Feeling within reason

How appraisals shape emotional responses to politics

Rebasso, I.

Publication date
2023

2 Affective Modeling of Emotion Labels in Politics

Abstract

We regularly ask participants how they feel about politics: angry, anxious, hopeful? Yet, we know fairly little about which considerations go into these answers, or which topics people think of when reporting general emotions about politics. In this chapter, we investigate the cognitive roots of political emotions by applying Scherer's Component Process Model and testing whether emotion labels are predicted by different appraisals in the political versus nonpolitical domains. We ask participants to recall and briefly describe an intense emotional event that is either related to politics or to their personal/professional lives. We then use a computational approach (Israel & Schönbrodt, 2019) to model distinct emotion labels based on 18 self-reported cognitive appraisals.

In the nonpolitical domain, we can clearly distinguish emotions based on their appraisal profiles, and appraisal profiles are largely in line with our theoretical predictions. However, in politics, negative emotions look strikingly similar in appraisal profiles. Anger, anxiety, despair, and guilt are all related to appraisal profiles that we would theoretically link to anger responses. While we can predict joy and despair with high accuracy in a nonpolitical setting, in a political setting, we can best predict joy and anger. The overall model performance in politics is low. The political and nonpolitical models differ in both the relationships between appraisals and emotions and their overall performance.

Emotion research in political science has relied on a core assumption: that emotions in politics work similarly as in the personal sphere. Over the last two decades, work that has applied psychological theories of emotions to politics has contributed – among other things – two important ideas: (1) emotions are not irrational, but rather they can be rational and justified responses to the environment (Marcus, 2010); and (2) anger and anxiety are distinct emotions with different effects on political behavior (Albertson & Gadarian, 2015; Valentino et al., 2011, 2008). In this chapter, we challenge these ideas. First, to determine whether emotions are justified responses to the environment, emotion labels must be systematically linked to a unique set of appraisals of the situation. Second, for anger and anxiety to be meaningfully different in the political domain, they should also be linked to different appraisals of the environment. While psychologists have extensively tested these core premises of appraisal theory (Israel & Schönbrodt, 2019; Roseman et al., 1990; Scherer & Meuleman, 2013; C. A. Smith & Ellsworth, 1985), emotion research in political science has engaged very little with the appraisals that predict distinct emotions. By testing whether emotion labels in politics are related to distinct appraisals of political events, we also question the core assumption that emotions work similarly in politics as they do in the personal sphere.

Early work in emotion research in political science outlined its own theory of emotions in politics: Affective Intelligence Theory (Marcus et al., 2000). This foundational approach primarily focused on enthusiasm, anxiety and anger and identified the novelty/familiarity of a stimulus or issue as the core appraisal that distinguishes the two emotions. More recently, other approaches have become prominent in political psychology.

Most importantly, appraisal theories (Lazarus, 1991; C. A. Smith & Ellsworth, 1985) have enabled scholars to broaden the range of studied emotions, including more complex emotions like guilt, shame, disgust, and contempt (Redlawsk & Mattes, 2022). These theories have also expanded the scope of appraisals beyond novelty/familiarity to, for instance, power/control (Rico, Guinjoan, & Anduiza, 2020; Valentino et al., 2011; Valentino, Gregorowicz, & Groenendyk, 2009), intentionality (Petersen, 2010), and blame attribution (Kühne et al., 2015)).

Yet, appraisal theories were originally developed and validated to study emotions in personal, rather than political, contexts. Social psychologists have examined how appraisals lead to emotions in social contexts, such as when group identities are made salient (Mackie & Smith, 2018; E. R. Smith & Mackie, 2016). Besides that, we usually assume that the way in which appraisals shape emotions in a political context is the same as in a personal context.

There are several reasons why we might expect that the process through which appraisals elicit distinct emotions works quite differently in the political compared to the nonpolitical domain. We argue that appraising political events and situations differs fundamentally from how we think about appraisal processes in the personal realm. Emotions in politics are directed at distant targets, which makes the emotion process more complex and abstract (Schiller et al., 2022). In addition, certain appraisals might be more relevant or more difficult to make in politics than in the personal domain. Politics (1) might be primarily about responsibility (Iyengar, 1994), (2) is often not immediately experienced but mediated by elites, and (3) is complex and often murky, leading people to use heuristics to make sense of it (Arceneaux & Stein, 2006).

To test the domain differences in appraisals processes, we rely on the Component Process Model (Scherer, 2022), a model of emotions that links a wide range of appraisals to distinct emotions. Empirically, we try to predict the basic emotion families based on appraisal profiles (see also Scherer & Meuleman, 2013). Following Israel and Schönbrodt (2019), we use a genetic *Differential Evolution* optimization algorithm suited for classification problems using high-dimensional data. We preregistered a simple hypothesis on OSF: that the model would perform better on nonpolitical data than on political data. Apart from this simple expectation, our approach was largely inductive, as we had little prior knowledge about the distribution of emotions or the appraisal prototypes within each emotion.

We find that appraisals and emotions do not link as well in the political domain as they do in the personal domain. We find that negative political emotions – anger, anxiety, despair, and guilt – are not distinguishable based on their appraisal profiles. Moreover, the model to predict emotion labels performs worse along all metrics when predicting political emotions than when predicting nonpolitical emotions. This calls into question the two core assumptions mentioned at the beginning of this introduction and suggests that we need to rethink how we theorize about appraisals, emotions, and the outcomes to which they have been linked in the domain of politics.

Methodologically, we argue that measuring key appraisals alongside emotion labels can provide important insights into the content of emotions. Emotions are multi-componential, and should be studied as such, particularly when theorizing about their behavioral consequences. An emotion's action tendency, which is assumed to be related to the underlying appraisals, can be better understood by analyzing the specific appraisals related to the emotional response.

2.1 Appraisals in the Component Process Model

Emotion theorists generally agree that emotions are elicited in response to our environment and serve to help us deal with situations and events. In other words, they agree that "emotions 1) consist of an episodic process in response to a perceived event or situation of major significance, 2) which is characterised by recursive causal effects (forward and backwards) between several components that include the evaluation of the event in terms of its significance for the goals and values of the individual, 3) creating physiological reactions, motor expressions, and action tendencies and 4) that this process is partially accessible to consciousness, resulting in feelings that 5) can be categorised and subsequently labelled by the individual in terms of its subjective conceptual structure" (Scherer, 2022, p.164). Scherer calls these ingredients the components of an emotion episode. While their relative importance differs by theoretical approach, they are present to some extent in all major theoretical accounts (Scherer, 2022; Scherer & Moors, 2019). In this chapter we focus on the first, second and last component: the situations and events that cause emotions, the evaluations that determine which emotion the event produces, and the categorization or labelling of the emotion episode. Particularly, we will contrast how these components play out in the political domain as opposed to the personal domain.

The appraisal of an event describes the process of detecting and evaluating its significance for one's well-being or the degree to which it helps or obstructs one's concerns (needs, attachments, values, goals and beliefs) (Moors et al., 2013). While most emotion theories grant a role to evaluative judgments in the emotion process, appraisal theories make them a central building block of an emotion episode. Appraisal theories argue that people constantly monitor and evaluate their surroundings to form emotions and prepare for action.

One commonality of all appraisal theories are appraisals of relevance and goal congruence. Relevance, according to Scherer, is the most central appraisal dimension (Scherer, 2022). A situation, such as a political event, must matter to us to invest in subsequent appraisals and emotional experience. Subsequently, we determine if an event helps or obstructs our goals, or put differently whether it is positive or negative. Then, a number of secondary appraisals are made. Although different appraisal theories disagree on the exact nature of these secondary appraisals, at least three of them can be found in most approaches: certainty, agency (who or what is the event caused by), and coping potential or control (Moors et al., 2013, p. 120). Other theories also include appraisals of novelty, expectancy, urgency, intentionality, legitimacy, or fairness, and/or norm compatibility.

Appraisals are made sequentially (Scherer et al., 2001, ch.5). As new information becomes available, the emotion we experience can change. Generally, the more appraisals are made, the more nuanced the emotional episode, and an emotional episode becomes more complex with each recursive cycle (Moors et al., 2013). Action tendencies and physiological responses may also feed back into appraisals. As a consequence, the emotions we measure through self-reports are the products of often very complex appraisal sequences, where all components of the emotion process model sequentially influence each other, consciously or unconsciously. In this chapter, we follow Scherer (2013) in considering a wide range of appraisals instead of limiting ourselves to a few appraisals that we select a priori.

Based on these appraisals individuals experience a certain emotion, which comes with a typical appraisal profile (Scherer & Moors, 2019). To give an example: appraisals of responsibility attribution, together with appraisals of high individual control and high certainty produce anger. If another person in the same situation attributes responsibility to themselves, they experience guilt. If people appraise the same situation differently, they also experience different emotions (Siemer, Mauss,

& Gross, 2007).

Following the Component Process Model appraisals are followed by physiological responses, motor expressions and action tendencies. The latter, also sometimes referred to as appraisal tendencies, describe “the preparation of action appropriate for dealing with or adapting to these events both mentally, in the form of states of action readiness or action tendencies, and physically, in the form of physiological responses”(Scherer & Moors, 2019, p.721). Most work in political science has focused on the (political) action tendencies of emotions, asking question such as: how does anger affect political participation (Valentino et al., 2011, 2009; C. Weber, 2013)? How do anger and anxiety about terrorism affect support for military action (Huddy et al., 2007)? Does anger drive partisan social division (Kim, 2016; MacKuen et al., 2010)?

Theoretically, appraisals come before an emotion label, but in practice it is not possible to establish a causal ordering in an observational study. An emotion episode unfolds quickly and is short-lived. Thus, when people report their emotions and appraisals after the fact, action tendencies have already fed into the self-reported appraisals through a feedback loop. Establishing a causal ordering of the components of an emotion episode is not the goal of this chapter. Rather, we aim to test a much more fundamental assumption of appraisal theory: can political emotions be differentiated and predicted by their distinct appraisal profiles?

Appraisal profiles are unique combinations of several appraisal dimensions, although not all appraisal dimensions are relevant in every distinct emotion’s appraisal profile. Scherer and Meuleman (2013) identify 25 distinct appraisals.[3]

Our expectations for the appraisal profiles are derived from Scherer and Meuleman (2013) and Scherer et al. (2001) and are shown in Table 1.

[3] Un/pleasantness, Suddeness, Predictability, Familiarity / Novelty, Goal / Need importance, Moral acceptability, Norm violation, Agency: chance, Agency: self, Intentionality: self, Image compatibility, Agency: other, Intentionality: other, Consequences: known, Consequences: expectedness, Consequences: near future, Consequences: far future, Goal conduciveness, Goal obstructiveness, Injustice, Coping, Urgency, Power, Adjustment.

2.2 Politics as a Separate Emotional Domain?

Appraisal theories have become one of the key frameworks through which political psychologists have studied emotions. Applying appraisal theory to politics has significantly fostered our understanding of how emotions function in politics. Anger, for instance, has been shown repeatedly to increase political participation (Valentino et al., 2011; C. Weber, 2013) and to decrease deliberation across party lines (MacKuen et al., 2010; Webster et al., 2022), while anxiety stimulates information seeking (Albertson & Gadarian, 2015; Valentino et al., 2008).

Appraisal theories have provided a framework for studying emotions in political science through three distinct approaches, namely (1) the integral observational approach, the (2) integral experimental approach, and (3) the incidental approach.

Integral emotions are emotions that are directly related to the primary event or task that is being studied. We can study integral emotions in two ways. First, an observational approach simply measures both emotion and outcome and controls for possible confounders. This has been done cross-sectionally (e.g. Petersen, 2010; Valentino et al., 2011; Vasilopoulou & Wagner, 2020), and in panel studies (Rico et al., 2017; Rudolph et al., 2000; Vasilopoulos, 2018; Wagner, 2014).

The strength of this approach is that it studies emotions directed at a specific political target and as they occur naturally. However, this approach usually only measures the emotion label, and the associated action tendency. As it is observational, this approach always suffers from problems of endogeneity (Ladd & Lenz, 2008, 2011), as we have no grasp of the feedback loops between action tendencies, behaviors and appraisals. Moreover, isolating the effect of emotions on behaviors, depends on whether we include the correct controls. If we fail to do so, we run the risk of building our knowledge on spurious relationships. Isolating the effects of discrete emotions is therefore a particularly tricky task, as we know little about the causes of and relationships between emotions (see for instance Jost, 2019).

The second approach to emotions, the experimental integral approach to emotions, tries to deal with the endogeneity problem of observational studies either through integral emotion induction tasks (Valentino et al., 2011, 2008), or by manipulating specific appraisals (Chadwick, 2015; Kühne et al., 2015). Hypotheses within this approach are often limited to a very specific selected case, such as populism (Hameleers et al., 2017) or manipulate only a single appraisal dimension, such as group identity (E. W. Groenendyk & Banks, 2014), message congruence (Bakker et al., 2021; Suhay & Erisen, 2018), or responsibility attribution (Hameleers et al., 2017; Kim, 2016; Kühne et al., 2015). However, there are some studies that have manipulated or measured several appraisals to study how they shape emotions in politics (e.g. Chadwick, 2015; Kang & Cappella, 2008)

The final approach uses incidental emotion induction to isolate the causal effect of emotions on behaviors. Incidental emotions are emotions that are unrelated to the outcome that is studied. The action tendencies carry over to the appraisal of other stimuli (Lerner & Keltner, 2000, 2001). For instance, people in an angry state, a high control emotion, tend to be more risk-taking, and tend to show approach behaviors (Lerner & Keltner, 2000). Incidental affect has been used to study political outcomes in the form of images (Erisen et al., 2014), emotion recall tasks (Banks & Valentino, 2012; Parker & Isbell, 2010), movie clips (Renshon et al., 2015), or "impossible tasks" that induce anger (Ambroziak et al., 2022)

Appraisal Check	ANGER	ANXIETY	DESPAIR	GUILT	JOY
Goal Relevance	medium	medium/high	high	high	high
Novelty					
Suddenness	open	medium	medium	low	medium
Familiarity	low	low	low	open	low
Predictability	low/medium	low	low	open	low
Pleasantness	low	low	open	open	high
Implication					
Cause: Other	high	medium	medium	low	low/open
Cause: Self	low	low	low	high	high
Intentionality	high	open	low	high	open/high
Cause: Chance	low	medium	high	low	low
Goal Conduciveness	low	low	low	low	high
Goal Obstruction	high	high	high	high	(very)low
Urgency	medium	high	medium	medium/high	low
Coping Potential					
Control	high	open	low	open	open
Power	medium	very low	very low	open	open
Adjustment	high	medium/low	low	medium	high
Normative Sign.					
Moral acceptability	low	open	open	low	open/high
Norm violation	high	open	open	high	open/low
Unjust consequences	medium/high	open	open	medium	low

Table 1: Theoretical predictions of appraisal profiles of five emotion families, based on Scherer and Meuleman 2013 (Table 1 supplementary material) and Scherer 2001. Open means an appraisal check is not relevant for a specific appraisal profile. The last column shows our final prediction for the anger emotion family.

These approaches rely on two core assumptions. The first assumption is that appraisals in emotions in politics work similarly in the political domain as in the personal domain. This means that we can use appraisal theory to make predictions about which political events elicit which emotion, and among whom, and that emotions predict specific action tendencies related to their core appraisal profiles. Moreover, this assumption leads to the idea that we can use nonpolitical emotion induction tasks, study their effects and generalize the findings to emotions in politics.

The second, and related, assumption is that there are meaningful differences between someone reporting that they are anxious and someone else reporting that they are angry. In practice, emotions with distinct appraisal profiles – anxiety and anger, or anger and disgust – are often highly correlated (Rhodes-Purdy et al., 2020) and experimental manipulations often affect more than one emotion (e.g. Aarøe, 2011; Brader et al., 2008; Valentino et al., 2008; C. Weber, 2013; Weeks, 2015; Wisneski et al., 2020). As a result, the integral emotion induction approach trusts that emotions have the same meanings for different people, i.e. that people have uniform appraisal profiles that drive effects across individuals.

While appraisal theories appear like a set of stringent rules where certain appraisals predict distinct emotions, it is important to understand that they do leave room for culture or domain differences (Ellsworth, 2013; Moors et al., 2013). Psychologists have mainly studied geographical and language differences finding that appraisal theories can predict emotions across several languages and cultures (Fontaine, Scherer, Roesch, & Ellsworth, 2007). Here, we argue that politics constitutes a realm that might be fundamentally different from the personal realm. There are several reasons for this.

Schiller et al. (2022) argue that emotions are always directed at a specific concern. Further, they describe emotions as allostatic. They help us to regulate and deal with concerns at different levels. Conceptually, these concerns are similar to what Frijda (1987) has called an emotion's action tendencies. They distinguish different levels at which such allostatic concerns take place, from immediate (internal, physiological), interpersonal, to distant (moral) concerns. The more distant a concern becomes, the higher the complexity of steps that need to be taken to deal with it, i.e., to go from object to action tendency. In contrast to emotions in the nonpolitical domain, which usually deal with immediate and interpersonal concerns, emotions in politics deal with more distal concerns. They are therefore situated between interpersonal and moral concerns, which makes them significantly more abstract and complex than the nonpolitical emotions that are usually discussed in psychological theories of emotions. In what follows, I discuss three specific reasons why emotions in politics are more complex than emotions in the nonpolitical domain.

First, appraisals in politics are always subject to debate. Most prominently, politics is inherently about responsibility and accountability (Iyengar, 1994). Who is responsible for any political event, good or bad, retrospective or prospective, is at the core of most political debates. Debates around climate change often revolve around appraisals of certainty (what do we know and how sure can we be?) and control (what can we do, as individuals, as collective?). The ambiguity of these appraisals might lead people to rely on just the most salient appraisals, i.e. appraisals of responsibility.

Second, and linking back to political emotions as distant allostatic concerns, appraisal theory was developed based on the appraisal of immediate situations, i.e. situations that people personally and directly experienced. The experience of politics, however, is often mediated, by news media, or discussion partners. This may mean that how we appraise a political problem is different from a personal issue.

Feeling an urgency to act, for instance, is important in light of a directly threatening situation but arguably much less important when we read about events that happened the past day in the news. Of course one could think of instances where politics becomes very immediately felt and intrudes personal lives, e.g. unemployment or poverty. But for the person making these appraisals to see such situations as political usually requires an additional step which links the event to a broader political and societal context.

Third, in politics, we frequently lack full information. We do not know the intentions of political actors and the exact spheres of responsibility, and can only assume the control different political actors have over a situation. Thus, people use various heuristics, such as party cues (Arceneaux & Stein, 2006), to navigate complex political events. It is therefore debatable to what extent people make differentiated appraisals about politics at all, or whether when they report political emotions, they merely repeat politicians' emotion cues.

This list of arguments why political emotions are different from emotions in the personal sphere is not exhaustive. Yet, it gives ample reason for broadly re-testing appraisal theories in the political context. Existing research has studied appraisals as causes of distinct emotions, has focused on only a few appraisals at a time and has usually manipulated appraisals rather than studying them as they occur in real-life (Kühne et al., 2015; Valentino et al., 2009). Following the Component Process Model, which suggests that we should look at the full appraisal profiles, we take a more comprehensive approach to political emotions by testing their appraisal profiles composed of 18 different appraisals.

2.3 Methodology: A Computational Approach to Modeling Emotions

To link appraisals and emotions, we will use an adapted version of the Geneva Emotion Analyst (GEA) task (Scherer & Meuleman, 2013). This task asks participants to recall an emotional event and write a short description of it. Afterwards, they are asked to label the emotion they felt and to appraise the event by answering a series of questions.[4] Scherer and Meuleman (2013) used GEA to investigate the match between emotions and theoretically expected levels of appraisals. In a follow-up study, Israel and Schönbrodt (2019) suggested a refined computational approach to derive the co-occurrence of appraisals and emotion labels and the importance of specific appraisals in predicting emotion labels. To analyze the differences between the nonpolitical and political domains we follow this latter approach by Israel and Schönbrodt (2019).

In sum, the goal of this chapter is to evaluate the similarities and differences between emotions in the political and nonpolitical domains. Using the Geneva Emotion Analyst (GEA) task, we will answer the research question by contrasting (1) the frequencies of each emotion label, (2) the co-occurrence of appraisals

[4]GEA is a follow-up from the Geneva Expert System on Emotions (GENESE). The latter is a similar task, but asks appraisals first and emotion labels second.

and emotion labels (i.e. the appraisal profiles), (3) the importance of appraisals to predict emotion labels in both domains and (4) the overall performances of models predicting emotion labels based on 18 appraisals in the political versus nonpolitical domain.

Task Procedure

Our research design closely follows Scherer and Meuleman (2013) and Israel and Schönbrodt (2019) but with one important difference: we include a political and a nonpolitical condition. The wording of the instructions and questions is almost identical to the original study. In the political condition, respondents saw the following text: "*Please recall a moment* ***related to politics*** *when you experienced an intense emotion, either positive or negative, as a consequence of an event. It might have been brought about by you, by someone else, or by natural causes. The only important thing is that it has something to do with politics (for instance: a discussion about politics you recently had, a politician, a specific political event or topic, etc.). Don't think for too long. Take the first strong emotion-producing event that came to your mind after reading my request and put yourself back into that situation, trying to recall the details of the eliciting events as clearly as possible.*"

The domain in the nonpolitical condition is the following: "*Please recall a moment in time when you experienced an intense emotion in your* ***personal or professional life****, either positive or negative, as a consequence of an event. It might have been brought about by you, by someone else, or by natural causes. Don't think for too long. Take the first strong emotion-producing event that came to your mind after reading my request and put yourself back into that situation, trying to recall the details of the eliciting events as clearly as possible.*

Next, respondents were asked to describe the event in as much detail as possible using an open text box: "*Please describe the event/situation as detailed and objectively as possible using the text box below. Everything you write is confidential. Don't worry about any typos. Please take your time to think and write about the event. You will have to write a minimum of 20 characters to continue.*"

On the next screen, respondents indicated the emotion that best describes how they felt in the situation selecting from a list of 15 emotions: Sadness, Joy, Anger, Anxiety, Surprise, Fear, Irritation, Shame, Guilt, Disgust, Pleasure, Despair, Pride, Contempt, Rage.[5] The emotion labels were presented in random order. Last, we presented an option to select: "None of these correspond to how I felt in the moment".

After respondents had reported the emotion label, they were asked to make more detailed appraisals of the emotion episode. The measures for all appraisal checks are based on the measures used by Scherer and Meuleman (2013) and the latest GEA version. We measured the extent to which an appraisal was present during the emotion episode or a description of the event that elicited the emotion episode using a scale ranging from "1 = not at all" through "3 = moderately" to "5 = extremely".

Table 2 and 3 show all appraisal checks and their question wordings.

Sample

We recruited an online sample of 1241 respondents through *Prolific*. The sample consisted of 49 percent men, 50 percent women and 1 percent identifying as non-binary. The sample had a mean age of 43.73 (SD = 15.57). In terms of education 43 percent had a College or Bachelor degree, 15.4 percent a Graduate Level Degree, 3.3 percent only primary education, 38 percent secondary education, and 0.2 percent without formal education. In sum, the sample had representative characteristics. Treatment randomization worked as intended with 637 participants in the nonpolitical domain and 592 participants in the political domain.

Measures

To construct our measures of interest we first took the following preprocessing steps. Following Israel and Schönbrodt (2019), we scaled all appraisal scores to range from 0 to 1. *Doesn't apply* responses were recoded to *Missing*. We then assigned the fifteen emotions to five emotion families: anger (anger, irritation, rage, contempt, disgust), anxiety (fear, anxiety), despair (despair, sadness), guilt (guilt, shame) and joy (joy, pleasure, pride, surprise). The reduction of emotions into broader emotion families also followed Israel and Schönbrodt (2019). Their com-

[5]The original study included the same measures except for surprise and disgust.

putational model was capable of predicting emotion families, but underperformed in predicting emotions. In contrast to Israel and Schönbrodt (2019) we distinguished 5 emotion families by differentiating between despair and anxiety. We did this because of the substantive interest in anxiety in political psychology.[6]

After preprocessing, we first calculated the frequencies of the emotion families in both domains. Then we calculated the appraisal profiles as the mean level of each appraisal per given emotion label. To evaluate the importance of each appraisal dimension in predicting the emotion family, we ran the emotion prediction model as described by Israel and Schönbrodt (2019). We distinguished four steps.

First, we split the samples (political and nonpolitical, respectively) in two random samples, a training and a test set. In the training set we oversampled some emotions, so that each emotion label is represented equally.

Second, we calculated the mean of each appraisal per emotion label in the training data. These means are called prototypes and they are different to the appraisal profiles presented above, which were calculated based on the full domain samples. To conduct an intial test of whether our model compares to the Israel and Schönbrodt (2019) model, we first calculated the prototypes for each emotion (not emotion family) and the correlation between our prototypes and the Israel and Schönbrodt (2019) prototypes. Thus, for each correlation, we had 12 data points, one for each emotion that was in both our data set and the Israel and Schönbrodt (2019) data set. Table A1 shows the correlations per appraisal dimension. Overall, the appraisal prototypes correlate highly with those reported in the original study. The prototypes for the five emotion families, which we used as input for our prediction models, are shown in the Appendix, Tables A2 and A3.

[6]We deviated in two other ways. (1) We did not aggregate appraisal dimensions. The only dimensions that correlate strongly are goal conduciveness and goal construction but combining them does not decrease dimensionality significantly. (2) The study included measures on action tendencies, e.g. a wish to withdraw or attack. As we were theoretically interested in the appraisal and not the action tendencies, we did not include them.

Appraisal	Question
Goal Relevance	Could the event have important consequences for you?
Novelty	
Suddeness	Did the event happen suddenly and abruptly?
Predictability	Could you have predicted the occurrence of the event?
Familiarity / Novelty	Were you familiar with this type of event?
Pleasantness	How would you evaluate this type of event in general, independent of your specific needs and desires in the situation you reported above?
Implications	
Cause: chance	Cause: Chance, special circumstances, or natural forces?
Cause: other	Cause: The behavior of one or more other person(s)?
Cause: self	Cause: Your own behavior?
Intentionality	Was the event caused intentionally (by you or others)?
Goal conduciveness	At the time of experiencing the emotion, did you think that real or potential consequences of the event… did or would bring about positive, desirable outcomes for you (e.g., helping you to reach a goal, giving pleasure, or terminating an unpleasant situation)?
Goal obstructiveness	At the time of experiencing the emotion, did you think that real or potential consequences of the event… did or would bring about negative, undesirable outcomes for you (e.g., preventing you from reaching a goal or satisfying a need, resulting in bodily harm, or producing unpleasant feelings)?
Urgency	After you had a good idea of what the probable consequences of the event would be, did you think that it was urgent to act immediately?

Table 2: The 18 Appraisal Checks used as input to our models.

Appraisal	Question
Coping Potential	
Coping	At the time of experiencing the emotion, did you think that real or potential consequences of the event could have been or could still be avoided or modified by appropriate human action?
Power	After you had a good idea of what the probable consequences of the event would be, did you think that you had the power to control the consequences?
Adjustment	After you had a good idea of what the probable consequences of the event would be, did you think that you were unable to deal with the consequences that could not possibly be avoided or modified?
Normative Significance	
Moral acceptability	Were the actions that produced the event morally or ethically acceptable?
Norm violation	Were the actions that produced the event morally unacceptable or violated social norms?
Injustice	At the time of experiencing the emotion, did you think that real or potential consequences of the event were or would be unjust or unfair?

Table 3: The 18 Appraisal Checks used as input to our models (continued).

Table 4: Correlation of prototypes with Israel and Schönbrodt 2019 prototypes.

Appraisal	Political	Nonpolitical
suddenness	0.82**	0.9***
familiarity	0.28	0.53
predictability	0.64*	0.69*
pleasantness	0.99***	0.99***
goal	0.48	0.36
chance	-0.3	0.67*
intention	-0.13	0.69*
urgency	0.41	0.33
coping	0.73**	0.9***
power	0.36	0.35
adjust	0.53	0.95***
goal.pos	0.82**	0.95***
moral	0.81**	0.93***
norm	0.9***	0.82***

Third, we used the prototypes to calculate weights for each appraisals dimension. Weights can be interpreted as the importance these appraisals have in distinguishing the emotions from each other, while the prototypes can be interpreted as the positions of the five emotions in a multi-dimensional space. The weights were produced as follows. The model takes the 18 appraisal dimensions of the balanced training set as input vector, and calculates the sum of squared differences between the input vector and the 5 emotion prototypes (means of each appraisal per emotion family in the unbalanced training set).

We used the Differential Evolution (DE) algorithm, which is a genetic algorithm suited for multidimensional classification tasks (Storn & Price, 1997). The model assigns weights to each appraisal dimension, i.e. their importance in predicting the emotion families and converges at a solution that minimizes the distance scores between an input vector and the prototypes. Other model parameters in the DEoptim function (from the DEoptim R package) were set in accordance with Israel and Schönbrodt (2019). We ran the optimization procedure ten times across different random seeds. Its output was thus a 10 by 18 appraisals-weights matrix.

In this chapter, we report the weights with the best out-of-sample performance. The ten remaining parameter matrices were used to calculate the variance across the ten solutions to assess the reliability of the appraisal weights. If the ten random seeds produced a similar solutions, the variance across models would be small and we can conclude that the model has high reliability.

Fourth, to validate our models, we report the overall accuracy (proportion of correct prediction by total cases) and precision scores per emotion families (proportion of correct positive predictions by all positive predictions, true and false positives).

We discuss the results as follows. First, we describe and contrast the frequencies of the five emotion families in the nonpolitical and political domains. Second, we contrast the appraisal profiles of each emotion family in both domains. Theoretical predictions are derived from Scherer et al. (2001) and Scherer and Meuleman (2013). We focus specifically on the appraisal dimensions that should help us distinguish between two distinct emotions and test whether they do so in each domain. We then proceed to the computational modeling of emotions. Thus, third, we contrast the importance of each appraisal dimension per domain by comparing the weights assigned to each appraisal dimension in the optimization procedure. The weights in the two models tell us something about the relative importance of the appraisals. It is possible that even if we do not find strong differences in appraisal profiles, single dimensions, such as moral considerations, determine discrete emotions in politics. Such a pattern could be uncovered using the model weights.

Finally, we present the model diagnostics. Here, we test a simple hypothesis, namely that the model performs better in the nonpolitical than the political domain. To assess model performance, we report overall accuracy of the models, precision, recall and F1 scores per emotion class.

2.4 Results

First, which emotions did respondents select? There were substantial differences in which emotions respondents selected depending on which domain they were assigned to. Figure 3 shows the proportions of selected emotions per domain grouped by emotion family.

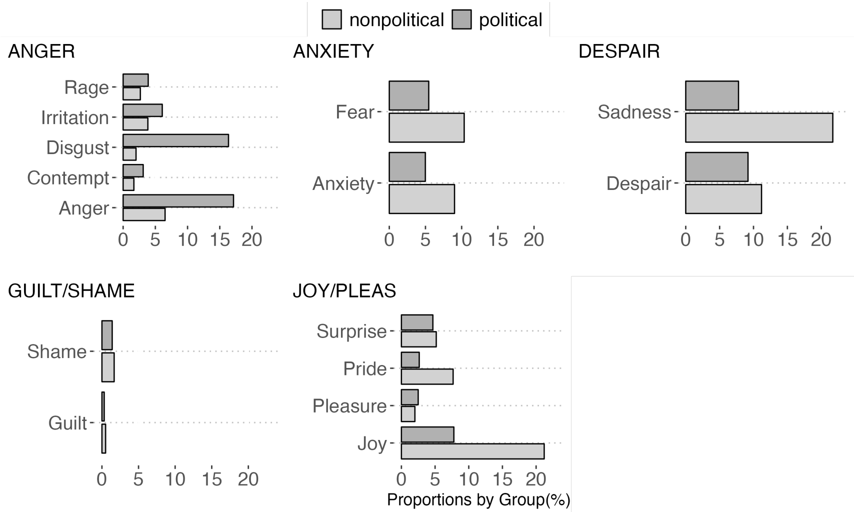

Figure 3. Selected emotions based on domain

Most notably, anger and disgust were the most frequently selected emotions in the political domain but rarely selected in the nonpolitical domain. In fact, all emotions that are part of the broader anger family were selected more frequently in the political domain than in the nonpolitical domain. For the other emotion families, the opposite pattern emerges. Emotions belonging to the anxiety, despair and joy families were selected often in the nonpolitical domain but less frequently in the political domain. Guilt and shame were selected rarely in either domain.

Next, we take a look at the targets of emotions, or the issues and situations that elicited an emotion. We asked participants to pick one of five (nonpolitical domain) or a multiple of eight (political domain) categories that best described the event[7]. This question was asked after the appraisal tasks had been completed. Figure 4 shows the proportions of targets in the political domain. A moment during a political campaign, a politician or candidate, a political issue and current domestic affairs

[7]The difference in choice options, single choice or multiple choices was due to a survey programming error. Ideally in the nonpolitical domain, participants would have been able to select multiple categories too but as the nonpolitical descriptions were much broader, we assume that participants were still able to meaningfully assign the events/situations to one of them.

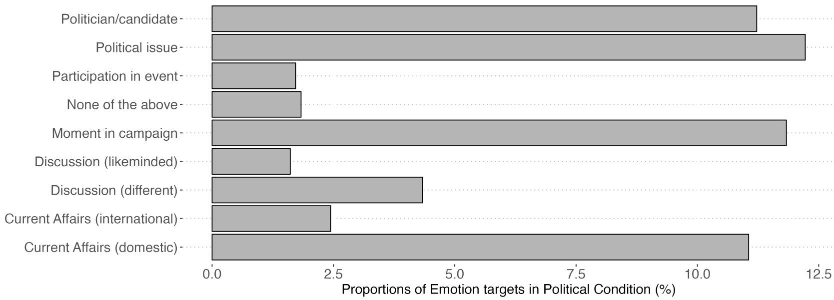

Figure 4. Emotion Targets in political domain

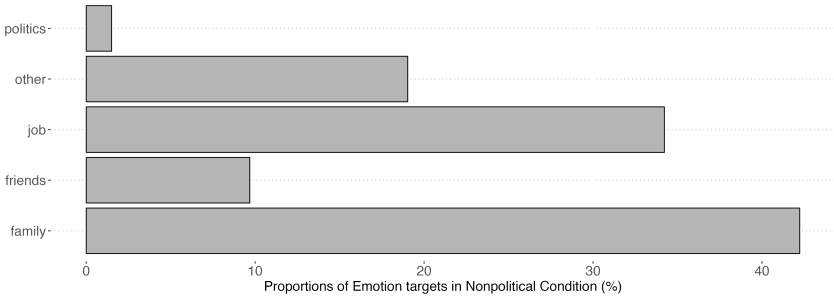

Figure 5. Emotion Targets in nonpolitical domain

were the most frequently selected categories. A moment during the campaign was most frequently selected together with other categories, with something a candidate had said (66 times), a political issue (50 times), or current affairs in domestic politics (49 times). The single most frequent combination was that of political issue and current affairs in domestic politics, which were selected together 97 times.

In the nonpolitical domain, family and respondents' jobs were named most frequently as descriptions of the emotion episode. Only 9 respondents in the nonpolitical domain described an event that was related to politics.

Appraisal Profiles in the Political and Nonpolitical Domains

We begin the discussion of our findings with the appraisal profiles of each emotion per domain. Appraisal profiles are the mean values of the 18 appraisals by emotion and domain of the full samples per domain. The appraisal profiles tell us which appraisals people have made per emotion, on average. We investigate whether we can distinguish the different emotion families based on their appraisal profiles in the political and nonpolitical domains. We focus on the distinction between anger and the other negative emotions but present the results for all tests in the online supplementary material (Figures B3–B5).

Figures 6-8 contain three pieces of core information: (1) the means, with 95% confidence intervals, of each appraisal dimension per domain (political and nonpolitical) and emotion, (2) the theoretical predictions based on Scherer et al. (2001) and Scherer and Meuleman (2013), and (3) the results of T-tests comparing the mean appraisal scores of the two emotions (corrected for multiple hypothesis testing).

Given the breadth of information included in Figure 6, we discuss a few key findings. Based on Scherer's theoretical predictions, we expect anger and anxiety to differ along 11 appraisal dimensions. In our sample, they differ along 8 of the expected dimensions in the nonpolitical domain. We do not find differences between anger and anxiety in terms of urgency appraisals (although mean urgency appraisal related to anxiety is higher than anger but the difference is not statistically significant), power and adjustment. The other appraisal dimensions differ as expected and in the expected direction. Anxiety is related to higher appraisals of goal relevance (M=0.85) than anger (M=0.75), moral acceptability (Anxiety: M=0.37, Anger: M=0.24) and chance as cause of the event (Anxiety: M=0.52, Anger: M=0.22). Anger, on the other hand, is related to higher appraisals of other behavior as cause (Anxiety: M=0.5, Anger: M=0.84), intentionality (Anxiety: M=0.34, Anger: M=0.64), control (Anxiety: M=0.5, Anger:M= 0.71), norm violations (Anxiety: M=0.35, Anger: M=0.71) and unjust outcomes (Anxiety: M=0.54, Anger: M=0.79).

Anger in the political domain follows the predicted pattern: moderate goal relevance (M=0.60) and suddenness(M=0.49), low pleasantness (M=0.12) and high goal obstruction (M=0.66), other responsibility (M=0.84), intentionality (M=0.73), control (M=0.72), and low moral acceptability (M=0.19), high norm violation (M=0.68) with unjust consequences (M=0.71). But anxiety largely follows the same appraisal pattern as anger, as indicated by the similarities in appraisal dimension means in Figure 6. In politics anxiety and anger are hardly distinguishable and only significantly differ in terms of one dimension: moral acceptability, which is lower for anger than anxiety. The other ten appraisal dimensions along which anger and anxiety should be differentiated do not differ. These null findings are striking especially in comparison to the nonpolitical domain. A few of them are particularly noteworthy. First, in politics, people report anxiety in response to events with highly unjust outcomes (M=0.64). Just like anger, political anxiety is related to appraisals of the responsibility of others (M=0.77), caused intentionally (M=0.71), and violated social norms (M=64). Based on these appraisal profiles, it appears essentially impossible to distinguish anger and anxiety based on the reported appraisal profiles in the political domain. In the nonpolitical domain, they can be clearly distinguished and the pattern closely follows the theoretical predictions laid out by the Component Process Model.

Next, we look at the appraisal profiles of anger and despair. For anger and despair, we would expect different appraisals on 11 dimensions. In the nonpolitical domain, we find statistically significant differences for nine of these dimensions. We do not find any differences for goal relevance (situations eliciting both emotions are appraised as high in goal relevance), predictability (although here, the predicted difference by Scherer is also small, from low-medium for anger to low for despair), and adjustment. Overall, anger and despair in the nonpolitical domain can be clearly distinguished and have distinct appraisal profiles.

In the political domain, we do not find distinguishable appraisal profiles. Anger and despair only differ in terms of goal obstruction (note that we do not theoretically predict a difference here) and goal relevance. The difference in goal relevance is, as predicted, higher for despair than for anger. Overall, however, the two emotions in the political domain are similar. Despair, like anxiety, looks much more like anger in terms of the appraisal profile. People reported feeling despair in politics even when a situation was intentionally caused by others, involved norm violations, had

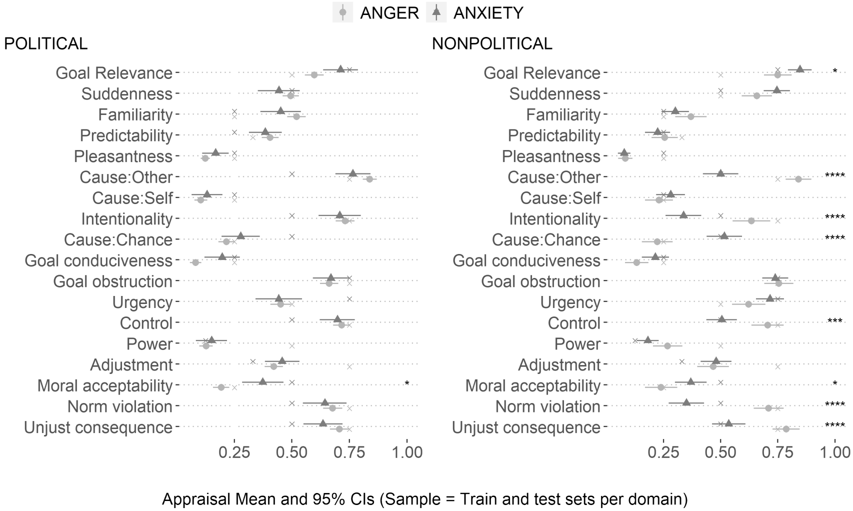

Figure 6. Appraisal profiles of anger and anxiety by domain (political vs. nonpolitical). Scherer's theoretical predictions are marked with an x (low = .25, medium/open = 0.5, high = .75) - see also Table 1. On the right side, we show the significance level of T-tests between the two emotions per domain correcting for multiple hypotheses testing using the Benjamini–Hochberg procedure.

unjust outcomes, and was not caused by chance.

Finally, we discuss the differences between anger and guilt's appraisal profiles, which are shown in Figure 8. We would expect anger and guilt to differ along ten dimensions but certainly the most striking difference should be in terms of the cause of the event. Guilt (and shame) should be related to appraisals of self-responsibility, while anger is related to appraisals of responsibility of others. In the nonpolitical realm, we find two differences that we do not theoretically expect (urgency and unjust consequences). The difference in the core differentiating dimension, whether the event was caused by one's own behavior or not, is statistically significant. Other differences are difficult to assess given the small number of guilt and shame responses in the nonpolitical domain. Strikingly, however, guilt and shame in the political domain are not related to appraisals of self-responsibility. In fact, the opposite is true: They are related to appraisals of others being responsible. In the

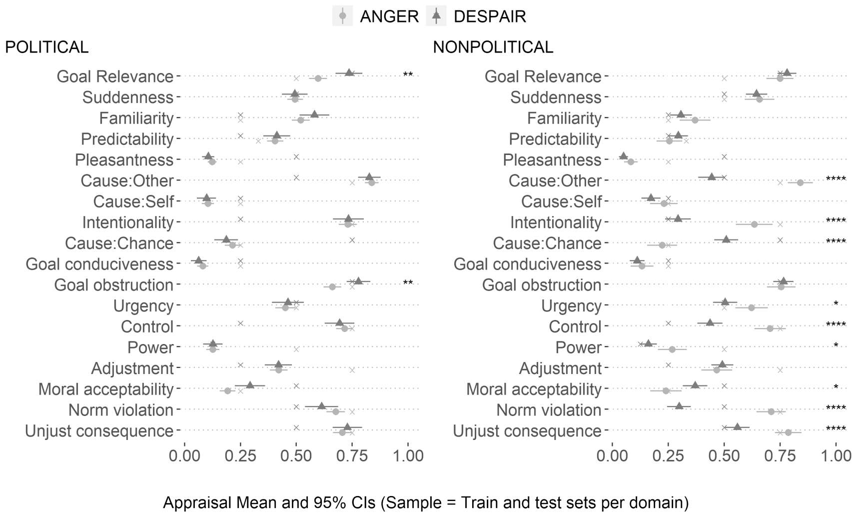

Figure 7. Appraisal profiles of anger and despair by domain (political vs. nonpolitical). Scherer's theoretical predictions are marked with an x (low = .25, medium/open = 0.5, high = .75) - see also Table 1. On the right side, we show the significance level of T-tests between the two emotions per domain correcting for multiple hypotheses testing using the Benjamini–Hochberg procedure.

political domain, guilt and anger differ on none of the appraisal dimensions. This strongly suggests that in politics people feel ashamed and guilty on behalf of others, likely a social group they identify with (E. R. Smith & Mackie, 2016), rather than for something they have done themselves.

The Relative Importance of Appraisals per Domain

Table 5 shows the weights for both nonpolitical and political models with the highest out of sample prediction. These weights are the outcomes of the optimization procedure described above and were ultimately used for our predictions of emotion labels. Weights can range from 0 to 10. We use two measures to contrast the political and nonpolitical domains. First, we look at the weights of each appraisal. A higher weight indicates a higher importance of that appraisal check in predicting the discrete emotion labels. Second, we look at the variance of weights across ten

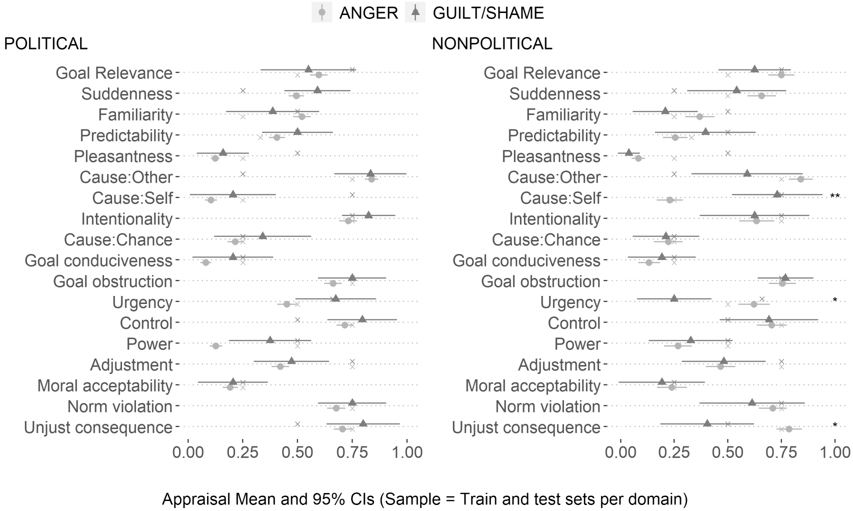

Figure 8. Appraisal profiles of anger and guilt by domain (political vs. nonpolitical). Scherer's theoretical predictions are marked with an x (low = .25, medium/open = 0.5, high = .75) - see also Table 1. On the right side, we show the significance level of T-tests between the two emotions per domain correcting for multiple hypotheses testing using the Benjamini–Hochberg procedure.

random seeds for the optimization algorithm. A higher variance is an indication of the lower reliability of the model.

The third column of Table 5 shows the differences in weights between the two domains. Values above 0 indicate that an appraisal check carries more importance in predicting nonpolitical emotions than political emotions, while values below 0 indicate a higher importance in the political domain.

The fourth column shows the differences in standard deviations of the nonpolitical weights. We first calculate the standard deviations per appraisal dimension across the ten iterations of the optimization procedure and then divide the difference between the final political and nonpolitical weights by the standard deviation of that appraisal dimension. We consider a difference between the two domains substantive if the political weight is either 2SDs larger or smaller than the nonpolitical weight.

The five most important appraisal checks in the nonpolitical domain are goal relevance, familiarity, goal obstruction, power, and urgency. In the political domain, the most important appraisal checks are goal relevance, goal conduciveness, predictabtility, moral acceptability, and pleasantness. Hence, in both domains, goal relevance is the most important appraisal dimension. This is in line with Scherer's (2001) Component Process Model, which grants a central role to this primary appraisal.

When it comes to secondary appraisals, however, there are substantial differences between the domains. We find the largest differences in familiarity, urgency and power. All three have high predictive weights in the nonpolitical model but not in the political model. The difference in weight of the power appraisal is in line with the appraisal profiles shown above. While power helped distinguish anger from anxiety or despair in the nonpolitical domain it did not do so in the political domain. In the political domain, power is generally lower than in the personal domain. Urgency is less important when events are mediated and are not happening as the emotion occurs.

At the bottom of Table 5 is predictability, which is important in the political domain but less so in the nonpolitical domain. To a lesser extent, we find differences in terms of control, and intentionality (both being more important in nonpolitical model), as well as goal conduciveness, other-responsibility, and suddenness (all three being more important in the political domain).

Next, we look at the variance in weights across the ten iterations of the optimization procedure, i.e. the reliability of our two models. The nonpolitical appraisal weights have variances between 0.435 and 4.647, the highest variance is found for the weights of power appraisals. Variances of the political appraisal weights range from 0.511 to 5.384, the highest being the weights for intentionality (var = 5.384) and control (var = 5.199). The variances for the political weights are overall slightly higher than those of the nonpolitical weights but the difference is negligible.

So while there are substantial differences in the relative importance of certain appraisals in the two domains, the models do not differ substantially in terms of their reliability, suggesting that the differences in weights we find is not random but indicates true differences in the importance of these appraisals in predicting

Table 5: Weights for 18 Appraisals with best out of sample prediction.

Appraisal	Political	Nonpolitical	Diff	SDsDiff
Power	1.15	8.3	7.15	3.31
Familiarity	3.42	8.7	5.28	4.80
Urgency	3.11	7.97	4.86	4.34
Intentionality	0.89	3.46	2.57	2.16
Goal obstruction	5.96	8.44	2.48	1.92
Norm violation	3.1	4.96	1.86	1.71
Control	1.95	3.7	1.75	2.36
Pleasantness	6.7	8.1	1.40	2.12
Cause:Chance	1.48	2.13	0.65	0.47
Goal Relevance	9.71	9.47	-0.24	-0.32
Cause:Self	3.81	3.05	-0.76	-0.87
Moral acceptability	6.71	5.34	-1.37	-0.95
Unjust consequence	6.52	4.7	-1.82	-1.46
Adjustment	6.13	4.07	-2.06	-1.03
Cause:Other	4.64	2.41	-2.23	-2.28
Suddenness	4.54	0.59	-3.95	-2.19
Goal conduciveness	9.1	4.72	-4.38	-2.81
Predictability	7.76	3.33	-4.43	-3.99

emotion labels.

In the final step, we evaluate the full model performances. We do this using several indices. Fist, we report overall model accuracy. Then, we discuss in more detail the precision, recall and F1 scores of each emotion class per model. As accuracy scores are not well suited to assess model predictions of unbalanced datasets, we focus on F1 scores to assess differences in model performance. Israel and Schönbrodt (2019) found an overall accuracy – proportion of correctly predicted cases by all cases – of 76.9% in their best performing model. Our nonpolitical model (M1) has an overall accuracy of 63.7%. The original study by Israel and Schönbrodt (2019) has a slightly higher accuracy. However, they allowed respondents to select two emotions and counted matches if either of the reported emotion matched their prediction, while we asked respondents to select only one emotion. Second, we included five emotion families, classifying anxiety and despair as distinct emotion families (see also Table 1 with theoretical predictions). Our approach is therefore significantly more conservative. Moreover, as accuracy is not the best measure to evaluate models predicting unbalanced data, we will mostly focus on class-wise precision, recall and F1 scores, which are presented below.

The nonpolitical model (M1) was best able to predict joy/pleasure with a precision (the proportion of true positives in all identified cases) of 95.9%, recall (the proportion of true positives by the sum of true positives and false negatives) of 91.7%, and an F1 score (the harmonic mean of precision and recall) of .93. Accuracy in the political model (M2) is significantly lower at only 45.8%. It is still better than chance, given that we predict five emotion classes but this is mostly driven by high scores of joy/pleasure (precision = 79.3%, recall = 77.2%, F1 = .76) and anger (precision = 74.10%). However, anger has low recall ad F1 scores. Overall, the political model performs worse than the nonpolitical model along all emotion classes.

The most frequent misclassification were cases classified as despair, although respondents had reported anger (77% of despair false positives were actually anger), and guilt (70.1% of guilt false positives were reported as anger), as well as anxiety (48% of false positives were self-reports of anger). As already indicated by the prototypes, the negative emotions in politics look too similar to be meaningfully distinguished by their appraisal profiles.

Despair and anxiety were selected much less frequently in the political domain than in the nonpolitical domain. Once the samples had been split into training and test data, the number of cases that could be used to train our political model was substantially reduced meaning that a few responses could affect the overall model performance. The nonpolitical domain contained significantly more despair and anxiety responses. Thus, to see if the distribution of emotions in the training set affected our results, we try to predict the political emotions based on the nonpolitical prototypes and weights. Additionally, this shows what happens when our assumptions are derived form a nonpolitical context and are applied to predict emotions in politics.

The resulting model (M3), predicting political emotions using weights derived from nonpolitical prototypes and the training set, shows an overall accuracy of 52.8%. While this overall accuracy is slightly better than that of the political model

(M2), it is still significantly worse than the nonpolitical model predicting nonpolitical emotions (M1). All precision scores of this model, with the exception of joy/pleasure, are lower than in the political model. Recall of anger and guilt/shame are higher than in the political model, but precision of guilt and recall scores of anxiety and despair are extremely low. It therefore does not show a substantive improvement over the political model (M2).

Table 6: Precision and Recall of three models

	Nonpolitical (M1)			Political (M2)			Mixed Model (M3)		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
ANGER	0.42	0.56	0.48	0.72	0.36	0.48	0.59	0.70	0.64
ANXIETY	0.45	0.52	0.48	0.23	0.21	0.22	0.44	0.12	0.19
DESPAIR	0.72	0.44	0.55	0.30	0.53	0.38	0.19	0.09	0.12
GUILT/SHAME	0.17	0.57	0.26	0.11	0.80	0.19	0.03	0.20	0.05
JOY/PLEAS	0.94	0.92	0.93	0.75	0.77	0.76	0.83	0.77	0.80

Exploratory Analysis

Finally, we explore what individual-level characteristics might predict how well we can classify emotions. As the test set is limited (299 observations in the political sample), we run simple T-tests instead of a full multivariate analysis. Above, we have outlined three reasons why emotions in politics are different from politics in the personal domain: (1) Politics is at its core about certain appraisals, such as blame attribution (Iyengar, 1994); (2) Politics is always mediated and not experienced directly; and (3) Politics is complex and people use heuristics to navigate it effectively. We explore four individual-level traits that are each related to these explanations: Political knowledge, political interest, internal efficacy and party identification. These tests are exploratory. We did not preregister any hypotheses and in some cases do not have strong theoretical expectations.

Political knowledge might help people make appraisals. If this is true, people whose emotions were correctly predicted should on average have higher political knowledge. Indeed, respondents that reported the same emotion that we predicted based on the appraisals had on average higher levels of political knowledge but the difference is not statistically significant ($t=-1.29$, $p=0.2$). Political interest could be related to our predictions in two ways: those high in political interest might make appraisals more often and know more about politics, increasing the likelihood of their emotions being predicted correctly. At the same time, they might be more likely to pick up emotion cues that tell them how they should feel but do not provide them with information about all appraisals. On average, participants who reported emotions in line with our model prediction had lower political interest ($t=2.08$, $p<0.05$). This means that politically interested participants more often have a mismatch between the appraisals they report and the emotion label they choose.

We have outlined above that internal efficacy might influence appraisals of coping potential and power but it could do so in both ways. In our sample, we find no difference in terms of internal efficacy ($t=0.86$, $p=0.4$). The relationship might be quadratic with only those with moderate levels of efficacy making control appraisals based on specific situations.

Finally, people with high levels of party identification are more likely to think of issues in terms of this group. Thus, we might suspect that people whose emotion

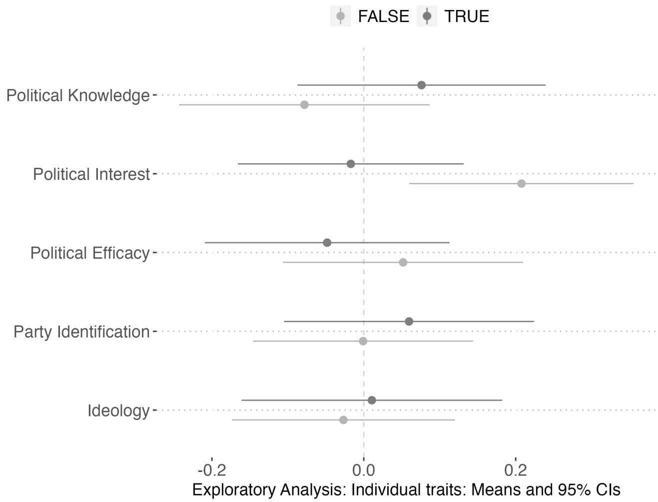

Figure 9. Exploratory Analysis: T-tests between individual level characteristics and predictions (correct = TRUE, or incorrect = FALSE) of our models predicting emotion labels.

was correctly predicted have, on average, lower levels of party identification. However, we find no differences between those with weaker or stronger party identification (t = -0.45, p=0.65), nor between those with more liberal or more conservative ideologies (t=-0.32, p-value = 0.75).

2.5 Discussion and Conclusion

Appraisal theories are one of the key frameworks through which we understand emotions in politics. Work on political emotions has relied on the core assumptions that emotions in politics work similarly as in the personal sphere – that they have similar appraisal profiles and action tendencies, and thus that we can use appraisal theory to make theoretical predictions about emotions in politics. Applying appraisal theory to politics has brought forward two key ideas (1) emotions are not irrational, but rather a response to our political environment; and (2) the discrete emotions anger and anxiety are distinct emotions with different effects on political behavior (Albertson & Gadarian, 2015; Valentino et al., 2011, 2008). However, this key assumption, that emotions work similarly in politics as in the personal domain, has not been systematically tested. In this chapter, we outline several reasons why the process of emotion elicitation differs in the political versus the nonpolitical domain. Empirically, we tested the two assumptions above by contrasting emotions, and their appraisal profiles and weights, in these two domains. We did this in two steps.

In the first part of the chapter, we assessed differences in appraisal profiles: the appraisals related to a distinct emotion. The Component Process Model provides clear hypotheses for which appraisal dimensions should help us distinguish between discrete emotions. We largely confirm these expectations in the nonpolitical sample but not in the political sample. One of the most striking findings of this chapter is that anger can not be distinguished from anxiety, despair or even guilt in terms of its appraisals. Anxiety is related to higher appraisals of moral acceptability of an event but apart from that, anger and anxiety appraisal profiles are largely identical. Despair is related to higher levels of goal relevance and goal obstruction, but otherwise also looks more like theoretical anger than theoretical despair. Guilt and shame do not differ from anger along any appraisal dimensions, suggesting that when people report guilt or shame in politics, they mean group-based but not individual guilt and shame.

This finding poses several questions for future research. Which other appraisals or individual characteristics determine whether people report anger or anxiety, when the underlying appraisals are not distinguishable? Can we expect the outcomes of political emotions to differ if the underlying appraisals are the same? Increasingly, scholars are looking at the differences between emotion self-reports and physiological responses to politics to move beyond studying how people themselves interpret and report their emotional state (Bakker et al., 2021; Karl, 2021). Similarly, an appraisal approach to emotions should distinguish between reported emotions and the underlying appraisal profiles and consider situations in which the two do not match and lead to contrasting behaviors. Should we, for instance, expect participants who report an anger appraisal profile but label their emotion as anxiety to show behaviors in line with anger's or anxiety's action tendencies? While appraisal theories suggest the former, other theories of emotions that put more emphasis on the individual interpretation, i.e. labelling of an emotion, would argue for the latter (Barrett & Russell, 2014). The answer to this question will tell us a lot about which political behaviors we can or cannot deduce from a person's self-reported emotion label.

In the second part of this chapter, we conceptually replicated past work that uses a computational approach to predict emotion labels. The second core finding of our study is that the model performs worse when predicting emotion labels in politics, both because the prototypes do not distinguish well between the different negative emotion families and because the relative predictive importance (weights) of the appraisals differ. This prompts the question to what extent appraisal theories can be applied to politics.

The most pronounced difference lies in appraisals of power and control, and their weights in predicting emotion labels. While power and control are central appraisals in appraisal theory, our findings suggest that they are not useful appraisals in politics: they neither help us distinguish emotions based on their appraisal profiles, nor play an important role in predicting emotions labels. Some existing work suggests that appraisals of power and control are perhaps better captured by stable feelings of political efficacy (Rico et al., 2020) or group identification (E. W. Groenendyk & Banks, 2014; Phoenix, 2019). Our exploratory analysis shows that we are neither more nor less likely to misclassify the emotions of people with higher political efficacy. This particular finding leaves ample room for future research into under which circumstances and at which level, individual or collective, power and control appraisals matter and take place.

Overall, this chapter has questioned the applicability of appraisal theory in politics. Past work has, if it tried to establish a causal link between emotions and political behaviors, relied largely on incidental emotions. Such emotions are deliberately unrelated to politics. However, the findings presented here suggest that political and nonpolitical emotions might differ substantially in their underlying appraisals. While incidental emotion manipulations are very well suited to isolate the effects of emotions, we suggest that future work should complement these findings by experimentally studying integral emotions (emotions related to a specific event stimulus) and emotions as they naturally occur. As this chapter has shown, naturally occurring emotions are much murkier, complicated by the political context, than clean experimental manipulations. However, studying them systematically will tell us a great deal about how emotions operate in the real world and will ultimately help the generalizability of established findings in the political emotions literature. Moreover, this chapter has shown that appraisals and emotion labels in politics do not always align. When studying integral emotions, future work should complement measures of emotion labels with measures of specific appraisals. Similar to work that complements emotion labels with the physiological component of an emotion episode (Bakker et al., 2021; Karl, 2021), work operating within the framework of appraisal theories should explicitly test its assumptions about the appraisals underlying an emotion label.