User Arousal in XAI Usage

Maurice L. Pape[1]

[1]Paderborn University, Organizational Behavior, Department of Management,
Warburger Str. 100, 33098 Paderborn, Germany

Abstract

As AI has become a ubiquitous technology in the past years, it is ever so important to foster better collaboration between human and machines. While standard AI has an opaque black box nature, XAI tries to circumvent this by offering explanations to users for its output. Many different aspects of this collaboration have been under research, while an important aspect, the emotions of the human users, have yet to be examined thoroughly. This PhD will try to provide an insight into the role of emotional arousal for the reception of explanations and try to develop systems that can take the arousal into consideration for their explanations. The first contribution will be to check the role of arousal stemming from context factors like time pressure and task importance in one-shot interaction. Further down the line, this will be used to try to build systems for repeated interaction that can recognize the human users´ arousal level and adapt correctly to foster the correct level of collaboration that humans want and need. A first contribution has been made showing the relationship between arousal and advice-taking indeed seems to follow a U-shaped pattern, where low and high levels seem suboptimal for correct reliability on the system.

Keywords

XAI, emotional arousal, human computer interaction, HCI, user study

1. Context & Motivation

In the past few years, AI has become a ubiquitous part of everyday life. The emergence of ChatGPT and subsequent other general AI agents has brought the topic into the eyes of the public. Other AI systems for expert use have been in use for a long time, be it in different branches of medicine [1], [2], in finance [3] or in education [4], just to name a few. While the usage of data-driven decision-making has its advantages and disadvantages [5], a big drawback comes from the opaqueness of many of these tools. To alleviate the black box nature coming with AI, explainable AI has been proposed to help humans with interpretability and transparency concerns [6]. While these provided local or global explanations on how the XAI derives at its answer or recommendation, it is important to consider the other side of the equation, namely the human receiving the explanation.

Many factors can influence the receptiveness to explanations by humans, one of them being the emotional state. Emotional arousal has been linked to performance for quite some time, with the Yerkes-Dodson law describing an inverted U-shaped relationship between the two [7]. Not only is arousal linked to performance, but in this context it is believed to also influence the reception of explanations. As many XAI applications are either in high arousal (important or fast decisions) or low arousal (repetitive decisions) environments, the influence of these context-related factors on explanations provided by the XAI is of immense importance. This doctoral thesis will be centered on the relationship between emotional arousal and XAI explanations. From looking at the relationship between those two, calibrating optimal levels, testing them in single-shot to repeated interactions and finally constructing XAI concepts where the goal and level of arousal can be negotiated, this work is far from finished. In the following parts, I will go deeper into the literature surrounding the research, the methodology that will help to fulfill those goals up until research we have already conducted in this field.

Late-breaking work, Demos and Doctoral Consortium, colocated with the 4th World Conference on eXplainable Artificial Intelligence: July 01–03, 2026, Fortaleza, Brazil

maurice.lutz.pape@uni-paderborn.de (M. L. Pape)
0009-0001-5627-8808 (M. L. Pape)

2. Related work

As the field of artificial intelligence emerged, one problem quickly became apparent, which is the black-boxed nature of AI agents [8]. To alleviate this problem, a new concept, explainable artificial intelligence, has been proposed [9]. While explaining the results is one side of the coin, the other is the reception of the explanation from the explainee.

For an XAI to create thorough and dynamic explanations, it must consider different context factors, for example, human knowledge, the surrounding context, the goals, and the emotional state of the human explainee. To ensure that XAI provides effective explanations, it is important to consider how emotions impact interactions between an AI explainer and a human receiving the explanation. Emotions are an integral part of everyday life as they accompany almost all humans at all times. Therefore, they also play a central role in human decision-making. Emotions are widely seen as having two dimensions, valence and arousal [10]. While valence captures the extent to which an emotion is positive or negative, arousal describes the degree of wakefulness or activation of an emotion. Therefore, arousal can be combined on all levels with positive and negative valence.

Prior research points out that arousal plays an important role in decision-making, more so than valence [11]. While arousal has long been considered a key determinant of performance [7], this U-shaped relationship is believed to also be found in advice-taking. Under very low arousal, individuals may not activate enough cognitive resources to prevent noisy or unstable decisions. On the other hand, high arousal may consume cognitive resources for emotion regulation and not for deliberating the task at hand, although it has been shown that heightened arousal may help facilitate adaptive and functional decision-making in some contexts [12]. Furthermore, the task at hand plays an important role regarding the arousal level. High arousal may be beneficial for simpler decisions to keep users engaged while intermediate levels of arousal are more conducive to complex tasks [13]. But high arousal may also hinder solid decisions as it may impair belief updating and may promote overgeneralization in the decision-maker [14]. Therefore, the role of higher arousal needs to be studied further under different situations, while low arousal seems to be undesirable in decision-making.

Regarding advice, there have been mixed results considering the level of arousal. It has been reported that increased levels of arousal may lead individuals to rather follow the advice given to them [15]. This has been contradicted by other studies, where higher arousal may lead to less advice-taking [16]. These studies have been conducted in human-human advice scenarios and not in scenarios regarding advice from a machine, which could be subject to, for example, algorithm aversion [17], which is another important factor to take into account. As the contemporary literature is very vague regarding the role of emotions in advice-taking, I want to narrow this gap in my research.

3. Research Goal & Question

The goal of this PhD thesis will be to investigate how the context factors time pressure and importance affect the goals and expectations of the explainee, and how to mitigate the effects of expectation mismatches between the XAI and the explainee. The assumption is that both contextual factors specifically affect arousal and the resulting level of engagement with the explanatory process. Whereas too little and too much time may result in individuals feeling that explanations are irrelevant given the social context, or that they may have too few cognitive resources for engagement in the explanatory process, the importance of a decision may result in a similar U-shape in which people are not aroused enough or choke under pressure. Furthermore, the goal will be to develop a dialog strategy to assess these context factors and to negotiate the resulting goals, for example, deep/shallow or enabling/comprehensive explanations. Another goal would be to develop explanation strategies to scaffold understanding in the presence of emotional arousal, based on an emotion-sensitive

grounding hierarchy assessing emotional and epistemic reactions, understanding, and nonacceptance of, acceptance of, or agreement with the explanatory item.
In particular, the research will address five related objectives to achieve the final goal:

1. To better understand the optimal level of arousal during explanations by differentiating context-dependent sources of arousal that may affect the explanatory goal (fast explanations vs deep comprehension).

2. To analyze arousal as a factor during the explanatory process. If the explainee needs to understand the AI system, they must stay engaged during the process driven by arousal. We analyze the effectiveness of intervention strategies to keep the explainee engaged.

3. To develop a computational model that negotiates the explanatory goals with the explainee and monitors their engagement to communicate accordingly during the process.

4. To research more realistic time contexts of XAI use. Previous research often focuses on one-time interactions in which there is little routinization of humans, low experience gains, and little learning. However, in many instances XAI will be used regularly at work, and XAI methods should take these repeated interactions into account by using the opportunity to explain consecutively. Also, they probably must avoid "obtuse" blind following, at least when the explainees goal is understanding.

5. To broaden the application area by employing not only lab studies with students but also lab-in-the-field experiments in more realistic settings and with high-stakes decisions.

Attaining these objectives will be an important step to improve the explainability of XAI. This will take human context in terms of diverse engagement levels into account, and reach the goal of making XAI explainable for humans not only in a technical but also in a social, co-constructive way.

4. Research approach

To realize the objectives of this research, multiple different approaches, experimental settings, and methods will be employed. The goal is to use lab, lab-in-the-field, and field experiments to develop XAI matching the various objectives listed above. All three types of data collection will be required.

In the first phase, regular lab methods will be used to analyze the effects of different sources of arousal on the propensity to engage in co-construction, advice-taking, and understanding. Initial research on emotional arousal was already conducted, and the goal is to broaden these results further in the first phase. The rationale behind lab experiments is to reach a sufficient sample size and generate some first results about the effects of social contexts in a controlled setting.

However, further down the line, repeated interactions within a laboratory setting will not help to analyze long-term decisions and secure external validity. In many real-world settings, individuals interact with an XAI system daily or at least regularly and repeatedly. Thus, the employment of lab-in-the-field methods should improve the understanding of repeated interactions and the problems and opportunities of co-construction these situations provide. The idea will be to use standard economic and psychological experiments and aiding the participants with an XAI advisor and repeated interactions with the XAI system. Further, lab-in-the-field experiments are an experimental setting with great control regarding the task, while providing a direct measure of arousal of "real", physically available individuals. These will be an important intermediary step before testing in the field under even more realistic conditions.

While these lab-in-the-field experiments will give us some more insights and more samples, the most important part will be the field experiments at the end of the project. For this, the goal would be to find a company or multiple companies with everyday AI or XAI users who then can test and use our developed tools. This would preferably happen over a longer time frame so that it will be possible to observe how not only one-shot interactions influence the perception of XAI explanations, but to see how these may also change over time in everyday, repeated interactions. Of great interest would be to see the system recognize and adapt to different levels of arousal over time coming from one user.

5. Preliminary results & contribution to date

As mentioned above, the first results have been obtained regarding the first of the five objectives. The research question asked was whether there was a significant relationship between arousal and reliance on AI-generated advice. As seen in the literature to date, it was hypothesized that there would be a U-shaped relationship between arousal and advice-taking, as seen in performance in general [7]. Therefore, we postulated the hypothesis:

H1: Participants displaying high or low levels of arousal will take advice more often than participants with a medium level of arousal.

Furthermore, the study was also aimed at taking a look at the relationship between different measures of arousal. There are two possibilities to assess the arousal level of a person: first is the objective measure, where physiological data is measured to detect arousal; in our example this was the heart rate variability (HRV). HRV is seen as the most direct and reliable physiological indicator of arousal [18]. The second measure of arousal is the subjective one, where the participants report their own perceived level of arousal. This can, for example, be measured on a scale like the Self-Assessment Manikin (SAM) [19]. There have been mixed results regarding the two measurement methods, with some reporting convergence of the two [20] and others seeing no correlation [21]. Therefore, the second hypothesis in the research concerned the relationship between the arousal measures, as this will be important for the further plans of the PhD thesis:

H2: There will be a low correlation between the subjective arousal reported by participants and the objective arousal measured via heart rate variability.

To test these two hypotheses, we conducted a total of two lab experiments. The participants HRV was measured by a wearable device, while they reported their arousal before and after the decision-making task. As a task, a Holt and Laury lottery was chosen, a well-established tool for assessing decision-making under risk [22]. The participants were aided by an XAI, which beforehand assessed their level of risk-taking and advised participants on the "correct" lottery for their risk level. In both experiments we induced non-task-related arousal, the difference being that in the first we induced light arousal while in the second induced stronger arousal. This was done in the first by the mental rotation task, which was always solvable, and in the second, where the participants solved anagram tasks, of which some had no correct answer. The dependent variable in both cases was the advice-taking index (WOA).

The experiments yielded similar results. Therefore, I will discuss both together. Regarding the first hypothesis, we saw promising results, as we indeed observed the expected u-shaped relationship between arousal and the WOA. While the two measures alone did not have a significant effect on the WOA, a highly significant effect was observed when using an interaction between the two. In this interaction effect, we can see that high and low levels of arousal reported in both measures meant a higher level of advice taking, while mixed arousal levels led to lower levels of the WOA. We suppose this is due to participants displaying high levels of arousal in both dimensions being subject to a lower level of information processing, thus resolving to simpler decision strategies [23]. On the other hand, participants with both measures reporting low arousal may be subject to them not discriminating between their options correctly and thus relying on AI advice [24].

Table 1
Multiple linear regression of WOA, Experiment 1

Variable	Model 1	Model 2	Model 3	Model 4
SAM	0.051		0.047	0.413**
	(0.065)		(0.065)	(0.174)
RMSSD		3.050	2.948	19.991**
		(2.818)	(2.842)	(8.062)
SAM × RMSSD				-6.427**
				(2.868)
Constant	0.233	0.160	0.042	-0.919*
	(0.188)	(0.205)	(0.263)	(0.495)

*S*ignificance levels denoted as $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses.

Regarding the second hypothesis, we saw a low and insignificant correlation between the two measures of arousal (experiment 1: $r(32) = -0.051, p = 0.773$, experiment 2: $r(37) = -0.266, p = 0.102$), in line with [21]. Therefore, these measures do not seem to be substitutes, but regarding the results of H1, complementary to each other. We suppose this may partly be due to the fact that neither high nor low-level aroused people can very well interpret their level themselves and therefore often return mixed results. This will need further investigation, which we will conduct in future parts of this thesis.

Overall, the results were very promising. While we saw strong divergence between subjective and objective measures of arousal, we saw that they certainly seemed to be complementing each other and together may play an important role in understanding arousal. As we have seen, there seems to be a level where arousal is indeed better and does not lead to blindly following the advice given by the XAI. While this was not yet a sophisticated XAI the participants worked with, we still saw the cooperation of a system giving advice to humans and explaining the reasoning behind the decision. Future research will need to dig deeper into the different arousal levels and cooperation between humans and XAI, which is outlined in the future plans for this doctoral thesis. But these early results give a first insight into the role that arousal may play in the hindrance or optimal cooperation between humans and XAI.

Table 2
Multiple linear regression of WOA, Experiment 2

Variable	Model 1	Model 2	Model 3	Model 4
SAM	0.074		0.098	0.358**
	(0.079)		(0.081)	(0.159)
RMSSD		2.604	3.622	17.505**
		(3.072)	(3.167)	(8.007)
SAM × RMSSD				-5.633*
				(3.002)
Constant	0.304	0.363**	0.098	-0.619
	(0.221)	(0.176)	(0.081)	(0.467)

*S*ignificance levels denoted as $^{**}p < 0.05$, $^{*}p < 0.1$. Standard errors in parentheses.

6. Next steps and future outlook

The future outlook for this project is relatively clear regarding the nearer future and mostly outlined as described in Section 3. In the near future, the next step will be to test the level of arousal stemming from the immediate contextual factors, which are the time pressure and importance of a task. For this, we are preparing three experiments. In a first experiment, we want to test the influence of task importance. For this, we prepare a visual search and decision task for participants in which they are to find specific allergens in prepared plates of food. They will be assisted by a Wizard-of-Oz style XAI in their task, in which the XAI will give a clue regarding whether the dish contains allergens or not. The participants then can receive further explanations if they want to. We will vary the level of importance in this experiment by varying the payout structure for different treatments, and will include multiple levels to check for the U-shaped curve described in the literature and seen in our first experiment. Arousal will then be measured by the Self-Assessment Manikin so that the study can gather a large sample size. In a second experiment, we want to test the influence of time pressure. We will use the same experimental setting and only vary the treatments, where there will be different levels of time pressure and time constraints on the participants. The rationale behind using only one measure of arousal is to gather a large sample over click-working sites to try to observe the effect of self-reported arousal there, as the first sample was small and homogeneous.

The third experiment will then aim to analyze the effects of both task-related arousal factors on the reception of explanations. We will again use the same task, but in a different setting. We aim to undertake this experiment on site in a lab. The rationale behind this is to also record the participants relevant psychophysiological data. As described above, we have seen the divergence of both measurements of arousal, and we want to again check whether there will be different results regarding the two different measurements when the arousal stems from contextual factors. It is also interesting to examine the combination of multiple arousing factors is an interesting factor, as often both of these factors co-occur in usual XAI tasks.

Another planned experiment will address the third part of the planned roadmap outlined above. As the goal of the project is to foster a scaffolding process between the explaining system and the explainee, another important step will be goal assessment to optimize the XAI explanation strategy for the humans situation. For this, we want to assess different users goals, from experts to novel users. This will be done in an experiment with different domains, one of them being chess. We will be asking different levels of chess players to solve given situations on a chessboard, where they will be supported by an XAI. The goal is to see whether they will just follow the XAI blindly or if they want to check the reasoning and explanation behind the moves put forward by the XAI. We expect different levels of users to have different goals, from simply putting the task behind them up to learning more about chess. This will be done in three to four different domains, although the experiment is still in its developmental phase.

While these will deal with the objectives of the first part of the PhD thesis and give a slight look into the third phase, there are several further parts that build on the first one. Depending on the results we see in the first three experiments, we will design further experiments to advance further goals. The second part will be to design experiments where we can measure arousal during the explanatory process and try to design optimal interventions to keep the participants, or later users, engaged with the explanation. When these intervention strategies have been developed, we will go further to create a system that can also negotiate goals with the explainee and incorporate the interventions to further foster optimal collaboration between the explainee and the XAI. In the fourth and fifth parts, the goal is to go away from the lab and perform experiments in the wild, either in lab-in-the-field experiments or complete field experiments to test the developed system. The end goal of this thesis will be to provide a system or strategies that will help to amplify explainability of the XAI and foster correct levels of reliance on the system. The contribution to knowledge then should be to broaden the understanding of the role of arousal in human-XAI interactions and to use this knowledge to design systems and strategies to use arousal and to also generate knowledge for these strategies and their functionality. To the best of our knowledge, this has not been researched in detail yet.

Declaration on Generative AI
The author has not employed any Generative AI tools.

References

[1] B. M. De Vries, G. J. Zwezerijnen, G. L. Burchell, F. H. van Velden, C. W. Menke-van der Houven van Oordt, R. Boellaard, Explainable artificial intelligence (xai) in radiology and nuclear medicine: a literature review, Frontiers in medicine 10 (2023) 1180773.

[2] C. de Margerie-Mellon, G. Chassagnon, Artificial intelligence: A critical review of applications for lung nodule and lung cancer, Diagnostic and interventional imaging 104 (2023) 11–17.

[3] J. Černevičienė, A. Kabašinskas, Explainable artificial intelligence (xai) in finance: a systematic literature review, Artificial Intelligence Review 57 (2024) 216.

[4] I. H. Y. Yim, J. Su, Artificial intelligence (ai) learning tools in k-12 education: A scoping review, Journal of Computers in Education 12 (2025) 93–131.

[5] S. D. Shaw, G. Nave, Thinking-fast, slow, and artificial: How ai is reshaping human reasoning and the rise of cognitive surrender, Available at SSRN 6097646 (2026).

[6] K. Kalasampath, K. Spoorthi, S. Sajeev, S. S. Kuppa, K. Ajay, M. Angulakshmi, A literature review on applications of explainable artificial intelligence (xai), IEEE access (2025).

[7] L. Beerendonk, J. F. Mejías, S. A. Nuiten, J. W. de Gee, J. B. Zantvoord, J. J. Fahrenfort, S. van Gaal, Adaptive arousal regulation: Pharmacologically shifting the peak of the yerkes–dodson curve by catecholaminergic enhancement of arousal, Proceedings of the National Academy of Sciences 122 (2025) e2419733122.

[8] A. Rai, Explainable ai: From black box to glass box, Journal of the academy of marketing science 48 (2020) 137–141.

[9] F. Doshi-Velez, B. Kim, Towards a rigorous science of interpretable machine learning, arXiv preprint arXiv:1702.08608 (2017).

[10] J. A. Russell, A circumplex model of affect., Journal of personality and social psychology 39 (1980) 1161.

[11] K. Thommes, O. Lammert, C. Schütze, B. Richter, B. Wrede, Human emotions in ai explanations, in: World Conference on Explainable Artificial Intelligence, Springer, 2024, pp. 270–293.

[12] Y. Hanoch, O. Vitouch, When less is more: Information, emotional arousal and the ecological reframing of the yerkes-dodson law, Theory & Psychology 14 (2004) 427–452.

[13] L. K. Sörensen, S. M. Bohté, H. A. Slagter, H. S. Scholte, Arousal state affects perceptual decision-making by modulating hierarchical sensory processing in a large-scale visual system model, PLoS Computational Biology 18 (2022) e1009976.

[14] L. F. Ciria, M. Suárez-Pinilla, A. G. Williams, S. R. Jagannathan, D. Sanabria, T. A. Bekinschtein, Different underlying mechanisms for high and low arousal in probabilistic learning in humans, Cortex 143 (2021) 180–194.

[15] F. Gino, A. W. Brooks, M. E. Schweitzer, Anxiety, advice, and the ability to discern: feeling anxious motivates individuals to seek and use advice., Journal of personality and social psychology 102 (2012) 497.

[16] F. Gino, M. E. Schweitzer, Blinded by anger or feeling the love: how emotions influence advice taking., Journal of Applied Psychology 93 (2008) 1165.

[17] B. J. Dietvorst, J. P. Simmons, C. Massey, Algorithm aversion: people erroneously avoid algorithms after seeing them err., Journal of experimental psychology: General 144 (2015) 114.

[18] M. Mohammadpoor Faskhodi, M. Fernández-Chimeno, M. A. García-González, Arousal detection

by using ultra-short-term heart rate variability (hrv) analysis, Frontiers in Medical Engineering 1 (2023) 1209252.
[19] M. M. Bradley, P. J. Lang, Measuring emotion: the self-assessment manikin and the semantic differential, Journal of behavior therapy and experimental psychiatry 25 (1994) 49–59.
[20] D. Ciuk, A. Troy, M. Jones, Measuring emotion: Self-reports vs. physiological indicators, Physiological Indicators (April 16, 2015) (2015).
[21] F. Bolinski, A. Etzelmüller, N. A. De Witte, C. van Beurden, G. Debard, B. Bonroy, P. Cuijpers, H. Riper, A. Kleiboer, Physiological and self-reported arousal in virtual reality versus face-to-face emotional activation and cognitive restructuring in university students: A crossover experimental study using wearable monitoring, Behaviour Research and Therapy 142 (2021) 103877.
[22] C. A. Holt, S. K. Laury, Risk aversion and incentive effects, American economic review 92 (2002) 1644–1655.
[23] S. Wichary, R. Mata, J. Rieskamp, Probabilistic inferences under emotional stress: how arousal affects decision processes, Journal of Behavioral Decision Making 29 (2016) 525–538.
[24] Y. Xu, M. Wokke, V. Noreika, C. Bareham, S. Jagannathan, S. Georgieva, C. Trentin, T. Bekinschtein, Effects of alertness on perceptual detection and discrimination, cortex (2025).