclusion that the facial emotional valence of readers corresponds to the lexical emotional valence of the different sections of the text. There is one result that attracts attention: The slope of the easy readable negative text suggests that facial emotional valence is predicted more strongly in this text. But as with the complete model, the explained variance of the fixed effects is very low at 1.4%. Even if this effect does not explain a large amount of the variance, it remains interesting as it fits to results of other studies. The finding that emotional reactions are suppressed with higher cognitive load (for instance a more hard-to-read text) have been reported for other tasks for example in Berggren et al. (2013), and in DeFraine (2016): "cognitive load reduced the intensity of negative emotions during passive-viewing of emotional images but not during emotion maintenance" (p. 459). Van Dillen et al. (2009) found a similar result, high cognitive load "eliminated all emotional expression differences" (p. 5).

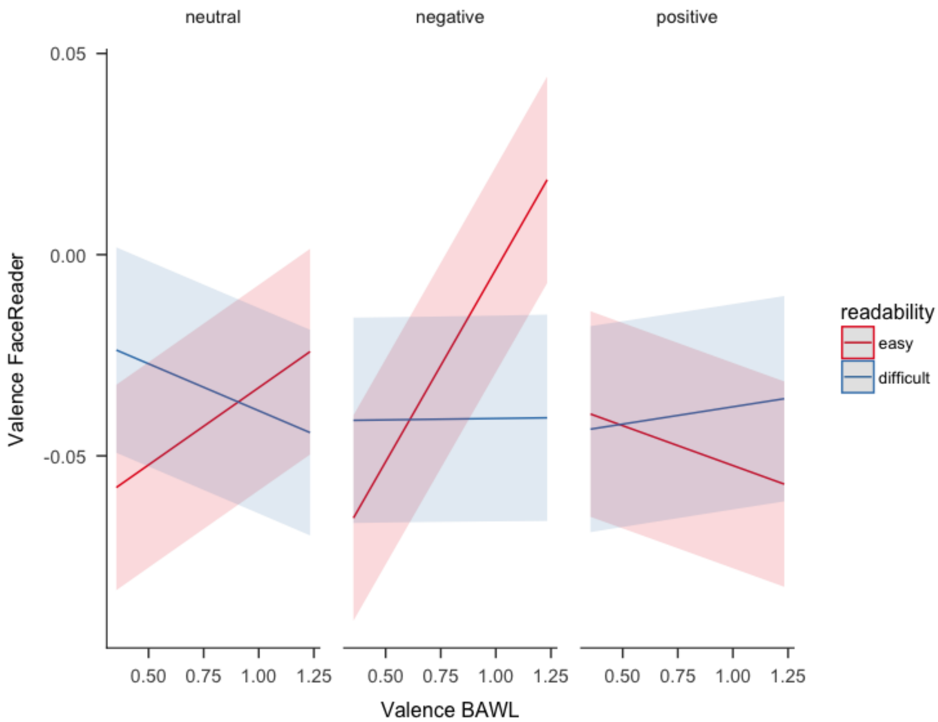

Figure 2: Regression lines of the full model: Prediction of facial emotional valence by lexical emotional valence (BAWL-R).

How can we explain the difference in patterns between lexical and facial emotional valence in this study?
It is highly unlikely that the subjects did not have a primary appraisal, that they had no emotional reaction to the text. The subjective feelings, also influenced by secondary appraisal, were reported just after reading each part of the text. The differences were statistically significant ($F[2,97]=8.32$; $p>.001$). A post hoc Tukey test revealed that the positive part of the text evokes significant more positive feelings than the neutral and negative parts. Other studies have predicted the subjective valence after reading short texts (poems, fiction) with text characteristics including lexical emotional valence and others like arousal, and psycholinguistic values (Ullrich et al., 2017; Hsu et al., 2015).
According to the Component Process Model (CPM) of Paul Scherer (2005), emotions are composed of five components: (1) cognitive (e.g. the appraisal itself), (2) neurophysiological (e.g. body symptoms), (3) motivational (e.g. action tendencies) (4) motor (e.g. facial or vocal expressions) and (5) subjective feelings (e.g. experiences of emotions). We only analysed the motor component with facial expressions. It's possible that, while reading a text, emotions are easier to measure on other components, for instance the subjective feelings.

Our measurement tool, the FaceReader®, has been validated with a manually scored database and been found to detect emotions with an accuracy of 95% (Noldus Information Technology, 2016). Therefore, we expect the FaceReader® to have scored correctly for facial emotional valence. The mean value over all three text parts of this experiment is –0.06 with a standard deviation of .18. That is, most values are near zero, i.e. neutral. But the minimum and maximum recorded values of -0.98 and 0.98 demonstrate a very strong negative and positive emotional valence.

A reason for the low explained variance in facial emotional valence could be the mostly neutral face expression of the participants, which might be due to the silent reading in front of a screen not being a social activity. Emotional expressions are more often shown in social contexts (e.g. Hess and Hareli 2015; Fischer et al. 2003; Jacobs et al. 1997). Further analyses should eventually concentrate on micro expressions (Pfister et al., 2011) or on distinct emotions.

It is of interest that 54% of the variance is explained when the random effects are included, i.e. also including the variance between subjects. The differences between the subjects are large and could be an indicator that the sample consists of groups with different emotional reactions, i.e. with different facial expression behavior.

There may be some problems in the research design. The estimation of the reading time for each section is based on the percentage of characters in each section. This presumes a regular reading speed. There may be some divergences that diminish the precision of the analyses, or the text sections may be too large. To get a more precise definition of the reading time of each section, or to obtain a more fine-grained analysis with a division in sentences or words we will use the eye tracking data that were collected during the experiment.

A critical aspect of the analysis model is that only the lexical emotional valence was entered without including other text characteristics as it was done by Ullrich et al. (2017) and Hsu et al. (2015).

6 Conclusion

Despite the significant results - the explained variance of the fixed effects is very small - we cannot conclude that the readers' facial emotional valence corresponds to the lexical emotional valence of a given text. Reading is not mirrored in the face. Our aim to predict the emotional reactions of distance learners presenting a text with known semantic characteristics is still far away. We have to work with other measurement methods or with other methods of analysis.

Acknowledgments

We thank Dr. Stéphanie McGarrity of the Institute for Research in Open-, Distance- and eLearning (IFeL) of the Swiss Distance University of Applied Sciences (FFHS) for proof reading and her valuable advice that greatly improved the manuscript. Also, many thanks to Dr. Fernando Benites of the ZHAW School of Engineering for his great and patient technical assistance in producing the paper.

References

Ulrike Altmann, Isabel C. Bohrn, Oliver Lubrich, Winfried Menninghaus, and Arthur M. Jacobs. 2012. The power of emotional valence—from cognitive to affective processes in reading. *Frontiers in Human Neuroscience* 6(June):1–15. https://doi.org/10.3389/fnhum.2012.00192.

Toni Amstad. 1978. *Wie verständlich sind unsere Zeitungen*. Dissertation, Universität Zürich.

R Harald Baayen, Richard Piepenbrock, and Hedderik van Rijn. 1993. *The CELEX lexical database [CD-ROM]*. University of Pennsylvania, Linguistic Data Consortium, Philadelphia.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2014. Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software* 67(1):1–48. https://doi.org/10.18637/jss.v067.i01.

Nick Berggren, Anne Richards, Joseph Taylor, and Nazanin Derakshan. 2013. Affective attention under cognitive load: reduced emotional biases but emergent anxiety-related costs to inhibitory control. *Frontiers in Human Neuroscience* 7(May):1–7. https://doi.org/10.3389/fnhum.2013.00188.

Margaret M Bradley and Peter J Lang. 1994. Measuring emotion: The Self-Assessment Manikin and the semantic differential. *Journal of Behavioral Therapy and Experimental Psychiatry* 25(1):49–59.

Margaret Mm Bradley and Pj Peter J Lang. 1999. Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. *Psychology* Technical(C-1):0. https://doi.org/10.1109/MIC.2008.114.

Francesca M M Citron. 2012. Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language* 122(3):211–226. https://doi.org/10.1016/j.bandl.2011.12.007.

William C DeFraine. 2016. Differential effects of cognitive load on emotion: Emotion maintenance versus passive experience. *Emotion* 16(4):459–467.

Paul Ekman and Wallace V Friesen. 1978. *The Facial Action Coding System*. Consulting Psychological Press, Palo Alto, CA.

Phoebe C Ellsworth and Klaus R Scherer. 2003. Ellsworth et al 1994.pdf. In A. R. Newman, J. P..; Lorenz, editor, *Handbook of affective sciences*, Oxford University, New york, chapter 29, pages 572–595.

Agneta Fischer, Antony Manstead, and Ruud Zaalberg. 2003. Social influences on the emotion process. *European Review of Social Psychology* 14(2003):171–201. https://doi.org/10.1080/10463280340000054.

Rudolf Flesch. 1948. A new readability yardstick. *Journal of applied psychology* 32(3):221–233.

Nico H Frijda. 1993. The place of appraisal in emotion. *Cognition & Emotion* 7(3-4):357–387.

Ursula Hess and Shlomo Hareli. 2015. The Role of Social Context for the Interpretation of Emotional Facial Expressions. In M K Mandal and A Awasthi, editors, *Understanding facial expressions in communication*, Springer, New Delhi, pages 119–141. https://doi.org/10.1007/978-81-322-1934-7.

Chun Ting Hsu, Arthur M. Jacobs, Francesca M.M. Citron, and Markus Conrad. 2015. The emotion potential of words and passages in reading Harry Potter - An fMRI study. *Brain and Language* 142(January):96–114. https://doi.org/10.1016/j.bandl.2015.01.011.

Arthur M. Jacobs. 2015. Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience* 9(April):1–22. https://doi.org/10.3389/fnhum.2015.00186.

Esther Jacobs, Agneta H Fischer, and Antony S R Manstead. 1997. Emotional experience as a function of social context : The role of the other. *Journal of Nonverbal Behaviour* 21(2):103–130.

Roman Jakobson. 1979. *Hölderlin, Klee, Brecht: Zur Wortkunst dreier Gedichte*. Suhrkamp, Frankfurt.

JASP-Team. 2018. JASP (Version 0.8.2). https://jasp-stats.org.

Johanna K. Kaakinen, Orsolya Papp-Zipernovszky, Egon Werlen, Nuria Castells, Per Bergamin, Thierry Baccino, and Arthur M. Jacobs. 2018. Emotional and motivational aspects of digital reading. In Mirit Barzillai, Jenny Thomson, Sascha Schroeder, and Paul van den Broek, editors, *Learning to Read in a Digital World*, John Benjamins Publishing Company, Amsterdam, chapter 6, pages 143–166. https://benjamins.com/catalog/swll.17.06kaa.

Stefan Koelsch, Arthur M. Jacobs, Winfried Menninghaus, Katja Liebal, Gisela Klann-Delius, Christian von Scheve, and Gunter Gebauer. 2015. The quartet theory of human emotions: An integrative and neurofunctional model. *Physics of Life Reviews* 13:1–27. https://doi.org/10.1016/j.plrev.2015.03.001.

Alexandra Kuznetsova, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13). https://doi.org/10.18637/jss.v082.i13.

Colin Martindale. 2008. *Regressive imagery dictionary. RID-pt. ykd*. https://provalisresearch.com/products/content-analysis-software/wordstat-dictionary/re gressive-imagery-dictionary/.

Daniella S McNamara and Arthur C Graesser. 2012. Coh-Metrix: An automated tool for theoretical and applied natural language processing. In P M McCarthy and C Boonthum, editors, *Applied natural language processing: Identification, investigation and resolution*, IGI Global, pages 189–205.

Shinichi Nakagawa and Holger Schielzeth. 2013. A general and simple method for obtaining R2 from Generalized Linear Mixed-effects Models. *Methods in Ecology and Evolution* 4:133–142.

Noldus Information Technology. 2016. FaceReader. Tool for automatic analysis of facial expressions. Technical report, Noldus Information Technology, Wageningen.

James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015*. University of Texas at Austin, Austin, Texas.

Tomas Pfister, Xiaobai Li, Guoying Zhao, and Matti Pietikäinen. 2011. Recognising spontaneous facial micro-expressions. In *Proceedings of the IEEE International Conference on Computer Vision*. pages 1449–1456. https://doi.org/10.1109/ICCV.2011.6126401.

Rosalind W. Picard. 1995. A ective Computing. Technical Report 321, Media Lab. Massachusetts Institute of Technology, Cambridge University.

Aurélie Ponz, Marie Montant, Catherine Liegeois-Chauvel, Catarina Silva, Mario Braun, Arthur M. Jacobs, and Johannes C. Ziegler. 2014. Emotion processing in words: A test of the neural re-use hypothesis using surface and intracranial EEG. *Social Cognitive and Affective Neuroscience* 9(5):619–627. https://doi.org/10.1093/scan/nst034.

R-Core-Team. 2017. R: A language and environment for statistical computing. https://www.r-project.org/.

Erika L Rosenberg. 2005. The study of spontaneous facial expressions in psychology. In P Ekman and E L Rosenberg, editors, *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*, Oxford University Press, Oxford, pages 3–18. 2 edition.

James A Russell and Lisa Barrett Feldman. 1999. Core Affect , Prototypical Emotional Episodes , and Other Things Called Emotion : Dissecting the Elephant. *Journal of Personality and Social Psychology* 76(5):805–819.

Urs Schallberger. 2005. Kurzskalen zur Erfassung der Positiven Aktivierung, Negativen Aktivierung und Valenz in Experience Sampling Studien (PANAVA-KS). Forschungsberichte Nr. 6 aus dem Projekt: "Qualität des Erlebens in Arbeit und Freizeit". Technical report, Psychologisches Institut der Universität Zürich, Zürich. https://www.laufbahndiagnostik.psychologie.zhaw.ch/download/panava_schallberger_2005.pdf.

Paul Scherer. 2005. What are emotions? And how can they be measured? *Social Science Information* 44(4):695–729.

Monika Schwarz-Friesel. 2015. Language and emotion: The cognitive linguistic perspecitve. In *Emotion in Language: Theory - research – application*, John Benjamins Publishing Company, Amsterdam, pages 157–174.

Susann Ullrich, Arash Aryani, Maria Kraxenberger, Arthur M. Jacobs, and Markus Conrad. 2017. On the relation between the general affective meaning and the basic sublexical, lexical, and inter-lexical features of poetic texts-a case study using 57 Poems of H. M. Enzensberger. *Frontiers in Psychology* 7(JAN). https://doi.org/10.3389/fpsyg.2016.02073.

Lotte F Van Dillen, Dirk J Heslenfeld, and Sander L Koole. 2009. Tuning down the emotional brain: an fMRI study of the effects of cognitive load on the processing of affective images. *Neuroimage* 45(4):1212–1219.

Melissa L H Võ, Markus Conrad, Lars Kuchinke, Karolina Urton, Markus J. Hofmann, and Arthur M. Jacobs. 2009. The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods* 41(2):534–538. https://doi.org/10.3758/BRM.41.2.534.

David Watson, Lee A Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54(6):1063–1070.

Rebekah Wegener, Christian Kohlschein, Sabine Jeschke, and Stella Neumann. 2017. EmoLiTe - A database for emotion detection during literary text reading. In *Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*.