Open access  •  Book Chapter  •  DOI:10.1007/978-3-319-66379-1_13

# EUStress: A Human Behaviour Analysis System for Monitoring and Assessing Stress During Exams — Source link 🔗

Filipe Gonçalves, Davide Carneiro, Davide Carneiro, Paulo Novais ...+1 more authors

**Institutions:** University of Minho, Polytechnic Institute of Porto

Related papers:

- Machine Learning Based Solutions for Real-Time Stress Monitoring

- Blueprint to Workplace Stress Detection Approaches

- IT-supported Stress Management: Mediating between Sensor Data and Individual Appraisal

- Characterization and Modeling of Acute Stress Under Dynamic Task Conditions

- A Systematic Review of Thermal and Cognitive Stress Indicators: Implications for Use Scenarios on Sensor-Based Stress Detection

Share this paper:  f  𝕏  in  ✉

View more about this paper here: https://typeset.io/papers/eustress-a-human-behaviour-analysis-system-for-monitoring-szx0ts0fov

# EUStress: a Human Behaviour Analysis System for Monitoring and Assessing Stress during Exams

Filipe Gonçalves[1], Davide Carneiro[1,2], Paulo Novais[1], and José Pêgo[3,4]

[1] Algoritmi Research Centre/Department of Informatics
University of Minho, Braga, Portugal
[1]`fgoncalves@algoritmi.uminho.pt`, {`dcarneiro,pjon`}`@di.uminho.pt`
[2] CIICESI, ESTG, Polytechnic Institute of Porto
Felgueiras, Portugal
[3] School of Health Sciences, Life and Health Sciences Research Institute (ICVS)
University of Minho, Braga, Portugal
[3,4]`jmpego@ecsaude.uminho.pt`
[4] ICVS/3B's - PT Government Associate Laboratory, Braga / Guimarães, Portugal

**Abstract.** In today's society, there is a compelling need for innovative approaches for the solution of many pressing problems, such as understanding the fluctuations in the performance of an individual when involved in complex and high-stake tasks. In these cases, individuals are under an increasing demand for performance, driving them to be under constant pressure, and consequently to present variations in their levels of stress. Human stress can be viewed as an agent, circumstance, situation, or variable that disturbs the normal functioning of an individual, that when not managed can bring mental problems, such as chronic stress or depression. In this paper, we propose a different approach for this problem. The EUStress application is a non-intrusive and non-invasive performance monitoring environment based on behavioural biometrics and real time analysis, used to quantify the level of stress of individuals during online exams.

**Keywords:** Stress, Human-computer interaction, Mouse dynamics, Decision making, Big Data Mining

## 1 Introduction

The term *stress* is commonly used to describe a set of physical and physiological responses that emerge as a reaction to a challenging stimulus that alter an organism's environment [18]. Perceiving an individual's physiological stress level is nowadays viewed as an important factor to manage individual performance, in a time when individual and team limits are pushed further.

Indeed, prolonged exposure to stress-inducing factors is a growing concern, specially in complex activities, which demands great responsibility and reliability. This can lead to states of emotional exhaustion (burnout) and other mental

disorders [13](e.g. depression, chronic stress, chronic diseases), with potential consequences at the personal, professional, family, social and economic levels [11].

Some good examples of stressful environments can be seen everywhere in our life, from workplaces [6] to academia [20]. In this paper we focus on academic stressors, which include the student's perception of the extensive knowledge base required and the perception of an inadequate time to develop it [7]. Students report experiencing academic stress at predictable times each semester with the greatest sources of academic stress resulting from taking and studying for exams, grade competition, and the large amount of content to master in a small amount of time [1]. This results in a high prevalence of anxiety disorders among higher education students.

Existing stress monitoring approaches rely on the use of complicated or expensive hardware, or in the collection of biological variables, all of which alter the routine of the student in some way, which is not desirable [10]. In this paper we propose a distributed system for monitoring human behaviour with the aim of measuring stress. Specifically, the paper discusses how groups of medical students can be monitored through their interaction with the computer [17], in order to study the effect of stress on performance during high-demand tasks. Students are monitored in terms of the efficiency of their interaction patterns with the computer and their decision making during the execution of an online exam. The aim is to provide a stress marker that can be effectively used in large numbers of students, without inconvenience. Information on how stress affects each student will make it possible to improve individualized teaching strategies as well as to empower these students with better coping strategies. All this will result in the development of better future professionals.

## 2   Background

Behavioural biometrics is a relatively new form of analysis, which defines a field that extracts users behavioral features from the use of the mouse and the keyboard [22]. These methods are mostly used for user identification and authentication (intelligent security systems) [2], which use multiple techniques for automatic recognition of individuals based on their physiological and/or behavioural characteristics. By using biometrics, it is possible to confirm or establish an individual's identity based on who the individual is, rather than by what the individual possesses (an ID card) or what the individual remembers (a password) [16].

Yet, behavioural biometrics is also applied in the monitoring of human performance analysis or human healthcare. For example, such techniques are used to assess distraction and fatigue in scenarios of classrooms or workplaces [19], quantify the stress level of an individual [5] or targeted as an assisted living applications, regarding the outbreak of the disease for adverse event prediction, and patient profiling in order to improve a patients everyday life [23].

Our work focuses in creating a monitoring distributed system of stress assessment, in which the software correlates the data and conclusions from different behavioural analysis softwares to assess the stress of an individual, through the use of mouse performance behaviour analysis (developed by our own research team in the past [5]) and the decision making performance analysis of individuals under pressure in workplace/study environments.

## 3   Architecture

As explained before, in order for the EUStress application to assess the stress level of an individual, it requires the analysis of two types of information: mouse behaviour and decision-making behaviour. As such, the EUStress application can be decomposed into the following components:

– **MedQuizz:** An e-assessment management system that enables trainers, educators and testing professionals to author, schedule, deliver, and report on surveys, quizzes, tests and exams, and an useful tool to create item banks. It allows the management of information about the quality of the items supporting the individual in the decision to design his/her assessments. It also has fail-safe features in the case of network failure and has functions capable of creating a log of the individuals actions[5]. This component is used to study the cognitive performance and the behaviour decision making patterns of the individual.
– **MouseDynamics:** A framework that includes not only the sheer acquisition and classification of the mouse input data, based on the biometric behaviour, but also a presentation tier that supports the human-based or autonomous decision-making mechanisms [5,4].

Another important factor is the use of saliva as a biomarker of exam stress and a predictor of exam performance. In the study published in the "journal of psychosomatic research", done by Miri Cohen and Rabia Khalaila, it is shown that pH levels of saliva may serve as a reliable, assessible and inexpensive means by which to assess the degree of physiological reactions to exams and other naturalistic stressors [9]. Also, salivary cortisol is routinely used as a biomarker of psychological stress and related mental or physical diseases [15].

In our case study, a sample of the individual's saliva is taken before and after each exam, as a mean to analyse his/her levels of cortisol in their biological system, and to predict the levels of stress of the individual. The variance of the levels of cortisol is calculated, using the following formula:

$$DeltaCortisol_{\alpha,\beta} = PreCortisol_{\alpha,\beta} - PosCortisol_{\alpha,\beta},$$

where $\alpha$ represents the identification of the individual, and $\beta$ represents the identification of the exam.

---

[5] The website of MedQuizz can be assessed at `http://www.medquizz.com/`

Since the study of behavioural features are mostly related to the individual's behavioural habits, the calculation of stress level considers the individual's ID, actions of mouse and decisions made during the exams (CyberPsychological computation methods), which are monitored and acquired by the computation system [24]. After pre-processing, those data are stored in a database for machine learning. With these techniques, it allows us to study the relationships between the ubiquitous individual's psychological reactions and their behavioral patterns in cyber space for the psychological assessment of their situations in the learning process [12]. The explained architecture is shown in figure 1.
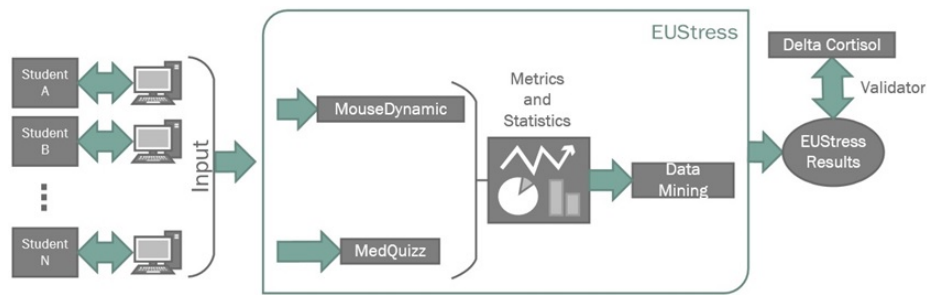


**Fig. 1.** EUStress Architecture

### 3.1 MouseDynamics Performance Analysis

MouseDynamics data collection output analyses the individual's mouse behaviour, and calculates his/her behavioural biometrics. These features aim at quantifying the individual mouse performance. Taking as example the movement of the mouse, one never moves it in a straight line between two points, there is always some degree of curve. The larger the curve, the less efficient the movement is [5,4]. Some of the most important calculated metrics are:

- **Absolute Sum of Degrees (ASD):** This feature seeks to find how much the mouse turned, independently of the direction to which it turned (in degree units).
- **Average Distance of the Mouse to the Straight Line (ADMSL):** This feature quantifies the average sum of the successive distances of the mouse to the straight line defined by two consecutive clicks (in pixels).
- **Average Excess of Distance Between Clicks (AED):** This feature measures the average excess of distance that the mouse travelled between each two consecutive mouse clicks (in pixels).

- **Click Duration (CD):** Measures the time span between two consecutive mouse clicks. The longer the clicks, the less efficient the interaction is (in milliseconds).
- **Distance Between Clicks (DBC):** Represents the total distance travelled by the mouse between two consecutive mouse clicks (in pixels).
- **Mouse Velocity:** The distance travelled by the mouse (in pixels) over the time (in milliseconds).
- **Mouse Acceleration:** The velocity of the mouse (in pixels/milliseconds) over the time (in milliseconds).
- **Time Between Clicks:**. The time span between two consecutive mouse clicks (in milliseconds).

### 3.2 MedQuizz Decision Making Analysis

**MedQuizz Data Output:** MedQuizz data collection output shows the individuals actions recorded during the exam into a log file. This file shows a list of actions, sorted by student and action time, making it possible to study and analyse the decision making behaviour of an individual. Some of the most important variables of this log are:

- **Exam Code:** Unique identification of the exam.
- **Student Code:** Unique identification of the individual/student.
- **Exam Question No:** Identifies the number of the question where the decision was made.
- **Action Time:** Defines the action time of the decision making, written in date and hour format;
- **Action Description:** Defines the type of the decision made by the individual. For example, action description can show when the individual entered or left a question, an answer was inserted or removed, among other actions.
- **Action Result:** When an answer is inserted or changed, shows if the decision made was right or wrong, based on the correct answer for the question.

**MedQuizz Analysed Features:** By monitoring the actions log file of each individual during the execution of an exam, it is possible to analyse his/her decision making behaviour. Everything an individual does, consciously or unconsciously, is the result of some decision. The information we gather is to help us understand occurrences, in order to develop good judgements to make decisions about these occurrences. To make a decision, an individual needs to know the problem, the need and purpose of the decision, the criteria of the decision, and the alternative actions to take. Then there's the need to determine the best alternative [21]. Decision making, for which we gather most of our information, has become a mathematical science today. It formalises the thinking we use so that, what we have to do to make better decisions is transparent in all its aspects.

With that in mind, decision making analysis aims to evaluate the performance of the individual, based on the time between decisions, the correctiveness of the

selected decisions, if those decisions serve the objectives of the decision maker, number of times a question was visualized, among other features. Some of the most important calculated behavioural features are:

- **Average Time Between Decision (ATBD):** This feature seeks to find the average time it takes each individual to take a decision. All decisions are taken into account. The decisions analysed vary from entering or leaving a question, inserting, changing or removing answers, marking or unmarking questions for review, among others (in milliseconds).
- **Median Time Between Decision (MTBD):** This feature quantifies the median of the time it takes each individual to take a decision (in milliseconds). The average is a measure greatly influenced by large or small number of values, even if these values appear in small numbers in the sample. These values are responsible for the misuse of the average in many situations where it would be more meaningful to use the median.
- **Standard Deviation / Variance Time Between Decision:** This feature measures the standard deviation/variance of the time between decisions for each individual (in milliseconds).
- **Average Time Between Questions (ATBQ):** This feature measures the average time the individual spent between all visualized question (in milliseconds).
- **Decision Making Ratio (DMR):** Verifies the ratio between the number of answers inserted, changed and removed and the total number of actions recorded (in percentage).
- **Correct Decision Making Ratio (CDMR):** Verifies the ratio between the number of decisions considered correct and the number of answers inserted, changed and removed (in percentage). A decision is considered correct when an individual inserts or changes an answer into a correct option or when he/she removes an incorrect answer from the question.

## 4 Study Design

In order to determine the levels of stress of a group of individuals, based on their mouse behaviour performance and decision making behaviour performance, data was collected from the participation of a group of medical students in computer-based high stake exams. Through the evaluation of exams, these students test their academic knowledge in a monthly period. In these exams, students are indicated their seats, and at the designated time they log in the exam platform using their personal credentials and the exam begins. The participation in the data-collection process does not imply any change in the student's routine, and all monitored metrics are calculated through background processes (using MedQuizz and MouseDynamics), making the collection data process completely transparent from the point of view of the students, just like a normal routine exam. These exams consists mostly of single-best-answer multiple choice questions, where the students only use the mouse as an interaction means. When

the exams end, students are allowed to leave the room. As explained before, a sample of the individual's saliva is taken before and after the execution of each exam. These cortisol values will be later used to compare with the predicted stress levels computed by the EUStress application, as a mean to validate its results and conclusions.

Specifically, the case study considers a group of 105 medical students, which are monitored to study the effect of stress/anxiety in the performance of high demand tasks, such as the execution of exams. The methods used for data collection were taken into account, since any external factor can influence variations in the decision making and mouse performance behaviour of the individuals. As such, it was important to include non-intrusive and non-invasive measures as essential requirements during the execution of the exams.

## 5   Data Collection and Preparation

In order to analyse the data received from both applications (MedQuizz and MouseDynamics), some conditions were required in the study.

### 5.1   Temporal Approaches

The first step of the preparation process was the verification of the variations in the decision making of the individual, during an exam. In order to study these variations, the set of events of each individual are ordered by action time, cloned and split into three different datasets:

- **Chronological Time:** The data collected is divided in intervals of five minutes;
- **Percentage Time:** Data is divided in intervals, each one comprising 5% of the total duration of the exam of the individual;
- **Quarter Time:** Data collected is divided in four intervals, each containing twenty five percentage of the total duration of the exam taken by the individual;

This preparation process is used as a way of presenting different approaches for our case study, and consequently to take advantage of the different conclusions for future data mining.

### 5.2   Data Dimension Reduction

The second step in the preparation of the data is the implementation of data transformation processes that can provide additional insights. Moreover, when dealing with real-world data, it is often necessary to reduce dimensionality to improve big data management. Dimensionality reduction is the transformation of high-dimensional data into a meaningful representation of reduced dimensionality [8]. Ideally, the reduced representation should have a dimensionality that

corresponds to the intrinsic dimensionality of the data. The intrinsic dimensionality of data is the minimum number of parameters needed to account for the observed properties of the data [14]. As a result, dimensionality reduction facilitates, among others, classification, visualization, prediction, and compression of high-dimensional data.

The data collected from from Mouse Dynamics comprises a significantly large amount. To reduce its dimensionality and given the usual shape of the data series (exemplified in Figure 2), we construct a linear fit and use the resulting quadratic function to represent the raw data. The model build of the linear and the quadratic model are represented as $f(x) = \alpha + \beta x$ for the linear model and as $g(x) = \alpha + \beta x + \gamma x^2$ for the quadratic model.

After calculating the coefficients of both models, a mean squared error (MSE) calculation is performed. This mathematical formula measures the quality of an estimator, that is, the difference between the estimator and what is estimated. In other words, values closer to zero are better. By comparing the MSE of both functions, we can choose which model is more accurate for the set of values. The MSE of the predictor can be estimated by:

$$MSE = \tfrac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2$$

in which $\hat{Y}_i$ is a vector of $n$ predictions, and Y is the vector of observed values corresponding to the inputs to the function which generated the predictions.

Figure 2 partially depicts the outcome of this process. In this figure, the dots represent the aggregated raw data at regular intervals while the lines represent the resulting quadratic model. After this reduction process, it is possible to use the parameters of the quadratic function instead of the raw data, which simplifies the posterior use of machine learning techniques.
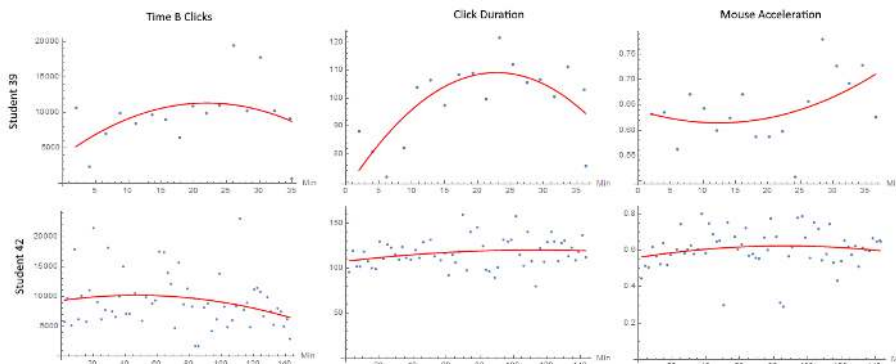


**Fig. 2.** Example of dimensionality reduction for three features in two students.

# 6 Conclusions and Future Work

In this paper we present a technological approach for a non-intrusive analysis of performance in groups of people. This approach is implemented in the form of a distributed system, that constantly collects, processes, stores and monitors data describing the behaviour of multiple individuals simultaneously, during the execution of online high-stakes exams. Through the metrics monitored during the execution of a set of exams, the platform will be able to correlate data between the mouse performance metrics (using MouseDynamics output) and decision making performance metrics (analysing MedQuizz features), in order to quantify the stress levels of an individual.

By monitoring human behaviour, our aim is to study the effect of stress in the performance of high demand tasks and to point out how each individual is affected by stress. This will allow the educational institution to act on each student, through personalized teaching and coping strategies, and thus improve the quality of the future professionals that are being trained.

As future work, we will apply data mining algorithms to analyse data from different perspectives and summarize it into useful information, as a mean to find correlations or patterns among dozens of features in our databases. Yet, it is required to define which of the mining methods and fields can influence positively the results to quantify the stress levels of an individual, during the execution of an exam.

# References

1. Abouserie, R.: Sources and levels of stress in relation to locus of control and self esteem in university students. Educational psychology 14(3), 323–330 (1994)
2. Bailey, K.O., Okolica, J.S., Peterson, G.L.: User identification and authentication using multi-modal behavioral biometrics. Computers & Security 43, 77–89 (2014)
3. Bicho, L.M.D., Pereira, S.R.: Stress ocupacional. Stress Ocupacional. Instituto Politécnico de Coimbra, Departamento de Engenharia Civil, Portugal [periódico na internet] (2007)
4. Carneiro, D., Novais, P.: Quantifying the effects of external factors on individual performance. Future Generation Computer Systems 66, 171–186 (2017)

5. Carneiro, D., Novais, P., Pêgo, J.M., Sousa, N., Neves, J.: Using mouse dynamics to assess stress during online exams. In: International Conference on Hybrid Artificial Intelligence Systems. pp. 345–356. Springer (2015)
6. Cartwright, S., Cooper, C.L.: Managing workplace stress, vol. 1. Sage (1997)
7. Carveth, J.A., Gesse, T., Moss, N.: Survival strategies for nurse-midwifery students. Journal of Nurse-Midwifery 41(1), 50–54 (1996)
8. Chen, H., Chiang, R.H., Storey, V.C.: Business intelligence and analytics: From big data to big impact. MIS quarterly 36(4), 1165–1188 (2012)
9. Cohen, M., Khalaila, R.: Saliva ph as a biomarker of exam stress and a predictor of exam performance. Journal of psychosomatic research 77(5), 420–425 (2014)
10. Cohen, S., Kessler, R.C., Gordon, L.U.: Measuring stress: A guide for health and social scientists. Oxford University Press on Demand (1997)
11. Colligan, T.W., Higgins, E.M.: Workplace stress: Etiology and consequences. Journal of workplace behavioral health 21(2), 89–97 (2006)
12. Dai, W., Duch, W., Abdullah, A.H., Xu, D., Chen, Y.S.: Recent advances in learning theory. Computational intelligence and neuroscience 2015, 14 (2015)
13. Fink, G.: Stress: definition and history. Stress science: neuroendocrinology pp. 3–9 (2010)
14. Fukunaga, K.: Introduction to statistical pattern recognition. Academic press (2013)
15. Hellhammer, D.H., Wüst, S., Kudielka, B.M.: Salivary cortisol as a biomarker in stress research. Psychoneuroendocrinology 34(2), 163–171 (2009)
16. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology 14(1), 4–20 (2004)
17. Novais, P., Carneiro, D.: The role of non-intrusive approaches in the development of people-aware systems. Progress in Artificial Intelligence 5(3), 215–220 (2016), http://dx.doi.org/10.1007/s13748-016-0085-1
18. OSullivan, G.: The relationship between hope, eustress, self-efficacy, and life satisfaction among undergraduates. Social indicators research 101(1), 155–172 (2011)
19. Pimenta, A., Carneiro, D., Novais, P., Neves, J.: Detection of distraction and fatigue in groups through the analysis of interaction patterns with computers. In: Intelligent Distributed Computing VIII, pp. 29–39. Springer (2015)
20. Ross, S.E., Niebling, B.C., Heckert, T.M.: Sources of stress among college students. Social psychology 61(5), 841–846 (1999)
21. Saaty, T.L.: Decision making with the analytic hierarchy process. International journal of services sciences 1(1), 83–98 (2008)
22. Wang, L.: Behavioral Biometrics for Human Identification: Intelligent Applications: Intelligent Applications. IGI Global (2009)
23. Xefteris, S., Andronikou, V., Tserpes, K., Varvarigou, T.: Case-based approach using behavioural biometrics aimed at assisted living. Journal of Ambient Intelligence and Humanized Computing 2(2), 73–80 (2011)
24. Zhou, X., Dai, G., Huang, S., Sun, X., Hu, F., Hu, H., Ivanović, M.: Cyberpsychological computation on social community of ubiquitous learning. Computational intelligence and neuroscience 2015, 12 (2015)