The correlation between Heart Rate Variability (HRV) and emotional arousal

André Castro Lundin
s141272
Technical University of Denmark

Patrick Leese
s141275
Technical University of Denmark

1. ABSTRACT (ANDRÉ)
It is widely accepted that LF/HF is correlated to emotional arousal. The specific emotions that affect it and how are still being researched. This paper documents the iterative work that we did in testing the hypothesis of LF/HF being correlated to emotional arousal. This was tested in a qualitative way, experimenting using both Lego and video clips whilst monitoring participants. The stacked resulting data was analysed and we found a correlation between time pressure and LF/HF, as well as a correlation between fear and LF/HF. In order to reliably conclude this, a larger base test would be required but we believe substantial evidence has been shown that could provide the basis for further interesting studies.

Keywords
Heart Rate Variability, gHRV, Emotional Response, Lego, Video, LF/HF

2. MOTIVATION (PATRICK)
Over the last two decades it has become increasingly apparent that heart rate variability (HRV) has the potential to represent an indicator of overall health through the eyes of the autonomic nervous system[2]. It is proven to provide valuable insights on a vast range of areas from patient reaction to medication to training optimisation for elite athletes and sports-persons[3]. There are a number of comprehensive studies which detail emotional responses to HRV such as listening to an emotionally intense part of a story [4] which shows HRV can be measured against corresponding physiological signs of arousal. Further studies research into the influence of mental stress on HRV [1] and show a tendency for increased sympathetic activity with a mental task. These findings clearly show a strong correlation between HRV and emotion suggesting that there is scope to pinpoint emotional triggers in our daily activities.

Heart rate (HR) can be defined as a number of heart beats per unit of time - at rest a normal HR ranges from 60 - 100bpm and it is widely considered that a more efficient heart adopts a lower heart rate. However, HR is not constant and presents variations due to both internal and external stress factors. It is these variations which are seen to give further insight, with HRV referring directly to the beat-to-beat alterations in HR [0].

There are two systems associated with HR: sympathetic and parasympathetic. The sympathetic system (often referred to as the ‘fight or flight system) can be seen as the accelerator option, which has the ability to provide and pump more blood in required situations. The parasympathetic system (often referred to as the ‘rest and digest’ system) offers the counter brake system which is needed to slow ourselves down.

Therefore increases in sympathetic activity should produce an increase in heart rate and an increase in parasympathetic activity will produce a decrease in heart rate. The measurement of HRV can work as an indicator towards working functionality of the parasympathetic nervous system. Therefore a high HRV indicates that the parasympathetic response has a dominance over the sympathetic response and possibly provide an indication that the individual is less likely to be affected by stressors.

Establishing the effect of these stressors on an individual daily basis was an initial assumption with regards to this project. Looking at stress indicators would allow us to examine whether there was anything objective we could gain from the results. We believed there was potential to see larger amounts of stress in particular situations and we were fascinated to look at a longer term study as to what trends we could derive from this. The scope for experimental design was endless and after much discussion it was decided to perform pre-testing experiments with Lego and further detailed experiments with the production of a short video.

We knew measurements were inexpensive, accurate and relatively easy to obtain and could be directly analysed through various user friendly software. We specifically focused on the LF / HF ratio as a good method for measuring the balance between the sympathetic and parasympathetic systems and explored what could be derived from this. The LF/HF ratio is calculated through the ratio of low frequency content (LF - in the band of 0.04 - 0.15 Hz) against the high frequency
With this report we will go on to detail how we explored the interaction between LF/HF and emotional arousal in a number of scenarios through our HRV calculations.

Our initial hypothesis was that, “we believe there is a correlation between a fluctuating LF/HF and certain situations in life” with a focus on establishing which situations are most likely to give an enhanced LF/HF signal. We did however accept early on that we would work on a very iterative basis due to the explorative nature of the study, aiming to gain immediate feedback and data that we could analyse through an app we would create and the gHRV software tool.

4. BACKGROUND (ANDRE)

Our first step was to create an app that we could use throughout the project to collect the data we would need for analysis. Considering the relatively short time-span we had for the project compared to it’s large scope it was imperative we worked from a very much lean perspective.

In order to help us understand the requirements and challenges we would face during the real experiment, we performed a number of pre-tests. The main aim of this was to perform the test in it’s entirety which would give us the knowledge of how to collect data, how to analyse it and any potential changes we should make. From a high level perspective we were looking to see whether we can, in controlled circumstances, see an expected pattern.

Our first data collection involved monitoring André, on his commute home from DTU. This was used as a test to see whether we could gain a signal, capture data and import it, so findings were insignificant but essential in order to lay the foundation for further experimental work.

After we had established a signal, naturally the next step was to conduct a low-key experiment that would give us more data to work with. This experiment involved the use of Lego. Research involving professors at our university had already established a project monitoring different aspects of children playing with Lego using eye tracking devices and it was felt that by monitoring the HRV this would also give further valuable insights. Tests in Germany and Denmark had already shown a difference in building technique, with German children much more likely to read the instructions before using external participants. Therefore the setup included a Lego set, step-by-step instructions, timer and video recording equipment and the experiment was performed in non-controlled surroundings. The timer and video recording were used in order to establish a synchronisation.

5. SOFTWARE (ANDRE)

5.1 Software and Device

To aid our understanding and data analysis we used the gHRV software which is a simple and easy to use graphical tool for HRV analysis. It is developed in python and includes a graphical user interface, the ability to import files in multiple formats, signal analysis of time intervals and ultimately export results [7].

At the start of the project we also received two Zephyr HxM BT wireless Bluetooth heart rate monitors. The monitors are not capable of storing the gathered data, but instead send a package via bluetooth after a one second interval, as seen in appendix 12.1.

The data package it sends is made up by 60 messages ranging from heart beat timestamps to instantaneous speed [8]. The relevant fields from the data package (each data package has 40 variables) for our experiment were:

- Heart Rate: Unsigned byte in the range of 30 to 240bmp. Zero means it has not detected a heart beat. This was used to detect any possible abnormalities in the heart beat timestamps when testing and developing the python script.
- Heart Beat number: Unsigned byte in the range 0-255. We used it to detect skipped beats in the received packages as well as duplicate packages.
- Heart beat timestamps 1 to 15: The timestamp is a 16 bit unsigned integer in the range 0 to 65535. It gives us a valuable timestamp of a heartbeats R wave. This was the most vital information which is used to calculate the HRV.
- Packaged timestamp: Timestamp of the package being received by the phone, so each package has a unique timestamp, which we used to help us manually locate errors in files that have thousands of entries - represented in current millis.

The functionalities we used in gHRV throughout the project can be seen in Appendix 12.2.

6. METHODS (PATRICK)

6.1 Data Acquisition

During each experiment all raw data was saved within an individual folder. Each file included hundreds of 40 variable
packages that would then be subsequently run through our python code in order to export the heart beat time stamps as well as detect any problems or anomalies in the data.

The output files were imported to gHRV, where it was filtered. The data was interpolated and a plot of the frame-based HRV was created. For our project, the window size was set to 40 seconds and the window shift to 20 which means every 20 seconds we take a new 40 second block of HRV measurements. It was advantageous for us to keep this window size and window shift as small as possible in order to pinpoint precisely where specific actions occurred. Everything else was set to default. From gHRV, we export the results which provided us with high understandable HRV data that was used for further analysis.

6.2 Pre-Testing

6.2.1 Individual

The graph gives an overview of Andrée’s trip home (see background), showing the various heart rate levels at different times through the coloured episodes. Here the focus as mentioned is just to determine whether we get a signal which we see through the Heart Rate (bpm). The graph confirms we have a strong signal which is proven through the increase in HR, when Andrée is active seen in episodes 1, 3 and 6.

Figure 1: Heart rate (bpm) during one and a half hour commute

6.2.2 Challenges

Over the course of setup and pre-testing, a number of challenges presented themselves. At an early stage, we received some code for a sample app from our lecturer Jakob which could allow us to receive and display data from the HR monitor. We kept partial elements of the code and expanded on it to establish a working app that could receive the data we were sending. After making this app our main problems came from the fact that we were not quite sure on the right way to process the raw data that we were getting. We had seen that data packages were sent every second and with HR’s beating faster than 1b/s this proposed an initial problem of where we would get this perceived missing data from. The opposite effect was also seen in slow heart beats which gave repeated data due to data being registered twice. This was solved as each package not only contained the information about the last heart beat but the timestamps of the last 15 beats which we took into account when updating our script.

Another issue that continuously occurred was the battery charge indicator which we found to be unreliable. The indicator was very bad, increasing and decreasing unexpectedly,

Figure 2: Method of extracting timestamps from data packages

which pointed to a calibration issue with the device. We fixed this adequately to what we needed to do by simply re-connecting the device until it was stable.

We also had to ensure that timestamps were synchronised and this was done through the use of a video recording as previously mentioned. A further challenge we were faced with was the limitation of phones and HR monitors we had at our disposal meaning only one participant could be used in an experiment at any one time.

Lastly, we encountered a further problem which made our code quite extensive as we had to write a script that could figure out in which of the skipped beats the buffer reset. We had to cover for the scenario of this happening at the same time as having skipped beats, because otherwise our output data would be flawed.

6.2.3 Lego

The following observations were made from performing a pre-test with Lego where participants were given a Lego set alongside some instructions and asked to assemble the set in the quickest possible way.

Participant 1: His LF/HF peaks towards the end and we can see from the video that this is when he looks at the timer for the first time. Direct feedback showed that he wanted to finish under 10 minutes, giving a clear connection to time pressure. Interestingly, by just looking at his HR it was relatively steady - however looking at the LF/HF which can be seen below, there is clear variance towards the end which may not have been interpreted from the initial HR findings. Therefore it would have not been clear that he struggled much more with the later stages on the instruction manual.

Figure 3: Participant 1 LF/HF ratio during a 11 minute Lego building session

Participant 2: We do not gain much knowledge in terms of results from the test with participant 2 but we learnt more from the process. It was clear that we needed measurements
at least one minute before the start and after of the test.

![Figure 4: Participant 2 LF/HF ratio during a 4 minute Lego building session](image)

Figure 4: Participant 2 LF/HF ratio during a 4 minute Lego building session

Participant 3: As can be seen from the graph below, the highest LF/HF is seen during the first half of the building process which he later confirmed was the most difficult. We also see a peak towards the end which is where, similar to participant 1, he looked at the timer and noticed he was over the current record for fastest completion time.

![Figure 5: Participant 3 LF/HF ratio during a 7 minute Lego building session](image)

Figure 5: Participant 3 LF/HF ratio during a 7 minute Lego building session

Although small, these quick experiments gave a substantial amount of knowledge and insight into how we would proceed further with this project.

So our hypothesis from these tests were:

- Pressure (time pressure) increases LF/HF
- Difficulty also increases LF/HF

### 6.2.4 What we learnt

The process of conducting these pre-tests was vital and we learnt a number of valuable lessons and considerations that we took forward into the main experiment.

- Repeatability: it could be advantageous to conduct a test that is repeatable for each participant which was not the case for the Lego experiment.
- Size of test: We need to test at least 10 people to gain a significant data set
- Variability in HRV between people: Average or stack multiple recordings from different people in similar conditions in order to remove random noise and fluctuations and also amplify the common parts for the signal that is related to or caused by the stimuli.
- Control group: We should consider having a separate control group in order to give us comparison opportunities.
- We should consider testing alternative scenarios that match our hypothesis to solidify findings
- We should be aware of personal bias to the experimental situation
- We should capture data for much longer before and after each experiment (at least 2 minutes) to give us a baseline.

- We should consider how much effect does the timer and video have and whether we should run separate experiments with and without it
- We should ensure we get as much data as possible - the more the better allowing us to establish mean values
- Alter window size, frequency bands and window shift when processing data - shorter window gives us more accuracy to establish precise points when something happens but does in turn produce more noise.
- Aim to stack the data to decrease noise and look for the common signal which will be amplified by adding them together
- Include self reporting
- Need a consistent way for capturing time and documenting it

Therefore these first tests were extremely valuable, they gave us a baseline to work from and circumstantial evidence than we can build on top of. We proceeded to perform a more robust analysis with a minimum of 10 people.

**Our hypothesis at this stage was that HRV was directly effected by how intense you perceive a situation to be.**

### 6.3 Final Experimental Design (André)

The experiment was split into three stages. First the user puts on the heart monitor. Then he/she sits down and watches a video for 12 minutes. Finally the user is asked a few questions.

The video consisted of a 3 minutes of a dark screen in order to obtain a baseline of the users LF/HF, a 2:20 minutes video of people doing acrobatics from heights without safety, a 30 seconds of dark screen in order to allow for the LF/HF to normalise, a 1:30 minute video of boring facts, followed by another 30 seconds of dark screen in order to allow for the LF/HF to normalise, a 1:50 minute horror movie trailer and a 3 minutes of a dark screen in order to have some extra data at the end. After watching the video, the test person was asked three questions.

1. Are you afraid of heights?
2. Do you often watch horror movies?
3. Anything in particular from the video that stood out to you?

This test was conducted on 4 individuals and the data analysed. The results demonstrated a few issues with our methodology. First of all, individuals would easily get distracted by their surroundings. It was also important to consistently ask the questions after the video, as increased expectations as well as conversation seemingly increased LF/HF. The tests had to be conducted in a similar environment such as an empty room which ensured the same noise level (See Results: Controlled vs. uncontrolled).
From the first videos it was noticeable that there was a high variance in-between videos as well as some users having some abnormal peaks. There also seemed to be a difference when the test individual was talking or silent.

This first part also brought a number of new hypotheses. Does talking increase LF/HF? Is the significant rise in LF/HF between videos due to expectations?

So for the second half, we carried on with the video experiment, but this time trying to remove all potential external influences as possible. The test person would be in a room by themselves. At this stage, the test was conducted on 9 individuals over the course of 2 days. After this stage we wanted to observe the effects of the same video on a person watching it repeatedly, so one of our test subjects watched the video 5 times in a row.

7. RESULTS (PATRICK AND ANDRE)

The batch analysis of all the tests was conducted using iPython Notebook where the following results and graphs were generated.

The first noticeable thing is that even though there are some outliers, the LF/HF tends to remain between 0-15 (Figure 6). We have not seen reference to a specific range like this in any of our readings. From the overall results there seems to be a huge peak while waiting for the first video to begin as well as the very end which is specifically when questions are asked.

Figure 6: Average LF/HF ratio over 12 minutes for 15 participants

As mentioned earlier, some later tests were conducted on a more controlled environment. The controlled tests tend to have smoother peaks which validates the need for a controlled environment, although a lot of the same peaks appear in both (Figure 7).

There seems to be a slight increase in LF/HF with individuals that are scared of heights at the very end of the first video which is where the men are doing unsafe gymnastics on top of the crane (figure 8).

Surprisingly, it also seems like the LF/HF of people who often watch horror movies was the most affected by the horror video (figure 9).

We also tested repeatability, having the same test person watch the video a number of times. The individual was afraid of heights, but it seemed like his LF/HF was less affected by the heights clip after watching it the first time (Figure 10).

Figure 7: Controlled vs uncontrolled surroundings

Figure 8: Participants who admitted to being scared of heights against those who were not

Figure 10: Repeatability test measuring the reaction to watching the same video 5 times in a row
Figure 9: Participants who admitted to not being used to horror movies against those who were

Although we did make a comparison between the results from male and female tested individuals, the small amount of female testers does not allow statistically significant conclusions. Our findings can however be viewed in Appendix 12.3.

Further results, code and all work related to this project can be found on: http://hrv.luanca.eu

8. DISCUSSION (PATRICK)

The main aim of this study was to explore the correlation between HRV and perceived emotional arousal. As an explorative study we consider the results given above to be of a very primitive nature but nonetheless provide some very interesting and intriguing results that should be further explored in future studies.

A key point for discussion is the expectation theory that has presented itself in our results. It is clear that participants commonly show a peak in the waiting period before the video starts. Our theory here lies in the fact that in some instances expectation of upcoming events can be more stressful than the event itself. By allowing the participant to sit and wait for 3 minutes before the video starts it gives adequate time for the participant to wonder what will they watch and this can directly impact the HRV. It would be fascinating to explore this notion further - do we tighten up when information of an upcoming event is unknown, only to relax when you realise it is not so bad after all? Is it an inherent human nature to expect the worst? Is it a defence mechanism? These are questions that could provide the base for further exploration.

In our results we see consistent clear indication of stronger HRV in time-frames where we speak directly with the participant. This has led us to the believe that we provide circumstantial evidence that interaction increases HRV and human interaction is perceived more intensely than that of the video.

Another interesting area for exploration that we touched upon was the learning effect. Could an individual or a group directly effect the HRV results from learning what would be seen on the screen. This could work in parallel with the exploration of expectation, supporting it if participants ratio levels decreased with each viewing.

Further exploration we discussed could also be to look into gender differences, the use of sound and even explore the correlation with skin conductance. We would also consider using a very intense video that would be sure to effect each participant. In this instance some participants were afraid of heights or horror movies and some were not so it was very divided and by providing a very strong emotional video this could give a much better baseline. Alternatively, we could permeate the order of trailers to see if the order has anything to do with the reaction.

It would have been beneficial to arrange the participants in a much more structured fashion instead of the random nature we did so on this occasion. We also feel it would be highly beneficial to give each and every participant the same instruction card before the first video would start to ensure continuity between participant information.

9. CONCLUSION (ANDRÉ)

This project was conducted in a field where research is still being done. Through every iteration, the results became more reliable and we were more clear on the hypothesis we were testing. Unfortunately there are some limitations to our work. Due to the iterative way of working, the user sample is small and more data would also be highly beneficial. The number of participants used in our experiment sufficed for what we needed to do but looking at the bigger picture we would require a much more substantial data set to work with.

This project has been a steep learning curve in a number of areas and we were both grateful and excited by that. We have learnt continuously from the beginning starting with developing the app and gaining a signal to implementing the final experiment. We could however have adapted a much leaner approach initially as the setup of project seemed relatively slow in comparison to the short amount of time we had. However it was also necessary to take the time to and do the research to gain an understanding of what HRV is and the scope of this project.

Due to our lack of knowledge in some areas we felt the pre-testing was completely invaluable to this project. That being said we still made a number of mistakes when approaching the later experiments and this would of course need to be rectified if we were to continue.

Overall it was a fascinating project to work with. It was a thoroughly enjoyable process to challenge ourselves and explore HRV in this manner and we hope to continue research in this area in future work.

10. ACKNOWLEDGEMENT

This work was conducted as part of the Advanced Project in Digital Media course at the Technical University of Denmark. The authors pay special thanks to Per Baekgaard for his support, advice and weekly meetings throughout the project. We also extend thanks to Jakob Eg Larsen and
Jakob Andreas Baerentzen the course responsible who’s weekly inputs have been vital in steering the direction of the project.

11. REFERENCES


12. APPENDIX

12.1 Communication device

Figure 11: HRM Bluetooth communication

12.2 Used from gHRV
Throughout the project we have utilised the following functionalities of gHRV:

- Statistical analysis
- Graphical Editor
- Time frequency and non-linear analysis

- Frame-shift and frame-length alterations
- Poincare Plots
- Text Editor
- Automatic and manual filtering
- Outlier filter
- Data export - results, data, plots

12.3 Men vs Woman Results

Figure 12: Men vs. women