Tap pressure on pressure sensitive touch screens and the correlation with emotion

Kevin Blom
University of Amsterdam
Science Park 904
xxxxxx@xxxxxx.xx

ABSTRACT
The detection of emotion has been subject to research for a long time and is key to affective computing; computers that assist humans or have an enhanced decision ability based on the user. One of the areas that recently has become available is pressure sensitive touch screens. Touch screens are a widely accepted and used technology where users deliberatly interact with multiple times per day. Utilizing such a touch screen and an affective picture database, this study tries to predict emotion from tap pressure. Unfortunately, while literature does suggest a correlation, this study finds none. The discussion raises some weaknesses as to the cause and proposes several solutions for further investigation.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation: User-centered design

Author Keywords
Human centered multimedia; emotion; tap pressure; motor expression

INTRODUCTION
I believe that emotion aware smart devices can provides users with a better experience, where better can be classified as an experience inline with how the user is feeling. Screens and information that present themselves in different ways, where the differences originate from current feelings of the user. This type of adaptation of user interfaces can tighten the social bond between user and smart devices making the devices feel more alive. Furthermore I believe the current state of technology, and in particular machine learning, capable of using the emotion of users in product designs that dynamically change. Picard [12] states that computers that are emotion aware not only possess the ability to be of the users' assistance, but also let that computer make decisions that are influenced by those emotions that are expressed by the user. Reeves et al. [15] have shown that humans have the unconscious and undeliberate response to computers to act in a natural and social way. Why not make computers, and smart devices, respond in natural, social way as well?

CHI'16, May 07–12, 2016, San Jose, CA, USA

ISBN 123-4567-24-567/08/06. . . $15.00
DOI: http://dx.doi.org/10.475/123_4

However, as Zimmerman et al. [23] rightfully write in their conclusion: What are adequate emotional reactions of a system that knows how the user feel? One posibility could be that your telecom provider app or website can contact customer support in the background already when it notices the user is upset about their bill. Another interesting case would be music applications promoting new and exciting songs when the user is happy, but slowed down and calm songs when the user is relaxed. Or how about games that have characters that can be helpful and friendly when progression is stalling, but can also respond to aggresion and anger in a suitable way, mitigating frustration or escalation. As a final example, educational systems that can detect if the user is confused can provide extra examples, or when it notices boredom can provide more challenging excercises.

Interestingly, most initiatives to measure emotion for use with computers either require intrusive sensors, or deliberate and willful interaction, as will be described in detail later. This study explores the use of pressure sensitive touch screens of smart devices as means for a less intrusive, more ubiquitous way of detecting emotion.

Key to affective computing, i.e. computers that have the means to detect, respond, and express emotion is the detection of emotion. Further explanation will be divided into two parts; models of emotion, and measuring of emotion. Models of emotion will help to grasp the different ways to approach and classify emotions, where as measuring of emotion will help understand how measuring is currently executed and how it can be improved. Subsequently, some practical applications are discussed and finally a research question and an hypothesis is presented.

Models of emotion
There exist several different models of emotion that attempt to classify them in varying ways. In 2000, Scherer [16] presents four models to represent emotion: (a) dimensional, (b) discrete emotion, (c) meaning, and (d) componential. Their focus lies respectively with subjective feeling, motor expression and adaptive behavior patterns, verbal descriptions of subjective feelings, and the link between emotion-antecedent evaluation and differentiated reaction patterns. The study tries to name often used elicitation mechanisms, however it fails to find spe-

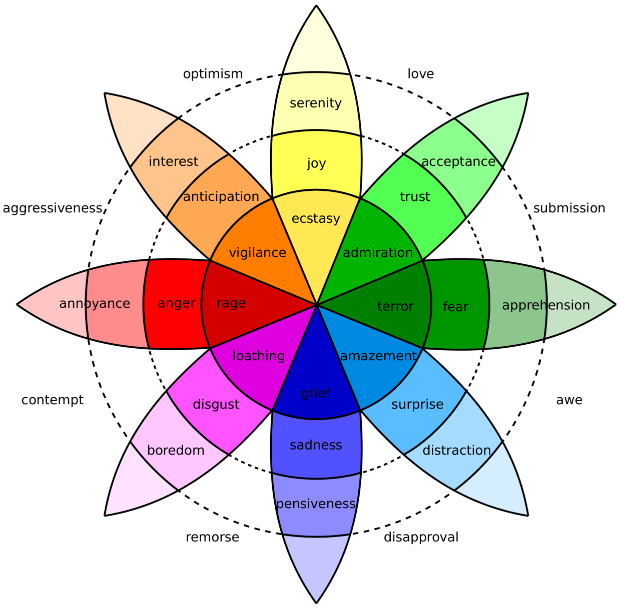


Figure 1: The discrete three dimensional model that R. Plutchik [13] introduced in 1980. It shows human emotions on a discrete scale, where the most intense emotion is displayed in the middle, and intensity drops when going outwards.

cific mechanisms for all but the compenenetial model, where an appraisal mechanism is often used. Appraisal is one of the six components that Fontaine et al. [5] incorporate in their model of emotions. The complete model consists of six components: (a) appraisals of events (b) psychophysiological changes, (c) motor expressions, (d) action tendencies, (e) subjective experiences, and (f) emotion regulation.

More recently, Shah et al. [17] state that there are in general two directions to represent emotion; discrete and continuous. The discrete model represent emotions that are measurable and physiologically distinct like angry, sad, happy, etc. [3]. A more detailed discrete model is the one proposed by Plutchik [13]. The continuous model represents emotions on a two-dimensional scale, where one axis represents *valence* and the other *arousal* [14] (Figure 2). Mauss et al. [10] suggest that using a dimensional, continuous framework is a better option when capturing emotion, relative to discrete frameworks.

Measuring emotion

The models of emotion also create the necessity to measure emotion. The measuring of emotion is a broad field that, in general, can be classified into three domains: (a) physiological, (b) facial, and (c) posture/gesture.

Physiological detection

The first domain uses physiological signals of the human body to measure and detect emotion. In a review by Wioleta[22], eight studies were collected that measure emotion using one or more physiological signals combined. These signals are (in no particular order): (a) EEG, (b) skin conductance, (c) blood volume pulse, (d) temperature, (e) heart rate, (f) blood pressure, (g) respiration, (h) EMG, and (i) ECG. Most of these physiological signals have the drawback that they need specialized sensors attached to the body, making unobtrusive measurements difficult. With the recent rise of smart wearables that include a heart rate monitor, heart rate is one of the signals that is more readily available to use in applications on smart devices.

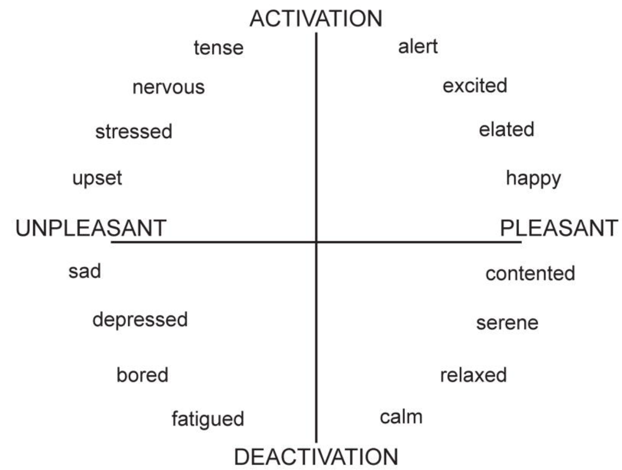


Figure 2: The two dimensional model as explained by Posner et al. [14]. Valence is expressed on the x-axis, arousal on the y-axis. Rather than classifying each emotion discretely, it can put emotion anywhere on the two axes

Facial detection

Facial detection of emotion incorporates the measurement of facial muscle movement, voice or speech [20], and furthermore includes the eye as point of detection, i.e. movement, blinking, and pupil dilation [19]. While one can argue that these methods are also physiological, they are research to such detail and extent that they can be classifed seperately. By connecting facial muscle movement to visual display of emotions, Ekman et al. [4] conclude with a basic set of six mutually exclusive emotions that could be recognized. Furthermore, De Silva et al. [18] found that several emotions are expressed by either visual or auditory cues, or both, meaning that some emotions can be recognized by visual cues alone, auditory cues alone, or need a combination of both to be detected accurately.

Posture/gestures emotion detection

Other means of detection emotions involve the tracking and interpretation of posture and gesture. Wallbott et al. [21] concluded in 1998 that there are, in some cases, distinctive patterns of movement and postural behavior that have a strong correlation to emotions. In other cases, they mention that in absence of patterns there are still distinctive features from which emotion could be inferred. Coulson et al. [1] researched static body postures and the recognition of emotions from these body postures by participants. It showed that disgust is a tough emotion to recognize but anger and sadness had over 90% correct detection rates. Furthermore, happiness and surprise were two emotions that were often confused.

Practical applications

Looking at a more practical and applied side of emotion detection, Gao et al. [6] used touchscreen devices, where the application of gestures on touch screens was successfully linked to emotional states with the use of a game. The emotional states that were tested for are: excited, relaxed, frustrated and bored, and accuracy of detection reached at minimum 69%. However, the research of Gao et al. was limited to gestures and did not incorporate data from taps. Furthermore, Lv et al. [7] have created means to detect emotion from keyboard pressure using feature extraction. This indicates that the use of a keyboard on a touch screen could also be used as means of detecting emotion, but one must keep in mind that a regular keyboard is not fully comparable to a touchscreen keyboard. It lays flat on a desk, and is often typed upon with more than one or two fingers, which means that the pressure exerted on the keyboard is likely not directly correlated with the pressure on a touchscreen keyboard. Moreover, Lee et al [8] propose an unobtrusive way of detecting emotion by analyzing smartphone usage patterns (not unlike LiKamWa et al. [9]) and social network status updates. However, this required that the user would post status updates through independently developed social networking applications, that are not officially supported by the social networks themselves.

Research question and hypotheses

From the related work can be concluded that most types of detection of emotions are invasive, either requiring constant monitoring, possibly with sensors attached to the body, or by constant recording of audio and visual data. The practical applications show that measuring is often usecase dependent. However, the practical applications have on thing in common: a touch screen. The touch screen is a technology a lot of people interact with every day, where they deliberately choose to participate in those interactions and is independent of a usecase. Using touch screen presses as indicators for emotional state could be an unintrusive way of detecting emotion without the need for constant monitoring through other sensors. With the introduction of pressure sensitive touchscreens in recent smart devices, an interesting new sensor is added to the plethora of sensors already available. Subsequently, this leads to the following research question:

Can pressure sensitive touch screen devices be used to tell more about the emotional state of the user?

The presented research question prompts thorough investigation of the connection between the pressure that a tap exerts on a touch screen and the emotional state of the user that performs the taps. The null hypothesis that springs from the research question is:

H_0: *Pressure of taps on touch screens exhibit no correlation with mood.*

and the alternative hypothesis:

H_1: *Pressure of taps on touch screens exhibit direct correlation with mood.*

The next section will present how the study was conducted before continuing to results.

METHODS

In order to test for the correlation between taps on a touch screen and emotion, there has to be a standardized way of eliciting different emotions. Fortunately, there exists a photo set that has been thoroughly tested for emotional response on a two dimensional scale that is called the Geneva Affective Picture Database (GAPED) [2]. Utilizing the emotional responses of this photo set as a baseline, touch screen taps and their pressure can be compared to emotional response. 51 Participants were selected using a convience sampling process at an office. The participants varied in age, educational level, current line of work, and background.

Emotional elicitation

Using a standardized photo set that has been thoroughly researched for emotional response when showed to participants, a ground truth for emotion was set. The GAPED photo set uses the continuous model of representing emotions, i.e. the two-dimensional valence and arousal model. The photo set counts 730 pictures and is divided into 6 categories: Animal, Human, Neutral, Positive, Snakes, Spiders. From each of the categories, 10 pictures were randomly selected, resulting in a set of 60 pictures used for the experiment. Each participant was presented with the same 60 pictures, but in random order in order to mitigate any side effects that might occur from presenting pictures in a particular order. Brown et al. [11] remark that 5 second exposure to pictures is often used for the International Affective Picture System photos. The GAPED photo set has been created because of two issues with the IAPS; extensive use decreases impact of the stimuli, and the limited number of pictures for specific themes. Both these issues are not exposure time related, so the choice of exposure time of the photo to the participant is 5 seconds.

Pressure detection

Taps were detected on an Apple iPhone 6s device with a 3D touch screen running iOS 10.3.1. The pressure of taps was registered on a floating point scale from 0.0 to 6.67 (Corresponding with 0 to ± 350 grams) and for every tap, several pressure measurements were registered in chronological order. Furthermore, the duration of a tap (in nanoseconds) was registered.

Data collection

In order to collect a larger data set, 4 taps per photo were required to advance to the next photo. These taps are directed with the use of gray colored buttons that are randomly shown on a 4 by 4 grid on the screen (Figure 3). The random pattern of the buttons ensures that the position of the tap on the screen does not matter for the pressure measurement. It uses a gray color because it is perceived as neutral. The buttons are random for every photo, and for every participant. In other words, no participant received the same grid for the same photo.

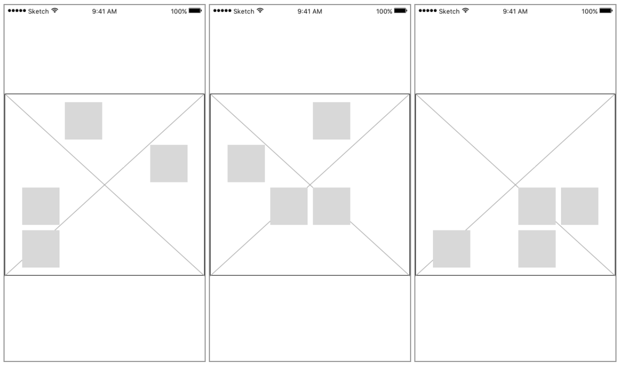


Figure 3: Three examples of the grid as presented over a picture. When a participant has pressed each of the grey squares (which disappear when pressed upon) the next picture is shown. After five seconds a new randomly generated grid is displayed on top of the picture. This process repeats 60 times.

All the data that was collected was anonymously and securely sent realtime to a Firebase [1] database. Firebase utilizes a JSON[2] tree structure that can be described as in Figure 4.

Experiment setup

Firstly, participants were told what the expirement entailed and were presented with a consent form. Subsequently, the participants continued the experiment on the smart device with test application. The test application is structured as follows:

1. Participant is presented with a screen that asks if they received and signed a consent form and if not, that they should contact the supervisor immediately. There is also a *start* button to start the experiment.
2. The participant is shown a picture.
3. After five seconds, four gray buttons are shown, overlaid on the picture in a random pattern (Figure 3).
4. When the participant pressed all the 4 buttons, the next picture is presented.
5. This process repeats untill al 60 pictures have been shown.
6. The participant is presented with a conclusive screen that has a thank you message and refers to the supervisor if there are questions.

The entirety of the experiment was completed in a room that contained no screens, speakers or other distractions and ensured the participants would fully focus on the experiment.

Data analysis

The collected data was exported as JSON from Firebase and subsequently mutated using Python[3] 2.7 on macOS[4] 10.12.4

[1]http://firebase.google.com/
[2]http://www.json.org
[3]https://www.python.org
[4]https://www.apple.com/lae/macos/sierra/

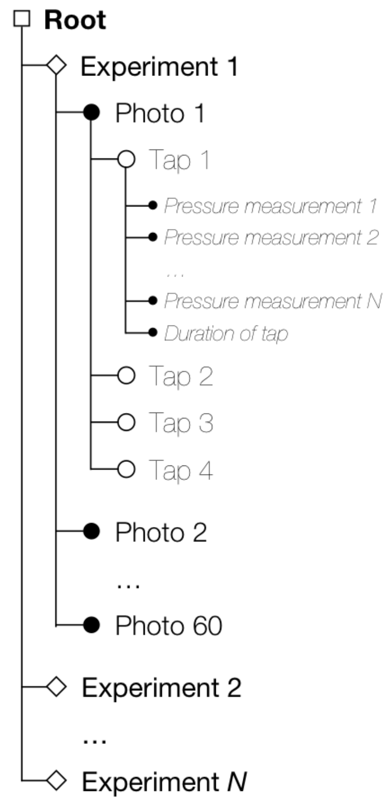


Figure 4: The database structure shows that for every experiment entry, there are 60 photo entries. Each photo entry contains four tap entries, that contain zero to *n* pressure measurements and a duration measurement

in order to create an *.csv* file that was readable by SPSS 20.0[5]. Because of a suspicion that either maximum exerted pressure or average exerted pressure of a tap might be of influence, these two variables were manually added using averaging. For maximum pressure, the maximum pressure value of each tap was extracted, and for each photo this was averaged. Regarding average tap pressure, the average pressure of a tap was calculated and subsequently all the average tap pressures were averaged again per photo. These averages make it possible to compare means. The result is six variables in SPSS;

1. **Photo filename** - String, containing the photo filename for identification purposes.
2. **Valence** - Numeric, decimal value on a scale from 0.0-100.0.
3. **Arousal** - Numeric, decimal value on a scale from 0.0-100.0.
4. **Maximum tap pressure average** - Numeric, decimal value on scale from 0.0-6.67.
5. **Average tap pressure average** - Numeric, decimal value on scale from 0.0-6.67.

[5]https://www.ibm.com/analytics/us/en/technology/spss/

6. **Duration** - Numeric, decimal value in nanoseconds.

In other words, for every photo there is a value for valence, arousal, maximum tap pressure average, average tap pressure average and duration.

Multiple linear regression
Testing for any correlation was completed with a multiple linear regression method. The advantage of this method is that both *duration* and *pressure* can be used as independent variables to check if there indeed is a relation to *valence* or *arousal* as dependent variables and that if there is a relationship, it also immediately produces a model to predict the dependent variables. A disadvantage of is that the relationship can not be checked for *valence* and *arousal* simultaneously, only for the seperate variables.

What follows is four seperate multiple linear regression tests, two with *valence* as dependent variable, and two with *arousal* as dependent variable. Both dependent variables were tested for any relation ship with either *maximum tap pressure & duration* or *average tap pressure & duration*.

Before proceeding to the results, several assumptions needed to be considered before concluding that the data could be analysed using multiple linear regression;

1. **Independence of observation** - Using Durbin-Watson to test for 1st-order autocorrelation. Value should be 2 ± 0.5.
2. **Linear relationships** - Visually inspecting a scatterplot of studentitized residuals and unstandardized predicted values, and partial regression plots can indicate a (non)linear relationship.
3. **Homoscedasticity of residuals** - Visually inspecting a scatter plot of studentitized residuals and unstandardized predicted values.
4. **No multicollinearity** - Inspection of correlation coefficients and Tolerance/VIF values for indication of correlation between independent variables.
5. **No unusual data points** - There should be no outliers, high leverage points or highly influential points.
6. **Normal distribution of errors** - Errors in prediction need to be normally distributed, otherwise determining significance can become problematic. By visually inspecting the histogram, Q-Q and P-P plots.

All these assumptions are considered for every of the four seperate multiple linear regression tests mentioned. Results of testing the assumptions are also presented in the next section.

RESULTS
This section starts with presenting the results of testing the assumptions. Subsequently, the model of the multiple linear regression is presented.

Assumptions
These are the results of checking each of the six assumptions. The section Methods explains the necessity of checking assumptions in more detail.

Independence of observation
As can be seen in Table 1, the Durbin-Watson values do not violate acceptable limits for each seperate test, indicating no autocorrelation.

Dependent variable	**Independent variables**	**d-w**
Valence	Max. tap pressure, duration	2.086
	Avg. tap pressure, duration	2.094
Arousal	Max. tap pressure, duration	2.246
	Avg. tap pressure, duration	2.257

Table 1: Durbin-Watson values (d-w) outside of 2 ± 0.5 indicate autocorrelation issues.

Linear relationships
By visually inspecting scatter plots of the *unstandardized predicted value* against *studentized residuals* it can be assumed there is linearity. Furthermore, by looking at partial regression plots of each of the independent variables for every dependent variable, it is again apparent that there is an approximate linear relationship. See Appendix Linear Relationships and Homoscedasticity for the graphs used for inspection.

Homoscedasticity
Using the scatter plots of *unstandardized predicted value* against *studentized residuals* for inspection, the random spread of values indicate homoscedasticity of values. See Appendix Linear Relationships and Homoscedasticity for the graphs used for inspection.

Multicollinearity
Inspection of correlation coefficients (Appendix Multicollinearity for full tables) show none of the correlations > 0.7. In Table 2, tolerance values are found. None of the tolerance values fall below the limit of 0.1.

Dependent variable	**Independent variables**	**Tolerance**
Valence	Maximum tap pressure	.728
	Duration	.728
	Average tap pressure	.758
	Duration	.758
Arousal	Maximum tap pressure	.728
	Duration	.728
	Average tap pressure	.758
	Duration	.758

Table 2: Tolerance values < 0.1 indicate collinearty issues.

Unusual data points
For each seperate test, *studentized residuals*, *Cook's Distances* and *Leverage values* were checked. There were no cases outside 3 Standard Deviations (SDs), no distances > 1.0 and no values > 0.2 respectively (Appendix Unusual Data Points for minimum and maximum values for each).

Normal distribution of errors

Visual inspection of histograms and P-P plots for each seperate test show no signs of significant violation of normality. (Appendix Normality for the inspected graphs and diagrams)

Findings

A multiple regression was run to predict either valence or arousal from maximum tap pressure and duration, or average tap pressure and duration. Table 4 shows the model summary of each of the four multiple linear regression tests, indicating that no model allows for significant prediction of valence or arousal, $p > .05$. Moreover, taking a look at Table 3, it can be seen that none of the seperate independent variables significantly add to the prediction model. (See Appendix Regression Coefficients for detailed coefficient information)

Dependent variable	Independent variables	Significance
Valence	Intercept	.702
	Max. tap pressure	.353
	Duration	.761
	Intercept	.669
	Avg. tap pressure	.306
	Duration	.750
Arousal	Intercept	.925
	Max. tap pressure	.095
	Duration	.722
	Intercept	.985
	Avg. tap pressure	.131
	Duration	.830

Table 3: Regression variables and their significance towards contribution to the model.

DISCUSSION

The main purpose of this research is to discover a meaningful and significant relationship between tap pressure on a touch screen and emotion. Using a quantitative approach with a sample size of 51 and multiple linear regression to explore this possible relationship, first results interestingly indicate no significant relationship.

Evaluation of findings

Taking into regard hypothesis H_0, Table 4 show us that none of the tests create a significant model, $p < .05$. This means that hypothesis H_0 is accepted, and conversely hypothesis H_1 is rejected. One of the indicators why the null hypothesis is accepted lies with the adjusted R^2. For valence tests R^2 values are negative, -.019 and -.015. This occurs when the model contains independent variables (*tap pressure*, *duration*) that do not contribute to a prediction of the dependent variable (*valence*). For arousal tests R^2 is barely positive, .021 and .012. While this indicates that the independent values do contribute to a prediction (in contrast with negative R^2), it only does so very slightly. The values show that the independent variables (*tap pressure*, *duration*) only explain 2.1% and 1.2% of the variability of the dependent variable (*arousal*). Interestingly, these findings are in direct opposition with the findings of Lv et al. [7] , where there is a a direct correlation between pressure and emotion. This leads to believe that there are some factors in the experiment set up are not taken into account or applied erroneously.

Weaknesses

One weakness in the experiment design that could influence the result is the random order of presentation of pictures. Every participant was shown 60 pictures in random order and the rationale behind that was to eliminate any unintentional side effect of several pictures eliciting the same emotion and in their turn strenghten or weaken the emotional response. However, this also means that none of the executed experiments were fully the same. Another possible weakness is again related to implementing randomness into the design. the grid overlay that was shown on every picture was also randomly generated in order to negate the effects of the position of the tap on the screen. This was decided for because with enough measurements the position of the tap starts to have less impact if the position is random. However, giving participants the same grids for the same photo might produce results that are signficant.

Furthermore, this research assumes that every individual has the same physical response to emotional elicitation. In other words, it assumes that every individual will exert a specifc amount of pressure on a touch screen.

Recommendations

If one is to venture into further research for this topic, it would be a wise decision to first try and repeat this experiment without the randomization that has been in place. Furthermore, rather than assuming every individual to respond the same to emotional elicitation and expecting a linear correlation, underlying patterns can be discovered. Especially looking for patterns specific to the participant rather than trying to find a pattern for the population. A suggestion to discover such patterns is the use of self-assessment as a methodology. Self-assessment by participants makes it possible to link tap pressure data to how the participant is feeling, rather then assuming a baseline set by visual stimuli.

REFERENCES

1. Mark Coulson. 2004. Attributing emotion to static body postures: recongition accuracy, confusion and view point dependence. *Journal of Nonverbal Behavior* 28, 2 (2004), 117–139. DOI: http://dx.doi.org/10.1023/B:JONB.0000023655.25550.be

2. Elise S Dan-glauser and Klaus R Scherer. 2011. The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. (2011), 468–477. DOI: http://dx.doi.org/10.3758/s13428-011-0064-1

3. Paul Ekman. 1992. An argument for basic emotions. (1992). DOI: http://dx.doi.org/10.1080/02699939208411068

4. Paul Ekman, E Richard Sorenson, and Wallace V Friesen. 1969. Pan-Cultural Elements in Facial Displays of

Predicted variable	Predictor variables	F(2, 57)	Adjusted R^2	Significance
Valence	Max. tap pressure, duration	.462	-.019	.632
	Avg. tap pressure, duration	.556	-.015	.577
Arousal	Max. tap pressure, duration	1.632	.021	.204
	Avg. tap pressure, duration	1.361	.012	.264

Table 4: Summary of main findings. F-Value: F(*regression degrees of freedom*, *residual degrees of freedom*). Significance: *p-value*.

Emotion. (1969). DOI: http://dx.doi.org/10.1126/science.164.3875.86

5. Johnny R J Fontaine, Klaus R Scherer, Etienne B Roesch, C Phoebe, Johnny R J Fontaine, Klaus R Scherer, Etienne B Roesch, and Phoebe C Ellsworth. 2007. The World of Emotions Is Not Two-Dimensional. 18, 12 (2007), 1050–1057. DOI: http://dx.doi.org/10.1111/j.1467-9280.2007.02024.x

6. Yuan Gao, Nadia Bianchi-Berthouze, and Hongying Meng. 2012. What Does Touch Tell Us about Emotions in Touchscreen-Based Gameplay? *ACM Transactions on Computer-Human Interaction* 19, 4 (2012), 1–30. DOI: http://dx.doi.org/10.1145/2395131.2395138

7. J. Dong H. R. Lv, Z. L. Lin, W. J. Yin. 2008. Emotion recognition based on pressure sensor keyboards. *2008 IEEE International Conference on Multimedia and Expo* (2008), 1089–1092. DOI: http://dx.doi.org/10.1109/ICME.2008.4607628

8. Hosub Lee, Young Sang Choi, Sunjae Lee, and I. P. Park. 2012. Towards unobtrusive emotion recognition for affective social communication. *2012 IEEE Consumer Communications and Networking Conference, CCNC'2012* (2012), 260–264. DOI: http://dx.doi.org/10.1109/CCNC.2012.6181098

9. Robert Likamwa, Yunxin Liu, Nicholas D. Lane, and Lin Zhong. 2013. MoodScope: Building a Mood Sensor from Smartphone Usage Patterns. *MobiSys '13 Proceeding of the 11th annual international conference on Mobile systems, applications, and services* April (2013), 389–402. DOI: http://dx.doi.org/10.1145/2462456.2464449

10. Iris B Mauss and Michael D Robinson. 2009. Measures of emotion: A review. *Cognition and Emotion* 23, 2 (2009), 209–237. DOI: http://dx.doi.org/10.1080/02699930802204677

11. Human Neuroscience, Stephen B R E Brown, Henk Van Steenbergen, P Guido, H Band, Mischa De Rover, and Sander Nieuwenhuis. 2012. Functional significance of the emotion-related late positive potential. 6, February (2012), 1–12. DOI: http://dx.doi.org/10.3389/fnhum.2012.00033

12. Rosalind W. Picard. 1995. Affective Computing. *MIT press* 321 (1995), 1–16. DOI: http://dx.doi.org/10.1007/BF01238028

13. Robert Plutchik. 1980. Psychoevolutionary Theory of Basic Emotions. *American Scientist* February (1980), 2007. http://web.archive.org/web/20010716082847/http://americanscientist.org/articles/01articles/Plutchik.html

14. Jonathan Posner, James A Russell, and Bradley S Peterson. 2005. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology* 17, 3 (2005), 715–34. DOI: http://dx.doi.org/10.1017/S0954579405050340

15. Byron Reeves and Clifford Nass. 1998. The media equation. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places* 34 (1998), 19–36. DOI: http://dx.doi.org/10.1109/MSPEC.1997.576013

16. Klaus R. Scherer. 2000. *Psychological models of emotion*. 137–162 pages. DOI: http://dx.doi.org/10.1097/00001504-198811000-00012

17. Sachin Shah, J. Narasimha Teja, and Samit Bhattacharya. 2015. Towards affective touch interaction: predicting mobile user emotion from finger strokes. *Journal of Interaction Science* 3, 1 (2015), 6. DOI: http://dx.doi.org/10.1186/s40166-015-0013-z

18. Liyanage C D E Silva and I Tsutomu Miyasato. 1997. Facial Emotion Recognition Using. September (1997), 9–12. DOI:http://dx.doi.org/10.1109/SMC.2015.387

19. Mohammad Soleymani, Maja Pantic, and Thierry Pun. 2015. Multimodal emotion recognition in response to videos (Extended abstract). *2015 International Conference on Affective Computing and Intelligent Interaction, ACII 2015* 3, 2 (2015), 491–497. DOI: http://dx.doi.org/10.1109/ACII.2015.7344615

20. Dimitrios Ververidis, Constantine Kotropoulos, and Pitas Ioannis. 2004. Automatic Emotional Speech Classification. *Artificial Intelligence and Information Analysis Laboratory* (2004), 593– 596. DOI: http://dx.doi.org/10.1109/ICASSP.2004.1326055

21. Harald G. Wallbott. 1998. Bodily Expression of Emotion. *European Journal of Social Psychology* 28, 6 (1998), 879–896. DOI: http://dx.doi.org/10.1002/(SICI)1099-0992(1998110)28:6<879::AID-EJSP901>3.0.CO;2-W

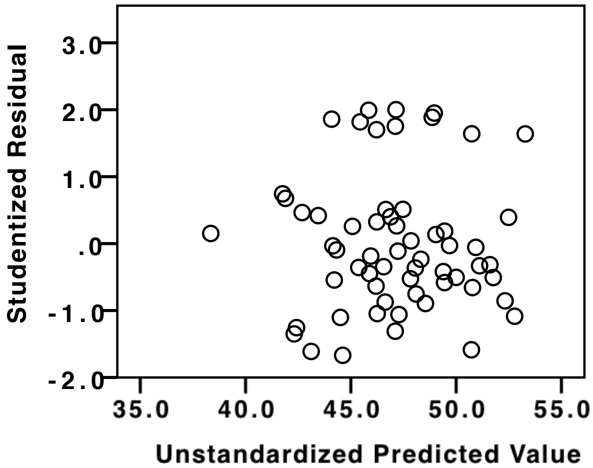
22. Szwoch Wioleta. 2013. Using physiological signals for emotion recognition. *Human System Interaction (HSI), 2013 The 6th ...* (2013), 556–561. DOI: http://dx.doi.org/10.1109/HSI.2013.6577880

23. Philippe Zimmermann, Sissel Guttormsen, Brigitta Danuser, and Patrick Gomez. 2003. Affective computingâĂŤa rationale for measuring mood with mouse and keyboard. *International Journal of Occupational Safety and Ergonomics* 9, 4 (2003), 539–551. DOI: http://dx.doi.org/10.1080/10803548.2003.11076589

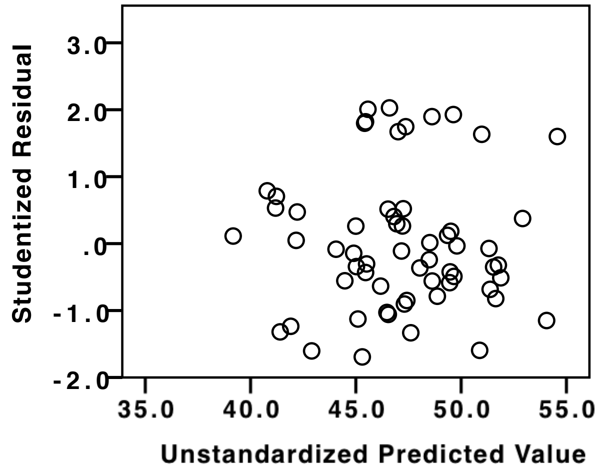
APPENDIX

LINEAR RELATIONSHIPS AND HOMOSCEDASTICITY

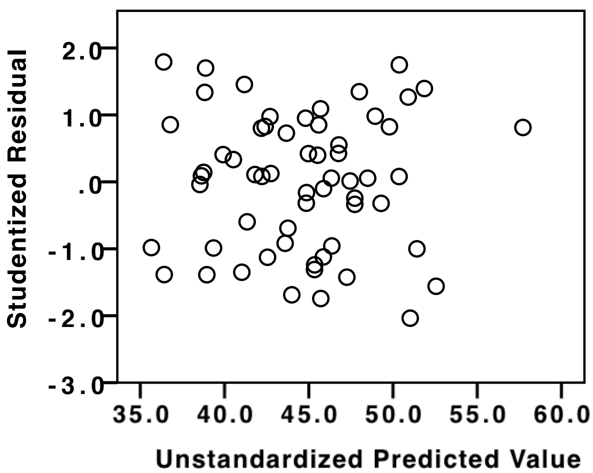
Scatter plot: unstandardized predicted value vs. studentized residuals



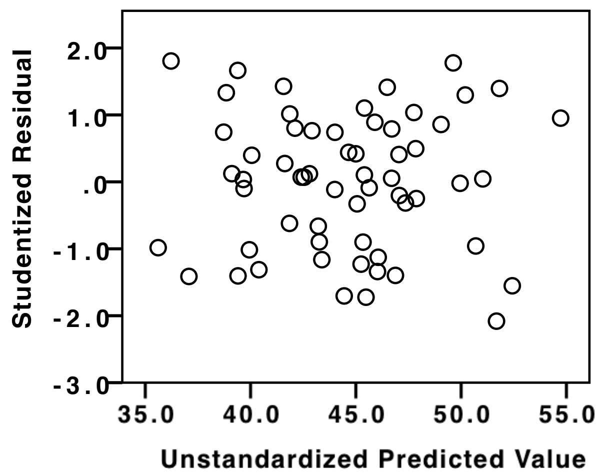
(a) Related to test of valence, maximum pressure, and duration



(b) Related to test of valence, average pressure, and duration



(c) Related to test of arousal, maximum pressure, and duration



(d) Related to test of arousal, average pressure, and duration

Figure 5: Scatter plots of predicted values against studentitized residuals. Note that because of random nature, linearity can still be assumed.

Partial regression plots

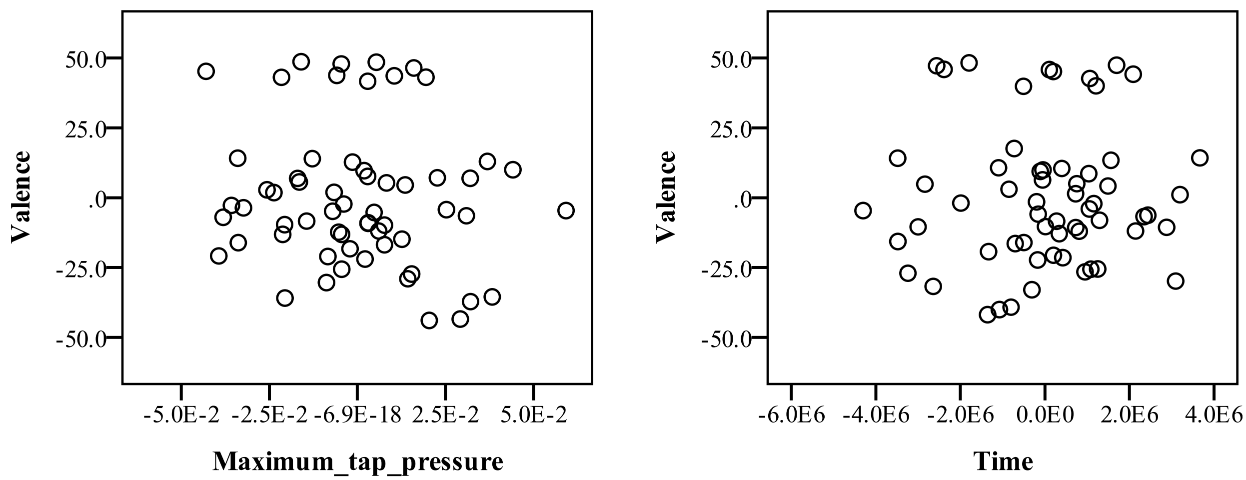


Figure 6: Partial regression plots with valence (dependent variable), maximum pressure and duration (independent variables). Note the approximate linearity.

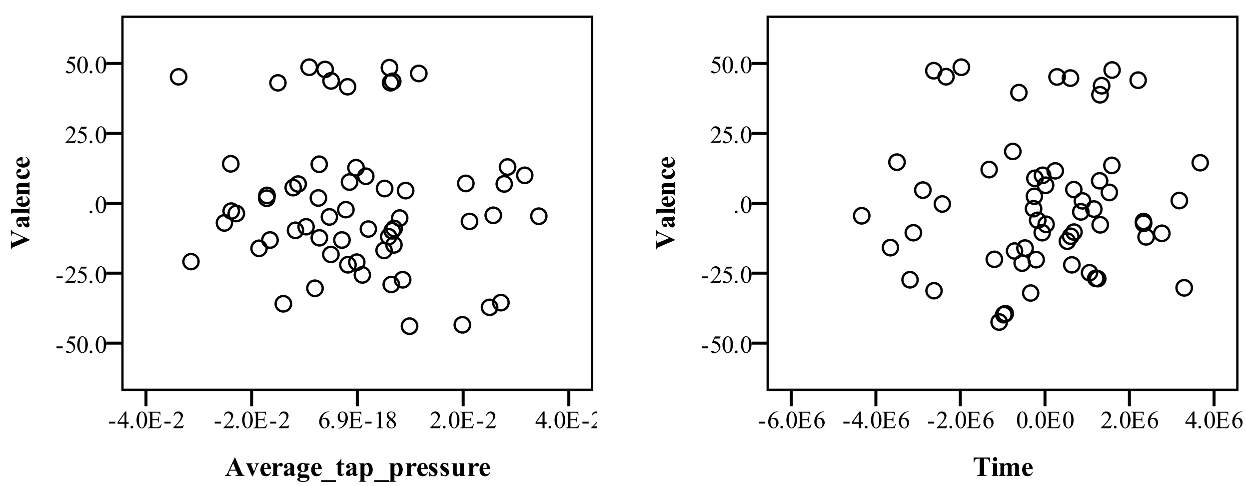


Figure 7: Partial regression plots with valence (dependent variable), average pressure and duration (independent variables). Note the approximate linearity.

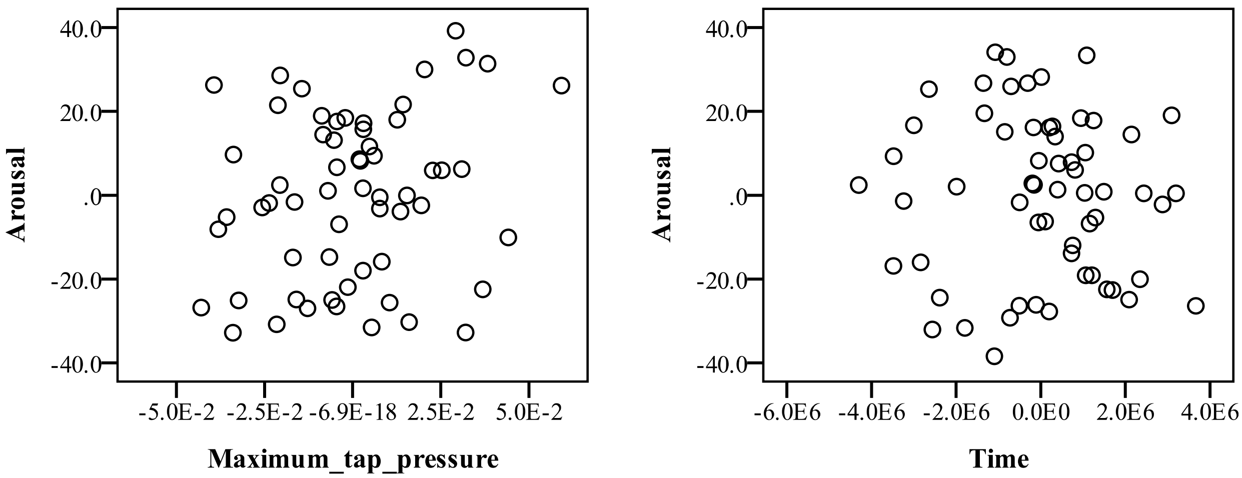


Figure 8: Partial regression plots with arousal (dependent variable), maximum pressure and duration (independent variables). Note the approximate linearity.

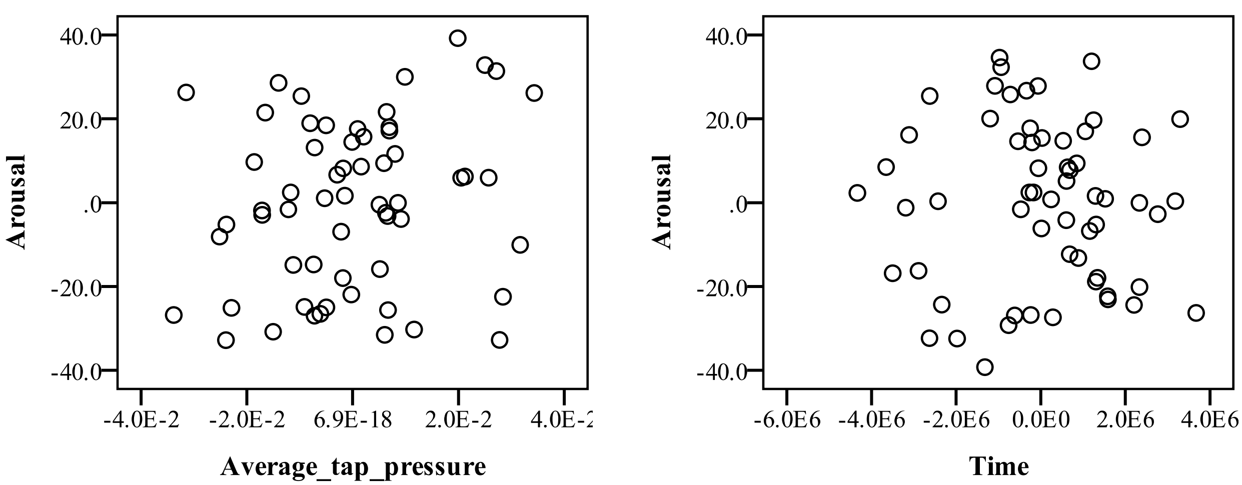


Figure 9: Partial regression plots with arousal (dependent variable), average pressure and duration (independent variables). Note the approximate linearity.

MULTICOLLINEARITY

		Valence	Duration	Max. tap pressure
Pearson Correlation	Valence	1.000	-.028	-.120
	Duration	-.028	1.000	.522
	Max. tap pressure	-.120	.522	1.000
Sig. (1-tailed)	Valence	.	.415	.181
	Duration	.415	.	.000
	Max. tap pressure	.181	.000	.

Table 5: Correlations of valence, maximum tap pressure and duration.

		Valence	Duration	Avg. tap pressure
Pearson Correlation	Valence	1.000	-.028	-.132
	Duration	-.028	1.000	.492
	Avg. tap pressure	-.132	.492	1.000
Sig. (1-tailed)	Valence	.	.415	.158
	Duration	.415	.	.000
	Avg. tap pressure	.158	.000	.

Table 6: Correlations of valence, average tap pressure and duration

		Arousal	Duration	Max. tap pressure
Pearson Correlation	Arousal	1.000	.080	.228
	Duration	.080	1.000	.522
	Max. tap pressure	.228	.522	1.000
Sig. (1-tailed)	Arousal	.	.272	.040
	Duration	.272	.	.000
	Max. tap pressure	.040	.000	.

Table 7: Correlations of arousal, maximum tap pressure and duration

		Arousal	Duration	Avg. tap pressure
Pearson Correlation	Arousal	1.000	.080	.212
	Duration	.080	1.000	.492
	Avg. tap pressure	.212	.492	1.000
Sig. (1-tailed)	Arousal	.	.272	.052
	Duration	.272	.	.000
	Avg. tap pressure	.052	.000	.

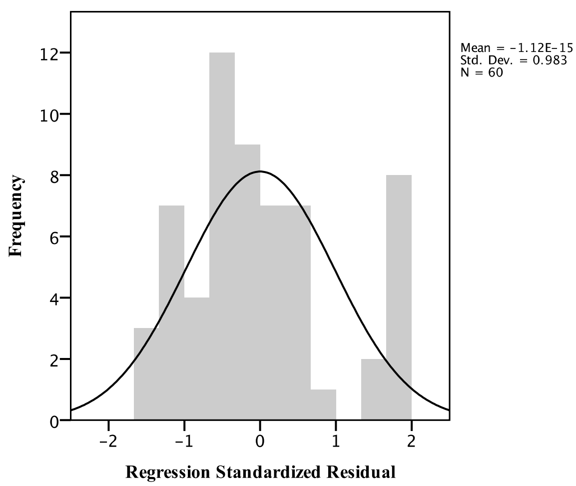
Table 8: Correlations of arousal, average tap pressure and duration

UNUSUAL DATA POINTS

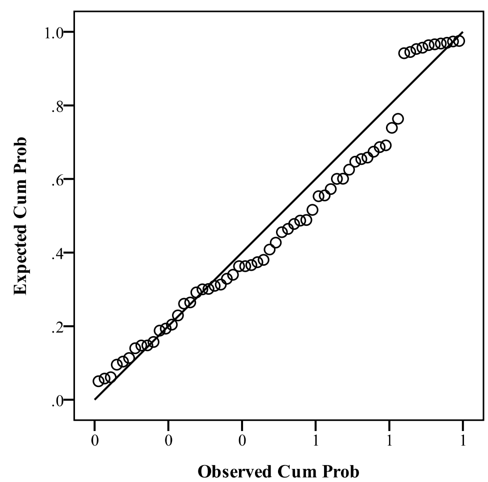
Dependent variable	Independent variables	Studentized residuals		Cook's Distance		Leverage values	
		Min	Max	Min	Max	Min	Max
Valence	Max. tap pressure, duration	-1.7	2.06	.0	.09	.0	.14
	Avg. tap pressure, duration	-1.72	2.09	.0	.09	.0	.1
Arousal	Max. tap pressure, duration	-2.1	1.83	.0	.08	.0	.14
	Avg. tap pressure, duration	-2.15	1.84	.0	.11	.0	.1

Table 9: Studentized residuals, Cook's values and leverage values. Studentized residuals are between -3 and +3 SDs. Cook's Distance values are < 1. Leverage values are $< .2$.

NORMALITY

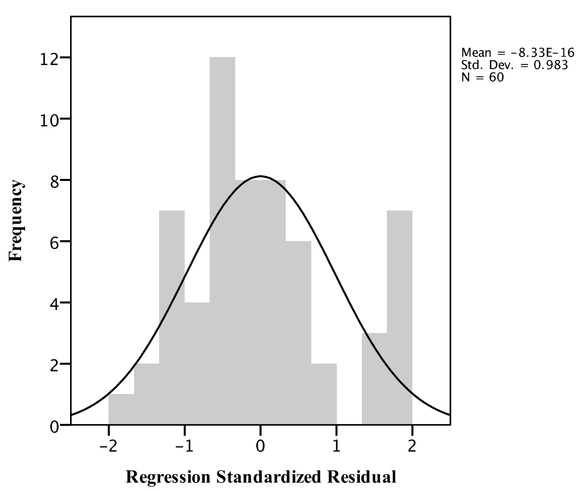


(a) Histogram

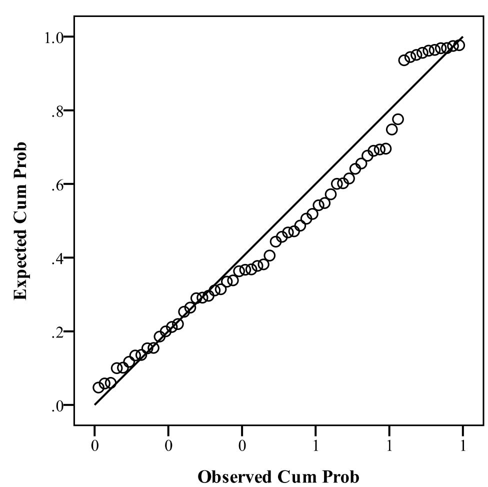


(b) P-P plot

Figure 10: Normality graphs for valence, maximum tap pressure and duration.

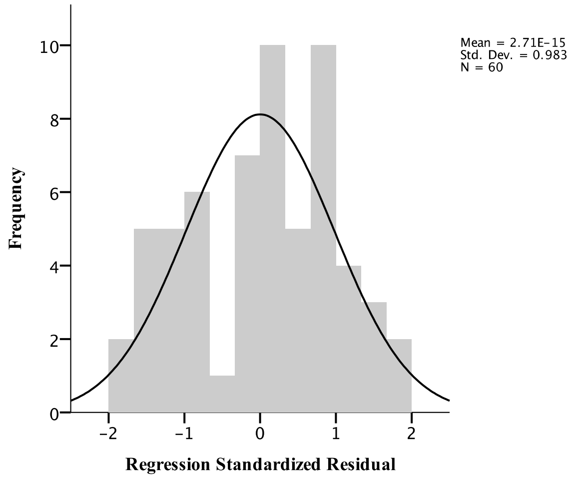


(a) Histogram

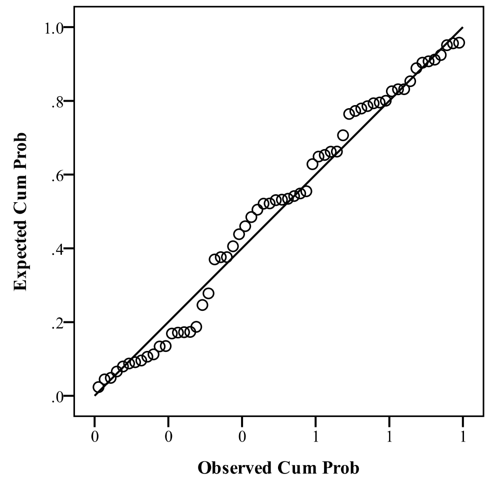


(b) P-P plot

Figure 11: Normality graphs for valence, average tap pressure and duration.

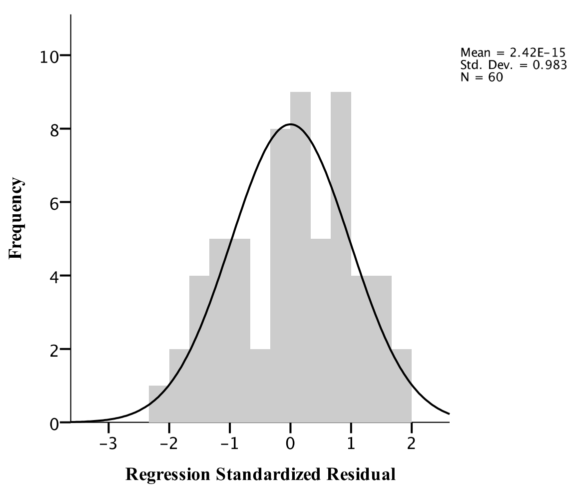


(a) Histogram

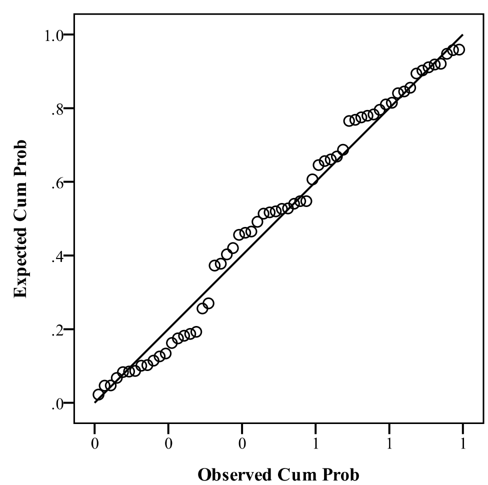


(b) P-P plot

Figure 12: Normality graphs for arousal, average tap pressure and duration.



(a) Histogram



(b) P-P plot

Figure 13: Normality graphs for arousal, average tap pressure and duration.

REGRESSION COEFFICIENTS

Dependent variable	**Independent variable***	B	*SE_b*	β	*Significance*
Valence	Intercept	53.722	139.848	n.a.	.702
	Max. tap pressure	-136.476	145.694	-.144	.353
	Duration	$5.566e^{-7}$	.000	.047	.761
	Intercept	59.459	138.468	n.a.	.669
	Avg. tap pressure	-211.353	204.722	-.156	.306
	Duration	$5.698e^{-7}$	.000	.048	.750
Arousal	Intercept	10.397	109.668	n.a.	.925
	Max. tap pressure	193.809	114.182	.256	.095
	Duration	$-5.1e^{-7}$	.000	-.054	.722
	Intercept	-2.076	109.254	n.a.	.985
	Avg. tap pressure	247.231	161.530	.227	.131
	Duration	$-3.027e^{-7}$	.000	-.032	.830

Table 10: Regression coefficients table with significance. B = Unstandardized coefficient. SE_b = Std. Error. β = Standardized coefficients. Significance = p-value. **Note*: Intercept should not be regarded as independent variable.