Signal Detection Theory for Emotional Faces

Grace Dalton

May 7, 2024

Abstract

This study explores emotional discrimination in epilepsy neurosurgical patients and amygdala lesion patients, shedding light on the neural mechanisms underlying emotion recognition. It is based on data from a former experiment (Wang et al., 2017), and uses fearful and happy faces that participants had to discern between. Signal detection theory analysis revealed both groups' ability to discriminate emotions, with lesion patients outperforming epilepsy patients. However, the exclusion of ambiguous expressions limits its generalizability. The human amygdala likely plays a role in discerning fearful facial expressions from happy facial expressions. This investigation emphasizes the significance of amygdala function in emotional processing and highlights implications for clinical populations' emotional cognition.

Introduction

In the realm of cognitive neuroscience, understanding the intricacies of facial recognition and emotional processing is paramount to human survival and engagement in daily life. This study delves into the realm of emotional discrimination, focusing on the responses of epilepsy neurosurgical patients and amygdala lesion patients. With a cohort comprising eight neurosurgical patients across thirteen sessions, and three amygdala lesion patients, this study investigates the mechanisms underlying emotion recognition and discrimination. The experimental design revolves around a task where participants are tasked with discriminating between two fundamental emotions: fear and happiness. The study uses faces selected from the STOIC database–which is known for its recognizable emotional expressions–and generates stimuli with fearful expressions and happy expressions. This manipulation allows for an exploration of emotional perception. Moreover, by incorporating control measures to equalize low-level image properties such as luminance and contrast, the study ensures that observed effects stem from genuine emotional processing rather than perceptual confounds.

The original experiment got its data through using a piecewise-cubic-spline transformation with the SHINE toolbox for precise image manipulation (Wang et al., 2017). Each session was treated as an independent sample for behavioral analysis, and offers insights into the neural underpinnings of emotion recognition. The task presented a face to the subject for one second, then prompted the participants to make the best guess of the facial emotion, either by pushing the left button to indicate a fearful expression, or by using the right button which indicated a happy expression (Wang et al., 2017). Participants had two

seconds to respond before a beep sounded and the trial was discarded (Wang et al., 2017). The participant was shown a blank screen for 500 milliseconds, then was asked to indicate the confidence level they had in their response within two seconds, otherwise the trial was discarded (Wang et al., 2017). By presenting faces in a randomized order and employing jittered inter-trial intervals, the study minimizes potential biases.

This study adopts an analytical framework grounded in signal detection theory, and compares the epilepsy neurosurgical patients with the amygdala lesion patients. This paper aims to find the participants' capacity to discern fearful expressions (signal) amidst emotional stimuli, such as happy expressions (noise). The data will be used to calculate the participants' errors in a two-choice decision-making environment between the stimuli. Furthermore, through examination of discriminability (d') and criterion (k), the study aims to highlight potential alterations in perceptual processes and decision-making strategies underlying emotional discrimination deficits. It also considers if there is a differential impact of neurological pathologies on emotion-processing mechanisms. This analytical paradigm not only affords a quantitative assessment of participants' perceptual activity, but also reveals the importance of shaping emotional cognition in clinical populations.

Data

The data in this study was taken from one task used in Wang et al. (2017). It can be found in the Confidence database on OSF in the format of a .csv file (Wang et al., n.d.). The data was loaded into a Jupyter Notebook using Python's "pandas" package. The original .csv dataset variables reported the subject ID, the type of face shown, the participant's response, the participant's confidence level, the response time in seconds to the face, the response time in seconds to the confidence rating, and the patient group the subject belonged to. The faces stimuli were based off a continuum of ambiguous expressions ranging from fearful to happy expressions; they were ranked as levels 1: 100% fearful, 2: 70% fearful, 3: 60% fearful, 4: 50% fearful, 5: 40% fearful, 6: 30% fearful, and 7: 0% fearful. The subjects' responses were represented by choice 1: judging face as fear, 2: judging face as happy, and NaN: missing button press. The confidence levels used were 3: very sure, 2: sure, and 1: unsure.

To prepare the data for interpretation under signal detection theory, both response times columns along with the confidence levels column were dropped from the dataframe. This left the subject ID, the response, and the group that the subject was from. The subject ID was used to track specific trials; the stimulus level variable was used as the ground truth since it reports what the actual stimulus presented was. The response variable was used as the decision since it was the choice that participants made to label the presented stimuli. The group the subject was from represented the condition–either epilepsy or lesion. Along with removing blank and duplicated rows, rows with stimuli levels from 30-70% were also deleted because they provided noise from the ambiguous faces they represented. The 100% fearful stimulus was coded as 1 to represent the signal, while 0% fearful was coded as 0 to represent the noise. The response that judged the face as fearful was coded as 1 to represent a subject's choice of "detect signal," while judging the face as happy (not fear) was coded as 0 to represent "detected no signal." The data were split into epilepsy and lesion groups, then ground truths were compared with decisions. This resulted in a new

dataframe (Table 1)–which consisted of two rows (one row for each group's hits, false alarms, misses, and correct rejections)–that was exported as a .csv file into JASP. The data were then analyzed using JAGS to calculate marginal d', k, and bias values (c and β), the joint d' and c values, and posterior predictive values. It is important to note that only "easy" stimuli (0% and 100% fearful) were used, which led to near ceiling performance. Because ambiguous faces were removed, the difficulty decreased and it was easier to discriminate between facial expressions; therefore subjects should make minimal errors. As a result, the amount of information learned from the data will be limited.

Table 1: Signal Detection Theory Table

Group	Hit	False Alarm	Miss	Correct Rejection
Epilepsy	283	6	6	281
Lesion	46	0	0	46

Results

The signal detection theory model was appropriate for analyzing this dataset because it can be used to analyze errors in two-choice decisions in which participants try to detect a signal from noise. In this model, there is an underlying continuum on which an individual experiences things, and uses a threshold criterion (k) to make a decision. The criterion is what differentiates signal detection theory from a Thurstonian approach: the person can set the criterion as a conscious effort to control how they make decisions and manage their errors. When the participant sets a high criterion, they need more evidence to say "signal", meaning that they detected a fearful expression; if they set a low criterion, they need less evidence to detect the signal. In the case of this experiment, if the amount of evidence a participant had fell above k, then they identified the facial expression as fearful. If the amount of evidence a participant had fell below the threshold, then they identified the facial expression as happy.

Table 2: The bias and discriminability mean, standard deviation, and 95% credible interval of both epilepsy and lesion patient conditions.

Parameter	Posterior		95% Credible Interval	
	Mean	SD	Lower	Upper
c_E	−0.002	0.120	−0.231	0.234
c_L	−0.214	0.772	−1.921	1.228
d'_E	4.092	0.249	3.636	4.601
d'_L	6.275	1.389	4.073	9.431

In Table 2, "E" subscripts represent the epilepsy patients and "L" subscripts represent the lesion patients. Both groups' discriminability values–which are measured using standard z-scores–explain how far apart the signal and noise distributions are for each group.

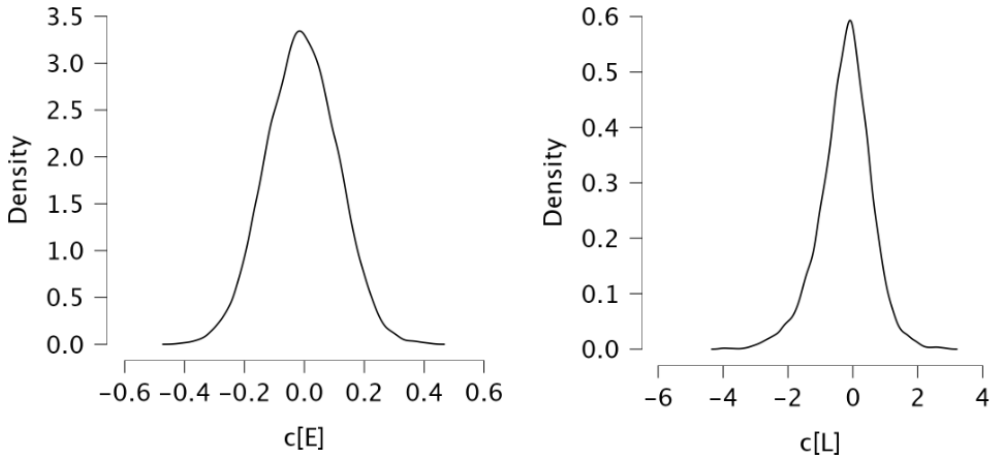

Figure 1: Marginal c density plots.

The epilepsy group's discriminability mean value is 4.092, which means that the signal and noise peaks are about four standard deviations away from each other; the lesion group's discriminability mean value is 6.275, which means that the peaks are closer to six standard deviations away from each other. Both discriminability parameters are considerably high, which represents little overlap between the signal and noise distributions. This signifies that there will not be many false alarms and misses for either group. The criterion, k, is analyzed in the context of d'. The criterion values for both the epilepsy and lesion patients are about half the size of the respective discriminability values. This means that the participants' thresholds in both groups are minimally biased. This is because when k equals half of d', the criterion is unbiased. Further examination of the c bias parameters displays how there is minimal bias in the both groups' criterion. The epilepsy group's c value is -.002, and the lesion group's c value is -.214. The marginal c density plots in Figure 1 display how the measures of bias in the participants' criterion value are both centered near 0. The 95% credible interval shows that the c value for epilepsy patients will most likely be between -.231 and .234, whereas the c value for lesion patients will most likely be between -1.921 and 1.228. The epilepsy group is closer to c=0, but because both credible intervals capture the value c=0, this insinuates the potential for the criterion to have no bias. In regards to the signal and noise distributions for each group, this means that the probability of false alarms and misses is almost evenly split, and participants are almost equally likely to make either errors. Although the c values are slightly negative, the credible intervals of the epilepsy and lesion groups overlap, so there is uncertainty as to where the true mean values lie. Furthermore, because neither group made many mistakes, it is unknown where their criterion is or how biased they are.

A visual inspection of the joint d' and c bivariate scatterplot and a correspondence of the marginals were made; these data did not contain additional information, as the joint is well approximated by the marginals, so an analysis of the mean and credible intervals was completed using Table 3. The mean values for both groups suggest that there is minimal bias in participant criterion, but the stimuli for fearful and happy faces are very discriminable from one another. However, the lesion patients are slightly more negatively biased than the

Table 3: Joint d' and c mean, standard deviation, and 95% credible interval of both epilepsy and lesion patient conditions.

Parameter	Posterior		95% Credible Interval	
	Mean	SD	Lower	Upper
c_E	−0.008	0.117	−0.236	0.228
c_L	−0.229	0.773	−1.888	1.149
d_E	4.077	0.230	3.643	4.555
d_L	6.209	1.375	4.076	9.290

epilepsy patients.

Table 4: Posterior Predictive distributions and observed counts for both epilepsy and lesion patient conditions. The 95% credible intervals are for the posterior predictive distributions.

Parameter	Values		95% Credible Interval	
	Mean	Observed Counts	Lower	Upper
Hit_E	282.681	283	275.000	288.000
Hit_L	45.710	46	43.000	46.000
$FalseAlarm_E$	6.439	6	1.000	15.000
$FalseAlarm_L$	0.404	0	0.000	3.000

The posterior predictive summary in Table 4 illustrates the model's ability to re-describe the data it just saw. Furthermore, the distributions in Table 4 reflect what the model expects to see. It expects a high hit to false alarm ratio, and its mean values are similar to the results from the measures in Table 1. Overall, this model appears to re-describe the collected results. This agreement between the observed data and the posterior predictive distributions validates the model and quantifies the uncertainty in predictions by considering the variability of the model parameters.

Discussion

This study used happy and fearful face stimuli to test people's ability to discriminate between different emotional expressions. It also compared the epilepsy neurosurgical patients with the amygdala lesion patients to discover if there is an impact of neurological pathologies on emotion-recognition. Both epilepsy and lesion patients were clearly able to discriminate between fearful and happy facial expressions. Although identical stimuli were used for both groups, the lesion patients were potentially better at discriminating between fearful and happy faces than the epilepsy patients, but there was uncertainty in this finding. While both groups reported minimal criterion bias, the ceiling effects of the data limit the certainty of the findings. These findings may indicate that neurons in the human amygdala are involved in the processing of fearful facial expressions. However, because the face stimulus levels between

the levels of 30-70% fearful were removed from the study, the results cannot account for the participants' ability to discriminate between ambiguous emotional expressions. This limits the ability of the study to describe how epilepsy and lesion patients are able to discriminate all fearful facial expressions in everyday life, as many naturally occurring facial expressions may exist in the middle of the fearful to happy expression continuum, rather than on the extremes. It also remains unclear whether the amygdala's involvement in the ability to discriminate fearful faces is generalizable to other facial emotions. The original study tested a second pair of emotions consisting of anger versus disgust (Wang et al., 2017), but these emotions were not analyzed under the scope of signal detection theory in this study. A notable follow-up experiment would be to analyze a third control group of healthy people to see if there are differences in the ability to discriminate between facial expressions. This could further confirm if the amygdala is involved in detecting fear from noise in facial expressions.

References

Wang, S., Yu, R., Tyszka, M. J., Zhen, S., Kovach, C., Sun, S., ... Rutishauser, U. (n.d.). *data-wang-2017-natcomm.csv.* Retrieved from `https://osf.io/ep3u8`

Wang, S., Yu, R., Tyszka, M. J., Zhen, S., Kovach, C., Sun, S., ... Rutishauser, U. (2017). The human amygdala parametrically encodes the intensity of specific facial emotions and their categorical ambiguity. *Nature Communications*, *8*(14821).