Leadership and Neurosciences: The Analysis of Emotional Arousal During Decision-Making Processes with Decision-Makers Exposed to Acute Stress

João Alexandre Lobo Marques[1], Filipa Magalhães de Moura e Mota Furtado[1], Pedro Crosara Motta[2], Bruno Silva[2] and Paulo Cortez[2]
[1]Laboratory of Applied Neuroscience, University of Saint Joseph, Macao, China
[2]Departament of Teleinformatics Engineering, Universidade Federal do Ceará, Fortaleza, Brazil

alexandre.lobo@usj.edu.mo
filipa.furtado@usj.edu.mo
pedro.motta@lesc.ufc.br
bruno@lesc.ufc.br
cortez@lesc.ufc.br

Abstract: Corporate leaders are constantly dealing with stress in parallel with continuous decision-making processes. The impact of acute stress on decision-making activities is a relevant area of study to evaluate the impact of the decisions made, and create tools and mechanisms to cope with the inevitable exposure to stress and better manage its impact. The intersection of leadership and neurosciences techniques is called Neuroleadership. In this work, an experiment is proposed to detect and measure the emotional arousal of two groups of business professionals, divided into two groups. The first one is the intervention/stress group, n=30, exposed to stressful conditions, and the control group, n=14, not exposed to stress. The participants are submitted to a sequence of computerized stimuli, such as watching videos, answering survey questions, and making decisions in a realistic office environment. The Galvanic Skin Response (GSR) biosensor monitors emotional arousal in real-time. The experiment design implemented stressors such as visual effects, defacement, unfairness, and time-constraint for the intervention group, followed by decision-making tasks. The results indicate that emotional arousal was statistically significantly higher for the intervention/stress group, considering Shapiro and Mann-Whitney tests. The work indicates that GSR is a reliable stress detector and may be useful to predict negative impacts on executive professionals during decision-making activities.

Keywords: Decision making, Stress, Emotional arousal, Galvanic Skin Response, Neurosciences

1. Introduction

Emotional states, such as stress, play a significant part in decision-making processes. Although there is plenty of evidence that feelings influence decision-making, the extent of the impact is unclear. While there is evidence that emotions have a biasing impact, there is also evidence that we use emotional signals in quick, automated decision-making and that they have a tangible benefit in daily decision-making (Bechara et al., 1997). According to a study on financial decision-making under risk, low levels of emotional experience led to higher levels of performance through greater risk neutrality due to a steadier correlation between rational reward and subjective value (Schunk and Betsch, 2006).

There is a wealth of evidence that emotions can lead to decision-making biases. Information processing may be skewed by emotions. There is evidence, for example, that it is easiest to remember experiences that are close to one's emotional state (Fenton-O'Creevy et al., 2011). Emotions may also influence decision-making directly; for example, fear and anger have opposite (but significant) impacts on risk judgments (Lerner et al., 2004). Emotions also influence the emphasis placed on outcomes. For example, considering the negative long-term effects, strong negative emotions maximize the importance of short-term results (Fenton-O'Creevy et al., 2011). Generally speaking, there's a relationship between positive affect and optimistic decision-making, whereas negative affect is connected with pessimistic decisions (Wright & Bower, 1992).

Emotions play a part in risky decision-making as well. According to a study on financial decision-making under risk, low levels of emotional experience led to higher levels of performance through greater risk neutrality due to a steadier correlation between rational reward and subjective value (Schunk & Betsch, 2006).

Although there is plenty of evidence that feelings influence decision-making, the extent of the impact is unclear. While there is evidence that emotions have a biasing impact, there is also evidence that we use emotional signals in quick, automated decision-making and that they have a tangible benefit in daily decision-making (Bechara et al., 1997).

Decision heuristics research often lacks background and is heavily dependent on experimentation with participants that are unfamiliar with the activities being tested. The literature provides a strong focus on domain-specific knowledge and skills, on the development of complex cognitive schema that allows for quick associative pattern recognition, and the stimulation of large and complex behavioral repertoires (Fenton-O'Creevy et al., 2011).

Decision-making under stress research discloses that a great number of theories were obtained from four different theoretical models. They are known as the decision-making conflict theory model, threat-rigidity model, crisis model, and finally under time pressure decision-making model (Etzion, 1984).

The conflict theory model theorizes that the stress specifically associated with decision-making originates from a level of danger or threat perceived by the decision-maker to its well-being and from decision context. Here, the decision maker's self-reputation is at stake, being defined as a good or bad decision-maker based on the individual's alternate choice (Atsan, 2016).

The "Threat-Rigidity Effect" model proposed by Staw et al. (1981) illustrates that, in threatening situations, organizations, groups, and individuals may generally tend to behave rigidly (Staw et al., 1981). The two sorts of existing effects are information processing restriction and limitation in control.

In 1963, Hermann was one of the first individuals to utilize a theoretical approach for decision-making under a situational crisis (Hermann, 1963). This relevant initial outlook impacts the future developments in decision-making study under crisis (Billings et al., 1980). A crisis is described as a device of change, which may be linked with extreme behavior. It has been separated from other concepts such as tension, stress, anxiety, disaster, and panic, by the concept of stimulus and response. It's understood to be a stimulus that develops certain possible responses, such as panic or anxiety.

Finally, acute stress plays a fast and time-dependent role in decision-making (Pabst et al., 2013). Decisions often must be made by individuals, under strict deadlines. This may be an intimidating task that develops stress and cognitive tension. Bronner (1982) developed the "Theory of Decision-Making Under Time Pressure", becoming the principal theory of decision-making under stress. He defines three variables of time pressure: decision time, sensitivity, and problem intensity. Apart from these three variables, he mentions that time pressure limits the interaction and coordination among the decision-making units.

More recently, the use of biosensors during experiments for stress monitoring during decision-making activities has been adopted. Nevertheless, one key challenge is the physiological signal that should be monitored. The most suitable biosignals selection is essential for the experiment's success because of two main reasons: the degree of voluntary regulation over physiological parameters and the selectivity with the desired arousal is measured. As an example, the Galvanic Skin Response (GSR), also known as Electrodermal Activity (EDA), is suitable for detecting stress since the sweat glands are innervated by the sympathetic branch of the autonomic nervous system, being under low voluntary regulation (Poh et al., 2011). In contrast, heart rate variability (HRV) is also widely considered a valid stress monitoring tool since it is influenced by both sympathetic and parasympathetic autonomic branches, as well as being under partial voluntary regulation (through respiratory sinus arrhythmia) (Raaijmakers et al., 2013). HRV is the end product and can be used as a stress indicator, but it is less selective than GSR since it results from the constant interaction between branches. These two (2) physiological modalities can be evaluated noninvasively using commercial wearable sensors, allowing researchers to investigate the tradeoffs between selectivity and voluntary regulation.

The main objective of this work is to evaluate the results of an experiment monitoring the emotional arousal of executive professionals during decision-making processes to determine if groups submitted or not to stress can be classified based on GSR physiological monitoring.

Finally, to further develop the analysis, the following hypothesis is then stated:

H1: The stress group is indeed more stressed than the control group.

2. Literature Review

There are significant studies indicating that emotions may lead to decision-making biases and information-processing skewness. For example, research shows that it is easier to remember experiences that are close to one's emotional state (Bower, 1981) (Fenton-O'Creevy et al., 2011). Emotions may also influence decision-making directly; for example, fear and anger have opposite (but significant) impacts on risk judgments (Lerner and Keltner, 2001) (Lerner, Small and Loewenstein, 2004). Emotions also influence the emphasis placed on outcomes. For example, considering the negative long-term effects, strong negative emotions maximize the importance of short-term results (Fenton-O'Creevy et al., 2011). Generally speaking, there's a relationship between positive affect and optimistic decision-making, whereas negative affect is connected with pessimistic decisions (Isen et al., 1978) (Kavanagh and Bower, 1985) (Schwarz and Clore, 2003) (Wright and Bower, 1992).

In research on the nature of expertise, intuition (an experientially dependent pattern recognition linked to the affectively cued) has been described as essential to expert performance. Dane and Pratt (2007) distinguish the decision heuristics literature, which largely promotes the opinion that intuitive decision-making is inferior to more rational ones, and expertise literature, which stresses the crucial role of intuition in expert performance. The distinct emphasis placed on the research process and context is a significant difference between these works of literature (Dane and Pratt, 2007). Decision heuristics research often lacks background and is heavily dependent on experimentation with participants that are unfamiliar with the activities being tested. There is also evidence that when decision-making is observed in the field or using methods that mimic real-life conditions, some of the primary cognitive biases known as maladaptive products of heuristic thinking in experimental environments either do not exist or contribute to better outcomes (Todd and Gigerenzer, 2007). In comparison, the expertise literature provides a strong focus on domain-specific knowledge and skills, on the development of complex cognitive schema that allows for quick associative pattern recognition, and the stimulation of large and complex behavioral repertoires (Fenton-O'creevy et al., 2011). As a result, Dane and Pratt (2007) conclude that intuition would be more likely to be an efficient component of decision-making in performance domains requiring extensive expertise and complex domain-relevant schema.

In addition, the research on emotions supports the belief that emotions have many important influences on decision-making. Any of these may be classified as biased, with the ability to impair decision-making performance. Nevertheless, it is clear that there is a counter-argument: rather than feelings themselves being counterproductive to decision-making performance, expertise, and emotion regulation dictate whether emotions have a positive or negative effect on decision-making performance (Fenton-O'creevy et al., 2011).

3. Research Methodology

This work aims to study the impact of acute stress on decision-making performance. For that, it is necessary to use an experimental model to accurately manipulate and measure the stress level in participants.

3.1 Sample and Experiment Modeling

The population comprises 44 healthy adults, split into a control group with 14 participants (32%) who were not submitted to stressors and a stress group with 30 participants (68%). From a demographics perspective, 64% were young adults (n=28), 36% were middle-aged adults (n=16), 59% were men (n=26), and 41% were women (n=18).

The proposed method considers the design elements and application specifics with a study to validate the proposed method's effectiveness in measuring arousal during the various stimuli and evaluating how it affects individual decision-making performance. Each participant follows the sequence of stimuli as the wearable biosensors collect their physiological data. These parameters are tracked and recorded in the iMotions software version 8.2 where participants' stress/ arousal level is computed. The system overview is illustrated in Figure 1.

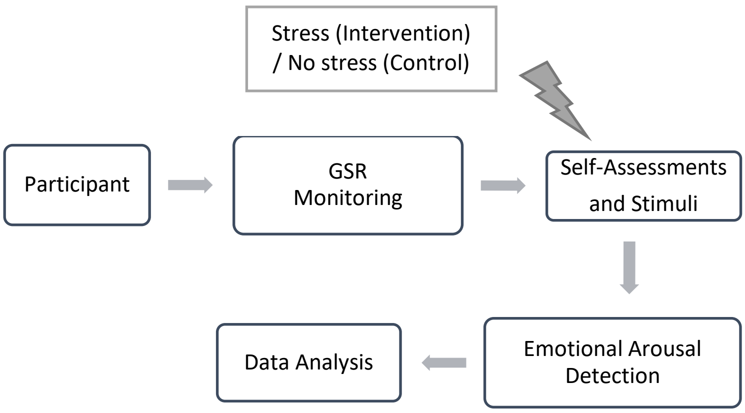

Figure 1: Overview of the study design

Several criteria were developed to determine the right decision-making activities. First and foremost, it was important to use reliable and valid decision-making activities. Choosing activities that could be done roughly in 20 minutes within which people are most overwhelmed was also critical. Finally, it was essential to use decision-making activities that could monitor naturalistic decision-making patterns, which could be repeated so that changes in preferences could be detected after stress induction. Within this research, to get to the simplest decision-making constructs, it was important to select activities that would allow the study of particular forms of decision-making, such as pure risk, pure uncertainty, or time constraint. Furthermore, since preferences for pre- and post-acute stress exposure had to be assessed, choosing tasks that could be conducted several times without experiencing practice consequences was critical.

The experiment was modeled with two (2) self-assessments, one in the beginning and the second after all the stimuli, and five (5) decision-making situations under a specific type of stressor. The model is presented as follows:

- Initial Self-Assessment
- Decision-Making (DM) under time pressure
- Decision-Making (DM) under misinformation
- Decision-Making (DM) under uncertainty
- Decision-Making (DM) under risk
- Framing of Outcomes Decision-Making (DM)
- Final Self-Assessment

3.2 Statistical Analysis

The current analysis will test normalization assumptions and nonparametric tests. Data normalization is estimated by the Shapiro-Wilk test. It examines how close the sample data is to a normal distribution. The null hypothesis is that the data are from a normal distribution, and the alternate hypothesis is that the data are not normally distributed (Ramachandran and Tsokos, 2021). Another test applied to the collected data is the Mann-Whitney. It is a nonparametric test equivalent to the t-Test for independent samples. Both tests are employed to determine if there are significant statistical differences between the two sample groups (MacFarland and Yates, 2016).

4. Results and Discussion

This section contains the findings collected during the stress exposition with the GSR monitoring.

4.1 Galvanic Skin Response (GSR) Monitoring Metrics

GSR originated from the sweat gland activation by the autonomic nervous system in the skin. The raw GSR data collection consists of two (2) characteristic components. The first component is known as peak detection, and it is measured in a binary code, (0) – if no peak is detected and (1) – if one or more peaks are detected. The second component refers to the peak count, which translates to the number of peaks present per stimuli for every participant. Both are relevant and presented as follows:

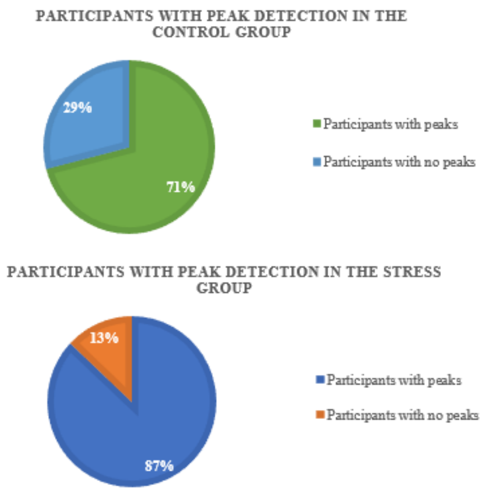

Figure 2: Percentage of participants with peak detection in the control and stress groups

4.1.1 Peak Detection

For peak detection, Figure 2 shows that in the control group, peaks were detected and counted in 71% of the participants (n=10), while there were no peaks detected in 29% (n=4) of the participants. For the stress group, peaks were detected and counted in 87% of the participants (n=26), while there were no peaks detected in 13% (n=4) of our participants. Hence there were many more participants with peaks detected and counted in the stress group than in the control group.

From Figure 3, it is possible to visualize the percentual comparison of peak detection per stimuli in both the control and stress groups. It is visually clear that the percentual number of peaks detected is higher for all the stimuli in the stressed group than in the controlled group. The percentual number of peaks detected varies more in both DM under misinformation and DM under risk, with a 34% difference between the two groups. For the three (3) stimuli, DM under time pressure, DM under uncertainty, and framing of outcomes DM, the variation is similar (around 30% difference) for both groups.

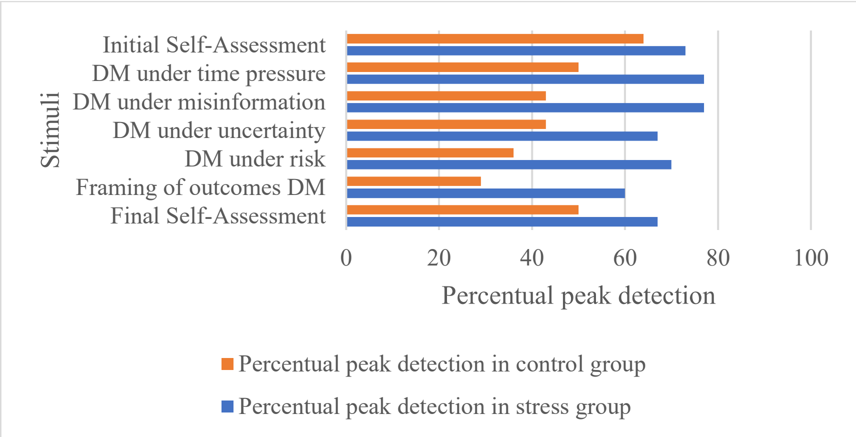

Figure 3: 1Percentual comparison of peak detection per stimuli in the control and stressed groups

The stimulus where the difference is less significant is the initial self-assessment, where the difference between the two (2) groups is only 9%. This lower difference may result from the natural stress created when a participant is submitted to self-perception questions [15], even for the control group, which was not exposed to stress information at the beginning of the experiment.

From Figure 3, it is also possible to infer that the stimuli that generated more peaks for the stress group were the stimulus where participants had to apply their DM skills under time pressure and after being misinformed about their task. For this stimulus, peaks were detected in twenty-three (23) out of thirty (30) participants (77%). The stimulus where fewer peaks were detected was for the framing of outcomes DM, with peaks detected for eighteen (18) out of thirty (30) participants (60%).

Regarding the control group, the stimulus that generated more peaks was the initial self-assessment, with peaks detected in nine (9) out of fourteen (14) participants (64%), followed by both the DM making under time pressure and the final self-assessment, with peaks detected in seven (7) out of fourteen (14) participants (50%) for both cases. The stimulus where fewer peaks were detected was also for the framing of outcomes DM, with peaks detected for four (4) out of fourteen (14) participants (29%).

It is important to highlight that the simple occurrence of GSR peaks is significantly relevant to the proposed research. Nevertheless, this parameter will denote emotional arousal and not necessarily a classification of stress. This means that excitement, joy, frustration, or other emotions may also generate these peaks. Because of that, the number of peaks detected during the whole experiment and each stimulus will provide a better comparison tool between the control and stress groups.

4.1.2 Peak count

Diving into the GSR's second metric, peak count, the results are presented as the number of peaks detected per participant in all seven stimuli. Figure 4 refers to the peak count per participant in the control group, and Figure 5 refers to the peak count per participant in the stress group.

From a general perspective, we can observe that the number of peak counts is much higher for the stress group than for the control group. The maximum peak count in the stress group is for Participant 43 (P43), with a sum of one hundred and ninety-five (195) peaks during the whole experiment. In the control group, the maximum peak count is attributed to participant 10 (P10) with a sum of only seventy-nine (79) peaks during the whole experiment. Furthermore, in the stress group for participants 16, 22, 31, and 34 (P16, P22, P31, P34) no peaks were detected, and therefore the count is equal to zero (0). In the control group, the participants 3, 8, 12, and 13 (P3, P8, P12, P13) no peaks were detected, and therefore the count is equal to zero (0).

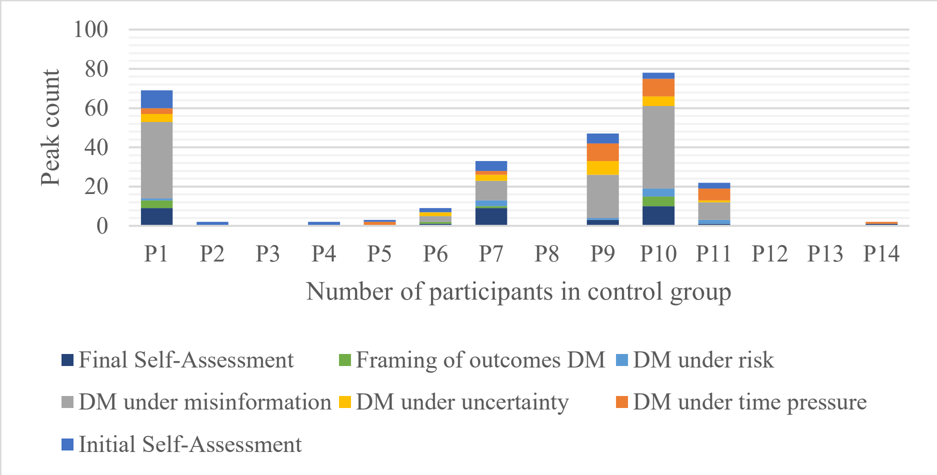

Figure 4: Peak count per participant in the control group for the whole experiment

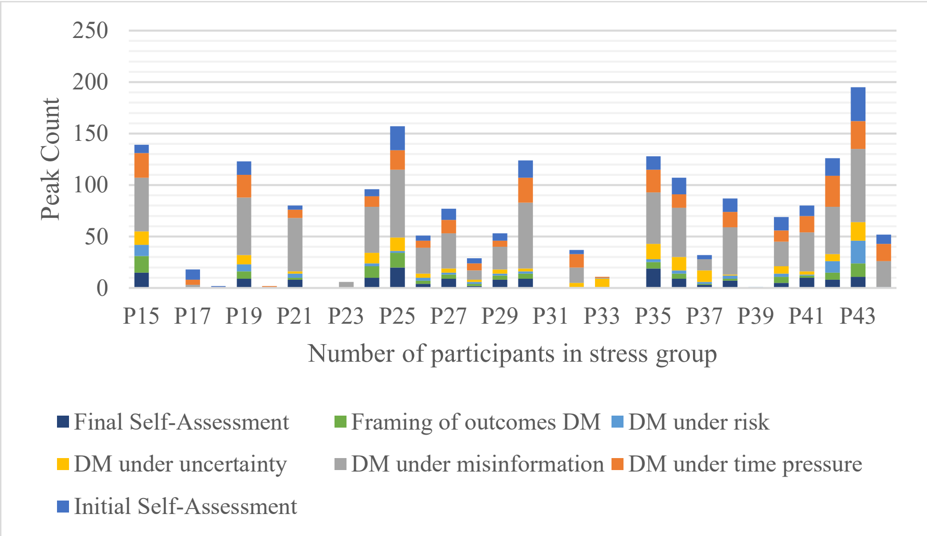

Figure 5: Peak count per participant in the stress group for the whole experiment

From Figure 4, we can also observe that even though, in the control group, the initial self-assessment was where peaks were detected for a higher number of participants (64%, n=8), the higher number of peaks is attributed to the DM under misinformation stimulus. For DM under misinformation, the average number of peak counts is nine (9) peaks, with a maximum peak count of forty-two (42) peaks (for P10) and zero (0) peaks in 57% (n=8) of control group participants. The lower number of peaks is attributed to DM under risk, where the average number of peaks is one (1), with a maximum peak count of four (4) peaks (P10) and zero (0) peaks counted for 64% (n=9) of control group participants (P2 until P6, P8, and P12 until P14).

From Figure 5, we can observe that even though, in the stress group, both DM under time pressure and DM under misinformation were the stimuli where peaks were detected for a higher number of participants (77%, n=23), the higher number of peaks is also singly attributed to the DM under misinformation. For DM under misinformation, the average number of peak counts is twenty-seven (27), with a maximum peak count of seventy-one (71) peaks (for P43) and zero (0) peaks in 23% (n=7) of stress group participants. The lower number of peaks is also attributed to the framing of outcomes DM, where the average number of peaks is four (4) peaks, with a maximum peak count of sixteen (16) peaks (for P15) and zero (0) peaks counted for 40% (n=12) of the control group participants.

4.2 Statistical Analysis

The first statistical test was Shapiro's to check whether each feature had a normal distribution, this being the null hypothesis. The features are Age, Gender, Peak Detection, and Peak Count. The results indicate that all features under analysis are nonparametric. With that, Hypothesis 1 can be tested.

H1 - Is the stress group indeed more stressed than the control group?

Testing this hypothesis is crucial to check if there is a difference between the groups. In other words, the control group presents different indicators to check if the stressors worked and if the stress group is indeed stressed. The features of Peak Detection and Peak Count are evaluated. In order to assess if features from the stress group have a different distribution from features from the control group, the Mann-Whitney test is used, with the null hypothesis being that features from both groups come from the same distribution.

The results show that the null hypothesis is rejected for the Peak Count feature, showing that the groups are statistically different, thus proving Hypothesis 1. It was expected that Gender and Age would have the same distribution, as there should not be an unbalance between groups.

5. Conclusion

The experiment conducted in this work indicates that the GSR, a real-time physiological monitoring tool, can be used to detect and quantify the presence of stress during decision-making activities. The conclusion is based on the exposure of experimental stimuli to two (2) groups of individuals (stress and control groups) and the corresponding recording of physiological reactions to understand stress's role in an individual's decision-making capacity.

The proposed experiment explores the possibility of acute stress affecting decision-making choices in healthy adults and to better understand the emotional expression of stress during the performance of decision-making tasks. We submitted adult participants to an acute laboratory stressor, followed by presenting them with five (5) tasks designed to assess different types of decision-making.

Future works might consider other biosensors, such as heart rate variability (HRV) and electroencephalograms (EEG), for recording brain activity. Furthermore, the conventional ways of studying human behavior have relied on a question-and-answer methodology (surveys and questionnaires) for a long time, which was used in this study to formulate the decision-making tasks. While this methodology can provide a lot of information, new approaches such as serious games and game biofeedback are being used to better evaluate the decision-making of individuals when submitted to single or multiple parameters that may originate from stress.

References

Atsan, N. (2016). Decision-Making under Stress and Its Implications for Managerial Decision-Making: A Review of Literature. International Journal of Business and Social Research, 6(3), 38. https://doi.org/10.18533/ijbsr.v6i3.936

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding Advantageously Before Knowing the Advantageous Strategy. *Science*, *275*(5304), 1293–1295.

Billings, R. S., Milburn, T. W., & Schaalman, M. L. (1980). A Model of Crisis Perception: A Theoretical and Empirical Analysis. Administrative Science Quarterly, 25(2), 300. https://doi.org/10.2307/2392456

Bower, G. H. (1981). Mood and Memory. American Psychologist, 36, 129–148.

Bronner, R. (1982) Decision Making Under Time Pressure: An Experimental Study of Stress Behavior in Business Management. Lexington Books, 186 pp. ISBN: 9780669046960.

Dane, E., & Pratt, M. G. (2007). Exploring Intuition and its Role in Managerial Decision Making. The Academy of Management Review, 32(1), 33–54. https://doi.org/10.5465/amr.2007.23463682

Etzion, D. (1984). Moderating effect of social support on the stress–burnout relationship. Journal of Applied Psychology, 69(4), 615–622. https://doi.org/10.1037/0021-9010.69.4.615

Fenton-O'Creevy, M., Soane, E., Nicholson, N., & Willman, P. (2011). Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders. Journal of Organizational Behavior, 32(8), 1044–1061. https://doi.org/10.1002/job.720

Hermann, C. F. (1963). Some consequences of crisis which limit the viability of organizations. Administrative Science Quarterly, 61–82.

Isen, A. M., Clark, M., Shalker, T. E., & Karp, L. (1978). Affect, Accessibility of Material in Memory, and Behavior: A Cognitive Loop? Journal of Personality and Social Psychology, 36(1), 12.

Kavanagh, D. J., & Bower, G. H. (1985). Mood and self-efficacy: Impact of joy and sadness on perceived capabilities. Cognitive Therapy and Research, 9(5), 507–525. https://doi.org/10.1007/BF01173005

Lerner, J. S., & Keltner, D. (2001). Fear, Anger, and Risk. Journal of Personality and Social Psychology, 81(1), 146–159.

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart Strings and Purse Strings. Carryover Effects of Emotions on Economic Decisions. Psychological Science, 15(5), 337–341. https://doi.org/10.1111/j.0956-7976.2004.00679.x

MacFarland Thomas W., Yates Jan M. (2016) Mann–Whitney U Test:103–132. Cham: Springer International Publishing.

Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress and decision-making: A few minutes make all the difference. Behavioural Brain Research, 250, 39–45. https://doi.org/10.1016/j.bbr.2013.04.046

Poh, M.-Z., McDuff, D. J., & Picard, R. W. (2011). Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam. IEEE Transactions on Biomedical Engineering, 58(1), 7–11. https://doi.org/10.1109/TBME.2010.208645.

Pope, N. (2010). The impact of stress in self- and peer assessment. Assessment and Evaluation in Higher Education, 30, 51–63. https://doi.org/10.1080/0260293042003243896

Raaijmakers, S. F., Steel, F. W., de Goede, M., van Wouwe, N. C., van Erp, J. B. F., & Brouwer, A.-M. (2013). Heart rate variability and skin conductance biofeedback: A triple-blind randomized controlled study. Affective Computing and Intelligent Interaction, Soesterberg, The Netherlands.

Ramachandran Kandethody M., Tsokos Chris P. (2021). Chapter 11 - Categor- ical data analysis and goodness-of-fit tests and applications in Mathe- matical Statistics with Applications in R (Third Edition) (Ramachan- dran Kandethody M., Tsokos Chris P.. , eds.):461-490. Academic Press. 3 ed.

Schunk, D., & Betsch, C. (2006). Explaining heterogeneity in utility functions by individual differences in decision modes. Journal of Economic Psychology, 27(3), 386–401. https://doi.org/10.1016/j.joep.2005.08.003

Schwarz, N., & Clore, G. L. (2003). Mood as Information: 20 Years Later. Psychological Inquiry, 14(3/4), 296–303.

Staw, B. M., Sandelands, L. E., & Dutton, J. E. (1981). Threat Rigidity Effects in Organizational Behavior: A Multilevel Analysis. Administrative Science Quarterly, 26(4), 501. https://doi.org/10.2307/2392337

Todd, P. M., & Gigerenzer, G. (2007). Environments That Make Us Smart: Ecological Rationality. Current Directions in Psychological Science, 16(3), 167–171. https://doi.org/10.1111/j.1467-8721.2007.00497.x

Wright, W. F., & Bower, G. H. (1992). Mood effects on Subjective Probability Assessment. Organizational Behavior and Human Decision Processes, 52, 276–291.