

The Influence of Emotion on Trust¹

Dan Myers² Dustin Tingley³

January 2, 2011

¹Research support from the Princeton Laboratory for Experimental Social Science made this research possible.

²Ph.D. Candidate, Department of Politics, Princeton University, Princeton NJ 08544. Phone: 609-258-6610, Email: cdmyers@princeton.edu, URL: <http://www.princeton.edu/cdmyers>

³Assistant Professor, Department of Government, Email: dtingley@gov.harvard.edu, URL: <http://scholar.harvard.edu/dtingley/>

Abstract

The willingness of individuals to trust each other is fundamental to economic and political life. In this paper we investigate the affective basis of trust. We explore whether and how manipulations of emotion influence behavior in the trust game. We also provide the first clear evaluation of a common method for manipulating emotions, the Autobiographical Emotional Memory Task (AEMT), to see how well it performs on different emotions as well as its ability to target specific emotions. Our findings suggest that negative emotions can decrease trust, but only if those negative emotions produce low certainty appraisals. Anxiety, a low certainty emotion, has a negative impact on trust while Anger and Guilt, two emotions that differ in their control-appraisals but induce the same high level of certainty, appear to have no clear effect on trusting behavior. We also find that the AEMT is ineffective with positive emotions and is not nearly as targeted as previous studies assume. To rectify this later problem we conduct mediation analyses to evaluate the influence of AEMT on trust via the targeted emotion. Our estimation strategy is the first to properly conduct mediation analysis for an outcome variable that is censored.

Key Words: autobiographical writing task, emotion induction, trust game, mediation analysis

1 Introduction

Political psychology has shown an increased interest in emotion in recent years, with a focus on the effects of emotion on cognition. While emotion has been a topic of study at least since Aristotle, much early research conceptualized emotion as the enemy of the sort of cool, calculated reasoning citizens would ideally engage in. This changed with more recent work that posited emotion as an aid to cognition (Marcus *et al.*, 2001). While other work has complicated this optimistic portrait of the effect of emotion (Verhulst and Sohlberg, 2009; Valentino *et al.*, 2008; Gadarian, 2010), the literature leaves no doubt that emotion has a powerful influence over how people acquire and process political information.

However, this focus on cognition has left little attention for other politically relevant processes¹. This paper aims to expand research in political science on emotion by studying the effect of emotion on trust. Political scientists have long been interested in trust as a basic element of political interaction (Levi and Stoker, 2000). Trust has been cited as an important element in everything from the success of new democratic institutions to Presidential approval (Mishler and Rose, 1997; Hetherington, 1998). Declining trust has been cited as evidence of political malease in the American public (Putnam, 2000).

There is reason to expect emotion to have a significant impact on trust. Emotions have been shown to affect a variety of decision making processes, and complex decisions, such as the decision whether to trust a stranger, a politician, or a potential competitor, are particularly likely to be influenced by a person's emotional state (Forgas, 2009). Emotion has been shown to affect related concepts like altruism, risk preferences, and the perceived likelihood of future events (Capra, 2004; Lerner and Keltner, 2001; DeSteno *et al.*, 2000). Finally, a variety of political processes that may be influenced by trust have been shown to be influenced by emotion, including the effects of frames and prejudice (Druckman and McDermott, 2008; Small and Lerner, 2008), the formation of attitudes about issues like immigration and terrorism (Lerner *et al.*, 2003; Brader *et al.*, 2008; Gadarian, 2010), and the way people choose sources from which to seek political information (Valentino *et al.*, 2008).

Nevertheless, the existing literature shows mixed support for the supposition that emotions can alter levels of trust. Capra (2004) finds that experimentally induced positive and negative moods have no effect on play in the trust game, a common behavior-based measure of trust that we use in this paper. However, Capra (2004) examines only the valance of emotions—whether they are generally positive or negative—

¹Though see, for example, Druckman and McDermott (2008) on risk preferences.

and ignores other dimensions of variation among emotions that may be as important as valance. Dunn and Schweitzer (2005) claim that emotional valance does affect trust, but only for emotions with a particular control appraisal. Specifically, emotions that induce a strong sense that another person is in control of the current situation increase or decrease trust, depending on their valance. However, this study relies on subjects' self-reported levels of trust, measures which have been shown to be poor predictors of experimental subjects' willingness to engage in trusting behavior (Glaeser *et al.*, 2000). Further, Dunn and Schweitzer (2005) examine only the effect of emotions' control appraisals on trust, ignoring the possibility that certainty appraisals—which represent a distinct dimension of emotional experience and describe the degree to which emotions make people certain or uncertain about their judgement of the present situation—can also affect trust. A complicating factor in these studies is that the manipulation used in both studies, the Autobiographical Emotional Memory Task (AEMT) may be poorly targeted, in that it tends to change subjects' levels of several distinct emotions in addition to the targeted emotion. Since several of these emotions may effect trust, this raises questions about the ability of studies using this manipulation to accurately measure the effects of specific emotions.

In this paper we build on these past studies in several ways. First, following Capra (2004) we measure the effect of distinct emotions using a behavioral measure of trust: the trust game. The trust game is a simple economic game that can be easily played by inexperienced laboratory subjects and is frequently used as a measure of the willingness to trust strangers. Second, following Dunn and Schweitzer (2005) we examine the effect of specific emotions on trust, but expand on their work by considering emotions that differ in the level of certainty they induce subjects to feel as well as the level of control they induce subjects to feel. Third, we are the first to properly estimate whether the AEMT has an effect across a range of emotions, as well as measure how targeted the task actually is. Fourth, because we find that the AEMT tends to change levels of several emotions in addition to the target emotion we use mediation analysis to measure the effect increased levels of the target emotion have on trusting behavior.

Contrary to Capra (2004), we find that emotions have an effect on trusting behavior. However, only one negative emotion, anxiety, reduces trusting behavior slightly, while others, including anger and guilt, have no significant effect. These results fail to support the findings of Dunn and Schweitzer (2005) that only emotions with an other-control appraisal tendency effect levels of trust. Instead, they suggest that the most important factor influencing the effect of an emotion on trust is the level of certainty that emotion induces people to feel. Interestingly, we find that the AEMT does a poor job of manipulating

levels of positive emotions in our subject population, suggesting the need for different manipulations that are able to target specific positive emotions.

The paper is structured as follows. Section 2 reviews previous work on the link between emotion and trust and charts out differences in emotions along the dimensions of control and certainty. Section 3 discusses the use of AEMT in experimental studies and relates this manipulation to naturally occurring political and economic processes. Section 4 introduces our experimental design and Section 5 presents our results which implements new algorithms for mediation analysis (Imai *et al.*, 2010a). Section 6 concludes.

2 Emotions and Trust

While emotion has long been recognized as one of the basic elements of human psychological experience, the cognitive revolution in psychology pushed the study of affect to the margins. However, in the last few decades a number of findings and theoretical models have suggested that emotion could have a powerful influence over cognition and decision-making (Lewis *et al.*, 2009, for a review see), with different emotions having different effects (Lerner and Keltner, 2000). While humans can experience a diverse range of emotions, researchers have identified several dimensions of emotional experience that can be used to characterize distinct emotional states. The effect of different emotions on behavior, including behavior in the trust game, depends on where the emotions sit on these dimensions.

2.1 Characterizing Emotions

Early research focused on the valance, or positivity or negativity, of different emotions. The Affect Infusion Model claims that a person's current emotional state alters their appraisal of new stimuli by nudging them into the direction of the valance of the emotion (positive or negative) that is already being experienced (Forgas, 1995). This happens through several processes. The first operates through memory, as a particular emotional state makes associated mental representations more accessible (Bower, 1981; Eich and Macaulay, 2000), and thus more likely to be used to judge new stimuli. The second process posits that affective state operates as a easily accessible heuristic that serves as information that is used to judge new objects (Schwarz and Clore, 1983). According to this Affect-As-Information theory, a person experiencing a negative mood may take his mood as a sign that he does not like a new stimuli, even if that stimuli is unrelated to his reason for being in a negative mood. Finally, affective states can

influence how people process information, with some negative affective states inducing more effortful and analytic processing strategies than positive affective states.

Affect infusion suggest that people in positive moods will be more trusting. People in positive moods will have an easier time accessing memories where trusting behavior led to a positive outcome. They may take their positive mood as information about the trustworthiness of others, and about the possible consequences of engaging in trusting behavior. These expectations are supported in several studies of the effects of mood on social interaction. People in positive mood are more altruistic and make more optimistic predictions about future events, such as the future actions of others (Hertel and Fiedler, 1994). Similarly, Forgas (1998) shows that positive mood makes experimental participants more cooperative in negotiating situations. However, as noted above, Capra (2004) finds no effect of mood on play in the trust game, which is puzzling given this earlier work.

While the Affect Infusion Model relies primarily on the distinction between positively and negatively valenced affect, other research has shown that distinct emotions with the same valance can produce wildly divergent effects on cognition and decision-making (DeSteno *et al.*, 2000; Lerner and Keltner, 2001). Drawing on the work of Smith and Ellsworth (1985), Lerner and Keltner (2000) argue that emotions differ in the way they lead people to appraise future events along several dimensions² in addition to the valance dimension. Emotions that share a valance may nevertheless vary on other dimensions that are important for determining their effect. While the way these dimensions have an impact on cognition and decision-making are similar to the way valance has an effect, the effects may lead similarly valenced emotions to have very different effects on how people think and act when these other dimensions are relevant to the stimuli at hand. For example, anger and anxiety are both negative emotions, but anxiety is characterized by a low degree of certainty while anger is characterized by a high degree of certainty (Smith and Ellsworth, 1985). Because of this difference in certainty, Lerner and Keltner (2001) finds that these emotions, both of which have a negative valance, have opposing effects on risk attitudes.

2.2 Valance, Control, Certainty and Trust

While studies have examined the impact of emotions on a variety of types of cognition and decision making, the effect that emotions that vary along these dimensions have on trust is almost unstudied. The exception is Dunn and Schweitzer (2005), which argues that emotions' control appraisals will determine

²These dimensions are : certainty, attentional activity, anticipated effort, situational control, responsibility and pleasantness.

their effect on trust. However, Dunn and Schweitzer (2005) limit their study to this control appraisals. We argue that in addition to control appraisals, emotions' certainty appraisals will be important determinants of their effect on trust.

Only one study that we are aware of examines the effect of affect on play in the trust game. Capra (2004) induces subjects to feel either positive or negative moods using the AEMT, and then has them play a series of economic games including the trust game. She finds no effect of the mood induction on play in the trust game. While this shows that positive or negative valence by themselves have no measurable effect on trust, it does not differential between specific negative and positive emotions. This leaves open the possibility that other dimensions of emotional experience can have an impact on trusting behavior.

While several of the dimensions identified by Smith and Ellsworth (1985) appear to be tangential to trusting behavior, Dunn and Schweitzer (2005) argue that those related to control appraisals are theoretically the most likely to influence trust. An emotion's control appraisal determines who, or what, a person experiencing that emotion will tend to believe is in control of a new situation. People experiencing emotions characterized Situational Control will tend to believe that the context of a situation or other factors that are not attributable to a particular person will determine the results of a new situation. Emotions characterized by Individual Control lead people to believe that an identifiable agent has control over a situation. This agent may be themselves (in which case Dunn and Schweitzer (2005) characterize the emotion as "Individual Control—Self") or another person (in which case Dunn and Schweitzer (2005) characterize the emotion as "Individual Control—Other"). For example, anxiety is characterized by situational control, while guilt is characterized by individual control-self and anger by individual control-other.

Based on these distinctions, Dunn and Schweitzer (2005) argues that emotions that lead people to believe that either themselves or impersonal contextual factors are in control will have no impact on trust. On the other hand, people experiencing emotions that lead them to believe that another person is in control will use their emotional state to decide whether or not to trust. "Individual Control—Other" emotions with a negative valence are expected to decrease trust as the potential trustee is viewed in a negative light; "Individual Control—Other" emotions with a positive valence are expected to increased trust as potential trustees are viewed in a more positive light. Dunn and Schweitzer (2005) tests this claim by inducing subjects to feel emotions that vary in their control appraisal and then measuring their self-

reported willingness to trust another person. Their findings support the claim that control and valance interact to effect trust. Specifically, they find that anger and gratitude, emotions with strong “Individual Control—Other” appraisal, influence self-reported levels of trust, while guilt and pride, similarly valanced emotions with strong “Individual Control—Self” appraisals have no effect on trust. However, Dunn and Schweitzer (2005) rely on self-reported levels of trust; as noted before, these self-reports have been shown to have little relation to trusting behavior (Glaeser *et al.*, 2000).

While Dunn and Schweitzer (2005, 's) findings about control are important, we believe that they ignore another dimension of emotional experience that is likely to influence trust: certainty. Certainty is “the degree to which future events seem predictable and comprehensible (high) vs. unpredictable and incomprehensible (low)” (Lerner and Keltner, 2000, pg.479). This dimension has been shown to have a variety effects on everything from cognition to risk preferences to political attitudes (Marcus *et al.*, 2001; Lerner and Keltner, 2001; Brader *et al.*, 2008). When a person experiences a high certainty emotion, we expect their affective state to have little impact on trust. However, when they experience a low certainty emotion, we expect them to use the valance of their current affective state to predict the outcome of future events. People experiencing a positive-valanced low-certainty emotion will expected better outcomes; those experiencing a negatively-valanced low-certainty emotion should expect worse outcomes. Thus low-certainty, negative valanced emotions should decrease trust, while low-certainty, positively valanced emotions should increase trust.

Note that while the logic of this agreement is similar to that in Dunn and Schweitzer (2005), expectations that control appraisals will influence trust and expectations that certainty appraisals will influence trust produce very different predictions about the effect of specific emotions. For example, anxiety is a negatively valanced, low-certainty, situational control emotion. While anxiety’s level of situational-control would lead us to expect it to have no effect on trust, its low level of certainty would lead us to expect a negative effect on trust. Similarly, anger shares a negative valance with anxiety, but is a high-certainty, personal control-other emotion. If control appraisals are the important determining factors for the effect of emotion on trust, anger should produce lower trust than anxiety, and its control appraisal leads the potential trustor to use their negative emotional state to evaluate the potential trustee. If, on the other hand, certainty appraisals are the important determining factor anxiety should produce lower trust than anger; in this case anxiety’s low certainty appraisal would lead the potential trustor to use their negative emotional state to evaluate the potential trustee.

Emotion	Type of Control	Certainty	Valance	Predicted Effect on Trust	
				From Control	From Certainty
Anger	Individual Control - Other	High	Negative	↓	None
Guilt	Individual Control - Self	High	Negative	None	None
Anxiety	Situational Control	Low	Negative	None	↓
Happiness	Individual Control - Other	High	Positive	↑	None
Self-Assuredness	Individual Control - Self	High	Positive	None	None

Table 1: Characterizing Emotions by Valance, Control and Certainty

Table 1 shows the five emotions that will be examined in this study, how they are characterized along the dimensions likely to effect trust, and the competing predictions made about effect on trust according to their control type and level of certainty. We test three negative emotions, anger, guilt and anxiety. According to the theory that an emotion’s type of control determines its effect on trust, anger should have a negative effect on trust while the other two emotions will have no effect. On the other hand, if an emotion’s level of certainty determines its effect on trust then anxiety will reduce trust. Using positive emotions we test only the predictions of the control theory, with happiness expected to produce a positive effect on trust. Combined, these emotions will allow us to test whether Dunn and Schweitzer’s (2005) findings about control hold in the context of the trust game and compare the effects of control to our predictions about the effect of certainty.

3 Manipulating Emotion

Our theory predicts very different effects from different but related emotions. Thus political and economic events that lead to these different emotions may have very different effects on how people behave. Testing this theory in a way that allows us to make causal claims about the role of emotion in changing levels of trust *ideally* requires an experimental manipulation that allows us to manipulate levels of a specific emotion while avoiding changing levels of related emotions. For example, testing the effect of anxiety requires that we be able to assign some experimental subjects to a high-anxiety condition. However, if such subjects also feel heightened levels of anger or depressed levels of self-assuredness it would be difficult to determine the effect of increased anxiety, since observing reduced levels of trust might be attributable to reduced levels of self-assuredness. Similarly, a null finding might mask a negative effect of anxiety on trust behind a positive effect of increased anger levels. Next we show that existing studies

generally fail on these grounds and how our new experimental design and analyses offer some important advantages.

Most studies have used the Autobiographical Emotional Memory Task (AEMT) to achieve this effect. A number of other methods are available to induce emotions in experimental subjects, including having them watch movie clips (Gross and Levenson, 1995), pictures (Mikels *et al.*, 2005), interact with confederates (Ax, 1953), and even undergo hypnosis (Houghton *et al.*, 2002). However, none of these techniques promise the specificity of the AEMT, which can ask subjects to recall a wide variety of emotions and focuses them on only the emotion that the researcher wants them to experience. Further, the AEMT induces emotion using a technique that mirrors an activity, remembering past experience, that is a frequent part of most people's life. This technique is particularly appropriate for testing the impact of emotion in the political arena, which frequently appeals to citizen's memories as a way of making appeals ("Are you better off than you were four years ago."). The use of rhetoric to remind people of their emotional memories has long been a powerful tool in politics (e.g. Lim, 2002), and the AEMT mimics this process albeit in a political neutral way.

The AEMT draws on research on how remembering past experiences can influence a person's current emotions. Strack *et al.* (1985) show that recalling past experiences can cause subjects to experience a similar type of affect as that originally generated by the experience. However, for this to happen subjects must describe the events vividly; brief recall of the events does not cause a change in affect. Further, a subject's description of the past events must focus on how the event happened, instead of speculating on why it happened. The AEMT achieves this kind of recall by asking subject to respond to some variation on the following prompt:

We'd like you to describe in detail the one situation that makes you (or has made you) most anxious. This could be something you are presently experiencing or something from the past. Begin by writing down what you remember of the anxiety-inducing event and continue by writing as detailed a description of the event as is possible. If you can, please write your description so that someone reading this might even feel anxious just from learning about the situation.

WHAT is the thing that makes you the most anxious? WHY does it make you so anxious?

Please write for several minutes.

By asking subjects to vividly recall their experiences in a way that focuss on what happened, and not why it happened, this task should cause subjects to relive the affective experience they are describing. Since this task asks subjects to think about the experience of the specific emotion desired by the experimenter, it seems possible that this task can pinpoint emotions in the way needed to test the effect of emotions on trust.

However, the standard manipulation check performed by most studies using the AEMT provides no evidence for a targeted effect (Lerner *et al.*, 2003; Lerner and Gonzalez, 2005; Small and Lerner., 2008). In general, these studies compare the effects of two emotions on a behavior or attitude. As a manipulation check, these studies solicit self-reported measures of the two emotions under study, then compare the level of emotion 1 among subjects induced to feel emotion 1 to the level of emotion 1 among subjects induced to feel emotion 2, and vice versa. For example, Lerner *et al.* (2003) induces subjects to feel either fear or anger, finds that subjects in the fear condition report more fear than subjects in the anger condition, and declare the fear manipulation successful. Likewise, they find that the subjects who wrote about an anger-inducing memory had higher levels of anger than subjects who wrote about a fear-inducing memory and conclude that the anger manipulation was successful.

Such a test demonstrates that the fear manipulation increases fear more than the anger manipulation increases fear, and vice versa. However, because this test lacks a control condition it does not prove that the effect is highly targeted. It is possible that the anger manipulation also increases (or decreases) fear; likewise it is possible that the fear manipulation changed levels of anger. If so, any effect of the manipulation can not be solely attributed to the target emotion, and a null finding might indicate cross-cutting effects of two emotions. It is also possible that one or both manipulation changed levels of a third emotion (for example, happiness), that might cause any observed effect or suppresses any real effect of the target emotion. In either case, the effect of the targeted emotion would be inaccurately estimated.

A few studies include a control condition in their manipulation checks, but the evidence they provide is mixed. DeSteno *et al.* (2004) find that writing about an anger-induced memory induced higher levels of anger than writing about a sadness-inducing memory, and vice-versa. Nevertheless, both writing tasks cause higher levels of anger and of sadness than a neutral control condition, indicating that the anger manipulation also increased sadness and the sadness manipulation also increased anger. However, Bodenhausen *et al.* (1994) conducts a similar test and finds no collateral effect of the anger manipulation on sadness or the sadness manipulation on anger. Neither of these studies tests for the possibility of

both manipulations affecting a third emotion. Lerner and Keltner (2001) does so by testing the effects of fear and anger versions of the AEMT on 16 emotional terms³ and found that, besides fear and anger, subjects in the two conditions reported no difference in the level of each emotion experienced. However, this manipulation check included only 13 subjects split between two conditions, raising doubts about the confidence we can place in this null finding. Further, since the study does not use a control condition, it can not rule out the possibility that both manipulations had a similar effect on a third emotion.

In the absence of pinpoint accuracy, estimating the effect of a manipulation on some behavior via a specific emotion by simply calculating the average treatment effect of the manipulation requires assuming pinpoint accuracy of the manipulation on the intervening emotion. Previous studies have not demonstrated that the standard mood inductions can do this. Similarly, previous studies often fail to capture a broad range of emotions. Our design attempts to address both of these problems. This proceeds in two steps. First, we manipulate the emotions identified above, emotions selected to allow us to test the effect of control and certainty appraisals on trust. Second, we use mediation analysis to examine the the manipulation's effect on trust behavior via particular emotional pathways while allowing for the manipulation to have an effect via other pathways as well. This lets us see how events that may induce emotions—here the common autobiographical writing task—have an effect on behavior through changes in an individual's affective state.

Finally, the ability of the AEMT to induce positive emotions is even less proven than its ability to induce negative emotions. Most studies employing the AEMT study the effects of emotions with negative valance; anger, sadness and anxiety are the most common. A few studies compare the effects of positive emotions against the effects of negative emotions and include pre-tests intended to measure the success of this manipulation at inducing these emotions. These pretests show that subjects who write about positively valanced memories show higher levels of positive affect than subjects who write about negatively valanced memories. However, none of these studies tests the effect on positive emotions against a neutral control condition, leaving open the possibility that the positively valanced manipulation has no effect while the negatively valanced manipulation decreases positive affect. Given this, a contribution of this paper is to systematically test the ability of the AEMT to induce positively valanced emotions.

³The exact number of distinct emotions these terms were used to measure is unclear.

4 Experimental Design

We used the autobiographical emotional memory task to induce subjects to feel five emotions: anger, anxiety, guilt, happiness and self-assurance. For example, subjects in the anxiety condition responded to the prompt listed in Section 3. Subjects in the control condition responded to a similar prompt that asked them to write about a novel they had recently read.

The experiment took place in the Princeton Laboratory for Experimental Social Science (PLESS) and involved 249 undergraduate students recruited through the PLESS subject pool. When subjects signed up for the study they completed a pre-survey questionnaire that included two survey-based measures of trust. The first, drawn from the General Social Survey, measured whether the subject thought that other people were, in general, trustworthy. The second battery, taken from Glaeser *et al.* (2000) asked subjects to report how frequently they engaged in a number of trusting behaviors.⁴ The pre-survey contained no mention of emotion.

Upon arrival in the lab, subjects completed the Autobiographical Emotional Memory task. After they finished this task, subjects completed a slightly abbreviated version of the PANAS-X battery. PANAS-X is an extensively validated measure of state emotion which presents subjects with a number of emotional adjectives and asks them to “indicate how much you feel this way right now” on a one to five scale.⁵ Responses to these items are combined into subscales of 3 to 10 adjectives that measure general positive and negative affect as well as 11 other distinct affective states.

This study uses a version of the PANAS-X with 40 items, including items needed to measure the five subscales associated with the emotions we were manipulating as well as all items needed to measure general positive and negative affect. The scales, their ranges and consistent items are listed in Table 2⁶. This Table also lists the theoretical construct described above that we are using the scale to measure, some of which have slightly different names than the names used by the PANAS-X’s authors. Most studies employing the autobiographical emotional memory task that include a manipulation check use subject self-rating on a single emotional adjective and do not offer any form of external validation of this measure; using the PANAS-X should give us a more accurate measure of the effect of the manipulation

⁴Both batteries are available in Appendix D.

⁵The PANAS-X is also commonly used as a measure of trait emotion; in this case the phrase “how much you feel this way right now” is changed to “how much you felt this way during the last few weeks” or some other time-duration.

⁶Further details of the PANAS-X procedure are available in Appendix E.

PANAS Scale	Construct Measured	Range	Items
General Positive	General Positive Affect	10-50	active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, strong
General Negative	General Negative Affect	10-50	afraid, scared, nervous, jittery, irritable, hostile, guilty, ashamed, upset, distressed
Fear	Anxiety	6-30	afraid, scared, frightened, nervous, jittery, shaky
Hostility	Anger	6-30	angry, hostile, irritable, scornful, disgusted, loathing
Guilt	Guilt	6-30	guilty, ashamed, blameworthy, angry at self, disgusted with self, dissatisfied with self
Joviality	Happiness	8-40	happy, joyful, delighted, cheerful, excited, enthusiastic, lively, energetic
Self-Assurance	Self-Assurance	6-30	proud, strong, confident, bold, daring, fearless

Table 2: PANAS Subscales

on subjects' affective state.

Once all subjects in a session completed the PANAS-X, the experimenter read them instructions for the trust game portion of the experiment, reproduced in Appendix F. In this version of the trust game, modeled after the one used in Cesarini *et al.* (2008), subjects in the first position (the “sender”) were given five dollars. They could then select a portion of this five dollars to send to the person in the second position (the “receiver”). Any money sent to the receiver was tripled. The receiver could then return any portion of the tripled amount to the sender. The Nash Equilibrium prediction for this game is for the sender to keep all five dollars, and for the receiver to keep any amount that the sender does send. However, as noted above, studies of the trust game have frequently found that subjects send and return significantly larger amounts.

Subjects were told that they were being randomly matched with another participant, and within this pair randomly assigned to be either the sender or the receiver. Senders and receivers selected their action simultaneously. Senders selected an amount to send to the receiver from a list of amounts from \$0.00 to \$5.00 in fifty cent increments. Receivers selected an amount to return using the strategy method by selecting an amount that they would return given each possible amount the sender could send to them. Using the strategy method gave us an opportunity to observe potentially heterogeneous effects of

emotions depending on the amount sent by the sender. Additionally, it allowed us to wait until the end of the experiment to tell subjects the action that their partner took. This ensured that the results of the trust game did not affect subjects behavior in the rest of the experiment.⁷

After the trust game, subjects completed a questionnaire that included the same measures of trust as the pre-study survey. They were then told the action of their partner in the Trust Game and paid in private. Average earnings, including a \$5 show-up fee, were \$11.72.

5 Empirical Analysis

5.1 Effects of writing tasks on emotions

We begin by examining the influence of the different treatment conditions on the various emotion scales. Figure 1 plots the mean and 95% confidence intervals for the various negative emotions and Figure 2 plots the results for positive emotions.⁸ Each plot represents a separate emotion with the treatment conditions listed along the x-axis. For example, the top left plot of Figure 1 represents the average levels of the *General Disposition-Negative* measure across the treatment conditions. For the Anxiety, Anger, Happy, Self-Assured, and Guilt treatments, the key comparison is to the Control condition.

The Anxiety, Anger, and Guilt conditions each produced significantly higher levels on the *General Disposition-Negative* scale compared to the control condition. The Happy and Self-Assured conditions produced slightly lower levels but these differences were not significant. The Fear scale saw similar result except that the Anxiety treatment was the only one with non-overlapping confidence intervals though on average the Anger and Guilt conditions produced higher levels of fear compared to the control. For the Hostility scale, the Anger treatment had the most profound effect. While the Anxiety treatment also produced elevated levels of hostility, the effect of the Anger treatment was the largest. Finally, while the Anxiety and Anger treatments produced higher levels of Guilt relative to the control, the largest effect was from the Guilt treatment. Interestingly, then, there was some degree of correspondence between the emotion scale and the type of treatment the produced the largest effect on the scale. The largest response

⁷For example, subjects in the sender position who recited no money back from their partner in the receiver position might feel angry. This feeling of anger might contaminate responses on the survey-based measures of trust that were asked after the trust game.

⁸Figures 11 and 12 in Appendix A present analogous violin plots to portray the distributions of these variables more completely.

on the Fear scale was from the Anxiety treatment, for Hostility the largest effect was the Anger treatment, and for Guilt the largest effect was the Guilt treatment. While promising, each of the treatments also had effects along other emotional scales. Importantly, this is despite the fact that each of these scales use different inputs.

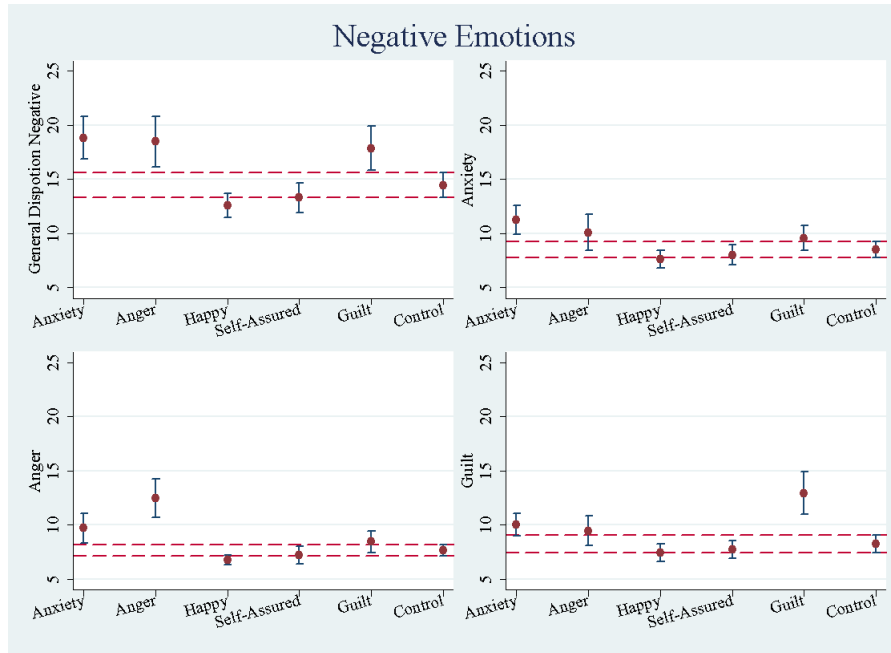


Figure 1: Effects of Emotion Manipulations on Negative Emotions. Means with 95% confidence intervals.

Turning to the positive emotions we see a very different picture in that the Happy and Self-Assured treatment conditions regularly have no effect compared to the control condition. These results are not likely the result of ceiling effects, as possible scores range from 10-50 on the general positive scale, 8-40 on the happiness scale, and 6-30 on the self-assuredness scale. While published research has previously reported success in using this manipulation to increase levels of positive emotions (e.g. Dunn and Schweitzer, 2005), communication with other researchers confirms that inducing positive emotions is frequently harder to accomplish, particularly among US student populations.⁹ Interestingly, most of the negative treatments have little impact on positive emotions either. Indeed, the Anger treatment had nearly identical averages to the Happy and Self-Assured treatments. The only significant comparison with the control for the negative treatments was the Guilt treatment on Joviality. Given that research has found levels of positive and negative affect to be only weakly correlated, this result is not surprising. Nevertheless, knowing that manipulations targeted at negative emotions are not also changing levels of

⁹Personal conversation, Jennifer Lerner via Rose McDermott

positive emotions should provide some comfort to researchers using this manipulation to test the effect of negative affective states

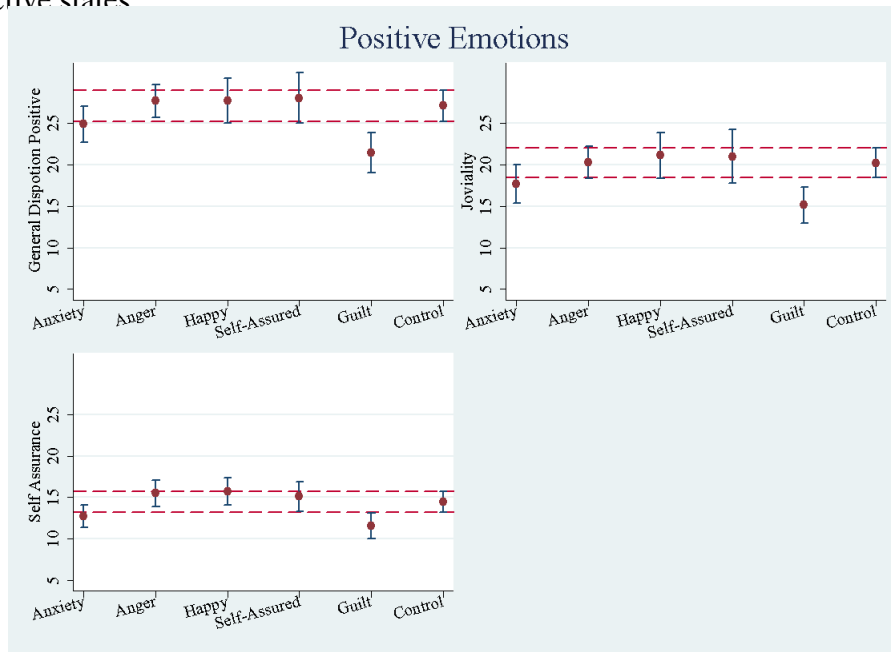


Figure 2: Effects of Emotion Manipulations on Positive Emotions. Means with 95% confidence intervals.

These results show mixed success in using the AEMT to manipulate subjects' emotions in the targeted way needed by most studies examining the effect of specific emotions. Looking at attempts to manipulate negative emotions, the AEMT succeeded in increasing levels of the targeted negative emotion. However, it also increased levels of negative emotions other than the targeted emotion and lowered levels of positive emotions. The means that any observed effect of the manipulation on an outcome variable of interest might work through the manipulated emotion, or might work through one of these collateral emotions. For example, the effect of having subjects write about an experience that made them feel angry might operate through increased levels of anxiety, or it might operate through increased levels of guilt or reduced levels of happiness. The AEMT failed to manipulate positive emotions in any meaningful way. This suggests that alternative ways of manipulating specific positive emotions are needed; if the AEMT is used, a through manipulation check is needed to demonstrate that the manipulation has succeeded. Previous studies have not done these things.

5.2 Effects of writing tasks on trust behavior mediated through emotion changes

5.2.1 Framework

Next we turn to whether the autobiographical writing task manipulations have any influence on behavior via changes in emotional states. Do the changes in the emotional states generated by the negative manipulations then have an influence on behavior? To answer this question we had our subjects play the simple trust game described above. Since the experimental manipulations changed subjects' levels of several different emotions, not just the targeted emotion, we conduct a causal mediation analysis to see how much of the effect of the manipulation on trust behavior is mediated by changes in the targeted emotional states, allowing for the effect to be transmitted through other pathways. Graphically, this can be represented in Figure 3. This shows that the treatment can effect the outcome either through the mechanism of interest, or through other mechanisms that are subsumed in the line directly connecting the treatment variable to the outcome. We will test the theories described above about the effects of anxiety, anger and guilt on trust by looking at how much of the manipulation's effect is mediated through these emotions. Since the manipulations failed to change subjects' levels of positive emotions we will not examine the effect of positive emotions on trust.

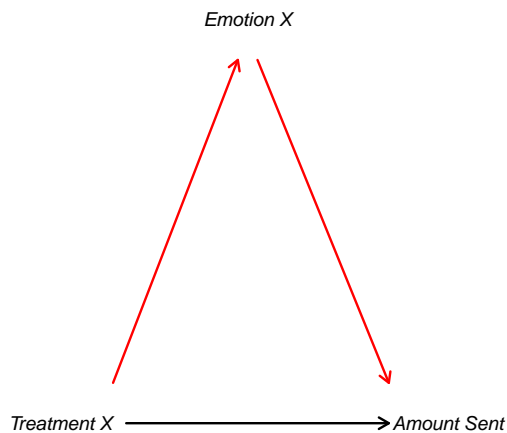


Figure 3: Graphical representation of mediation relationship allowing for treatment to influence amount sent through changes in a particular emotion and through other pathways.

Before beginning formal estimation we anticipate our estimated results with some descriptive statistics for the anxiety, anger, and guilt treatments in Figures 4, 5, and 6, respectively. The left pane plots

the distribution of amount sent by whether the emotion was induced or was in the control condition. The differences in the means of these distributions would give the average treatment effect of the induction. The middle pane plots the distribution of the targeted emotion by the treatment condition, and the right pane plots the distribution of emotions by whether the full amount was sent versus some lower amount. Figure 4 shows there was relatively little difference in amount sent between the anxiety induction and control condition. However, the induction did have an impact on anxiety and those who sent less than the full amount tended to have higher levels of anxiety. Figure 5 shows a similar pattern, except that there was little difference in anger levels between those who sent the maximum amount and those who sent less. Finally, Figure 6 shows that the amount sent by those in the guilt condition was less than the control condition. However, the right pane shows relatively similar levels of guilt between those who sent the maximal amount and those who sent less. These results provide suggestive evidence that the anxiety induction may have had a small effect on trust behavior through changes in anxiety, but that there was no total treatment effect. Conversely, the guilt induction appears to have had a total treatment effect but this effect does not appear to be due to changes in guilt emotions per se. To investigate these patterns more thoroughly we must move to a causal mediation analysis. In particular, we need to estimate the effect of the mood induction that goes through the particular targeted emotion.

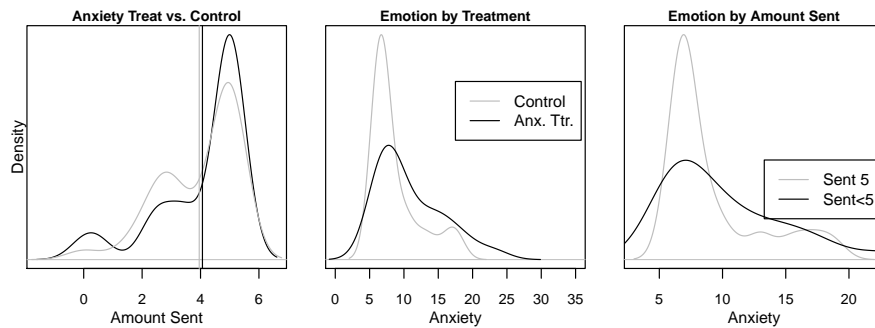


Figure 4: Distributions of amount sent and emotions by amount sent and treatment status for the anxiety induction treatment versus control condition.

While an extensive review of causal mediation analysis is beyond the scope of the current paper (Imai *et al.*, 2010d,a), we briefly review some relevant details and provide a technical appendix for interested readers.¹⁰ A key point of this discussion is that this approach allows for the treatment variable to have an

¹⁰Mediation analysis (Judd and Kenny, 1981; Baron and Kenny, 1986; MacKinnon *et al.*, 2002; Kenny, 2008) is commonly conducted in psychology experiments but rarely outside of that discipline. Imai *et al.* (2010d,a) show that these earlier methods face substantial problems including erroneous estimation techniques in non-linear models and unstated identification assumptions.

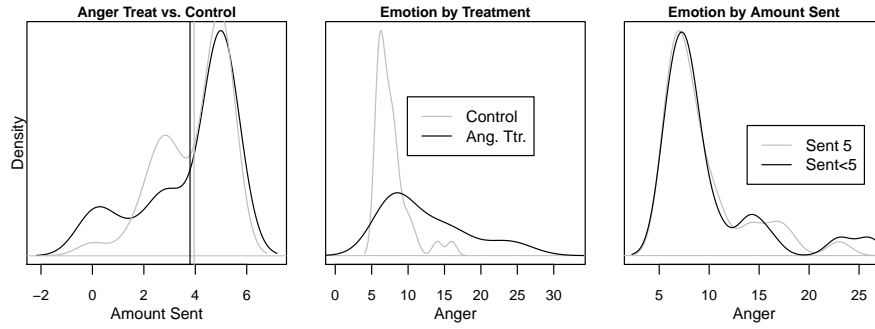


Figure 5: Distributions of amount sent and emotions by amount sent and treatment status for the anger induction treatment versus control condition.



Figure 6: Distributions of amount sent and emotions by amount sent and treatment status for the guilt induction treatment versus control condition.

effect both through the hypothesized mechanism (i.e., trust behavior is influenced by an event inducing emotion, here a writing task, via some particular emotional mechanism) but also through other pathways (e.g., other emotions or other variables). Because of this it is possible that while average treatment effects might be zero, this does not mean that the treatment did not have some effect via the mechanism of interest.¹¹

Consider the contrast between one of our treatment conditions, the anxiety writing task ($T = 1$), and the control condition ($T = 0$). Let $M_i(0)$ represent the value of the mediating emotion that subject i would take on if they were assigned to the control condition and $M_i(1)$ under the treatment condition. In the typical experiment we only observe one of these values for each subject, as we do not also observe them in the other condition. Using this setup, the causal mediation effects for each unit i can be defined as,

$$\delta_i(t) \equiv Y_i(t, M_i(1)) - Y_i(t, M_i(0)), \quad (1)$$

¹¹Alternative approaches to estimation, such as instrumental variables, rule out the possibility of a direct effect or other mechanism influencing the outcome. This is a major limitation to the highly popular instrumental variable approach (Imai *et al.*, 2010a).

for $t = 0, 1$. Here the causal mediation effect is simply the indirect effect of the treatment on the outcome through the mediating variable (Pearl, 2001; Robins, 2003).

Equation (1) asks the following *counterfactual* question: How would trust behavior change if one changes the emotion level (the mediator) from the value that would occur in the treatment condition ($M_i(1)$), to the value that would be observed if the subject did not receive the anxiety task ($M_i(0)$), while holding the treatment status at a constant level t ? Note that if the writing task has no effect on the emotions, i.e., $M_i(1) = M_i(0)$, then the causal mediation effects are zero. For example, this gives us immediately that the the mediation effects for the positive emotion writing tasks are zero.

A similar set of definitions allows us to define the direct effects of the treatment for each unit as follows,

$$\zeta_i(t) \equiv Y_i(1, M_i(t)) - Y_i(0, M_i(t)), \quad (2)$$

for $t = 0, 1$. Direct effects in the present context can either be thought of the direct effect of the treatment on trust behavior, or the influence of the treatment on the outcome via other mediating variables. Finally, the total effect, τ_i , can be decomposed into the sum of the direct and indirect effects.¹² If the causal mediation and direct effects do not vary with the treatment status (i.e., $\delta_i = \delta_i(1) = \delta_i(0)$ and $\zeta_i = \zeta_i(1) = \zeta_i(0)$), then the mediation and direct effects sum to the total effect, i.e., $\tau_i = \delta_i + \zeta_i$. This *no-interaction assumption* can easily be relaxed and we explore this below. In practice researchers are typically interested in the *average causal mediation effects* (ACME) which is,

$$\bar{\delta}(t) \equiv \mathbb{E}(Y_i(t, M_i(1)) - Y_i(t, M_i(0))),$$

for $t = 0, 1$. This is simply the average indirect effect over the subjects (each of whom can have different pre-treatment covariates) in the experiment. Similar average quantities can be defined for the direct and total effects.

Because in the standard type of experiment, where individuals are assigned either to the treatment or control, we do not observe the counter-factual values of the mediator nor do we manipulate them, an identification assumption must be made. In particular, Imai *et al.* (2010d,a) show that these quantities are non-parametrically point identified by assuming that there is no confounding variable that influences both the mediating variable and outcome variable once the treatment assignment is controlled for. Failure

¹²Symbolically, $\tau_i \equiv Y_i(1, M_i(1)) - Y_i(0, M_i(0)) = \frac{1}{2} \sum_{t=0}^1 \{\delta_i(t) + \zeta_i(t)\}$, where we simply average over the treatment assignment.

to control for such a variable leads to the violation of the identification assumption. Additional details are provided in Section B.

Non-parametric point identification of the ACME gives us considerable flexibility because we can estimate the quantities of interest under a variety of statistical models to measure the relationship between the treatment, mediator and outcome variables. In the present case this is very helpful, because as shown in Figures 4, 5, and 6, many individuals sent the maximal amount making the use of linear regression problematic. Thus in our formal analysis we rely on tobit or quantile regression models. However, before turning to our flexible way to calculate the ACME, we review the necessary building blocks with reference to linear regression models.

The conventional way to calculate mediation effects is to estimate two linear regression models:

$$M_i = \alpha_2 + \beta_2 T_i + \xi_2^\top X_i + \epsilon_{i2}, \quad (3)$$

$$Y_i = \alpha_3 + \beta_3 T_i + \gamma M_i + \xi_3^\top X_i + \epsilon_{i3}, \quad (4)$$

The first estimates the relationship between the mediating variable and treatment. The second estimates the outcome as a function of both the mediator and the treatment. Both condition on the same set of pre-treatment covariates. In this case Imai *et al.* (2010d) show that under the sequential ignorability assumption the average causal mediation and direct effects are identified as $\bar{\delta}(t) = \beta_2 \gamma$ and $\bar{\zeta}(t) = \beta_3$, respectively, for $t = 0, 1$.¹³ In this case a no-interaction assumption is made and with linear models $\bar{\delta}(0) = \bar{\delta}(1) = \beta_2 \gamma$.¹⁴ The no-interaction assumption simply restricts the ACME to be the same for those in the treatment and control conditions. The no-interaction assumption can be relaxed and several scholars argue that the assumption is often unrealistic (Kraemer *et al.*, 2002, 2008).¹⁵

Unfortunately life is not like a box of linear regressions and instead alternative statistical models are necessary due to the nature of the data generation process. For example, most studies of the trust game use a tobit model to estimate the amounts sent (Scharlemann *et al.*, 2001; Tingley, 2010) or a probit model for binary trust games (Eckel and Wilson, 2004). In this case the formulas described above will not yield valid estimates of the causal parameters, something that was previously shown by Imai *et al.* (2010a,c) for dichotomous and ordered variables. A similar criticism applies to situations with a

¹³The average total effect is given by β_1 which equals $\beta_2 \gamma + \beta_3$.

¹⁴Hence the common practice of multiplying slope coefficients from linear models (Baron and Kenny, 1986) to generate mediation effects is predicated on the sequential ignorability assumption.

¹⁵See Section B for additional details.

censored outcome that are modeled with tobit.

Fortunately, Imai *et al.* (2010a) show that Equation 7 implies a general computational approach to calculating causal quantities of interest and uncertainty estimates. This generality is due to the non-parametric nature of the identification result. The intuition of the algorithm is simple. First, estimate Equations 3 and 4 using the user’s choice of statistical model. For example, below we use regression for the mediating emotion variable and tobit for the amount sent model. Next, calculate *predicted* values of the mediator under the treatment and control conditions using Equation 3. Next, calculate expected values of the outcome variable under both the treatment and control, but each time using the predictions of $M(1)$ and $M(0)$ obtained in the previous step, to calculate $\bar{\delta}(t) \equiv \mathbb{E}(Y_i(t, M_i(1)) - Y_i(t, M_i(0)))$. Uncertainty estimates are then calculated using either the model variance-covariance matrices to conduct a quasi-Bayesian parametric bootstrap (Tomz *et al.*, 2003) or by completely bootstrapping the estimation.

For this paper we extended the software provided by Imai *et al.* (2010b) to accommodate a tobit model for the outcome variable, here the amount sent in the trust game. We also estimate quantile regression model for the outcome variable to further relax the parametric assumptions of the tobit model. This lets us ask if the mediation effects are stronger for different quantiles of the amount sent. Finally, it is also possible to conduct sensitivity analyses to deviations from the sequential ignorability assumption when particular parametric models are used.

5.2.2 Estimation

With this framework in mind we now turn to the estimation of the mediating role of emotions which may transmit the effects of the writing manipulations. We focus on the three mood induction tasks that had any influence on emotions: anxiety, anger, and guilt. Inductions that had no first stage effect on emotions are unlikely to generate any mediation effects.¹⁶ For each induction we estimate the mediating role of their targeted emotion, allowing for the fact that the treatment might effect the outcome in other ways.

The relationship between the mediator and the treatment is modeled with linear regression and the relationship between the amount sent and the treatment/mediator are estimated using a tobit model. We include controls for gender, political ideology, and pre-treatment measures of generalized trust (*insert cite) and estimate models assuming no interaction between the treatment and mediator. The supplementary materials in Appendix C relax this assumption. Figure 7 presents the results for each of the three mediation analyses.

¹⁶In supplementary analyses we did not find non-zero ACME or total effects for the positive emotion treatments.

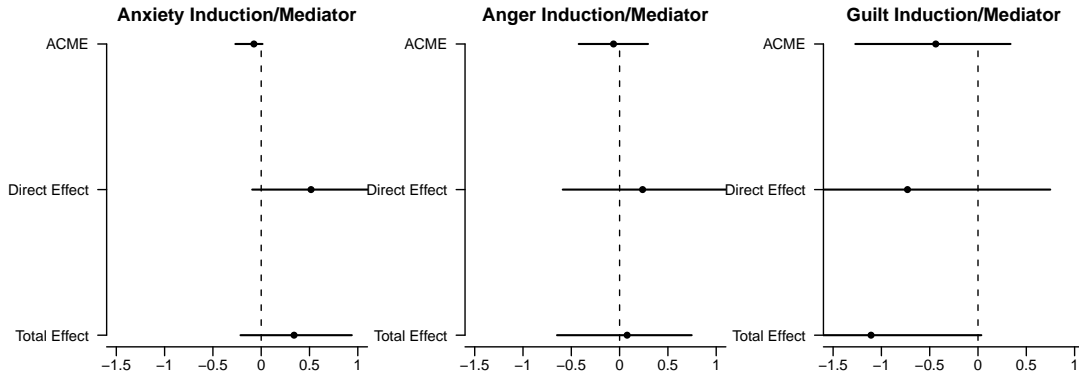


Figure 7: ACME, Direct, and Total Effects of anxiety, anger, and guilt manipulations on their respective targeted emotions. Point estimates with 95% confidence intervals. Models assume no interaction between treatment condition and mediator.

Figure 7 reports the ACME, direct, and total effects for each manipulation condition using the respective targeted emotion. We focus first on the anxiety writing task that is used for the left pane in Figure 7. According to our theory about the effect of certainty on trust, we should see a negative mediation effect. Individuals that take on a *more* negative emotional disposition from the writing task (a positive relationship) should trust their opponents less and hence send them less (a negative relationship). We observe a small negative ACME. While the confidence interval slightly overlap with zero this overlap is relatively small. Consistent with Figure 4 the total effect was not different from zero.

The middle and right panes of Figure 7 provide analogous plots for the anger and guilt inductions. The ACME and total effect for the anger manipulation was not significantly different from zero. The ACME for the guilt manipulation is negative but has confidence intervals that overlap slightly more with 0. As expected, the total effect is negative and nearly significant. Interestingly, the the largest estimated total effect across the inductions was for the guilt condition. This estimate makes clear the importance of mediation analysis; naively examining the total effect of the manipulation would suggest that guilt, not anxiety, has the largest negative impact on trust.

The causal mediation effects presented in Figure 7 are estimated using a combination of linear regression (for the mediator) and tobit (for the amount sent). Under the sequential ignorability assumption the mediation effects are non-parametrically identified. Because of this, alternative parametric and semi-parametric techniques are available for calculating causal mediation effects (Imai *et al.*, 2010a). One approach is to estimate mediation effects across different quantiles of the amount sent distribution. For example, we might be interested in the ACME around the median, or some lower-mid quantile. By using

quantile regression (Koenker, 2008) instead of linear regression for Equation (4) we generate point and uncertainty estimates for the ACME and direct effects.¹⁷ Here we suspect that the negative mediating effect should be strongest for those who sent a lower amount. Individuals may form more negative expectations about the type of Role 2 player they face if they are negatively aroused emotionally. The transmission then of the arousal manipulation on amount sent should be strongest for individuals that subsequently choose to send less. In Figures 8, 9, 10 we plot the ACME and direct effect across the 25th through 75th quantiles of the amount sent for the anxiety, anger and guilt treatments.

For the anxiety manipulation the ACME remains negative but is increasing over the amount sent. Interestingly, the anger and guilt conditions do not display this dynamic, instead staying relatively constant over the quantiles. Of course, the confidence intervals are much wider across much of the distribution of amount sent. This is common for semi-parametric models and not surprising given the amount sent distribution. Finally, Appendix G provides a sensitivity analysis to violations of the sequential ignorability assumption.

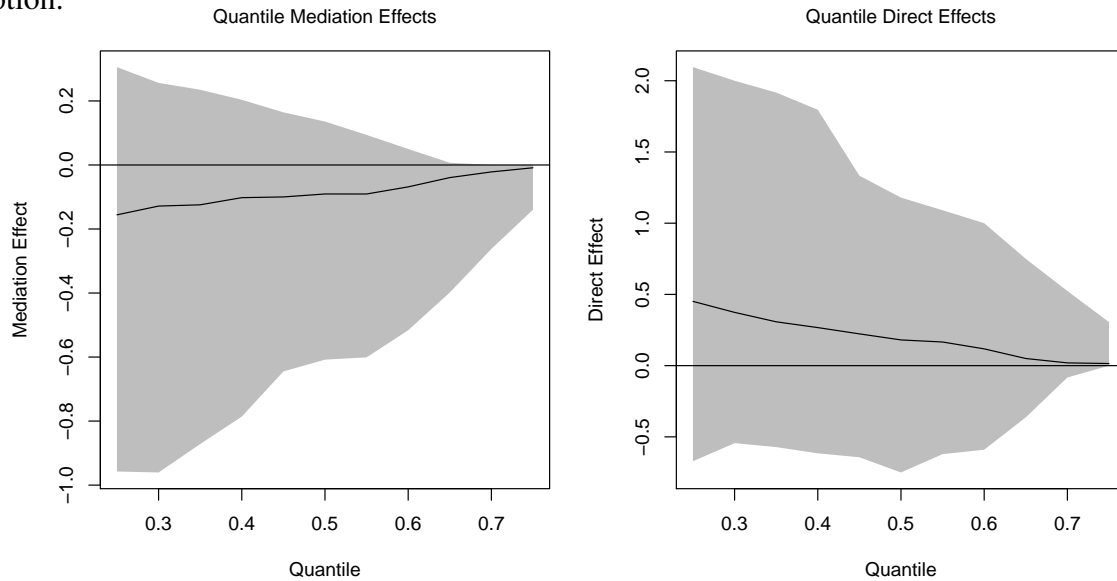


Figure 8: ACME and Direct Effects of anxiety manipulation on amount sent using Anxiety mediator across the 25th through 75th quantiles of the amount sent. Point estimates with 95% confidence intervals shaded. The negative mediating effect is strongest for lower levels of the amount sent.

¹⁷See Section B for technical details.

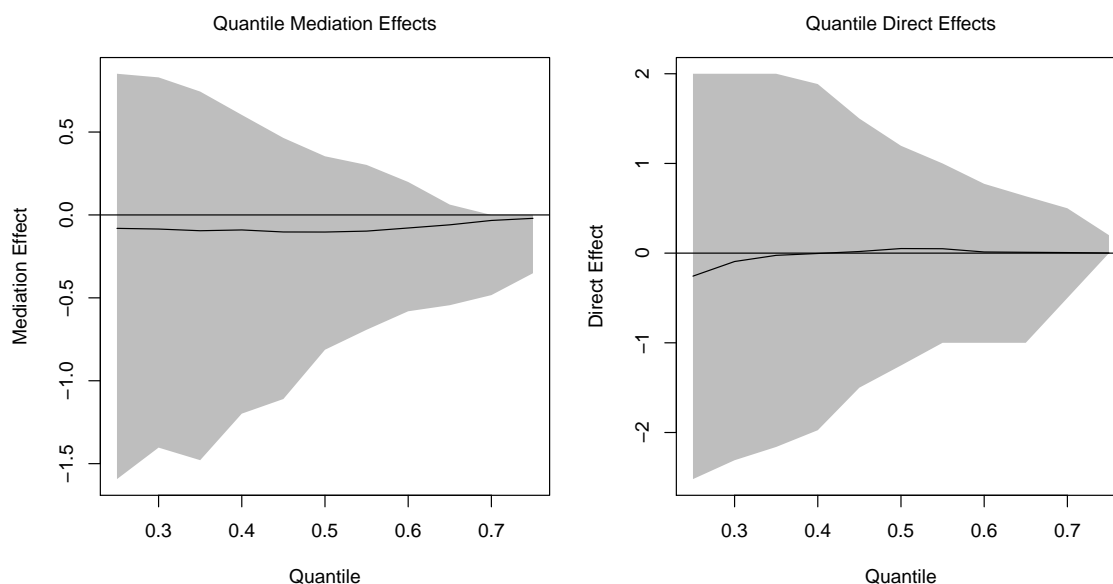


Figure 9: ACME and Direct Effects of anger manipulation on amount sent using the Anger mediator across the 25th through 75th quantiles of the amount sent. Point estimates with 95% confidence intervals shaded. The effect stays close to zero across the quantiles.

6 Conclusion

The willingness of individuals to trust each other is fundamental to economic and political life. In this paper we investigate the impact of affect on trust. We explore whether and how emotional states influence behavior in the trust game. While our manipulations use a very specific task—the AEMT—the manipulation of emotion is a common feature of political life, perhaps especially in the media-charged environment we live in.

Our findings suggest that negative emotions can decrease trust, but only if those negative emotions produce low certainty appraisals. Anxiety, a low certainty emotion, has a negative impact on trust while Anger and Guilt, two emotions that differ in their control-appraisals but induce the same high level of certainty, appear to have no clear effect on trusting behavior. We do not directly test the mechanism behind this relationship, but it is likely that low-certainty emotions cause people to use their affective state to judge a potential trustee. While we were unable to test the effect of certainty appraisals on trust among positive emotions, we see no reason why a similar process would lead people experiencing low-certainty but positive emotional states to act in a more trusting manner. This research shows the importance of going beyond the valence-based approach of Capra (2004) when examining the impact of emotions on political and economic processes.

Our findings do not support the findings by Dunn and Schweitzer (2005) that control appraisals

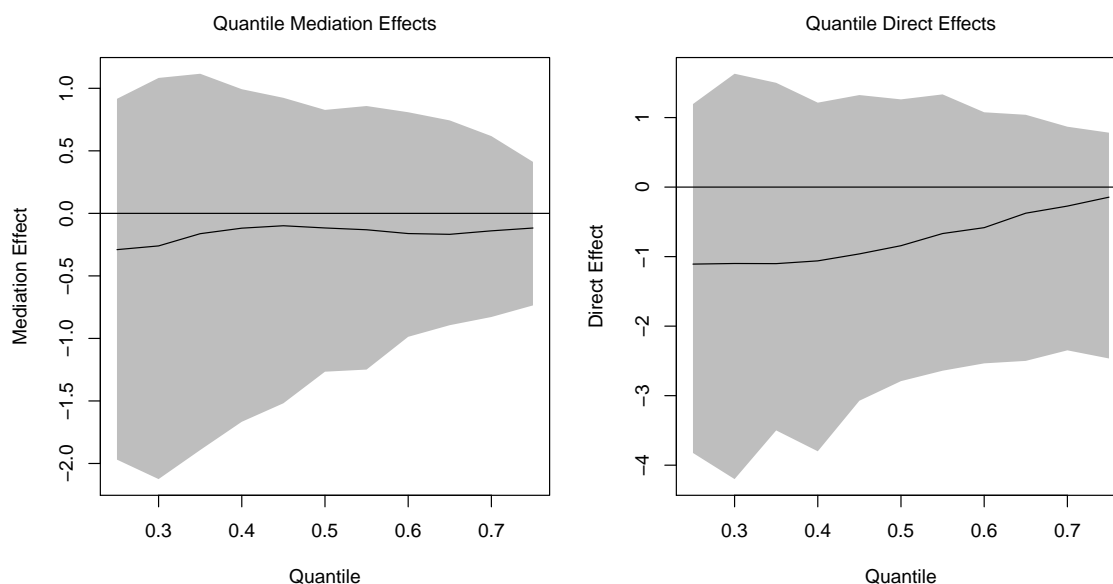


Figure 10: ACME and Direct Effects of guilt manipulation on amount sent using Guilt mediator across the 25th through 75th quantiles of the amount sent. Point estimates with 95% confidence intervals shaded. A small negative mediating effect is relatively constant across the quantiles of the amount sent.

determine whether an emotion will influence trust. In particular, we find no effect from anger, an emotion with an “personal control - other” appraisal, on trust; Dunn and Schweitzer (2005) found that anger had a strongly negative influence on trust. While there were several differences across our experimental setups that might account for this difference, two stand out as the most likely. First, Dunn and Schweitzer (2005) measure trust using a survey instrument, while we measure trust using the trust game. It is likely that these two techniques actually measure different constructs; as noted before survey-measures of trust seem to be more predictive of trustworthy behavior than of trusting behavior in the trust game (Glaeser *et al.*, 2000). While we have no strong theoretical reasons for believing that control-appraisals would matter for trustworthiness, it is not surprising that different emotions would be important for different aspects of trusting relationships. Secondly, Dunn and Schweitzer (2005) do not measure the effect of the manipulation on their subjects or perform mediation analysis. This leaves open the possibility that while writing about an anger-inducing experience causes subjects to show less trust, this effect is not caused by an increase in anger but instead through the effect of the task on other emotions.

Future research will proceed along several lines. In recent work Tingley (2010) explores how individuals choose human faces with very particular face structures to represent them in a trust game. Building on earlier work (Scharlemann *et al.*, 2001) this provides new evidence about the way people infer intentions of others based on appearance. An open question, though, is *how* does appearance influence trust behavior. Does being confronted with a “trustworthy” face change affective dispositions? Our results

suggests that emotions related to certainty, such as anxiety, are likely to transmit at least some of this effect. We plan to investigate these questions in the near future. Second, while we plan to use other forms of stimuli that more directly reflect what individuals are likely to receive in the “real” world.

References

- Ax, A. F. (1953). The physiological differentiation between fear and anger in humans. *Psychosomatic Medicine* **15**, 2, 433–442.
- Baron, R. and Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* **51**, 6, 1173–1182.
- Bodenhausen, G. V., Sheppard, L. A., and Kramer, G. P. (1994). Negative affect and social judgment: the differential impact of anger and sadness. *European Journal of Social Psychology*, **24**, 1, 45–62.
- Bower, G. H. (1981). Mood and memory. *American Psychologist* **36**, 2, 129–148.
- Brader, T., Valentino, N., and Suhay, E. (2008). What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration Threat. *American Journal of Political Science* **52**, 4, 959–978.
- Capra, C. M. (2004). Mood-driven behavior in strategic interactions. *AEA Papers and Proceedings* **94**, 2, 367–372.
- Cesarini, D., Dawes, C. T., Fowler, J. H., Johannesson, M., Lichtenstein, P., and Wallace, B. (2008). Heritability of cooperative behavior in the trust game. *Proceedings of the National Academy of Sciences* **105**, 10, 3721–3726.
- DeSteno, D., Dasgupta, N., Bartlett, M. Y., and Cajdric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science* **15**, 5, 319–324.
- DeSteno, D., Petty, R. E., Wegener, D. T., and Rucker, D. D. (2000). Beyond valence in the perception of likelihood: The role of emotion-specificity. *Journal of Personality and Social Psychology* **78**, 3, 397–416.
- Druckman, J. N. and McDermott, R. (2008). Emotion and the framing of risky choice. *Political Behavior* **30**, 3, 297–321.
- Dunn, J. R. and Schweitzer, M. E. (2005). Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology* **88**, 5, 736–748.

- Eckel, C. and Wilson, R. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization* **55**, 4, 447–465.
- Eich, E. and Macaulay, D. (2000). Fundamental factors in mood-dependent memory. In J. P. Forgas, ed., *Feeling and Thinking: The role of affect in social cognition*. Cambridge University Press, Cambridge.
- Forgas, J. P. (1995). Mood and judgment: The affect infusion model (aim). *Psychological Bulletin* **117**, 1, 39–66.
- Forgas, J. P. (1998). On feeling good and getting your way: Mood effects on negotiating strategies and outcomes. *Personality and Social Psychological Bulletin* **24**, 2, 173–185.
- Forgas, J. P. (2009). Affective influences on attitudes and judgements. In R. J. Davidson, K. R. Sherer, and H. H. Goldsmith, eds., *Handbook of Affective Sciences*. Oxford University Press, Oxford.
- Gadarian, S. K. (2010). The politics of threat: How terrorism news shapes foreign policy attitudes. *The Journal of Politics* **72**, 02, 469–483.
- Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, **113**, 3, 811–846.
- Gross, J. J. and Levenson, R. W. (1995). Eliciting emotions using films. *Cognition and Emotion* **9**, 1, 87–108.
- Hertel, G. and Fiedler, K. (1994). Affective and cognitive influences in social dilemma game. *European Journal of Social Psychology* **24**, 1, 131–145.
- Hetherington, M. J. (1998). The political relevance of political trust. *American Political Science Review* **92**, 4, 791–808.
- Houghton, L. A., Calvert, E. L., Jackson, N. A., Cooper, P., and Whorwell, P. J. (2002). Visceral sensation and emotion: a study using hypnosis. *GUT: An International Journal of Gastroenterology and Hepatology* **51**, 2, 701–704.
- Imai, K., Keele, L., and Tingley, D. (2010a). A general approach to causal mediation analysis. *Psychological Methods* **15**, 309–334.

- Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2010b). Causal mediation analysis using R. *Advances in Social Science Research Using R* 129–154.
- Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2010c). Unpacking the Black Box: Learning about Causal Mechanisms from Experimental and Observational Studies. *Unpublished manuscript* .
- Imai, K., Keele, L., and Yamamoto, T. (2010d). Identification, inference, and sensitivity analysis for causal mediation effects. *Statistical Science* **25**, 1, 51–71.
- Judd, C. and Kenny, D. (1981). Estimating mediation in treatment evaluations. *Evaluation Review* **5**, 5, 602–619.
- Kenny, D. A. (2008). Reflections on Mediation. *Organizational Research Methods* **11**, 2, 353–358.
- Koenker, R. (2008). *Quantile Regression*. Cambridge University Press, Cambridge.
- Kraemer, H., Kiernan, M., Essex, M., and Kupfer, D. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology* **27**, 2, S101–S108.
- Kraemer, H., Wilson, G., Fairburn, C., and Agras, W. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry* **59**, 877–883.
- Lerner, J. S. and Gonzalez, R. M. (2005). Forecasting ones future based on fleeting subjective experiences. *Personality and Social Psychology Bulletin* **31**, 4, 454–466.
- Lerner, J. S., Gonzalez, R. M., Small, D. A., and Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism : A national field experiment. *Psychological Science* **14**, 2, 144–150.
- Lerner, J. S. and Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion* **14**, 3, 473–493.
- Lerner, J. S. and Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology* **81**, 1, 146–159.
- Levi, M. and Stoker, L. (2000). Political trust and trustworthiness. *Annual Reviews of Political Science* **3**, 1, 475–507.

- Lewis, M., Phd, J. M. H.-J., and Barrett, L. F. (2009). *Handbook of Affective Sciences*. Oxford University Press, Oxford.
- Lim, E. T. (2002). Five trends in presidential rhetoric: An analysis of rhetoric from George Washington to Bill Clinton. *Presidential Studies Quarterly* **32**, 2, 328–348.
- MacKinnon, D., Lockwood, C., Hoffman, J., West, S., and Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods* **7**, 1, 83–104.
- Marcus, G. E., Neuman, W. R., and MacKuen, M. (2001). *Affective Intelligence and Political Judgment*. University of Chicago Press, Chicago.
- Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., and Reuter-Lorenz, P. A. (2005). Emotional category data on images from the International Affective Picture System. *Behavioral Research Methods* **37**, 4, 626–630.
- Mishler, W. and Rose, R. (1997). Trust, distrust and skepticism: Popular evaluations of civil and political institutions in post-communist societies. *Journal of Politics* **59**, 2, 418–451.
- Pearl, J. (2001). Direct and indirect effects. In *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*, 411–420. Citeseer.
- Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster, New York.
- Robins, J. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. *OXFORD STATISTICAL SCIENCE SERIES* 70–82.
- Scharlemann, J., Eckel, C., Kacelnik, A., and Wilson, R. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology* **22**, 5, 617–640.
- Schwarz, N. and Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology* **45**, 3, 513–523.
- Small, D. A. and Lerner, J. S. (2008). Emotional policy: Personal sadness and anger shape judgments about a welfare case. *Personal Psychology* **29**, 2, 149–168.

- Smith, C. A. and Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology* **48**, 4, 813–838.
- Strack, F., Schwarz, N., and Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology* **49**, 6, 1460–1469.
- Tingley, D. (2010). Face-off: Face Features and Strategic Choice. *working paper* .
- Tomz, M., Wittenberg, J., and King, G. (2003). CLARIFY: Software for interpreting and presenting statistical results.
- Valentino, N. A., Hutchings, V. L., and Banks, A. J. (2008). Is a worried citizen a good citizen? emotions, political information seeking, and learning via the internet. *Political Psychology* **29**, 2, 247–273.
- Verhulst, B. and Sohlberg, J. (2009). The double-edged sword of political anxiety. Stony Brook University Typescript.

A Violin plots of emotion distributions

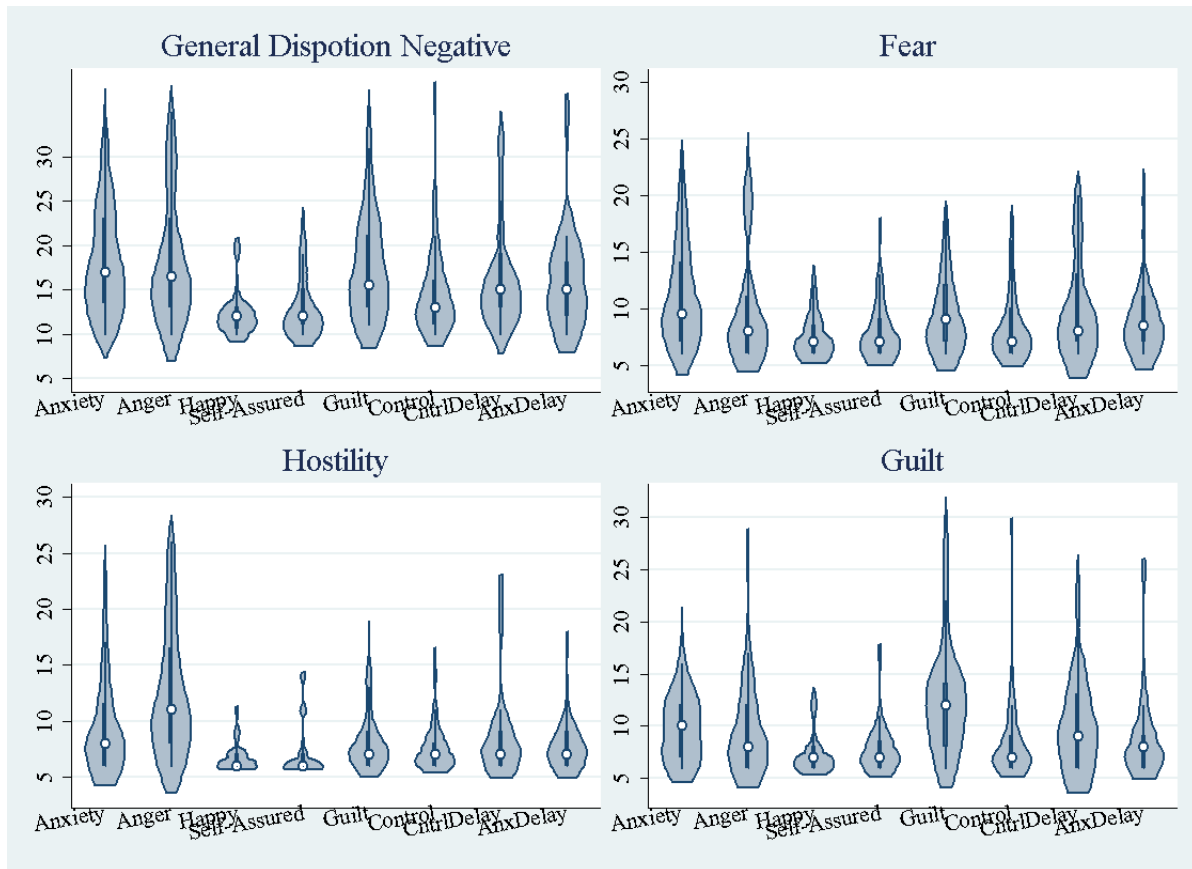


Figure 11: Effects of Emotion Manipulations on Negative Emotions. Means with 95% confidence intervals.

B Technical Appendix for Mediation Analysis

B.1 Identification of mediation effects

In order for the ACME to have a causal interpretation one must make some sort of assumption for the simple reason that the values of the mediator are not randomly assigned. In particular, Imai *et al.* (2010d) work with an assumption that there is no omitted variable that has a causal effect on both the mediator and the outcome variable. Formally this assumption is known as sequential ignorability and, once made,

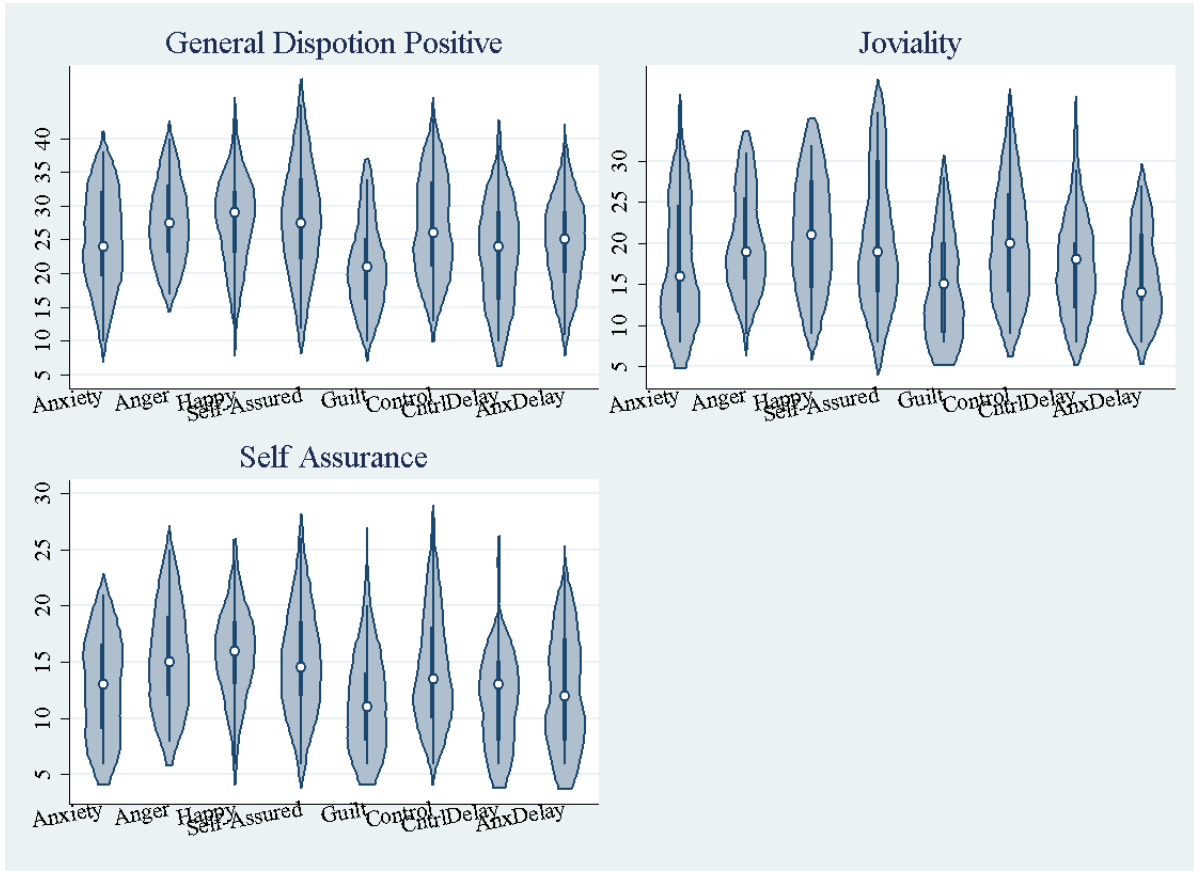


Figure 12: Effects of Emotion Manipulations on Positive Emotions. Means with 95% confidence intervals.

allows for the non-parametric point identification of causal mediation effects.¹⁸

Appendix G presents a formal sensitivity analysis for the anxiety manipulation that plots the ACME allowing for violations of the sequential ignorability assumption.

¹⁸Sequential Ignorability (Imai *et al.*, 2010d) assumes that the following two statements of conditional independence hold,

$$\{Y_i(t', m), M_i(t)\} \perp\!\!\!\perp T_i \mid X_i = x, \quad (5)$$

$$Y_i(t', m) \perp\!\!\!\perp M_i \mid T_i = t, X_i = x, \quad (6)$$

where $0 < \Pr(T_i = t \mid X_i = x)$ and $0 < p(M_i = m \mid T_i = t, X_i = x)$ for $t = 0, 1$, and all $x \in \mathcal{X}$ and $m \in \mathcal{M}$. Under this assumption the ACME is point identified as

$$f(Y_i(t, M_i(t')) \mid X_i = x) = \int_{\mathcal{M}} f(Y_i \mid M_i = m, T_i = t, X_i = x) dF_{M_i}(m \mid T_i = t', X_i = x),$$

for any pre-treatment covariates $x \in \mathcal{X}$ and $t, t' = 0, 1$.

B.2 Relaxing the no-interaction assumption

Relaxing the no-interaction assumption is straightforward. Here we illustrate things using linear models for simplicity. Here we can replace equation (4) with,

$$Y_i = \alpha_3 + \beta_3 T_i + \gamma M_i + \kappa T_i M_i + \xi_3^\top X_i + \epsilon_{i3}. \quad (7)$$

which simply adds an interaction term between the treatment and mediator. Similar calculations are then done using the estimated parameters to recover the mediation effects. Imai *et al.* (2010a) show that under sequential ignorability assumption, the average causal mediation effects are given by,

$$\bar{\delta}(t) = \beta_2(\gamma + \kappa t), \quad (8)$$

for $t = 0, 1$. Similarly, the average direct and total effects can be calculated as,

$$\bar{\zeta}(t) = \beta_3 + \kappa\{\alpha_2 + \beta_2 t + \xi_2^\top \mathbb{E}(X_i)\}, \quad (9)$$

$$\bar{\tau} = \beta_2 \gamma + \beta_3 + \kappa\{\alpha_2 + \beta_2 + \xi_2^\top \mathbb{E}(X_i)\}, \quad (10)$$

for $t = 0, 1$. Extension to non-linear models is straightforward using the algorithm discussed earlier (Imai *et al.*, 2010a). The key result is that this procedure allows for mediation effects to differ depending on the treatment condition. Our robustness checks allow for this possibility.

B.3 Quantile mediation effects

Formally, α -quantile causal mediation effects are defined as, $\tilde{\delta}_\alpha(t) \equiv q_{t1}(\alpha) - q_{t0}(\alpha)$, for $t = 0, 1$ and $0 < \alpha < 1$ where $q_{tt'}(\alpha) \equiv \inf\{y; F(Y_i(t, M_i(t'))) \leq y\} \geq \alpha\}$ is the quantile function for the distribution of $Y_i(t, M_i(t'))$. Similarly, we can define quantile direct and total effects as, $\tilde{\zeta}_\alpha(t) \equiv q_{1t}(\alpha) - q_{0t}(\alpha)$.

C Mediation Effects Allowing for Treatment-Mediator Interaction

Figure 13 provide mediation estimates using models that permit an interaction between the treatment condition and mediator. The left pane shows that the mediation effect from the anxiety induction is highest amongst those in the treatment group. Conversely, for the guilt condition the mediation effect is highest for those in the control condition. Future research will explore this discrepancy.

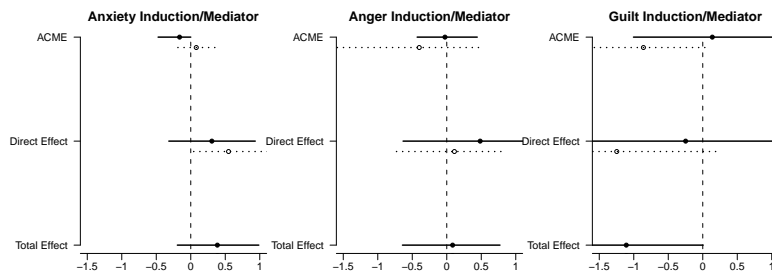


Figure 13: ACME, Direct, and Total Effects of anxiety, anger, and guilt manipulations on their respective targeted emotions. Point estimates with 95% confidence intervals. Models allow interaction between treatment condition and mediator.

D Survey Instruments

The following survey items were used in this paper’s analyses. The first is a standard seven-point measurement of ideology, the second a three question battery measuring trust that comes from the GSS. Both were administered as part of the pre-survey completed when subjects signed up for the experiment (on average greater than 24 hours prior to the experiment), and thus should neither contaminate nor be contaminated by the results of the experiment

Ideology We hear a lot of talk these days about liberals and conservatives. On a 7-point scale, where 1 is very liberal and 7 is very conservative, where would you place yourself on this scale, or haven’t you thought much about this?

Very Liberal — Liberal — Slightly Liberal — Moderate — Slightly Conservative — Conservative — Very Conservative— Haven’t Thought Much About This

Generalized Trust

Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?

Would try to take advantage of you — Would try to be fair

Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?

Try to be helpful — Just look out for themselves

Generally speaking, would you say that most people can be trusted or that you cant be too careful in dealing with people?

Most people can be helpful — Can’t be too careful

E PANAS-X Details

The version of PANAS-X this study used consisted of 40 items. Subjects were asked to respond to the following prompt with respect to each item:

This scale consists of a number of words that describe different feelings and emotions. Read each item and select the appropriate option next to that word. Indicate to what extent you feel this right now, (that is, at the present moment).

Very slightly or not at all — A little — Moderately — Quite a bit — Extremely

Responses were coded from 1 (very slightly or not at all) to 5 (extremely). The responses were aggregated into the scales shown in Table 2. Items were displayed in a random order, to minimize possible order effects.

F Experimental Instructions

To introduce the Trust Game, the following instructions were read aloud to all subjects:

We will now pair subjects together and have you make decisions that determine how much money you will make from today's experiment. You will be paid privately at the end of the experiment.

You will be randomly paired with one other person and what you earn depends only on your decision and the decision of the other person. One of you will be randomly placed in role 1 and the other in role 2. The person in role 1 is given \$5 and they must decide how much, if any, of this to transfer to the person in role 2. The money you give to the other person will be then be tripled, so if the role 1 person gives \$2 the role 2 person will receive \$6. The person in role 2 then decides how much, if any, of the tripled money to return to the role 1 person and how much to keep for themselves. The person in role 1 earns whatever money is returned plus the share of the \$5 they decided to keep.

After the instructions were read subjects were randomly assigned to role one or role two. Instructions specific to their roles appeared on their screens. For role one subjects, these read:

You are given \$5 and must decide how much, if any, of this to transfer to the person in role 2. The money you give to the other person will be then be tripled, so if the role 1 person

gives \$2 the role 2 person will receive \$6. The person in role 2 then decides how much, if any, of the tripled money to return to you and how much to keep for themselves. You earn whatever money is returned plus the share of the \$5 you decided to keep.

For role two players these read:

The person you have been paired with has been given \$5 and must decide how much, if any, of this to transfer to you. The money they give to you will then be tripled, so if they give you \$2 this will be tripled to \$6. You then decide how much, if any, of the tripled money to return and how much to keep for yourself.

G Sensitivity Analysis

Given the strong nature of the sequential ignorability assumption—and the fact that it cannot be directly tested, Imai *et al.* (2010a) recommend conducting sensitivity analyses. Of course, the assumption is a priori plausible in that pre-treatment trust levels are used as a control, but there could always exist some omitted confounder. Heretofore such sensitivity analyses are only available for a limited set of parametric models. In particular the necessary math has not yet been worked out for the case where the outcome variable is censored and so we rely here on models that use linear regressions for Equations 3 and 4.¹⁹ Figure 14 presents two such tests. The first, displayed in the top row, plots the ACME as a function of the correlation in error terms between the mediator and outcome models (ρ). For the estimate of $\tilde{\delta}(1)$ to be positive this correlation must be nearly $-.5$. An alternative representation is to plot the ACME in terms of the total variance of the mediator and outcome variables. The bottom row plots the ACME in a contour plot. For the ACME to be 0, the omitted confounder would need to explain roughly 40% of the mediator and outcome variables, for example. These results suggest that the negative mediation effect is quite robust to an omitted confounding variable.

¹⁹Perhaps not surprisingly ACME estimates using the linear outcome model tended to be slightly more extreme than those based on the tobit model.

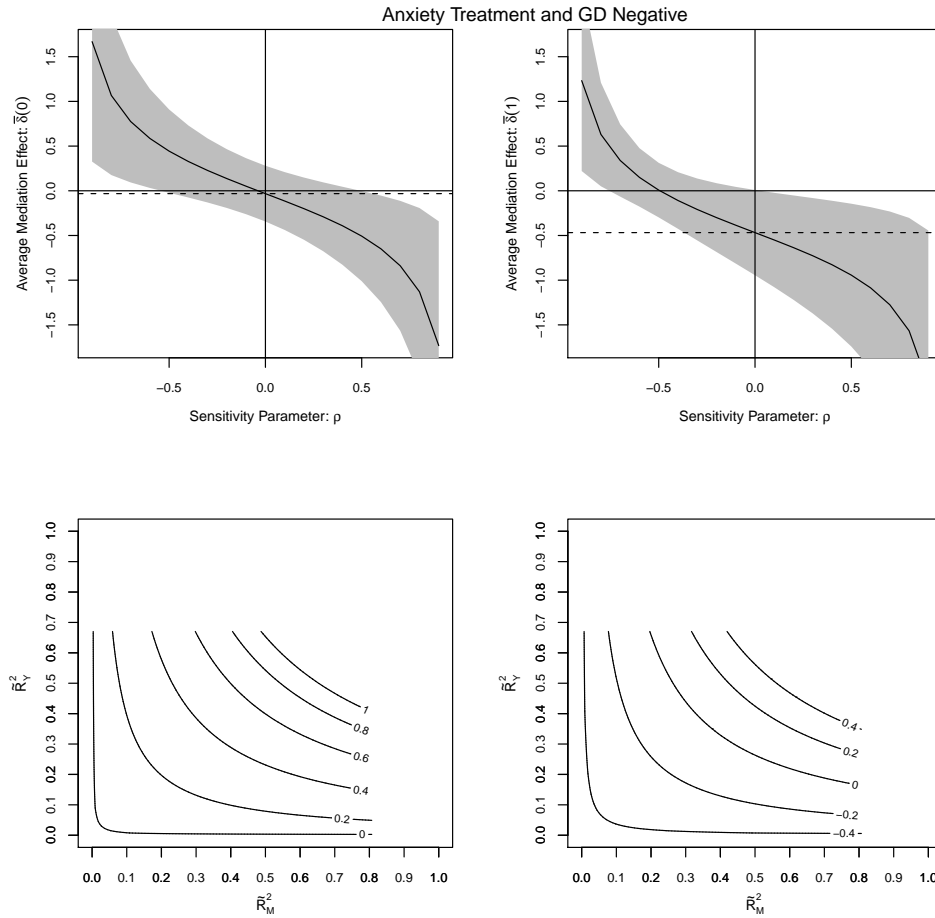


Figure 14: Sensitivity analysis for the anxiety treatment and GD Negative mediator. Top row presents sensitivity analysis in terms of the ρ (correlation between model error terms) and the bottom row presents in terms of the proportion of total variation in the mediator and outcome models. Left column for the control and right column for the treatment condition. The mediation result is quite robust, with extreme violations of the sequential ignorability assumption required for the hypothesized negative sign to be reversed.