Analysis of Multiple Physiological Sensor Data

Lauren Reinerman-Jones[1], Grant Taylor[1], Keryl Cosenzo[2], and Stephanie Lackey[1]

[1] University of Central Florida, Institute for Simulation and Training
3100 Technology Pkwy, Orlando, FL 32826
[2] Army Research Laboratory
Aberdeen Proving Ground, MD
lreinerm@ist.ucf.edu

Abstract. Physiological measures offer many benefits to psychological research including objective, non-intrusive assessment of affective and cognitive states. However, this utility is limited by analysis techniques available for testing data recorded by multiple physiological sensors. The present paper presents one set of data that was attained from a repeated measures design with a nominal independent variable for analysis. Specifically, the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008), a series of images known to convey seven different emotions, was presented to participants while measures of their neurological activity (Electroencephalogram; EEG), heart rate (Electrocardiogram; ECG), skin conductance (Galvanic Skin Respond; GSR), and pupillary response were taken. Subsequently, a discussion of statistics available for analyzing responses attained from the various sensors is presented. Such statistics include correlation, ANOVA, MANOVA, regression, and discriminant function analysis. The details on design limitations are addressed and recommendations are given for employing each statistical option.

Keywords: EEG, ECG, Eye Tracking, Statistical Analyses, Emotion.

1 Introduction

Psychological research benefits from the implementation of physiological measurement techniques as a way to assess and predict performance. Physiological measures index cognitive state and resources used. Along those lines, resource theory is supported and cognitive states indexed include workload, stress, fatigue, and emotion. Unlike surveys and questionnaires that require task interruption to be administered, physiological measures are continuous and tend to be relatively non-intrusive. The advantage is that state is assessed throughout an entire task, not just before or after task completion. Thus, dynamic changes are recorded and a detailed understanding of the phenomenon under investigation is provided. Additionally, physiological measures provide an objective method for evaluating state, unlike questionnaires that fall subject to bias. Ultimately a multi-dimensional approach employing physiological, subjective, and performance measures is best to account for the most variance, but a couple challenges need to be addressed with regard to physiological assessment.

D.D. Schmorrow and C.M. Fidopiastis (Eds.): FAC 2011, HCII 2011, LNAI 6780, pp. 112–119, 2011.

Physiological measures are currently limited by two factors. First, different physiological measures tend to not strongly correlate with one another. For example, results attained using Electroencephalogram (EEG) typically have low correlations with Electrocardiogram (ECG) responses recorded simultaneously on the same task. This makes it difficult to determine which measure to discuss for interpretation and which is accurate. Thus, there is no standardization throughout literature and meta-analysis is challenging. Second, individual difference is an influential factor across physiological measures and is the reason for taking baseline readings. Eye tracking is the exception to the rule of needing a baseline. The restriction is that a system employing physiological measures must be calibrated to the individual user and every time the operator changes, this calibration process must occur. Physiological response during a task is compared to the baseline reading taken before a task to determine the amount of resources utilized and change in state. The goal is to find an analysis technique that allows researchers to create a physiological profile of overall state that can then be used as a model for all individuals, enabling a percentage or type of change to inform a closed-loop system for any person entering the system. To clarify, a physiological profile for workload might be described as an increase in certain EEG activity, perspiration, heart rate, and pupil dilation.

Keeping with that effort, the current paper provides a review of statistics as applied to one physiological data set. Knowledge certainly is gained by analyzing physiological variables separately, but their greatest potential for usefulness lies in the ability to perform multivariate analyses. Univariate analysis determines the effect of the independent variable (IV) on the dependent variable (DV) in isolation, but multivariate analysis lends itself to investigating potentially complex interactions that occur between the many variables, thus moving closer to identifying state profiles. Easily accomplished with some study designs, other data structures do not lend themselves to multivariate analyses. The aim for the present discussion is to investigate the potential solutions to analyzing data collected from one experimental design – a repeated measures design with a nominal independent variable.

2 Method

2.1 Participants

Forty-six participants ranging in age from 18-40 years volunteered from the University of Central Florida.

2.2 Procedure

The present study examined the influence of emotion on various physiological responses. The International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008), a series of images known to convey seven different emotions, was presented to participants while measures of their neurological activity (Electroencephalogram; EEG), heart rate (Electrocardiogram; ECG), skin conductance (Galvanic Skin Respond; GSR), and pupillary response were taken. Specifically, participants were required to watch a computer monitor as 42 images were presented, meaning that participants saw six images of each of the emotional

categories – amusement, awe, contentment, disgust, excitement, fear, and sadness. The presentation of the images was randomized, with each participant observing a uniquely randomized order. No physical or verbal response was required. Each image was presented for six seconds with a six second inter-stimulus interval.

Before the study began, each participant completed a series of baseline tasks to account for individual differences in physiological activity. The baseline for EEG was Advanced Brain Monitoring's (ABM) three task baseline battery consisting of a visual stimulus-response eyes open task, auditory stimulus-response eyes closed task, and a short vigil. Analyses used data computed from the change in power measured during the task compared to the power measured during baseline. A five minute resting baseline (with eyes closed) was recorded for ECG and GSR for use as a comparison of the recordings obtained during the experimental task. A baseline was not required for pupil diameter.

3 Results and Discussion

To attack the problem of analyzing physiological data, several statistical analyses were conducted. The two most common methods for revealing trends and effects were performed: correlation and Analysis of Variance (ANOVA). As previously mentioned these often provide interesting results, but are limited. Additional analyses were explored for comparing the multiple DVs present when using physiological measures. Multivariate Analysis of Variance (MANOVA), regression, and discriminant function analysis (DFA) were discussed as potential options with the understanding that the application of such analyses would need to be performed using a different experimental design.

Table 1. Intercorrelations between physiological measures

*correlation is significant at $\alpha = .05$.

	IBI	HRV	GSR
HRV	$r = 0.944$* $p < 0.001$		
GSR	$r = 0.014$ $p = 0.547$	$r = 0.010$ $p = 0.660$	
Pupil	$r = 0.003$ $p = 0.885$	$r = 0.040$ $p = 0.084$	$r = 0.179$* $p < 0.001$

3.1 Correlation

Correlation data is often a valuable starting point for analysis. However, this approach is unlikely to reveal insight from the physiological measures, as week correlations often result. In this particular study a total of 31 data points were obtained for each stimulus presentation [27 EEG values (Nine Sensor Sites: F3, Fz, F4, C3, Cz, C4, P3, P0, P4), interbeat interval (IBI), heart rate variability (HRV), galvanic skin response (GSR), and pupil diameter]. The average intercorrelation between these variables,

taking the absolute value of all 31 points so as to eliminate a possible artifact of averaging positive and negative numbers, was found to be moderately weak (r=0.2870). Even the intercorrelations within the EEG data produced a moderately weak average (r=0.346). The remaining correlations can be seen in *Table 1*.

3.2 Analysis of Variance (ANOVA)

Repeated measures ANOVA is arguably the most practical analysis to conduct for the given dataset because the design and assumption requirements are met. By conducting a series of ANOVAs with each of the measurements (EEG values, interbeat interval (IBI), heart rate variability (HRV), galvanic skin response (GSR), and pupil diameter) obtained by the physiological sensors as the dependent variable and emotion type as the independent variable, a basic understanding of how the different physiological systems respond to emotions is obtained.

ECG. The ECG data was processed to form two separate variables: inter-beat interval (IBI, the inverse of heart rate) and heart rate variability (HRV). As evidenced by the very strong correlation between these two variables (r = .944, p < .001), they are roughly equivalent. This equivalence is likely due to the brief time periods over which each data point is sampled. Over longer time periods, these two methods of processing ECG data would likely produce two distinct values, but it seems six seconds is not sufficient to adequately differentiate between the two. The ANOVA results show virtually identical results for both variables. Significant main effects for emotion were found for both IBI [F(6, 264) = 7.435, p < .001] and HRV [F(6, 264) = 7.406, p < .001]. Subsequent pairwise comparisons found identical patterns for both variables, the only difference being the direction of the group differences, which is a result of the algorithms used to process the two values. The direction of results of the IBI pairwise comparisons are presented in *Figure 1*.

- Sadness > Amusement, Awe, Contentment, Excitement, and Fear
- Disgust > Awe, Contentment, Excitement, and Fear
- Amusement > Excitement
- Amusement < Sadness
- Awe, Contentment, and Fear < Disgust and Sadness
- Excitement < Amusement, Disgust, and Sadness

Fig. 1. IBI pairwise comparison results. All listed comparisons are significant at α = .05.

GSR. GSR data was computed as the average skin conductance measured over the six seconds of each image presentation. A repeated measures ANOVA found no significant effect of emotion category on GSR [F(6, 264) = 0.958, p = .454].

Pupillary Response. The pupil response was measured using a head-mounted camera, which recorded the average pupil diameter for each image presentation. The ANOVA showed that pupil diameter was significantly influenced by emotion category [F(6, 270) = 19.903, p < .001]. Details from the pairwise comparisons are presented in *Figure 2*.

- Awe < Amusement, Contentment, Disgust, Excitement, Fear, and Sadness
- Excitement < Amusement, Disgust, Fear, and Sadness
- Excitement > Awe
- Contentment < Amusement, Disgust, and Fear
- Contentment > Awe
- Sadness < Disgust and Fear
- Sadness > Awe and Excitement
- Amusement > Awe, Contentment, and Excitement
- Disgust and Fear > Awe, Contentment, Excitement, and Sadness

Fig. 2. Pupil diameter pairwise comparison results. All listed comparisons are significant at $\alpha = .05$.

EEG. The large amount of EEG data required a much more complex analysis. Each image generated 27 separate EEG values (power within three separate wavelengths across nine separate sensors), therefore simple one-way ANOVAs were not sufficient. A 7 (emotion category) x 3 (anterior/posterior sensor position) x 3 (lateral sensor position) x 3 (wavelength) repeated measures ANOVA was conducted. Anterior/posterior sensor position indicates the location of the sensor on the head with three sensors each covering frontal, central (prefrontal), and parietal areas. Lateral sensor position further defines the sensor's location, with three sensors each covering left, central (mid-sagital), and right areas. This analysis matrix enabled testing complex interactions, which could potentially determine how different wavelengths within separate brain areas were affected by various emotions. Unfortunately increasing the complexity of the analysis, while improving its potential insight, also increases the difficulty of interpretation.

The results of the EEG ANOVA showed a significant main effect for emotion [$F(6, 150) = 4.399$, $p < .001$], such that Fear resulted in less EEG power than Amusement, Awe, Contentment, Excitement, and Sadness, while Disgust was less than Amusement, Contentment, and Sadness. A significant interaction was shown between emotion, wavelength, and lateral position [$F(24, 600) = 1.792$, $p = .012$]. To further analyze this interaction, an emotion x lateral position ANOVA was conducted for each wavelength. Within the alpha band, there was a significant main effect for emotion [$F(6, 192) = 4.116$, $p = .001$], such that Fear had significantly less power than all other emotions and there was no significant emotion by lateral position interaction. Significant emotion by lateral position interactions were found for both beta [$F(12, 360) = 2.709$, $p = .002$] and theta activity [$F(12, 408) = 2.143$, $p = .014$]. These significant interactions required the completion of one-way ANOVAs for emotion within each category of lateral position for both beta and theta activity. No significant effect was found for emotion on either left or central sensors for either beta or theta activity. However, as shown in *Figure* 3, there was a significant main effect for emotion on the right sensors within both beta [$F(6, 216) = 9.018$, $p < .001$], and theta [$F(6, 252) = 4.307$, $p < .001$].

- Beta
- Fear and Disgust < Amusement, Awe, Contentment, Excitement, and Sadness
- Theta
- Fear < Amusement, Contentment, Excitement, and Sadness
- Disgust and Awe < Amusement, Contentment, and Sadness

Fig. 3. EEG activity recorded from the right hemisphere. All listed comparisons are significant at $\alpha = .05$.

To summarize the complexity of these results, EEG power does reflect changes in emotion. Specifically, Fear, Disgust, and Awe tend to result in less EEG power, but this effect only occurs in the right hemisphere and only within the beta and theta wavelengths with the effect varying slightly between the two wavelengths.

3.3 Multivariate Analysis of Variance (MANOVA)

At a glance, MANOVA appears to be the best option for analyzing multiple physiological measures. However, the use of MANOVA does not yield a great amount of additional information over the use of individual ANOVAs. MANOVA creates a linear combination of all of the dependent variables and then a traditional ANOVA is conducted on this newly calculated DV. Therefore, the analysis only reveals whether the IV (emotion) has an effect on the combination of all of the DVs (physiological response from each sensor). Assuming this analysis is significant, the only real knowledge gained is that emotion has some effect on some aspect of the physiological measures, but individual ANOVAs (as discussed above) must still be conducted to determine the specifics of this effect.

3.4 Regression

Regression could be capable of providing a great deal of information about the data, but was unable to be used to analyze the current dataset because emotion is a nominal variable. Regression could still be used with emotion as a variable by dummy coding it into a series of dichotomous variables acting as predictors and the physiological measures as dependent variables. However, analysis would be limited to one DV at a time, which does not provide an understanding of how the physiological responses interact or vary together for a given state. The true potential for regression would only be possible if emotion were comprised of interval or ratio data. This would allow emotion to be entered into the model as if it were the dependent variable (to be predicted) with the many physiological variables included simultaneously as predictors. In addition to showing the relationships between each physiological measure and emotion, this analysis would enable the evaluation of complex interactions between physiological measures (though these terms are not computed automatically and so would need to be generated manually by the researcher) and would test for mediation between variables. Given the extent of the additional

analyses possible when employing regression, it is unfortunate that it is not an option for the current dataset. Nevertheless, any study in which the independent variable can be quantified on some continuous scale, for example if the same study were conducted by varying image size, regression should be considered as one of the primary analysis techniques.

3.5 Discriminant Function Analysis

Thus far, regression affords the greatest potential for understanding the true underlying physiological processes of emotion. However, DFA is likely to provide the most benefit toward more practical applications because the result of DFA is a simple classification algorithm, which weights each of the predictor variables according to their relationship with the predicted value (emotion in this case). This algorithm could then be used to predict the emotion a person is experiencing based solely on their physiological responses. As with regression, DFA is again incapable of being used to analyze the current dataset, this time being restricted by the repeated-measures design of the study. Traditional DFA assumes that each observation is independent of all others, an assumption that is clearly violated by the current study. Fortunately, Roy and Khattree (2005) have begun adapting traditional DFA methods for analyzing repeated-measures data. No single set of standards has yet been agreed upon and as such, no statistical analysis software has adopted this new implementation of DFA. As a result, it would be both difficult and inadvisable for researchers to attempt to utilize these new methods at this time, but their continued development should be monitored. DFA for a repeated-measures design will truly be a powerful tool.

4 Conclusion

The endeavor to effectively capitalize on the potential offered using physiological measures for assessing human state and performance is complicated. Correlation and ANOVA provide direct methods for analyzing physiological data. However, the limitations for explaining in greater depth overall physiological response recorded by multiple sensors led to the search for more comprehensive analyses. Regression and DFA appear to provide the greatest utility for analyzing physiological measures, but are also limited by design and scale requirements. Therefore, a carefully designed study should be conducted that addresses evaluating the effects of multiple physiological measures used to classify overall physiological response for phenomena such as emotion, workload, stress, and other states. The rise of physiological assessment implementation into all types of human research demands effort put forth to discovering and using the best analyses. The present paper is the start of that journey and provides one study example that clarifies the types of analyses available for physiological measures and when to use each. The challenge to researchers utilizing physiological measures, whether for Brain-Computer Interface (BCI), Augmented Cognition, or Neuroergonomics, is to stretch the limits and attain deeper insight through the best analyses possible.

Acknowledgments. This work was funded by Alion IDIQ Subcontract Agreement No. 8005.00X.10 in support of Army Research Laboratory Prime Contract DAAD19-01-C-0065, Delivery Order #119 along with Alion P.O. #STM1196782.

References

1. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL (2008)
2. Roy, A., Khattree, R.: On discrimination and classification with multivariate repeated measures data. Journal of Statistical Planning and Inference 134, 462–485 (2005)