Middle East Technical University
Informatics Institute

Investigating Emotions over the Monetary Outcomes of Wheel-Spin Game using Bayesian Models

Advisor Name: Doç. Dr. Barbaros Yet
(METU)

Student Name: İrem Arıcı
(Cognitive Science)

January 2024

TECHNICAL REPORT
METU/II-TR-2024-

Orta Doğu Teknik Üniversitesi
Enformatik Enstitüsü

Bayes Modellerini Kullanarak Çark Döndürme Oyununun Parasal Sonuçları Üzerine Duyguların Araştırılması

Danışman Adı: Doç. Dr. Barbaros Yet
(ODTÜ)

Öğrenci Adı: İrem Arıcı
(Bilişsel Bilimler)

Ocak 2024

TEKNİK RAPOR
ODTÜ/II-TR-2024-

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Internal Use)	**2. REPORT DATE** 19.01.2024
3. TITLE AND SUBTITLE **Investigating Emotions over the Monetary Outcomes of Wheel-Spin Game using Bayesian Models**	
4. AUTHOR (S) İrem Arıcı	**5. REPORT NUMBER (Internal Use)** **METU/II-TR-2024-**
6. SPONSORING/ MONITORING AGENCY NAME(S) AND SIGNATURE(S) Cognitive Science Master's Programme, Informatics Institute, METU Advisor: Doç. Dr. Barbaros Yet	Signature:
7. SUPPLEMENTARY NOTES	

8. ABSTRACT (MAXIMUM 200 WORDS)

This study investigates the relationship between emotions and outcomes of wheel-spin game with monetary reward. The inference of emotions by a probabilistic programming language account for both the overall emotional responses and individual emotions for different outcomes. The study examined previously published data with Bayesian models. With different probabilities of winning for each outcome, the emotional reaction for the amount received and the pointer position on the outcome slice are included in the models. The findings conclude that, although the positional inferences have failed, they feel less surprised in the event of high probability win and while players experience more positive emotions as the amount of monetary outcome increases, their negative emotional responses decrease.

9. SUBJECT TERMS	**10. NUMBER OF PAGES**
emotion inference, affective cognition, appraisal theory of emotion, prediction	43

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS / ABBREVIATIONS

PP: predictive processing

NN: neural networks

CNN: convolutional neural network

RNN: recurrent neural network

FER: face emotion recognition

OCC: Ortony, Clore, Collins theory

HMM: hidden Markov models

AU: action unit

CHAPTER 1

INTRODUCTION

Emotion inference has been a crucial part of human social interaction as it is highly essential to understand and interpret emotions correctly in a social context, and consequently, enable observers to act upon them appropriately (Ekman, 2004). Especially, facial expression analysis date back to Darwin (1872), with suggestion of the universality of human facial expression, and later, with Ekman and Friesen's 1971 study, six primary emotions has been hypothesized, each having unique and distinctive qualities, namely, joy, sad, fear, disgust, anger and surprise. Moreover, previous literature reviews have also shown high performances on emotion recognition of humans in a limited range of emotions, indicating universal pattern of emotion expressions over such emotions (Scherer et al., 2011). Another point of emotions is their substantial impact on human decision-making, attention, action taking, and memory (Gratch & Marsella, 2001). Therefore, humans are intuitive in the sense that they are likely to act on their, or for anothers' predicted emotions to plan their future behavior. Specifically, in PP framework, top-down and bottom-up prediction errors, when integrated, represent feelings with error dynamics as emotional valence (Joffily & Coricelli, 2013): Van de Cruys (2017) claims emotions act as continuous feedback of uncertainties of the predictions that are made.

Many theoretical frameworks have been proposed to investigate emotions scientifically; lay theories of emotion (Gopnik & Wellman, 1992; Ong et al., 2015), which offers optimal reasoning of others' mental states, behaviors, beliefs, and intentions by a human observer, while appraisal theories (e.g., OCC, Ortony et al., 1988) suggests a categories of emotions in which they are related to the causing event, actions of others and objects.

Inspired from previous research of emotions, this project aims to implement a probabilistic model for emotion inference in the existence of differential monetary outcomes of a wheel-spinning game. For such goal, previously collected dataset by Ong and colleagues (2015) is used to create a probabilistic inference model for each of 8 emotions (happy, sad, anger, fear, surprise, disgust, content, disappointed) in PyMC, a probabilistic programming language for Bayesian modeling in Python (Salvatier et al., 2016). The dataset contains outcomes of a wheel-spinning game where workers win one of three monetary rewards with different occurrence probabilities, and each of ten spins by a worker is evaluated for expected emotional response. Besides the emotion rating for each wheel-spinning trial outcome (amount won), the probability of winning that certain monetary prize and the distance to bordering prizes measured in angle proportions has been included in each Bayesian model. Additionally, since the dataset consists of 100 participants (workers) that have spun the wheel 10 times each, another Bayesian model has been built for each of 8 emotions mentioned, also including the win amount, the probability of winning, and the angle proportion of the wheel section. This study consists of as follows: in Chapter 2, the functional definition of emotion has been made within the boundaries of predictive processing (PP) framework, specifically, how emotions are regulated as a part of human sensory experience, and in turn, how they influence our decision-making and action-taking. In Chapter 3, the relevant emotion frameworks from the literature have been examined. The appraisal dimensions of emotion have been studied extensively and many theories of such nature has been reviewed here. The Chapters 4 and 5 give an insight on previous emotion recognition and Bayesian emotion models using several different techniques of computation, machine learning, and neural networks are widely deployed for emotion recognition, and, similarly, Bayesian models of emotion have been implemented to account for reasoning about emotions, such as appraisal evaluation. From Chapter 6 and on, the report focuses on the current research aim of this study, including the description of procedures, dataset, a deeper look into the nature of dataset, and the result of Bayesian models. Finally, Chapter 7 comments on the main findings of the study and possible implications of the results, and it concludes this paper.

CHAPTER 2

THEORETICAL BACKGROUND

2.1. Predictive Processing

Predictive processing (PP) is an in-demand framework for cognitive psychology studies, as Clark states it as “the perfect neuro-computational partner for recent work on the embodied mind.” PP posits that, the brain computes the external perceptual input (bottom-up) regarding higher cognitive processes (top-down) such as a knowledge-based probabilistic inference, since the brain has no access to external world, and it is bound to learn patterns and regularities of the world through actions (Seth, 2013; Clark, 2013). Specifically, this inference is thought to governed by Bayesian brain hypothesis, in which the brain aims to optimize weighting of incoming perceptual evidence against prior set of knowledge, thus resulting in inference of “hidden cause that make the current sensory data most likely” (Clark, 2015, p. 8).

The discrepancies between the expected sensory data (the prior model) and the actual incoming perceptual information arise as prediction errors, which brain continuously attempts to minimize and thus, "explaining away" to fit the data into the agent's current model of world (Seth, 2013). Moreover, another strategy to resolve prediction errors go through *interoceptive processes:* 1) the agent regulates (almost reflexively) their internal sensation (e.g., heart rate, glucose levels, temperature etc.) against the stimulus, and 2) agent moves their body (i.e., engages in action to reduce prediction error by modification (Barrett & Simmons, 2015). Consequently, as agents are unable to modulate their internal states, they are bound to generate a model of these states and their "expected" causes (van de Cruys, 2017). Hence, the agent goes through its environment engaging in action through perception to optimize its interoceptive states. So, emotions are not the (in)ability of correct inference of internal states, as argued by Barrett and Simmons (2015), and Seth (2013) against "perception of body" theories (James & Lange, 1922), yet they are defined as the brain's ability to regulate prediction errors over time. In short, emotional inferences are rooted in the PP model.

Particularly, many scholars agree on that facial expressions hints the behavioral tendencies of expressing person (Lazarus, 1991; Ekman, 1973; Russel, 1997). In their 2019 article, Albohn and colleagues combine this claim with Gibson's (1979) ecological psychology approach, as "behavioral affordances, or the opportunities a visual stimulus in our environment has to act on or be acted upon by the observer." (p. 28). More interestingly, Albohn and colleagues further suggest that there is no "true-neutral" facial expression, as the face is the dominant social cue, which observers try to "mind read" and predict further actions from the actor.

2.2. The Appraisal Theories of Emotion

Emotions prepare and bestow control to humans in the next event they need to deal, and engages the person entirely, assuring an action readiness (Frijda, 2007). As emotions are elicited in response to the relevant stimuli in the environment, which have direct relation to a perceiver's needs, goals and values, relevancy of a stimuli is based on the appraisal of a couple factors, such as its pleasantness, the novelty or surprisal of stimuli, and its influence on achievement of perceiver's goal (Scherer, 2001). Interestingly, in similar psychological contexts, humans are universally found to display 16 facial expressions, spanning in 12 world regions and 144 countries (Cowen et al. 2021). Based from emotion expressions, one can draw many information, for instance, the competency of a third person could be inferred by the instructor's surprisal, similarly by both adults and children (Asaba et al., 2020).

Inference of emotion situated with an "intuitive theory of mind" model involves the prediction of others' intentions, and consequently, their future behavior, from their appraised beliefs and desires (Baker et al., 2009; Saxe & Houlihan, 2017). Affective cognition, as termed by Ong and colleagues (2015), suggests the lay theories of emotion comprises of application of domain-general cognitive models to domain-specific emotion knowledge, adopting an ideal observer model approach (Baker et al., 2009). Mostly, people are attuned to infer the emotions of others (Zaki & Ochsner, 2011), and consequently, intuitive theories of emotion involve inferring agents and what are they likely to do based on the emotion cues observed, and given an agent's behavior, belief of the world and desires (de Melo et al., 2014; Saxe & Houlihan, 2017; Wu et al., 2018).

Consequently, it is also possible to model to how agents infer others' latent emotional states (i.e., affective cognition) under Bayesian inference. Interpretation of emotional expressions explains the probability of expected beliefs, desires of the person and which events that caused those emotions. As such, emotions serve rich information for both linking causal relationships, and followingly, observational causal learning, in the environment through reasoning others' mind (Ong et al., 2019; Houlihan et al., 2022; Teo et al., 2022). In short, theories of emotion relate the mental states of an agent (e.g., beliefs, desires) and the experienced event to emotion expressions of the agent, and it is through *appraisals,* which agents infer their and others' emotions. Appraisal theories suggest that agents evaluate the event outcomes according to their prior beliefs, such as their influence on prior goals and expectations (Ong et al., 2019). In addition, the same inference is applied for others as well to infer their intentions, appraisals and predict future behaviors by integrating multiple sources of cues (de Melo et al., 2014; Houlihan et al. 2022; for more information on appraisal theory within PP, see van de Cruys, 2017).

A meta-analysis study by Jie and Ong (2023) illustrates appraisal profiles for 63 emotions and affective states across 300 publications and organized into 47 dimensions (see Appendix A). Their extensive work includes associations of several emotions to a range of appraisal dimensions. Further, they also crafted a condensed list with 24 appraisals divided into further 4 sets of Likert-scale measurements in their paper: the first group is agency/control appraisals (Accountability-Circumstances, Accountability-Other, Accountability-Self, Control-Circumstances, Control-Other, Control-Self), the second is related to values, norms and goal consistency (Fairness, Goal conduciveness, Goal relevance, Normative significance (external), Normative significance (internal) and Pleasantness), the third is related to negative events and ability of coping with consequences of such events, (Challenge, Emotion-focused coping potential, Future expectancy, Loss, Problem-focused coping potential, and Threat) and fourthly, appraisals related to expectedness and critical information of an event (Attentional activity, Certainty, Effort, Expectedness, Novelty, Future predictability).

In addition to the forward appraisal inference, reverse appraisals serve as a process where agents make their decisions via facial expressions of other agents, which create an expectation of a further event (i.e., cooperation or competition) (de Melo et al., 2014b; de Melo et al., 2019). Several appraisal models have been suggested which involve goal attributions, actions and objects, from which observers are said to infer inputs (i.e., cause of an emotion) through the output (i.e., electrophysiological measures, facial expressions etc.) (Lazarus, 1991; Clore & Ortony, 2013; Ortony et al., 1988/2022). OCC model (Ortony et al., 1988/2022) is proposed as an appraisal model which distinguished 22 emotions by the psychological context of the situation. These emotions differ from each other by their outcomes, agency, and attribution dimensions and it provides a core framework for several emotion inferring models (de Silva et al., 2016; Conati & Maclaren, 2009; Joffily & Coricelli, 2013). For instance, El-Nasr and colleagues (2000) proposed an earlier computational model of emotion, namely FLAME (Fuzzy Logic Adaptive Model of Emotions) based on the appraisal theory of OCC. Even though there are also other models that adopt Lazarus' (1991; EMA, Gratch & Marsella, 2004) and Scherer's theories of appraisal (2001; WASABI, Becker-Asano & Wachsmuth, 2008), OCC is likely to be the most prevalent affect-derivation model in the literature (Marsella et al., 2010).

CHAPTER 3

COMPUTATIONAL MODELS OF EMOTION RECOGNITION

Recently, several computational models have been implemented for emotion inference, both for human and non-human agents using Bayesian networks (Ong et al., 2019; Houlihan et al., 2022; for neurocomputational models Hesp et al., 2020; for social robotics Kowalczuk & Czubenko, 2016). While some human agent-based models use neurophysiological and behavioral data to build a generative, others deploy deep learning algorithms via sensory data (e.g., audio, computer vision, video).

More recently, Houlihan and colleagues (2023) proposed a model of emotion prediction using inverse planning to infer an agent's belief and desires, which is coupled with appraisals to "reverse engineer people's intuitive theory of emotions". Working on dataset from "Split or Steal" game, their computational model simulates (1) how observers infer preferences and beliefs of a player, (2) how an outcome of an event is evaluated by the player and (3) predict player's emotion from their likely appraisals. Their model formalizes the causal structure and prior beliefs on inferring unobserved mental contents which cause unobserved behavior with Bayesian inversion of forward planning models.

Reverse appraisal is specifically implemented in (2), using “mental state representations inferred via inverse planning to generate probabilistic representations that reflect observers’ latent reasoning about how a target will appraise a situation” (Houlihan et al., 2023; p. 5). More broadly, Bayesian intuitive theory of emotion is widely used to formalize a computational model for emotion prediction in previous literature (also Ong et al., 2015; Wu et al., 2018).

Emotion recognition agent models, especially Facial Emotion Recognition (FER), are implemented in both machine learning and deep learning algorithms. Most commonly, FER models are built on CNNs (Cakmak & Develi, 2023). while other models aiming to process continuous variables as time-series include RNN, applied to LSTM variants (Jadhav & Sugandi, 2018; Ong et al., 2019). CNNs coupled with RNNs are used widely to learn emotion from continuous events. Although RNNs and LSTMs are found to be more powerful for continuous emotion recognition. Yet, those deep architecture methods require large amounts of data for a good performance. For instance, Kosti and colleagues (2020) developed an annotated emotion recognition dataset of emotional states of people in natural scenarios, EMOTions in Context (EMOTIC), combining 26 discrete emotion categories along with Valence, Arousal and Dominance dimensions. Training CNNs for achieving automatic emotion recognition, their work includes the context of scene (i.e., whole image), as well as emotional information of person in body bounding box. The dataset contains 23,571 unconstrained images with 34,320 people in it.

There also exist model-based fusions, which employ both ML methods with a generative causal model. Recently, Pei and colleagues (2022) also proposed a continuous affective state estimation model with Bayesian Filtering coupled with RNNs. Another popular model for FER inference is Hidden Markov Models (HMM), which “uses the transition probabilities between the hidden states and learns the conditional probabilities of the observations given the state of the model." (Cohen et al., 2000, p. 2) In their work, Cohen and colleagues have modeled a multilevel HMM to examine the temporal sequences of facial expressions for emotion recognition, which each expressions having an HMM model (in total, six emotions: *happy(1), angry(2) , surprise(3), disgust(4), fear(5), sad(6)*). Another widely used method for building such agents is via the optimization of HMM with Kalman filters to deal with continuous data, since it enables multimodal inferences by providing real time values. Kalman filter is applied to model the dynamics of the affective states as a state space model (SSM) (Wang et al., 2022; Pei et al., 2022).

Emotion inference from facial expressions can utilize facial action units (AU) as well, where their patterns signify valence and appraisal of the subject. One of such studies by Scherer and colleagues (2021) concludes facial movements, as indexed by AUs can be products of specific appraisal and outcomes, specifically novelty, valence and coping potential categories. Their results are aligned similarly with the claims of facial expressions being influenced by appraisal outcomes simultaneously with the emotion process. Moreover, in computational psychiatry, decision-making under uncertainty, reward prediction error and such during behavioral tasks are measured to detect subjective feelings, which can serve as clinical symptom detection, as investigated by Kao and colleagues (2023).

CHAPTER 4

BAYESIAN EMOTION INFERENCE MODELS

Apart from their emotion classification features, emotion recognition models compute appraisal computations via probabilistic frameworks for simulating the inference about latent emotion through emotion expressions (Houlihan et al., 2023). The probabilistic framework for social cognition (Griffiths et al., 2008) suggest three claims for human reasoning: 1) construction of generative models of the world, 2) these models are probabilistic (i.e., causal), and 3) the inference process involves the application of Bayes' theorem. Bayesian models have been a fundamental part of the research of how we infer others' states and predict as it proves useful in identification of beliefs and desires to predict the future action of another agent (Baker et al., 2009; Ong et al., 2019). More specifically Bayesian hierarchical generative models and similar derivatives are proposed to formalize emotion and computational appraisal via non-human agents (Saxe & Houlihan, 2017; Chung & Yoon, 2012). Hence, computational affective studies mostly have foundations in Bayesian inference (see for a comprehensive review of computational emotion models, Yanagisawa et al. 2019; Ojha et al., 2021).

Emotion and Adaptation Model (EMA) by Gratch and Marsella (2004), computes mood as a compound of the emotions experienced to alter intensities of emotion in the next appraisal round. EMA is based on Smith and Lazarus' work (1990) on "cognitive-motivational-emotive theory" as an appraisal framework. Their model includes both appraisal and coping processes, which is suggested to give rise to emotions (Gratch & Marsella, 2005). Another computational model of emotion is WASABI (Affect Simulation Architecture for Believable Interactivity) (Becker-Asano, 2008), which includes appraisal inference on a continuous 3-dimensional emotion theory of Pleasure, Arousal and Dominance (PAD) (Mahrabian & Russell, 1974). In the model, there are two parallel processes of appraisal, conscious and non-conscious. Non-conscious appraisal process contains the determination of an agent's mood, pleasure (P), and arousal (A), while conscious appraisal process involves goal conduciveness and memory update, and expectation generation. Specifically, conscious appraisal process determines the level of dominance (D), and it generates secondary emotions - the "prospect-based emotions" cluster of the OCC theory (Becker-Asano & Wachsmuth, 2010).

Ong and colleagues (2015) propose a causal model of affective cognition using Bayesian inference for emotion reasoning through expressions, inferring from emotion to outcome, or vice versa. Further, their lay theory of emotions model involves an emotional cue integration (Zaki, 2013) process, where observes draws cues from multiple domains (i.e., contextual, facial) and weigh and combine them via Bayesian principles. The results of the model points to the flexibility and optimality of a human observer over backward and forward inference, in addition to the integration of multiple cues in emotion reasoning. One interesting suggestion of the authors is a "valence dominance" effect, which involves higher weighing of positively-valenced cues compared to the negative emotional cues.

Specifically, the Bayesian model weighs reliability of emotional cues over multiple cue sources, in contrast to previous findings on single cue dominance (i.e., face versus context). A more recent work (Ong et al., 2023) furthers this argument by stating that, although people tend to rely on linguistic cues when available, they can flexibly and optimally reason about presented cues, according to their informational values across contexts. Following this, Ong and colleagues (2023) developed machine learning models, using CNN for FER, and following, classification of face is achieved by applying “fine tuning”, through addition of new NNs that maps classification categories (i.e., emotion labels) with model’s internal representations of emotion via backpropagation. Similarly, another study by Wu and colleagues (2018) proposed a causal generative inference model that shows how people relate observed events and emotional expressions to the observed agent’s beliefs and desires differ or converge . More specifically, the model contains a backward inference of beliefs and desires, given observed events, action, outcome, reactions are computed corresponding to each forward causal dependencies based on Bayes’ rule.

Recently, Doan and collaborators (2023) proposed a third-person emotion understanding motivated by appraisal theories and captured third-person appraisals in computational models. Previous work by Ong et al. (2015) models third-person appraisals in a gambling game, which builds upon people’s prediction error and expectancy of reward. de Melo and colleagues (2014) similarly modelled a computational appraisal inference in a Prisoner’s Dilemma game given the outcome and emotional displays of the opponent. Wu et al. (2018) proposed a computational model of inference of belief and desires from a short passage about an agent, an action, outcome, and their emotional expressions using reverse inference of an appraisal. Specifically, Doan and colleagues’ framework (2023) used categories of intrinsic (un)pleasantness; goals and desires; expectation-related dimensions; agency; coping ability; social dimensions; and self-consistency to model children's emotion understanding, especially in their third-person appraisals.

Regarding current report, humans reduce the multiple possibilities and outcomes into lower dimensions of emotion-related features using appraisals (Ortony et al., 1988), and an event's valence (i.e., whether it is positive or negative) impact the way and magnitude of humans' reasoning about others' emotions. People, most commonly, assess the events they have experience according to their expectations of them. Those deviations from expectations, also known as *prediction errors*, influence the specific response: people are predisposed to react more strongly to negative prediction errors rather than positive prediction errors (Ong et al., 2015). This phenomenon, referred as loss aversion (Kahneman & Tversky, 1984), along with reward expectation and prediction error, are the key features that observers deploy to theorize about others' emotions. Ong and colleagues' 2015 gambling outcome study investigates such constructs, even further, via appraisal process and backward-inference models, the researchers included the distance an agent in a better or worse outcome and how likely they were to achieve a better outcome (i.e., closeness). Their results indicate a valence dominance effect, where "positively-valenced emotional cues tended to have higher reliabilities and were weighted more so than negatively-valenced cues (Ong et al., 2015; p. 156)." Another interesting claim of the study is that observers also rely on same set of features while assessing their own emotional states, leading to the conclusion that people do not have "privileged" access to their true emotion that differs from others'.

CHAPTER 5

RELATIONSHIP BETWEEN EMOTIONS AND WHEEL-SPINNING GAME OUTCOMES

This study aims to examine how emotion of a third person (worker) changes with respect to the monetary outcome of the wheel-spinning which consists of three different amounts in dollars. Using the probability of occurrence for that certain outcome, in addition to the closeness or distance to the bounding outcomes.

This report uses the dataset from Ong et al. (2015), specifically from their first experiment. Two different Bayesian models have been implemented to (1) see the relationship of a certain emotion to the amounts won by workers (in $), the probability of landing on the outcome slice in the wheel, and the location in which the pointer has landed between the boundaries of the outcome slice; (2) investigate the emotion effect for each worker, including the outcome amount, the outcome probability, and the angle proportion of the outcome boundary. The models have been built in PyMC, a library for Python for probabilistic programming, and uses MCMC. The models describe the influence of the amount won, the probability of outcome and the angle proportion.

5.1 Dataset

The dataset consists of ID for each worker, totaling to 100 workers, where they have spun the wheel 10 times each. Each of the 18 pre-generated wheels is composed of three different payoff amounts with different probabilities according to their proportion of the wheel. Each worker has won one of these three payoffs, denoted as win amount, and its corresponding probability. In addition, dataset also includes the position of pointer for each wheel slice of amount won, angle proportion, denoted as a formalized value between 0 and 1, 0.5 being the middle of the slice the pointer has landed finally. Specifically, a value closer to 0 indicates the closeness of the edge of the previous slice for payoff, and vice versa for the values near 1. Each outcome for worker has been assessed for each likely emotion outcome for the winner by evaluators in a 9-point Likert scale, namely for emotions 'happy', 'sad', 'anger', 'surprise', 'disgust', 'fear', 'content', 'disappointment'. The first six emotions are adapted from the "basic emotions theory" of Ekman and colleagues (1982), and the last two were included by researchers to "capture emotion concepts related to counterfactual comparisons with outcomes that could have, but did not occur" (Gilovich & Medvec, 1995; Sweeny & Vohs, 2012, cited from Ong et al., 2015, p. 145).

5.2 Descriptive Data Analysis

In the current research, each emotion has been assigned values for every trial and the maximum valued emotion has been defined as another variable (N_{happy} = 642, $N_{disappointment}$ = 166, $N_{surprise}$ = 59, $N_{content}$ = 56, N_{sad} = 55, N_{anger} = 15, $N_{disgust}$ = 5, N_{fear} = 2), with a total of 1000 evaluations. For each emotion rating, the mean and standard deviation values are as follows: M_{happy} = 6.21, SD_{happy} = 2.44; $M_{disappointment}$ = 3.61, $SD_{disappointment}$ =

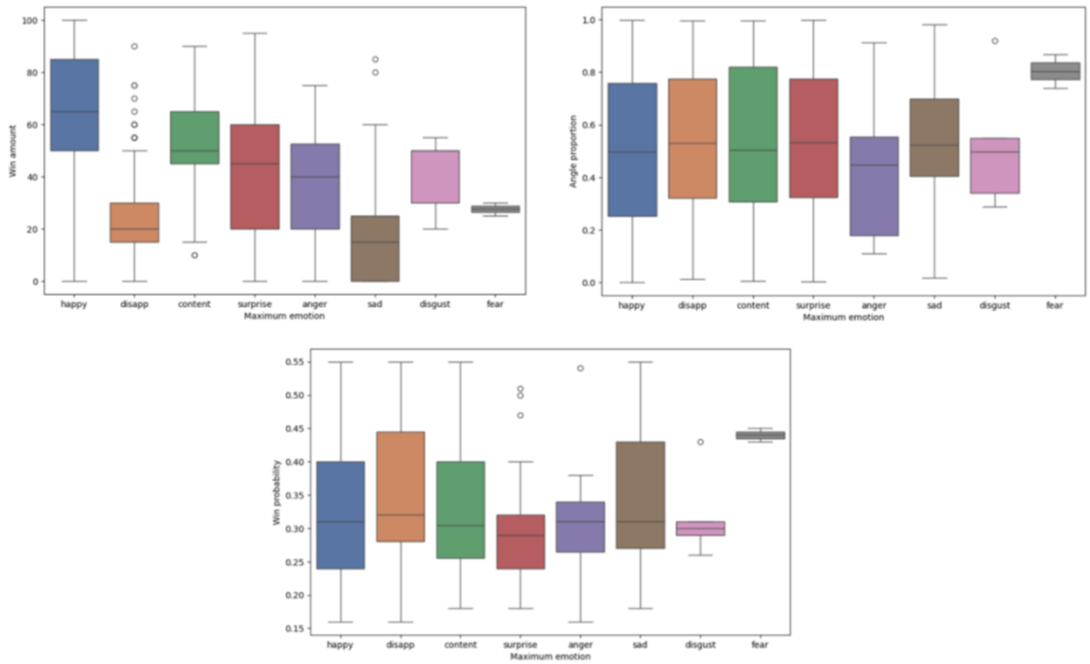

Figure 1. Plot of maximum rated emotions and win amount, angle proportion, and win probability.

2.66; $M_{surprise} = 5.29$, $SD_{surprise} = 2.31$; $M_{content} = 5.16$, $SD_{content} = 2.62$; $M_{sad} = 2.65$, $SD_{sad} = 2.18$; $M_{anger} = 2.1$, $SD_{anger} = 1.83$; $M_{disgust} = 1.95$, $SD_{disgust} = 1.66$; $M_{fear} = 1.34$, $SD_{fear} = 1.02$. The average amount won by participants across every trial is 53.29 $ (SD = 8.72), and the average win probability is 0.33 (SD=0.03). The angle proportion of payoff won by participant is 0.51 in average, with a standard deviation of 0.28.

In the correlation analyses, some strong relationships have been found between the emotions. Namely, “happy” emotion is found to be associated with “content” emotion ($\rho = 0.62$), and conversely, a negative correlation is present between “sad” ($\rho = -0.71$), and “disappointed” ($\rho = -0.79$). Moreover, there are also moderate relationships between “happy” and other emotions: “anger” ($\rho = -0.57$), and “disgust” ($\rho = -0.54$) emotions showed a negative association, while “surprise” is moderately associated with “happy” emotion ($\rho = 0.45$). “Sad” emotion also displayed significant relationship with “disappointed” ($\rho = 0.74$), “anger” ($\rho = 0.76$), and “disgust” ($\rho = 0.64$). On the other hand, “anger” and “disgust” showed also high association ($\rho = 0.74$) in the analyses. It is also noteworthy that “disappointment” and “disgust” has a relatively strong relationship ($\rho = 0.63$).

5.3 Bayesian Models

The Bayesian emotion model has been built to investigate the effects of amount won, the probability of occurrence of amount won, and the angle of pointer which landed in the amount of won on the emotion experienced. For each of 8 emotions, a different Bayesian model has been implemented. In every model, win amount, win probability, and angle proportion has been added as continuous variables to the equation:

$$\mu_{Emotion} = \alpha + \beta_{winAmount} \text{ x win} + \beta_{winProbability} \text{ x winProb} + \beta_{angleProportion} \text{ x angleProp}$$

The priors were defined for each coefficient, winAmount, winProbability and angleProportion:

$$\beta \sim \text{Normal}(0,10)$$

$$\sigma \sim \text{Uniform}(0,10)$$

In each emotion model, the likelihood distribution is defined as:

$$Emotion \sim \text{Normal}(\mu_{Emotion}, \sigma)$$

Additional Bayesian models have been built to see the effects mentioned above on a worker basis. These models are constructed with the same priors for β coefficients as the previous models, except, for each emotion,

$$Emotion \sim \text{Normal}(\bar{\mu}, \sigma)$$

$$\bar{\mu} \sim \text{Normal}(0,1)$$

$$\sigma \sim \text{Exponential}(1)$$

Also, emotion ratings have been standardized, and the equation for this new model has been defined as below:

$$\mu_{Emotion} = \bar{\alpha}_{[worker]} + \beta_{winAmount} \times win + \beta_{winProbability} \times winProb + \beta_{angleProportion} \times angleProp$$

5.4. Results

5.4.1. Relationship of happy with win amount, win probability and angle proportion

In happy emotion model, the posterior of α has a mean of 2.84, with a value of 0.22 standard deviation. 97% HDI values are found to be α = [2.43, 3.26]. The posterior of σ value has a mean of 1.52 and standard deviation of 0.04. For $\beta_{winAmount}$, a mean value of 0.07 with a standard deviation of 0.002 is reported, while the 97% HDI interval is equal to [0.066, 0.072]. $\beta_{winProbability}$ coefficient (M = -0.53, SD = 0.45) has yielded 97% HDI of [-1.44, 0.40], while $\beta_{angleProportion}$ (M=-0.25, SD = 0.17) showed 97% HDI equal to [-0.55, 0.08].

According to the model results, happiness is positively influenced by the amount of winnings. In addition, the angle proportion and win probability seems to negatively affect happiness of worker, yet, since their 97% HDIs span both negative and positive values, there is uncertainty about their direction of effect.

In the worker-specific model of happiness emotion, $\bar{\alpha}$ value (M = -1.36, SD = 0.09) has 97% HDI equal to [-1.53, -1.19], while σ (M = 0.54, SD = 0.01) showed 97% HDI of [0.51, 0.56]. The coefficients $\beta_{winAmount}$ (M = 0.028, SD = 0.001) , $\beta_{winProbability}$, (M = -0.32, SD = 0.18) and $\beta_{angleProportion}$ (M = -0.02, SD = 0.07) have 97% HDIs equal to [0.027, 0.029], [-0.66, 0.03], and [-0.15, 0.09], respectively.

The result of worker-specific model for happiness shows a decrease in happy emotion in workers when the angle proportion increases, although the direction of the

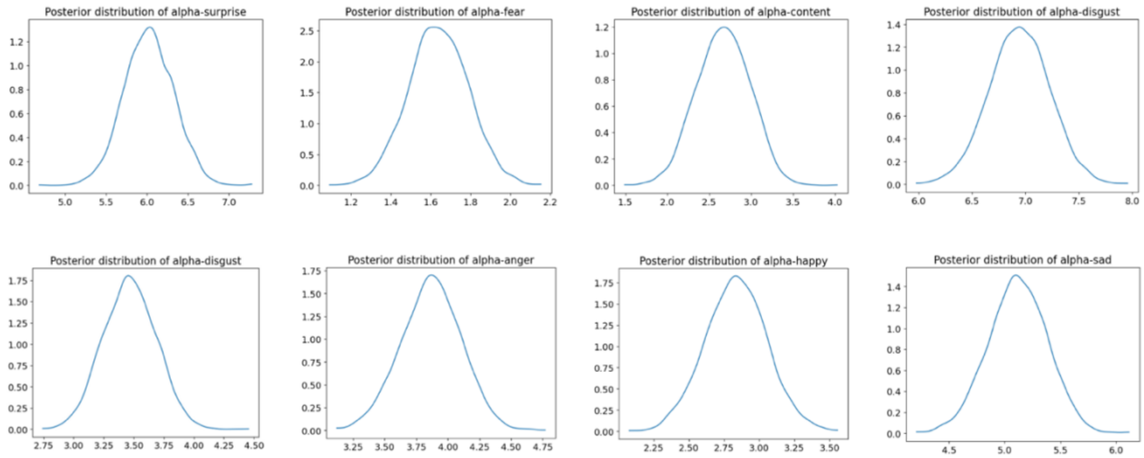

Figure 2. Posterior plots for α from the Bayesian emotion model.

relation is not specific. Moreover, workers are less happy by the influence of increasing the winning probability, yet this impact is also uncertain. Lastly, the happiness of workers is increased when the amount won by the worker is greater, and it has a credible direction, although somewhat low in magnitude.

5.4.2. Relationship of sad with win amount, win probability and angle proportion

α value for sad emotion model has a mean of 5.12 and a standard deviation of 0.27, while 97% HDI is found [4.63, 5.63]. σ has a mean of 1.79 and a standard deviation of 0.04, with 97% HDI equal to [1.72, 1.87]. The coefficient $\beta_{winAmount}$ (M= -0.045, SD = 0.002) has 97% HDI equal to [-0.049, -0.042]. For $\beta_{winProbability}$, (M = -0.47, SD = 0.60) 97% HDI values were [-1.56, 0.70]. Lastly, $\beta_{angleProportion}$ (M = 0.19, SD = 0.21) yielded 97% HDI as [-0.21, 0.59].

While a higher intercept value points that workers are more likely to be sad as a baseline mood, the higher the amount won and the probability of winning, workers are less likely to experience sad emotion. Conversely, angle proportion influences sadness positively, yet both for win probability and angle proportion, these relationships are not certain due to wide range of their 97% HDIs.

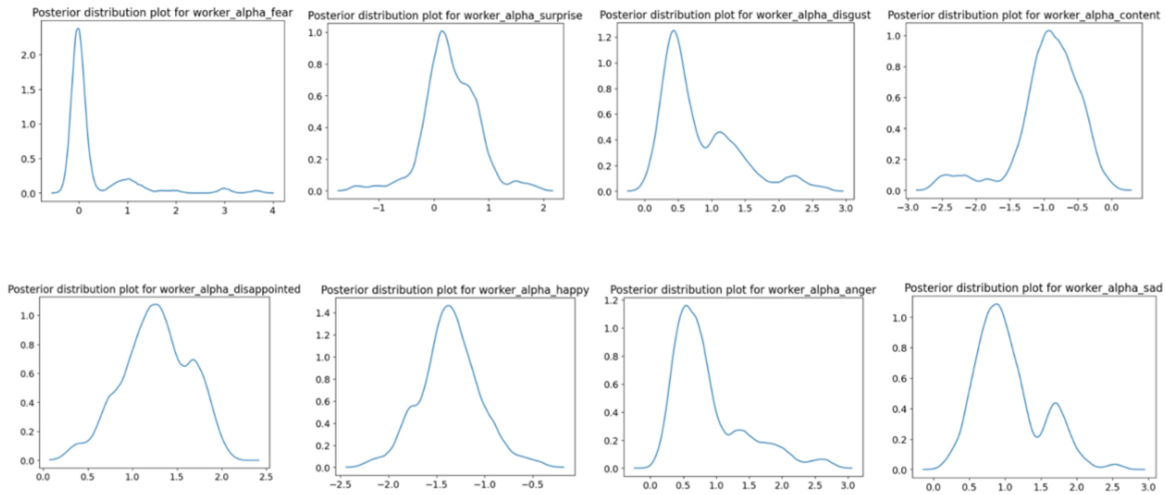

Figure 3. Posterior plots for $\bar{\alpha}$ from worker-specific Bayesian emotion model.

In the worker-based model of sadness emotion, $\bar{\alpha}$ (M = 1.04, SD = 0.13) has 97% HDI = 0.81, 1.28], while σ (M = 0.68, SD = 0.02) has a 97% HDI equal to [0.65, 0.71]. This Bayesian emotion model show that the coefficients of $\beta_{winAmount}$ (M = -0.02, SD = 0.001),

$\beta_{winProbability}$, (M = -0.09, SD = 0.22) and $\beta_{angleProportion}$ (M = 0.04, SD = 0.08) have 97% HDIs of [-0.02, -0.018], [-0.49, 0.35], and [-0.10, 0.19] respectively.

Worker-specific model result implies a low negative influence of the increase of winnings amount, and since the 97% HDI interval spans from negative to positive, this result is not credible. Similarly, as win probability increase, workers tend to feel less sad, nevertheless, this relationship is also uncertain. In contrast, the angle proportion increase leads to sadness in workers, yet this relationship is also not exactly credible.

5.4.3. Relationship of anger with win amount, win probability and angle proportion

In the anger model, α (M =3.87, SD = 0.24) had a 97% HDI interval of [3.41, 4.31], while σ (M = 1.66, SD = 0.04) had a 97% HDI interval equal to [1.59, 1.73]. The

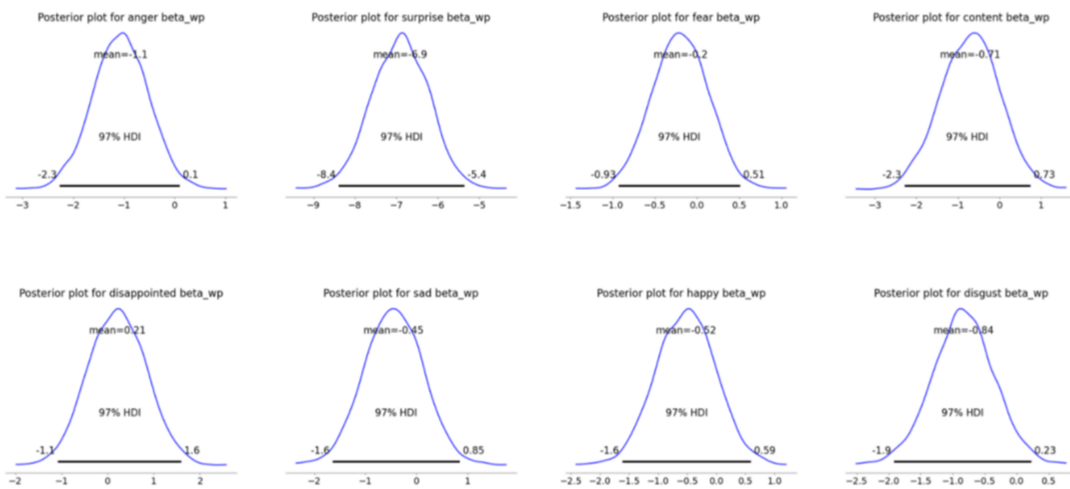

Figure 4. Posterior plots for $\beta_{winProbability}$ from the Bayesian emotion model.

$\beta_{winAmount}$ coefficient (M= -0.03, SD = 0.002) has 97% HDI equal to [-0.032, -0.025]. 97% HDI values of $\beta_{winProbability}$ (M = -1.05, SD = 0.54) were [-2.04, 0.02], and $\beta_{angleProportion}$ (M = 0.15, SD = 0.19) showed 97% HDI of [-0.18, 0.52].

For anger emotion, the participants are predicted to be less angry in proportion to increasing outcome amount, and win probability, although the relationship is not certain between the probability of winning. Similarly, the angle proportion is positively related to anger, suggesting that workers get angrier while pointer lands further from the boundary, yet there is also uncertainty about the direction of the relationship.

In the worker-based model of anger, $\bar{\alpha}$ (M = 0.87, SD = 0.13) resulted in 97% HDI = [0.63, 1.11], while σ value (M = 0.71, SD = 0.02) has a 97% HDI equal to [0.68, 0.74]. The results of worker-specific model show that, for anger emotion, $\beta_{winAmount}$ (M = -0.01, SD = 0.001), $\beta_{winProbability}$, (M = -0.46, SD = 0.24) and $\beta_{angleProportion}$ (M = 0.01, SD = 0.08) have 97% HDIs of [-0.015, -0.012], [-0.90, -0.01], and [-0.14, 0.17], respectively.

The worker-specific model suggests that workers have higher anger scores as the angle probability increases in low magnitudes, yet 97% HDI interval shows no exact

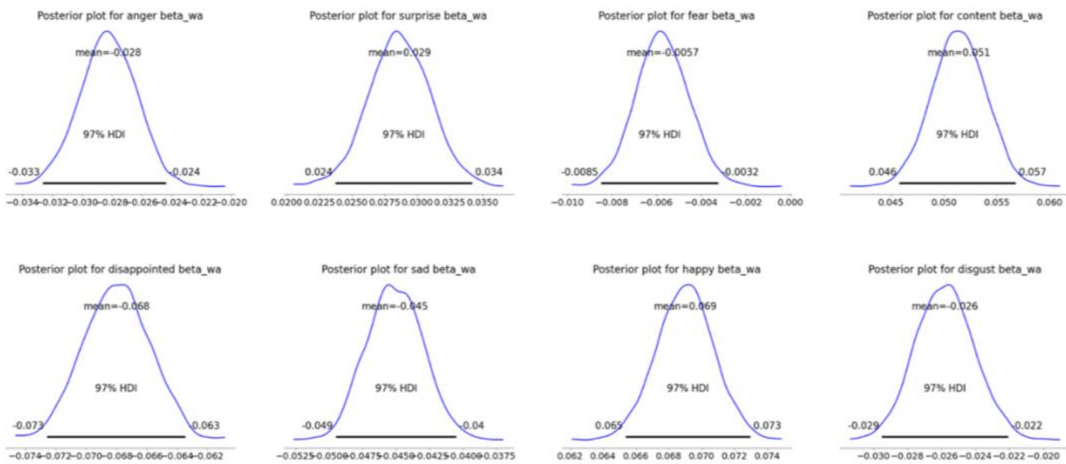

Figure 5. Posterior plots for $\beta_{winAmount}$ from the Bayesian emotion model.

direction for this relationship. In contrast, workers are less angry when the amount they have won increases, and this relationship is supported in a negative direction. Similarly, as the win probability is higher, anger is negatively influenced, but this direction is not backed up with a credible interval.

5.4.4. Relationship of fear with win amount, win probability and angle proportion

In this Bayesian model, α (M = 1.64, SD = 0.15) 97% HDI values are [1.37, 1.95]. σ value (M = 1.01, SD = 0.02), yielded 97% HDI scores of [0.97, 1.06]. The coefficient $\beta_{winAmount}$ (M = -0.006, SD = 0.001) has the 97% HDI interval equal to [-0.008, -0.004]. $\beta_{\text{winProbability}}$ coefficient (M = -0.20, SD = 0.34) has a 97% HDI of [-0.81, 0.43], and $\beta_{\text{angleProportion}}$ (M=0.13, SD = 0.11) showed 97% HDI of [-0.08, 0.35].

Although win amount increases in opposite to fear, it has a lower influence on the emotion experience than the probability of winning. The angle proportion, on the other hand, leads worker to fear as it increases, yet, similar as the probability of winning, 97% HDI values indicate an ambiguity in the direction of their relationship with fear emotion.

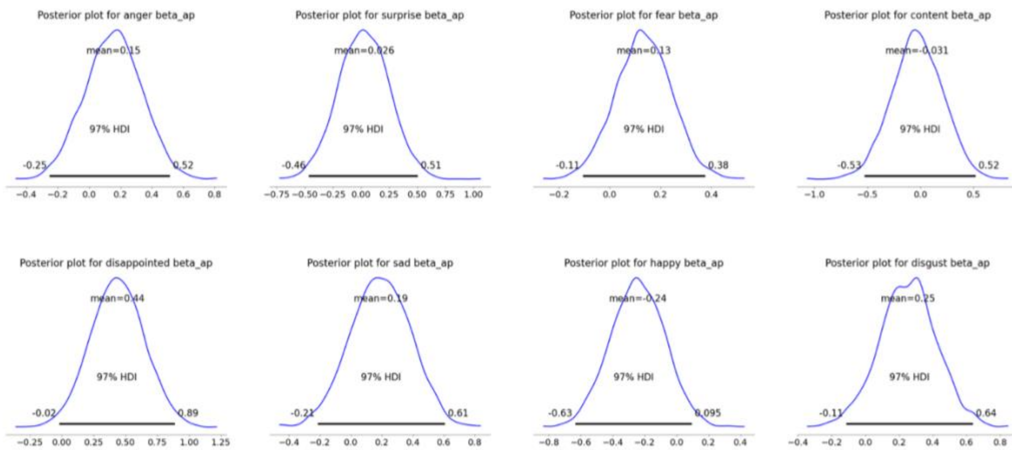

Figure 6. Posterior plots for $\beta_{angleProportion}$ from the Bayesian emotion model.

In the worker-based model of fear, $\bar{\alpha}$ (M = 0.25, SD = 0.12) has 97% HDI = [0.04, 0.49], while σ (M = 0.69, SD = 0.02) has a 97% HDI of [0.66, 0.72]. For fear, the coefficients $\beta_{winAmount}$ (M = -0.005, SD = 0.001), $\beta_{winProbability}$, (M = -0.1, SD = 0.24) and $\beta_{angleProportion}$ (M = 0.08, SD = 0.08) have 97% HDIs of [-0.006, -0.003], [-0.54, 0.33], and [-0.06, 0.24], respectively.

The worker-specific model of fear shows that higher win probability is associated with less fear experienced by workers, yet the direction of impact is not supported. The increase in winnings lead to less fearful workers that is significant. In turn, the angle proportion increase causes fear, although the interval is not credible to detect a specific direction.

5.4.5. Relationship of surprise with win amount, win probability and angle proportion

For α value in surprise model (M = 6.02, SD = 0.30), 97% HDI is [5.50, 6.63], while σ (M = 2.04, SD = 0.05) has 97% HDI of [1.95, 2.13]. The $\beta_{winAmount}$ coefficient (M= 0.03, SD = 0.002) showed 97% HDI of [0.02, 0.03].

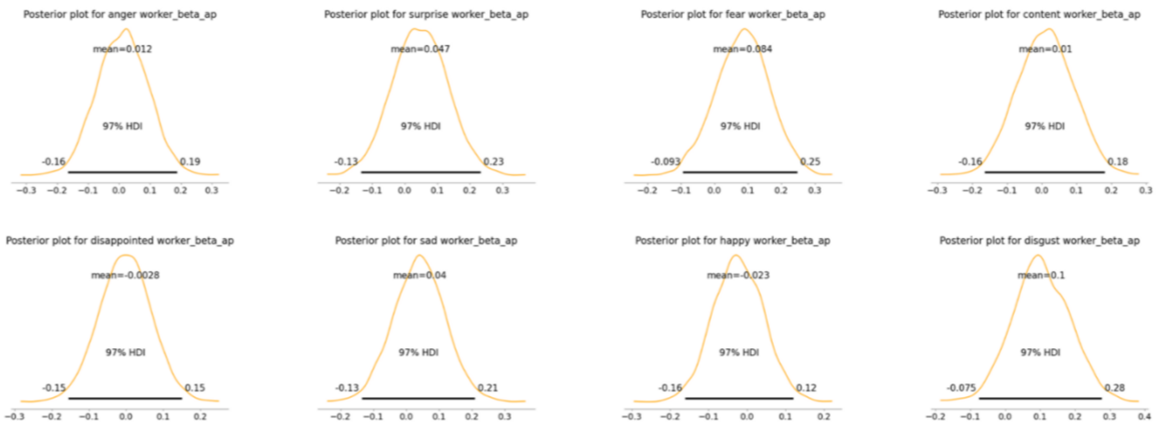

Figure 7. Posterior plots for $\beta_{angleProportion}$ from the worker-specific Bayesian emotion model.

$\beta_{winProbability}$ (M = -6.92, SD = 0.69) has 97% HDI equal to [-8.15, -5.60], and $\beta_{angleProportion}$ (M = 0.03, SD = 0.23) has 97% HDI of [-0.40, 0.47].

In Bayesian surprise model, while surprisal is a probable baseline emotion, the amount won increases this emotion with a high confidence level. As the win probability increases, the workers are less surprised, which is great in magnitude and certainty. Yet, angle proportion of the outcome slice reflects an increase in surprise, the direction of the relationship is uncertain, as inferred from the wide-spanning 97% HDI.

In the worker-based model, $\bar{\alpha}$ (M = 0.30, SD = 0.13) has a 97% HDI of [0.05, 0.54], and σ (M = 0.72, SD = 0.02) has a 97% HDI of [0.69, 0.75]. The worker-specific coefficients of surprisal, $\beta_{winAmount}$ (M = 0.01, SD = 0.001), $\beta_{winProbability}$ (M = -3.08, SD = 0.25) and $\beta_{angleProportion}$ (M = 0.05, SD = 0.09), have 97% HDIs that are [0.011, 0.015], [-3.54, -2.61], and [-0.11, 0.21], respectively.

The increase at the angle proportion is linked to workers experiencing surprise, although the interval of this variable does not allow for a specific direction. Yet, it is apparent that the greater the monetary outcome is, the more likely the worker to be

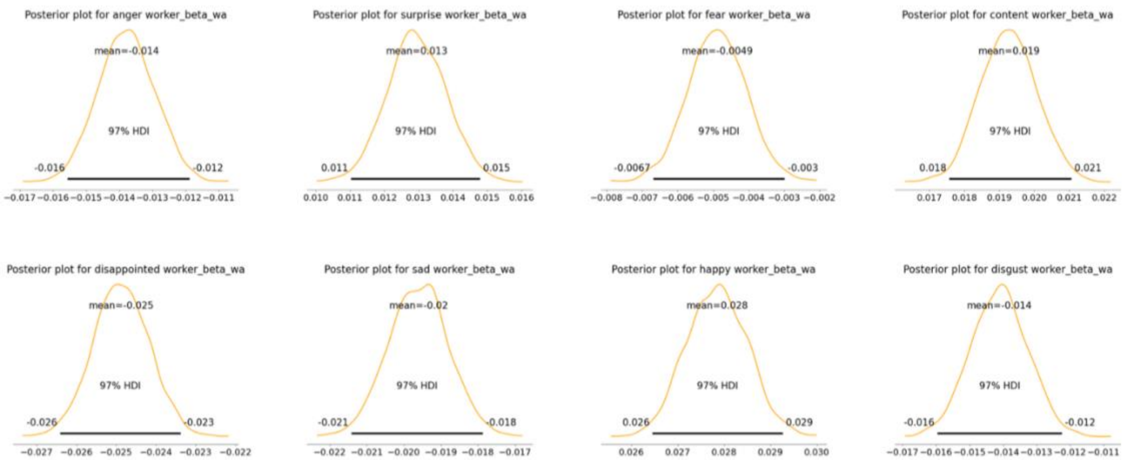

Figure 8. Posterior plots for $\beta_{winAmount}$ from the worker-specific Bayesian emotion model.

surprised. Similarly, in great magnitudes, higher win probability induces less surprise in workers in a clear direction.

5.4.6. Relationship of disgust with win amount, win probability and angle proportion

In this model, α value (M = 3.47, SD = 0.22) has 97% HDI equal to [3.06, 3.87], while σ (M = 1.51, SD = 0.03) has 97% HDI scores of [1.44, 1.57].The $\beta_{winAmount}$ coefficient (M = -0.03, SD = 0.002) has the 97% HDI interval equal to [-0.029, -0.023]. Meanwhile, $\beta_{\text{winProbability}}$ coefficient (M = -0.83, SD = 0.48) has a 97% HDI of [-1.80, 0.006], and $\beta_{\text{angleProportion}}$ (M=0.25, SD = 0.17) showed 97% HDI of [-0.07, 0.57].

The workers tend to feel less disgust when the amount they have won is greater, and in accordance, increasing win probability has a negative influence on feeling of disgust, but with less certainty. In contrary, increasing angle proportions seem to induce disgust in workers, yet there is not a strong indication of direction for this effect.

In the worker-based model of disgust, $\bar{\alpha}$ (M = 0.80, SD = 0.13) has a 97% HDI of [0.57, 1.03], while σ (M = 0.69, SD = 0.02) has a 97% HDI of [0.67, 0.72]. The worker-specific coefficients $\beta_{winAmount}$ (M = -0.01, SD = 0.001), $\beta_{winProbability}$, (M = -0.39, SD =

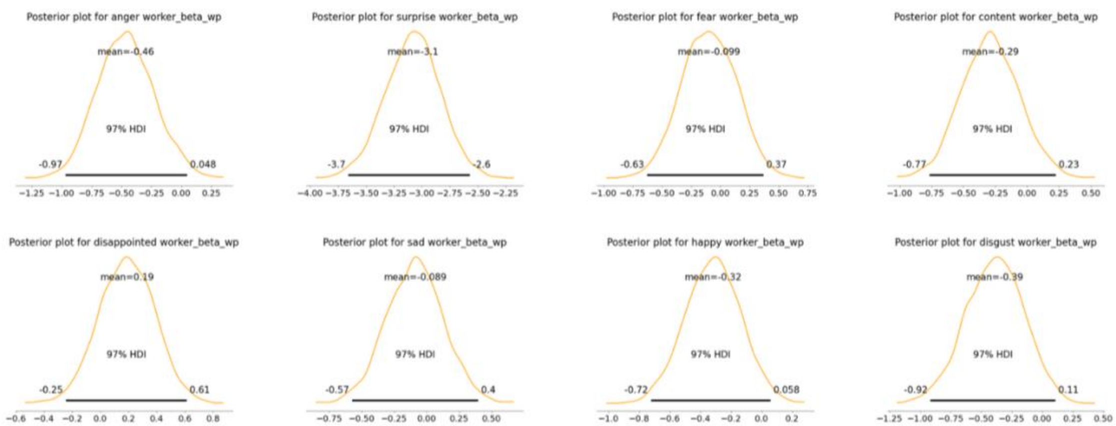

Figure 9. Posterior plots for $\beta_{winProbability}$ from the worker-specific Bayesian emotion model.

0.24) and $\beta_{angleProportion}$ (M = 0.10, SD = 0.08) have 97% HDIs equal to [-0.016, -0.013], [-0.84, 0.06], and [-0.05, 0.26], respectively.

The worker-specific results for feeling of disgust point to an uncertain, negative influence of the win probability in workers. As the amount of monetary prize increase, workers are feeling less disgust, with the direction being the exact. On the other hand, as the pointer is further from the edge, the workers tend to feel more disgusted, yet it is not clear that this relationship is established.

5.4.7. Relationship of content with win amount, win probability and angle proportion

The α (M = 2.67, SD = 0.32) of this model has 97% HDI values of [2.09, 3.25], and σ value (M = 2.20, SD = 0.05) has 97% HDI scores equal to [2.11, 2.30]. The $\beta_{winAmount}$ coefficient (M = -0.006, SD = 0.001) has the 97% HDI interval of [-0.008, -0.004]. $\beta_{\text{winProbability}}$ (M = -0.70, SD = 0.73) has a 97% HDI of [-2.01, 0.69], and $\beta_{\text{angleProportion}}$ coefficient (M=-0.03, SD = 0.24) has yielded a 97% HDI equal to [-0.48, 0.41].

In the worker-based model of contentment, $\bar{\alpha}$ (M = -0.93, SD = 0.12) has 97% HDI = [-1.15, -0.70], while σ (M = 0.67, SD = 0.02) has a 97% HDI of [0.64, 0.69]. The worker-specific results show that, for contentment, $\beta_{winAmount}$ (M = 0.02, SD = 0.001), $\beta_{winProbability}$, (M = -0.29, SD = 0.23) and $\beta_{angleProportion}$ (M = 0.01, SD = 0.08) have 97% HDIs of [0.018, 0.021], [-0.70, 0.14], and [-0.15, 0.15], respectively.

For contentment emotion, the results on worker-basis Bayesian model show that, greater angle proportion makes workers to feel more content, yet the direction of this association is not clear. Higher win probabilities, on the other hand, decrease the contentment of the worker, which might be due to a wide interval that includes both negative and positive values. For the amount of reward, the workers are clearly more content when winning more money. Surprisingly, the workers feel less content when the amount they have won is increased, and the direction of this effect is more credible than the probability of winning and the proportion of angle. Conversely, the last two variables have a reverse relationship with emotion of contentment.

5.4.8. Relationship of disappointment with win amount, win probability and angle proportion

In this Bayesian emotion model, for α (M = 6.94, SD = 0.28), 97% HDI values are equal to [6.38, 7.44], and σ value (M = 1.89, SD = 0.04) has 97% HDI of [1.81, 1.97]. For $\beta_{winAmount}$ (M = -0.07, SD = 0.002), the 97% HDI interval is equal to [-0.07, -0.06]. $\beta_{\text{winProbability}}$ coefficient (M = 0.21, SD = 0.65) has 97% HDI of [-1.00, 1.43], and the $\beta_{\text{angleProportion}}$ coefficient (M = 0.44, SD = 0.21) showed 97% HDI equal to [0.05, 0.83].

The results indicate high baseline levels for disappointment and increasing angle proportion and a higher probability of winning lead to more disappointment. To note, for the probability variable, the direction of relationship is uncertain . As for the amount won the increase in winnings points that the workers are less likely to experience disappointment.

In the worker-based model, $\bar{\alpha}$ value (M = 1.26, SD = 0.11) resulted in 97% HDI = [1.05, 1.45], and σ (M = 0.60, SD = 0.01) has a 97% HDI of [0.57, 0.62]. The coefficients of the worker-specific disappointment model, $\beta_{winAmount}$ (M = -0.03, SD = 0.001) , $\beta_{winProbability}$, (M = 0.19, SD = 0.20) and $\beta_{angleProportion}$ (M = -0.003, SD = 0.07) have 97% HDIs equal to [-0.026, -0.023], [-0.20, 0.54], and [-0.14, 0.12], respectively.

The disappointment score is lower for workers that gains greater monetary outcomes, and this influence is specifically directed. In contrast, they feel more disappointed when the winning probability is greater, but the results do not provide a direction for this relationship. Similarly, while the disappointment is lower for workers when the angle proportion is increasing, no specific direction could be determined by the interval.

CHAPTER 6

DISCUSSION AND CONCLUSION

In this study, the influence of the monetary reward from a wheel-spin game on the emotions experienced by the players of the game is investigated by Bayesian models over a dataset by Ong and colleagues (2015). The probabilistic model for eight emotions (happy, anger, sad, disappointed, disgust, content, fear and surprise) has been deployed, including the probability of the outcome won out of all three monetary prizes, and the angle proportion of the wheel section in which the pointer lands is included. These variables have been studied to see whether the closeness of another reward (measured by the angle proportion), and the odds of winning a certain amount have a significant impact on the player's emotions. Two Bayesian emotion models in PyMC have been built to research the overall emotional outcomes, and the emotion for each player of the wheel-spin game.

The significant findings for the first emotion model conclude that, in overall, happiness increases as the monetary reward increase. Conversely, sadness, fear, disgust is less experienced when the amount of winnings increases. Players tend to show high levels of surprise, and their levels of surprisal immensely decrease as their probability for winning gets higher, in opposite to the increase in the monetary outcome, in which a higher reward tremendously impacts the surprise of the player positively. On the other hand, the disappointment has a high valence among all outcomes, and as expected, higher monetary rewards lead to less disappointment. Another interesting finding is that the contentment of the workers decreases as the outcome amount is greater, which is not expected.

The second model, which accounts for individual differences in the outcome, probability, and proportion impacts on their emotional responses, conclude such significant results that, as it is also observed in the first model, the amount of the monetary prize has great influence on the workers' happiness, contentment, and their surprisal. Similarly, workers feel less angry, less fearful, less disgusted, and less disappointed when they win more money. In addition, as expected, workers are less surprised when they have won the outcome with higher probabilities of occurring.

Some results, although many of them interesting, did not converge on a specific direction of influence. For instance, in the first model, uncertain directions are found for all emotions, except for surprisal, for their relationship with winning probability of the outcome. This might be due to obtaining a lower outcome amount out of all three possible outcomes, since some relationships are expectedly negative (i.e., sad, anger, fear, disgust), and some were not (i.e., content, happy, disappointed). Similarly, in worker-specific model, higher probability of winning showed a positive effect on disappointment, and a negative effect on happiness & contentment. Nevertheless, although a reverse relation has been observed for anger, sadness, disgust and fear between win probability, the findings were not supported of their direction.

For the angle proportion variable, across the both models, due to wide-span of their intervals, no specific direction of influence was found for any emotion. Hence, both the overall model and the individual model does not suggest any evidence for the effect of angle proportion. Yet, the direction of influence is certain for the monetary outcome changes across every emotion for both models.

Compared to the results of Experiment 1 from Ong and colleagues (2015), similarly, the amount won is a significant predictor of emotional responses of the workers. Moreover, surprisal is also lower when the probability of outcome won increases, which is consistent with their results. The main difference in the current result is that, interestingly, the contentment significantly decreases as the amount of money won increase, which is not the matter in Ong and colleagues' research. Their study does not conclude a significant finding for the amount of winning on fear, disgust and anger emotions, and this is also the case in this report, except for fear and disgust.

To conclude, the current study has produced significant findings for the influence of monetary reward on the emotion of players of wheel-spinning game. While this study failed to capture the impact of the portion of section of the wheel which the pointer has landed, the current research also provides an invaluable insight into the prediction of emotional outcomes from monetary prize received and, to an extent, the emotional response to having a higher probability in a contest where, although varying, the outcome never involves a loss.

In a greater perspective, this study shows the universality of the emotion reasoning, and more importantly, the accuracy of prediction of others' emotions in response to the outcome of an event. In short, computation over affective cognition provides valuable highlights to understand how people reason about others' emotions. Moving from this point, future behavior prediction is also theoretically possible to generate for any agent, which could improve decision-making in many contexts.

REFERENCES

Albohn, D., Brandenburg, J., & Adams, R. (2019). *Perceiving Emotion in the "Neutral" Face: A Powerful Mechanism of Person Perception* (pp. 25–47). https://doi.org/10.1007/978-3-030-32968-6_3

Asaba, M., Wu, Y., Carrillo, B., & Gweon, H. (2020). *You're surprised at her success? Inferring competence from emotional responses to performance outcomes*.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*(3), 329-349.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, *16*(7), 419-429.

Becker-Asano, C. (2008). WASABI: Affect simulation for agents with believable interactivity, 39. IOS Press.

Becker-Asano, C., & Wachsmuth, I. (2010). Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems, 20*(1), 32–49. https://doi.org/10.1007/s10458-009-9094-9

Cakmak, B., & Develi, I. (2023). st International Conference on Frontiers in Academic Research Convolutional Neural Network-Based Classification of Facial Emotional Expressions and Computational Complexity Analysis. https://www.icfarconf.com/

Chung, S. Y., & Yoon, H. J. (2012, October). Affective classification using Bayesian classifier and supervised learning. In *2012 12th International Conference on Control, Automation and Systems* (pp. 1768-1771). IEEE.

Clark, A. (2013). Expecting the world: Perception, prediction, and the origins of human knowledge. *The Journal of Philosophy*, *110*(9), 469-496.

Clark, A. (2015). Radical predictive processing. *The Southern Journal of Philosophy*, *53*, 3-27.

Clore G. L., & Ortony, A. (2013). Psychological construction in the OCC model of emotion. *Emotion Review, 5*(4), 335–343. https://doi.org/10.1177/1754073913489751

Conati, C., & MacLaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, *19*(3), 267–303. https://doi.org/10.1007/s11257-009-9062-8

Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2021). Sixteen facial expressions occur in similar contexts worldwide. *Nature*, *589*(7841), 251–257. https://doi.org/10.1038/s41586-020-3037-7

da Silva, F. N., Werhli, A. V., & Adamatti, D. F. (2016). Using Bayesian Networks to Structure the OCC Emotions Model. *Journal of Intelligent Computing, 7*(4), 157.

Darwin, C. (1872). On the origin of species.

de Melo, C. M., Carnevale, P. J., Read, S. J., & Gratch, J. (2014a). Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology, 106*(1), 73.

de Melo, C. M., Gratch, J., & Carnevale, P. J. (2014b). Humans versus computers: Impact of emotion expressions on people's decision making. *IEEE Transactions on Affective Computing*, *6*(2), 127-136.

de Melo, C. M., Marsella, S., & Gratch, J. (2019). Human cooperation when acting through autonomous machines. *Proceedings of the National Academy of Sciences*, *116*(9), 3482-3487.

Doan, T., Ong, D. C., & Wu, Y. (2023). *Emotion understanding as third-person appraisals: Integrating appraisal theories with*.

El-Nasr, M. S., Yen, J., & Ioerger, T. R. (2000). Flame—fuzzy logic adaptive model of emotions. *Autonomous Agents and Multi-agent systems, 3*, 219-257.

Ekman, P. (1973). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review* (pp. 169–222). New York: Academic.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? In Emotion in the human face (pp. 39–55). New York: Cambridge University Press.

Frijda, N. H., & Scherer, K. R. (2009). Emotion definitions (psychological perspectives). The Oxford companion to emotion and the affective sciences, 142-144.

Gibson, J. (1979). An ecological approach to visual perception.

Gilovich, T., & Medvec, V. H. (1995). The experience of regret: what, when, and why. *Psychological Review*, *102*(2), 379.

Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research, 5*(4), 269-306.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory.

Hesp, C., Smith, R., Parr, T., Allen, M., Friston, K. J., & Ramstead, M. J. (2021). Deeply felt affect: The emergence of valence in deep active inference. *Neural Computation*, *33*(2), 398-446.

Houlihan, S. D., Kleiman-Weiner, M., Hewitt, L. B., Tenenbaum, J. B., & Saxe, R. (2023). Emotion prediction as computation over a generative theory of mind. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *381*(2251), 20220047. https://doi.org/10.1098/rsta.2022.0047

Houlihan, S. D., Ong, D., Cusimano, M., & Saxe, R. (2022). Reasoning about the antecedents of emotions: Bayesian causal inference over an intuitive theory of mind. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44).

James, W., & Lange, C. (1922). The emotions (Vol. 1).

Jie, G. C. Y. Z., & Ong, D. C. (2023). *A meta-analytic review of the associations between cognitive appraisals and emotions* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/ystxc

Joffily, M., & Coricelli, G. (2013). Emotional Valence and the Free-Energy Principle. *PLoS Computational Biology, 9*(6). https://doi.org/10.1371/journal.pcbi.1003094

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*(4), 341.

Kao, C.-H., Feng, G. W., Hur, J. K., Jarvis, H., & Rutledge, R. B. (2023). Computational models of subjective feelings in psychiatry. *Neuroscience & Biobehavioral Reviews*, *145*, 105008. https://doi.org/10.1016/j.neubiorev.2022.105008

Kosti, R., Alvarez, J., Recasens, A., & Lapedriza, A. (2019). Context Based Emotion Recognition using EMOTIC Dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. https://doi.org/10.1109/TPAMI.2019.2916866

Kowalczuk, Z., & Czubenko, M. (2016). Computational approaches to modeling artificial emotion–an overview of the proposed solutions. *Frontiers in Robotics and AI*, *3*, 21.

Griffiths, T., Kemp, C., & B Tenenbaum, J. (2008). Bayesian models of cognition.

Lazarus, R. S. (1991). Cognition and motivation in emotion. *American Psychologist*, 46(4), 352.

Marsella, S., Gratch, J., & Petta, P. (2010). Computational models of emotion. A Blueprint for Affective Computing-A sourcebook and manual, *11*(1), 21-46.

Mehrabian, A., Russell, J. A. (1974). An approach to environmental psychology (1 ed.). Cambridge, Mass.: MIT Press.

Ojha, S., Vitale, J., & Williams, M. A. (2021). Computational Emotion Models: A Thematic Review. In *International Journal of Social Robotics, 13*(6), 1253–1279. Springer Science and Business Media B.V. https://doi.org/10.1007/s12369-020-00713-1

Ong, D. C., Jospe, K., Reddan, M., Wu, Z., Kahhale, I., Chen, P., & Perry, A. (2023). People optimally and flexibly process emotional information across multiple modalities.

Ong, D. C., Soh, H., Zaki, J., & Goodman, N. D. (2021). Applying Probabilistic Programming to Affective Computing. *IEEE Transactions on Affective Computing*, *12*(2), 306–317. https://doi.org/10.1109/TAFFC.2019.2905211

Ong, D. C., Zaki, J., & Goodman, N. D. (2015). Affective cognition: Exploring lay theories of emotion. *Cognition*, *143*, 141-162.

Ong, D. C., Zaki, J., & Goodman, N. D. (2019). Computational models of emotion inference in theory of mind: A review and roadmap. *Topics in Cognitive Science*, *11*(2), 338-357

Ortony, A., Clore, G. L., & Collins, A. (1988/2022). The cognitive structure of emotions. Cambridge university press.

Pei, E., Zhao, Y., Oveneke, M. C., Jiang, D., & Sahli, H. (2022). A Bayesian Filtering Framework for Continuous Affect Recognition from Facial Images. *IEEE Transactions on Multimedia*. https://doi.org/10.1109/TMM.2022.3164248

Russell, J. A. (1997). Reading emotion from and into faces: Resurrecting a dimensional-contextual perspective. In J.A. Russell & J.M. Fernández-Dols' (Eds.), The psychology of facial expression, Cambridge University Press, New York 295–320. https://doi.org/10.1017/ cbo9780511659911.015.

Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, *2*, e55.

Saxe, R., & Houlihan, S. D. (2017). Formalizing emotion concepts within a Bayesian model of theory of mind. *Current Opinion in Psychology*, 17, 15-21.

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. Appraisal processes in emotion: *Theory, Methods, Research, 92*(120), 57.

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology*, *46*(6), 401-435.

Scherer, K. R., Dieckmann, A., Unfried, M., Ellgring, H., & Mortillaro, M. (2021). Investigating appraisal-driven facial expression and inference in emotion communication. *Emotion*, *21*(1), 73–95. https://doi.org/10.1037/emo0000693

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, *17*(11), 565-573.

Sweeny, K., & Vohs, K. D. (2012). On near misses and completed tasks: The nature of relief. *Psychological Science*, *23*(5), 464-468.

Teo, D. W., Ang, Z. Y., & Ong, D. (2022). Modeling causal inference from emotional displays. In *Proceedings of the Annual Meeting of the Cognitive Science Society, 44*(44).

Van de Cruys, S. (2017). Affective value in the predictive mind.

Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., Gao, S., Sun, Y., Ge, W., Zhang, W., & Zhang, W. (2022). A systematic review on affective computing: emotion models, databases, and recent advances. *Information Fusion*, 83–84, 19–52. https://doi.org/10.1016/j.inffus.2022.03.009

Wu, Y., & Schulz, L. E. (2018). Inferring beliefs and desires from emotional reactions to anticipated and observed events. *Child Development*, *89*(2), 649–662. https://doi.org/10.1111/cdev.12759

Yanagisawa, H., Kawamata, O., & Ueda, K. (2019). Modeling emotions associated with novelty at variable uncertainty levels: A Bayesian approach. *Frontiers in Computational Neuroscience*, *13*. https://doi.org/10.3389/fncom.2019.00002

Zaki, J. (2013). Cue integration: A common framework for social cognition and physical perception. *Perspectives on Psychological Science*, *8*(3), 296-312.

Zaki, J., & Ochsner, K. (2011). Reintegrating the study of accuracy into social cognition research. *Psychological Inquiry*, *22*(3), 159-182.

Cohen, I., Garg, A., & Huang, T. S. (2000, November). Emotion recognition from facial expressions using multilevel HMM. In *Neural information processing systems, 2*. State College, PA, USA: Citeseer.

Gratch, J., & Marsella, S. (2001, May). Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In *Proceedings of the fifth international conference on Autonomous agents* (pp. 278-285).

Ekman, P. (2004, April). What we become emotional about. In *Feelings and Emotions. The Amsterdam symposium* (pp. 119-135). New York: Oxford University Press.

Jadhav, N., & Sugandhi, R. (2018, November). Survey on human behavior recognition using affective computing. In *2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN)* (pp. 98-103). IEEE.

APPENDIX A

Table 1. The 47 appraisal dimensions identified by Jie and Ong, 2023.

Appraisal	Description	Related terms
	Appraisals related to basic desires	
Pleasantness	Whether one thinks that the situation is intrinsically pleasant. This does not include goals, but only whether the event by itself is pleasant or not.	valence, intrinsic pleasantness
Threat	Whether one's physical or emotional well-being is being threatened by an imminent danger.	danger, risk, susceptibility, vulnerability, intimidated
Loss	Whether something has been lost and could not be retrieved / whether one thinks something irreversible has happened that could not be returned to its original state.	reversibility
Harm	Whether one perceives that someone/something has harmed or is currently harming them.	
Severity	How severe are the consequences of an event to one's well-being.	impact
Easing of threat	*Whether one perceives that a threat or harm has been removed from the situation (k=1)*	
	Goal-related Appraisals	
Goal conduciveness	Whether the situation and its outcomes are consistent with one's goals, desires, wants, needs or whatever one cares about. This also refers to whether the event has benefit or harm in relation to one's goals.	goal congruence, goal hindrance, goal obstacle, desirability, situational state, benefit
Goal relevance	Whether the situation is relevant or important to oneself. Relevance could be in terms of whether the situation is relevant to one's goals, well-being, desires, wants, needs or whatever one cares about.	relevance, importance, self-concern, centrality
Value	Whether the situation has any value in achieving one's goals.	attainment value, intrinsic value, extrinsic value, outcome value
Difficulty	How difficult is it to achieve one's goals. This also refers to how difficult is it to deal with the task/situation at hand.	situational demand, skill demand
Perceived obstacle	The extent to which one perceives problems or obstacles in the situation that are hampering attainment of a desired goal.	problems
Attainability	Whether one could obtain something that one desires.	possibility

Desire for object	*Whether one desires an object that another person has. (k=3)*	
	Agency-related Appraisals	
Accountability-Circumstances	How much are impersonal circumstances (beyond one's, any other person's, or object's control) perceived to be causally responsible for the situation. Examples of impersonal circumstances include luck or other factors unaccounted for by any person or object.	circumstances-agency, circumstance-attribution, situational-agency
Accountability-Other	Whether another person (or object) is causally responsible for the situation. The other person (or object) is also accountable (deserved to be blamed or credited) for bringing about the situation.	other-agency, other-caused, other-attribution, other-responsibility, other-blame, blameworthy, external agency
Accountability-Self	How much one feels they are causally responsible for the situation, and are accountable (deserved to be blamed or credited) for bringing about the situation.	self-agency, self-caused, self-attribution, self-responsibility, self-blame, internal agency, internal locus
Intentionality	Whether the event was intentionally caused by any person.	self-intentionality, other-intentionality
	Control and coping-related Appraisals	
Control-Circumstances	Whether one thinks circumstances (or situational factors) beyond anyone's control have the ability to control and influence the current situation.	situational control
Control-Other	Whether another person (or object) has the ability to control and influence the situation.	external control
Control-Self	Whether one has the ability to control and influence the current situation.	control, personal control, power, helplessness
Emotion focused coping potential	Whether one is able to emotionally cope and adapt to the situation at hand or with its consequences. This also refers to evaluating whether one could regulate one's emotional state to deal with the situation.	
Problem focused coping potential	Whether one is able to cope and adapt to the situation at hand and with its consequences. This also refers to whether one thinks one can influence the situation to make it better.	coping, competence
Effortful success	*How well is one able to achieve what one wants if one tries hard enough. (k=2)*	*effortful optimism*
Modifiability	*Whether the situation or its outcomes can be modified (k=2)*	
	Expectations-related Appraisals	
Novelty	Whether one thinks that the situation is new and has not been experienced before. The opposite of this is familiarity—Whether one thinks the situation has been experienced before and is familiar.	familiarity
Expectedness	Whether one expected the particular situation to occur	unexpectedness
Remarkable	How remarkable is it to have gotten this outcome?	exceed expectations

Uniqueness	*Whether a similar event had happened before, or if this event is a unique event that would not ever happen again. (k=3)*	*pattern/unique*
	Current and Future Expectations-related Appraisals	
Future expectancy	Whether one thinks that the situation would turn for the better or worse in the future.	positive future expectancy, negative future expectancy, optimism, outcome expectancy, expectancy, expected success, negative likelihood
Attentional activity	Whether one needs to take time or effort to attend and consider the situation further, or whether one tries to shut it out.	attention, consideration, urgency, safety
Certainty	Whether one knows what is happening in the situation. This also refers to how well one understands what was going on in the current situation and its subsequent consequences.	understandability, clarity
Future predictability	Whether one is able to predict future events as a consequence of the current situation.	outcome probability, probability
Effort	Whether one needs to exert some effort (mental or physical) to deal with the situation.	
Stability	*Whether the effects of the situation will be temporary or permanent / whether one thinks the effects are going to change over time or are constant (long-lasting). (k=3)*	
	Appraisals related to norms and value judgments	
Fairness	To what extent is the situation fair to all parties involved.	deservingness, legitimacy, justice
Normative significance (External)	Whether the situation was consistent with one's external and social norms.	external self-compatibility, fail to live up to external standards, norm violation, norm compatibility, compatibility with external standards: norms
Normative significance (Internal)	Whether the situation is consistent with one's personal values and ideals.	compatibility with individual norms, compatibility with internal standards, internal self-compatibility, self-consistency, moral congruence, immorality
	Appraisals related to social others	
Concern for others	Whether one is concerned with the well-being of another person.	other-significance, other-involved, other-concern
Closeness	*How close is the relationship between another person to oneself. (k=3)*	

Evil character	*Whether one perceives others (e.g., outgroup members) as dispositionally evil (k=1)*
Liking	*How much does one like another person. (k=2)*
	Other appraisals
Challenge	Whether one anticipates a potential obstacle in the situation, but also perceives a future gain.
Globality	*The extent to which one perceives that the situation is relevant to all aspects of one's life. (k=2)*
Reality	*How real the event is perceived to be (k=1)*
Self-esteem decreased	*Did the event enhance or decrease one's self-esteem (k=1)*
Temporal distance	*Whether the situation is perceived as temporally close or distant to present time (k=1)*
Vastness	*How conceptually and perceptually extensive is the stimulus relative to oneself (k=1)*

Note: Italicized appraisals appear in 3 or fewer studies (k < 3), (Jie & Ong, 2023).