# Sad But Accurate? Examining the Role of Discrete Emotions on Prediction Performance

Thomas Y. Li
*University of St. Gallen*

**Abstract**

How do discrete emotions influence prediction accuracy? Drawing on appraisal-tendency framework and approach-avoidance models, we hypothesized that approach-oriented emotions with high certainty (anger, happiness) would enhance prediction accuracy through heuristic processing, while avoidance-oriented emotions with low certainty appraisals (fear, sadness) would trigger over-deliberation and impair performance. We further tested whether trust-in-feelings (TIF) moderates these relationships. Participants ($N$ = 740) underwent emotion induction and TIF manipulation (high vs. none) before making predictions about weather, movie box office performance and stock market movements. Results contradicted predictions: fear and sadness *improved* weather prediction accuracy (ORs = 1.67-2.28), performing comparably to anger and happiness. Effects were domain-specific, emerging only for weather. TIF showed minimal effects (one of seven predictions reached significance), representing a failure to replicate Pham et al.'s (2012) "emotional oracle effect". These findings challenge assumptions about negative emotions in forecasting and suggest emotional engagement – rather than discrete emotions or TIF – drives accuracy.

**Keywords:**

*Discrete emotions, prediction accuracy, forecasting, trust-in-feelings, emotional engagement, appraisal-tendency framework, approach-avoidance motivation, dual-process theory, affect-as-information, emotional oracle effect, decision-making, judgment*

# Introduction

Human beings have always been interested in predicting the future. Whether it is forecasting tomorrow's weather, next year's crop yield in early societies, or organizational outcomes (Aboody et al., 1999; Bayazit & Mannix, 2003), humans are genuinely drawn to understanding and reducing uncertainty about what comes next. To facilitate the prediction process, many scholars, and practitioners alike, have attempted to uncover methods of prediction that foster the highest level of accuracy. While many advances in prediction have focused on computer-driven models (e.g., machine learning and deep learning) (e.g., Collins & Moons, 2019), there have also been developments to maximize the human potential. These approaches range from the introduction of new formats, such as hosting crowd forecasting tournaments that facilitate and aggregate peoples' predictions (e.g., Tetlock et al., 2014; Tetlock & Gardner, 2015), to fine-tuning an individual forecaster's ability, such as engaging in statistical and probability training (e.g., Mellers et al., 2014). For example, studies have shown that the aggregation of individual opinions through prediction markets has led to higher accuracy of election forecasts (compared to expert polls), printer sales (compared to company projections), or even the weather (compared to National Weather Service) (Arrow et al., 2008).

Thus far, the dominant methods shown to increase prediction accuracy have been focused on fine-tuning rational thought processes, particularly debiasing and training (e.g., Arkes, 1991; Benson & Önkal, 1992; Larrick, 2004, p. 319; Mellers et al., 2014). However, emerging evidence suggests that emotional processes may also play a crucial role in improving predictive accuracy (e.g., Pham et al., 2012). While this may seem counterintuitive, many prominent scholars (e.g., Epstein et al., 1996) have hypothesized that humans have two dominant modes of cognitive processing, one leveraging logic and reason (comparable to system 2 [Kahneman, 2003]), whereas the other focuses on utilizing experiences, emotions, and feelings (comparable to system 1 [Kahneman, 2003]) to make judgments and decisions. The potential contribution of this second, emotions-based mode of processing to prediction accuracy remains largely unexplored, despite its fundamental role in human decision-making. This raises an intriguing question: could this second mode

of processing be equally effective for generating accurate predictions? We seek to explore this further in this paper.

The role of emotions in human behavior has been debated since the era of Ancient Greek and Roman Philosophers (Lazarus, 1999). While historical views toward emotions have often been negative (e.g., Lazarus, 1999; Graver, 2007; Seneca, 1928), contemporary research has documented their positive functions (e.g., Fredrickson, 2001; Isen, 2001; Mikels et al., 2010; Schwarz & Clore, 2003). Recent studies demonstrate that emotions, rather than hindering rational thought, complement and interact extensively with cognitive processes – an effect that has been historically underappreciated (George & Dane, 2016). For instance, emotional deficits, whether innate (Damasio, 1994), or experimentally induced (e.g., part of the amygdala was removed) (e.g., Bechara et al. 1999), have been shown to severely degrade decision-making quality. Emotions can substantially influence risk perception (Johnson & Tversky, 1983), motivate morally relevant behavior (Huebner et al., 2009), or influence choice when cognitive information points toward another option (Loewenstein, 1996). While debate continues about the inherent value of emotions, many prominent researchers (e.g., Ekman, 2007; Gilbert, 2006) now recognize emotions' centrality in human life and as potent and pervasive drivers of decision-making (Lerner et al., 2015), creating a constant "dance of affect and reason" (Finucane et al., 2003, p. 327).

Within the growing body of research on emotions and judgment and decision-making (JDM), making predictions – a higher-level cognitive process within JDM (Brust-Renck, 2021) – appears equally susceptible to emotional influences. The specific role of emotions in prediction accuracy remains particularly intriguing. Surprising findings by Seo and Barrett (2007) demonstrate that members of investment clubs who experienced more intense emotions outperformed their peers in simulated market returns – a process fundamentally reliant on future-state predictions. While these findings highlight the potential utility of emotions – albeit in one dimension of "intensity" – for making accurate predictions, little is known about which specific emotions can drive improved predictions. This paper addresses two primary goals: 1) examining the "black box" of human emotions to identify specific emotions that enhance prediction accuracy, and 2) determining whether and how to leverage these emotions when making predictions.

## Theory and Hypothesis Development

The ubiquity of emotions in human experience provides a critical foundation for examining their role in prediction accuracy. Zajonc (1980) argued that feelings are fundamentally inescapable from cognition, a finding supported by contemporary research showing that individuals experience at least one emotion 90% of the time (Trampe et al., 2015). From a large two-wave survey of over 1,000 adults, the most reported basic emotion categories were happiness and anger, followed by sadness (Scherer et al., 2004). This empirical foundation has evolved alongside theoretical frameworks of basic emotions. While Ekman (1992) initially proposed six basic emotions (later revised to seven [Ekman, 2007]), recent studies (Jack et al., 2014) suggest that anger, fear, happiness and sadness capture the four broad emotion categories. This selected list of basic emotions, characterized by their universality across cultures and their evolutionary significance in triggering behavior with high survival value, forms the foundation of the study.

This research adopts the theory of dual-process model of decision-making as its theoretical framework. While scholars employ various nomenclature ("system 1" and "system 2", "rational" and "experiential"), these labels often refer to a broadly similar distinction (for a review, see Evans, 2008). Across many accounts, the underlying contrast is between a slower, deliberative analytic[1] mode and a faster, intuitive experiential mode that is often affect-linked (e.g., Epstein, 1994; Kahneman, 2003; Loewenstein et al., 2001; Reyna, 2004; Slovic et al., 2002). When faced with decisions, individuals draw upon both information sources, with judgments guided by parallel "affect-laden experiential" (emotional) and "rational-analytic" (analytical) systems (e.g., Epstein et al., 1996).

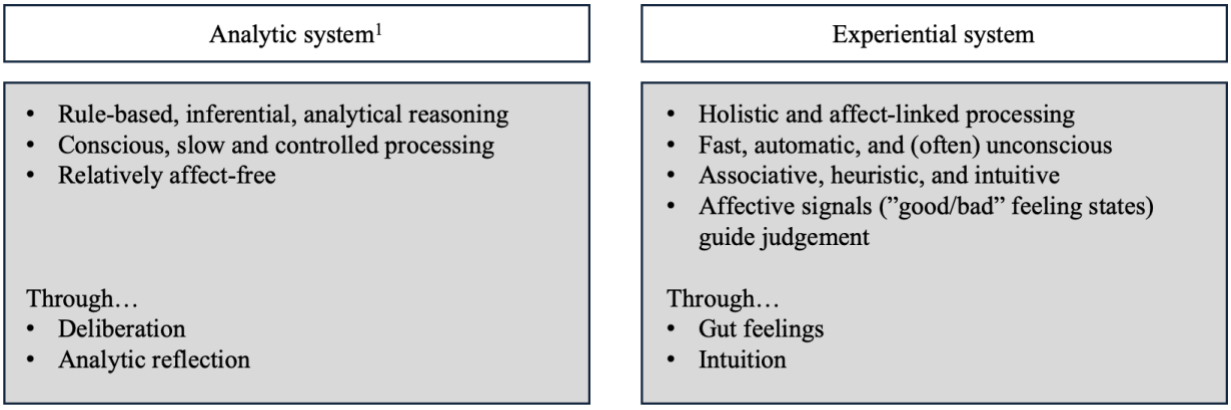
It is crucial to understand that these systems are not mutually exclusive. While individuals may tend to rely more heavily on one system in certain contexts, both systems typically operate in parallel and interactively, with their relative influence varying across contexts (Epstein, 1994; Lieberman, 2003). The degree to which one system dominates can vary based on many factors (e.g., emotional state, task complexity, time pressure, individual differences in processing, etc.). This dynamic interplay suggests that optimal

[1] Analytic is used instead of "rational", as there are some elements of rationality in the experiential model of processing (e.g., Slovic et al., 2002)

JDM, including making predictions, would involve appropriate collaboration between both systems rather than exclusive reliance on either one.

The analytical system is characterized predominantly by logical connections and reason-based processing. Its operations are typically slower and more detailed oriented, with behavior often motivated through conscious processing and systematic evaluation. In contrast, the experiential system demonstrates more association-based and holistic processing patterns. Notably, the experiential system can facilitate comprehensive information processing by distilling situations to their gist and "capturing its essence" (Stephen & Pham, 2008, p. 5). While processing in this system occurs more rapidly, the resulting behaviors are informed by accumulated experiential knowledge and past events. Experiential processing is generally considered heuristic in nature (Frankish, 2010, p. 915) (see Figure 1 for a summary). As both analytic (rational thought) and experiential (feelings and emotions) systems operate in parallel during JDM, the distinction lies in the relative dominance of one system over the other. Prior research suggests, perhaps counterintuitively, that the experiential system can be equally, if not more, beneficial compared to the analytical system (Seo & Barrett, 2007).

**Figure 1:** *Summary of Dual-Process Theory*




[1]Analytic is used instead of "rational", as there are strong elements of rationality in both systems

Based on Epstein & Pacini (1999); Slovic et al. (2002); Frankish (2010)

While dual-process theories often contrast the experiential (System 1) and analytic (System 2) modes of thinking (Evans, 2008; Kahneman, 2003), such work typically treats the "emotional" side in broad terms, rather than distinguishing between specific emotions.

However, discrete emotions (e.g., anger, sadness, fear, and happiness) can shape JDM in distinct ways (e.g., Lerner et al., 2015). By examining which specific emotions enhance or hinder prediction accuracy, this research extends beyond the simple dichotomy of "emotional vs. rational" and offers a more nuanced perspective on how particular emotions within System 1 might influence the quality of future-state judgements.

Research on emotional influences in JDM distinguishes between two types of emotions: integral and incidental. While some studies examine emotions emerging from a specific target (integral) (e.g., a felt attraction toward a new pet), most research has focused on emotions from unrelated sources that are independent of the target being evaluated (incidental) (e.g., an emotional state triggered by background classical music) (Bodenhausen, 1993). In fact, research has shown that incidental emotions elicited in one context can shape judgments in subsequent, unrelated context (e.g., Johnson & Tversky, 1983; Lerner et al., 1998). For example, Avnet et al. (2012) demonstrated how positive emotions experienced in one situation change cognition in a completely unrelated situation. Remarkably, people are often unaware that an emotion elicited in one situation has influenced their judgment in another situation (e.g., Johnson & Tversky, 1983, p. 30-31). Even if individuals are aware of the source of priming, they have difficulty discounting these influences (Schwarz & Clore, 1983) without over- or under-correcting (Wegener & Petty, 1997, p. 143). Consistent with prior research, this paper focuses solely on incidental emotions.

**Emotions and Prediction Accuracy**

To the best of my knowledge, there is no extant research on how specific emotions impact prediction accuracy of future events. However, characteristics of specific discrete emotions have been identified as potential levers that facilitate the processing of making predictions, and thus, can impact the accuracy of said predictions. One key distinction between different emotions is the type of information processing that the emotion encourages (Lerner et al., 2015), and affective states can show how information is processed and used in judgment (Clore et al., 1994).

Processing style accounts (Keltner & Lerner, 2010, p. 330) posit that different emotions trigger qualitatively different types of cognitive processing, broadly categorized

as systematic processing and heuristic processing. Systematic processing involves deep thinking and intensive reasoning, whereby one carefully evaluates information in detail. In contrast, heuristic processing leverages well-learned mental and often intuitive judgmental shortcuts (heuristics) that offer a more efficient way of processing (Chaiken, 1980; Chaiken & Ledgerwood, 2012, p. 247). Here, one considers only specific decision-relevant cues.

Recent research has demonstrated that simple heuristics can sometimes be more accurate in predicting the future than standard statistical methods (Gigerenzer & Gaissmaier, 2011). The authors have shown by using these simple, often innate, heuristics, predictions about upcoming election results, potential investment returns, and geographic locations of crimes were of similar accuracy – or more accurate – than those provided by advanced statistical models. For example, using the recognition heuristic [2] alone, Gaissmaier and Marewski (2011) showed that the forecasting of federal and state election winners was relatively accurate compared to extensive polls. Moreover, using the recognition heuristic, Ortmann et al. (2008) showed that recognition-based portfolios (the set of most recognized options), on average, outperformed the market, and other popular managed funds. In another example, Alter and Oppenheimer (2006) showed that using fluency heuristic[3] alone, people were able to accurately predict the performance of stocks. Additionally, the simple one-clever-clue heuristic allowed laypeople to be as accurate as an advanced computer algorithm to determine where a specific crime scene is (Snook et al., 2004), the caveat to this example is that this heuristic was taught to the laypeople. The aforementioned heuristics may be particularly accessible under certain emotional states that favor quick, holistic judgments.

Adopting Shah & Oppenheimer (2008), Gigerenzer and Gaissmaier (2011) described a heuristic as a strategy that ignores part of the available information (p. 454), with the goal of making decisions more quickly, frugally (limited number of cues being searched for), and sometimes, as accurately or more accurately than complex methods. Phrased differently, heuristic processing can improve prediction by filtering information

[2] The recognition heuristic uses one simple rule: if one of two alternatives is recognized and the other is not, then infer that the recognized alternative has the higher value with the respect to the criterion.
[3] If one of two objects is processed (i.e., retrieved or recognized) easier and faster than the other, then the object is inferred to have a higher value with respect to the criterion (Schooler & Hertwig, 2005, p. 611)

and focusing on a small set of high-utility signals (e.g., Silver, 2012) conducive to making accurate predictions. Gigerenzer & Gaissmaier (2011) review "less-is-more" findings, including inverse-U-shaped relations between accuracy and the amount of information used for prediction, such that beyond a point additional information can lead to more inaccurate predictions. In contrast, studies have shown that increased systematic processing through negative affect (e.g., sadness) can aggravate anchoring effects owing to increased focus on the anchor (Bodenhausen et al., 2000), rumination (Lyubomirsky & Nolen-Hoeksema, 1995), as well as negatively impact the accuracy of judgment (Ambady & Gray, 2002). It is imperative to note, again, that systematic processing and heuristic processing are not mutually exclusive, but rather a tendency for an individual to rely more on one type of processing compared to the other (e.g., Epstein, 1994).

Just as important, a classical understanding of heuristics typically assumes a trade-off between accuracy and effort; however, recent research into heuristics challenges the prior assumption. Contrary to traditional views, heuristic processing can be both conscious in nature, and precise in its outcomes (Gigerenzer & Gaissmaier, 2011). Additionally, the effort required for heuristic processing, compared to systematic processing, may not necessarily be less, but rather the way effort is required changes, as effective heuristic use can still require effort to identify and learn which cues are diagnostic in each environment (e.g., determining what the "one-clever-one" is rather than computing statistics).

**Trust-in-Feelings**

Work by Pham et al. (2012) demonstrates that trust in one's feelings[4] can systematically enhance prediction accuracy. The authors' research shows that individuals with high *trust*-in-their-feelings (TIF) make more accurate predictions about a variety of topics at a specific point in time in the future (e.g., politics, stock market behavior, and weather) at specific points in the future. TIF is manipulated in a simple way that increases or decreases reliance on one's gut feeling or what "feels right" (Avnet et al., 2012, p. 727; see also Avnet & Pham, 2005); thereby, altering the extent to which individuals treat their

[4] For the purposes of this paper, and consistent with the trust-in-feelings literature (e.g., Avnet & Pham, 2005; Pham et al., 2012), the term "feelings" is used broadly to refer to subjective affective experiences, including emotions and mood states, that can serve as inputs to judgment.

feelings as valid inputs to judgment. Importantly, the effect appears to hold primarily when individuals have sufficient domain knowledge and when outcomes are at least somewhat predictable (i.e., not purely random outcome).

Coined as the "emotional oracle effect" (p. 461), Pham et al. (2012) argue that feelings provide a "privileged window into all we tacitly know about the environment [around us]" (p. 468). Similar suggestions have previously been made by Bechara et al. (1997) who proposed that people can generate advantageous responses prior to consciously identifying an explicit decision strategy. It is also crucial to distinguish TIF from the low-effort "how-do-I-feel-about-it?" heuristic (e.g., Forgas, 1995; Schwarz & Clore, 1988): TIF is better conceptualized as a *metacognitive assessment* of whether one's feelings are diagnostic and should be used, rather than a mindless reliance on affect (Avnet, Pham & Stephen, 2012).

The relationship between TIF and prediction accuracy can be understood through two linked mechanisms: 1) *how* the TIF manipulation increases reliance on feelings, and 2) *why* greater reliance on feelings can improve prediction accuracy. Pham et al. (2012, p. 463) manipulated TIF by asking participants to recall either two (high TIF) or ten (low TIF) situations in which they trusted their feelings and it was the right thing to do. Recalling two instances is typically easy, leading participants to infer that trusting feelings is generally useful. Recalling ten instances is more difficult, leading participants to infer that feelings may be less reliable. This difference is consistent with an ease-of-retrieval mechanism (Schwarz et al., 1991; Wänke, 2013), closely tied to the availability heuristic (Tversky & Kahneman, 1973): people consider not only *what* they recall, but also *how* easily they recall it, and sometimes the affective tone of what comes to mind (Slovic et al., 2002). When examples come to mind fluently, feelings are more likely to be treated as valid and diagnostic (Avnet & Pham, 2005; Avnet et al., 2012); when retrieval is difficult, people may conclude their feelings are not consistently trustworthy (Pham et al., 2012).

Second, Pham et al. (2012) argue *why* TIF is useful for predicting future events is because feelings summarize large amounts of information, both subconsciously and consciously, about the world around us (e.g., Greifeneder et al., 2011), and such information is continuously absorbed in the background (Hogarth, 2005). In this view, the

predictive validity of feelings may lie in the fact that feelings can provide "relevant meta-summaries" of the vast amount of knowledge individuals accumulate, consciously and unconsciously, about their environment (Pham et al., 2012, p. 468) This logic aligns with affect-as-information accounts (Clore & Tamir, 2002; Schwarz, 1990), which treat affect (emotions) experienced as informational input to judgment.

**Trust in Feelings and Emotions**

Pham et al. (2012) have demonstrated a possible pattern: groups of individuals with higher reliance on their feelings (i.e., high TIF) achieve greater predictive accuracy compared to groups with low reliance (either low TIF, or no TIF). The TIF manipulation is said to be non-selective in nature, suggesting that it amplifies reliance on all feelings and emotions (Avnet & Pham, 2005; Avnet et al., 2012). For example, Avnet et al. (2012) found that participants who were induced into positive or negative moods via incidental video clips (e.g., a stand-up comedy excerpt vs. scenes from a tragic film) were then prompted to trust their feelings made evaluations more congruent with their induced mood state in an unrelated task (e.g., more *positive* ratings of a book synopsis following *positive* mood induction). Such experiments underscore the potential value of optimizing baseline emotions for predictive tasks. Because TIF concerns the extent to which people treat their subjective feelings as valid inputs to judgment, the question of whether the usefulness of "trusting feelings" depends on the discrete emotions shaping those feelings is crucial. However, Pham et al. (2012)'s studies did not examine which specific emotions might be most beneficial when relied upon; given our understanding of how emotions influence cognitive processing, it seems plausible that certain emotions may prove more valuable than others in generating accurate predictions.

As specific discrete emotions shape information processing and types of information accessed, they can significantly influence how predictions are formed (e.g., the way judgments are made during the prediction process), and ultimately, their accuracy. While emotional experiences typically involve multiple concurrent emotions (Larsen et al., 2004; Schimmack et al., 2002), increasing the salience of one specific emotion may heighten its influence without necessarily eliminating the affective signals or the

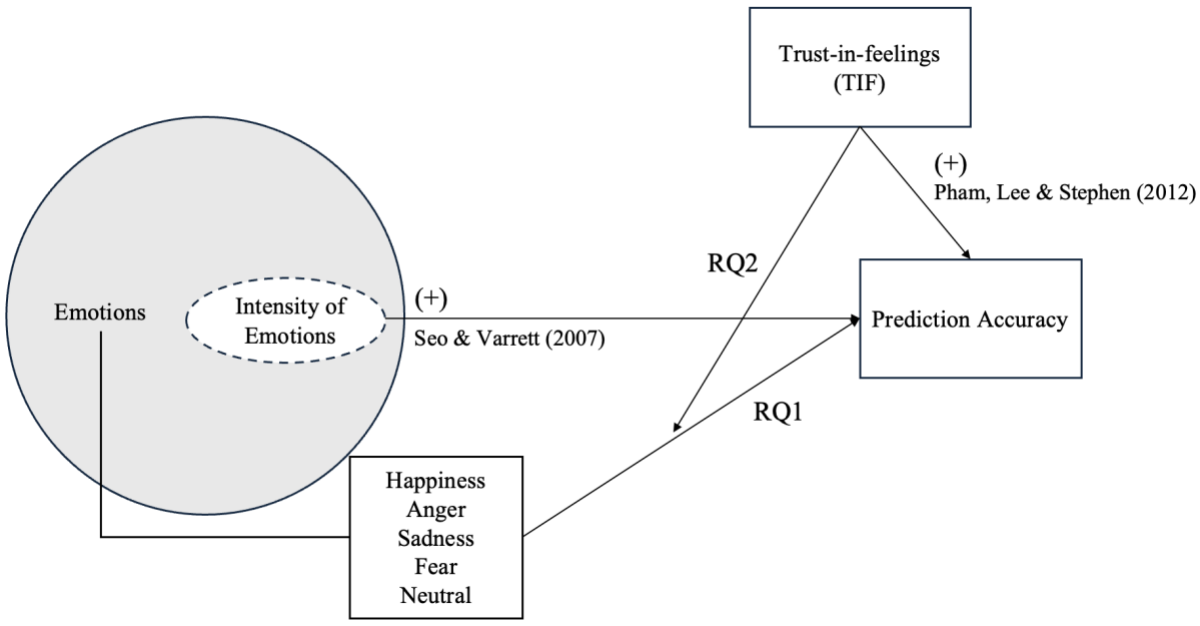
informational value of other concurrent emotional states (Cacioppo & Berntson, 1994; Larsen et al., 2004; Schimmack et al., 2002).

Beyond the four discrete emotions, it is also important consider a neutral affective state. If discrete emotions influence the formation of predictions in ways that affect their accuracy, then a state characterized by the relative absence of emotional experience provides a necessary point of comparison. Therefore, this paper addresses two critical research questions (see Figure 2 for a summary):

1) *How does anger, sadness, happiness, fear and/or neutrality impact prediction accuracy?*

2) *How does one's reliance on specific emotions (i.e., trust-in-feelings) influence the relationship between the emotions (anger, sadness, happiness, fear, neutral) and prediction accuracy?*

**Figure 2:** *Summary of Research Gaps*



## Hypotheses

### *Hypotheses for Emotion Effect (H1-H5)*

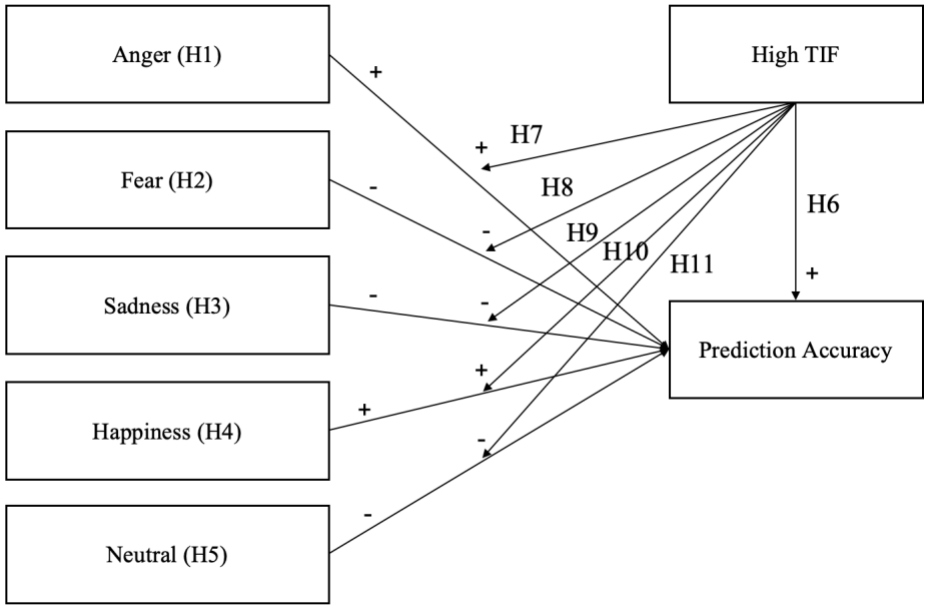
Building on Ekman (1992, 2007), we focus on four core emotions, happiness, sadness, anger, and fear, consistent with evidence that facial emotion communication may be organized around only four broad categories (Jack et al., 2014), alongside a neutral affective state as a theoretically motivated comparison condition. While early research suggested that positive emotions generally encourage heuristic processing and negative emotions promote systematic processing (e.g., Batra & Stayman, 1990; Forgas, 1995; Mackie & Worth, 1989), more recent findings reveal this valence-based approach (i.e., classifying emotions as either positive or negative) as insufficient. For instance, although anger and sadness are both negative states, their effects on information processing and judgment differ substantially (e.g., Lerner et al., 1998). Some emotions can promote heuristic processing regardless of their valence (Lerner & Keltner, 2000).

The appraisal-tendency framework[5] (Lerner & Keltner, 2000, 2001) proposes that discrete emotions are accompanied by specific appraisal patterns that carry over to later judgments and shape cognitive processing. These appraisals vary along dimensions such as certainty (the perceived predictability of future events), control (whether outcomes feel personally vs. situationally determined), and responsibility (whether causes are attributed to the self vs. others). For example, anger is associated with high certainty and perceived individual control, whereas fear elicits appraisals of low certainty and external control (Lerner & Keltner, 2001; Lerner & Tiedens, 2006). Because these appraisal tendencies can persist beyond the initial eliciting event, they may shape later JDM. In the present study, we focus on *certainty* as the main pathway, because certainty appraisals have shown to shift reliance to more heuristic versus more systematic processing (Tiedens & Linton, 2001), which may in turn influence forecast accuracy.

[5]The ATF extends beyond immediate emotional responses to explain how emotions influence unrelated decisions through persistent cognitive predispositions. For instance, anger's high certainty appraisals typically lead to more optimistic risk assessments across domains.

Moreover, approach-avoidance motivations[6] (Carver & Harmon-Jones, 2009) offer another lens for understanding how emotions influence processing style and behavior. Anger, despite negative valence, and many positive states (e.g., happiness) are often approach-oriented, promoting engagement and action, whereas fear and sadness are more avoidance-oriented, fostering withdrawal and caution. These motivational orientations can bias information processing: approach-oriented states tend to support faster, cue-driven judgments with narrower attentional scope, whereas avoidance-oriented states are more likely to heighten vigilance and risk sensitivity, encouraging more deliberative, systematic processing (Hengstler et al., 2014) – a distinction likely to matter most when forecasts are more uncertain (e.g., at longer horizons), where reliance on simple cues versus deeper evaluation may differentially affect accuracy. We propose the following hypotheses (see Figure 3 for an overview).

**Figure 3:** *Overview of Hypotheses*




Anger is a negative emotion with unique appraisals of high certainty and confidence in one's understanding of the situation (Lerner & Keltner, 2001). These appraisals create a strong belief that the individual understands what is happening and what is likely to happen

[6] Approach/avoidance motivations reflect fundamental action tendencies that operate at both behavioural and cognitive levels, influencing both physical responses, as well as the depth and style of information processing.

next. Anger also aligns with approach-oriented motivation (Carver & Harmon-Jones, 2009), which can encourage greater risk-taking and direct engagement with the environment. This combination of certainty and engagement promotes reliance on heuristic processing (Tiedens & Linton, 2001), as individuals feel confident in their ability to make quick judgments without extensive deliberation.

In prediction tasks, this reliance on heuristics can be advantageous especially in uncertain environments where key cues must be identified quickly. For example, an angry individual might quickly identify the most relevant cues and dismiss irrelevant distractions, allowing for faster and potentially more accurate decisions. Although anger might reduce the depth of cognitive analysis, its support of focus and efficient decision-making can lead to better prediction outcomes. Thus, anger is expected to enhance predictive performance by fostering confidence, engagement, and heuristic-based reasoning.

*Hypothesis 1. Anger during the process of making a prediction increases the accuracy of the prediction compared to control*

Fear is a negative emotion that is characterized by appraisals of low certainty and heightened perceived threat/uncertainty (Lerner & Keltner, 2001). These appraisals create a sense of doubt and vulnerability, prompting individuals to heighten vigilance and avoid risks. Fear is also an avoidance-oriented emotion (Carver & Harmon-Jones, 2009), motivating cautious withdrawal from potential threats and intensifying systematic or detail-oriented processing (Tiedens & Linton, 2001).

Although systematic processing can sometimes improve decision quality, in prediction tasks, fear's heightened vigilance can lead to over analysis of worst-case scenarios – even if many of those scenarios have low relevance or likelihood. This future-focused vigilance can overwhelm cognitive resources, causing individuals to second-guess more adaptive heuristics that might otherwise boost accuracy. Consequently, fear is expected to undermine predictive performance by increasing cognitive load, dwelling on potential threats, and distracting from core predictive cues.

*Hypothesis 2. Fear during the process of making a prediction decreases the accuracy of the prediction compared to control*

Both anger and fear are negative emotions, yet they differ in motivational orientation. Anger is typically approach-oriented, associated with heightened confidence and a greater likelihood of risk-taking (Lerner & Keltner, 2001), whereas fear is primarily avoidance-oriented, marked by uncertainty and caution (Carver & Harmon-Jones, 2009). This fundamental distinction helps explain why anger can sometimes expedite decisive judgment (and thus bolster prediction accuracy), whereas fear's heightened caution and vigilance can slow the decision process and compromise predictive performance (Carver & Harmon-Jones, 2009; Lerner & Keltner, 2001).

Sadness is a negative emotion that is commonly associated with appraisals of low certainty and diminished efficacy or helplessness (Tiedens & Linton, 2001; Smith & Ellsworth, 1985). Unlike fear's future-oriented vigilance, sadness often involves rumination[7] – a backward-focused, repetitive dwelling on negative thoughts or past events (Lyubomirsky & Nolen-Hoeksema, 1995). This inward focus can lead to overanalysis of irrelevant details rather than scanning for salient cues in the environment (Whitmer & Gotlib, 2013).

Although a certain amount of deliberation can be beneficial for decision-making, sadness-induced rumination can create cognitive overload and reduce attention to real-time signals. In a prediction context, the tendency to replay or overfit past failures may diminish the ability to identify the most diagnostic information. Thus, sadness' ruminative nature can ultimately detract from accurate forecasting.

*Hypothesis 3. Sadness during the process of making a prediction decreases the accuracy of the prediction compared to control*

Happiness is a positive emotion characterized by appraisals of high certainty and a sense of situational comfort (Lerner & Keltner, 2001; Tiedens & Linton, 2001). Like anger, happiness is an approach-oriented emotion (Carver & Harmon-Jones, 2009), encouraging active engagement with the environment (Fredrickson, 2001). However, unlike anger, which may fuel urgency or risk-taking, happiness often triggers a broaden-and-build

[7] Fear emphasizes vigilance toward future uncertainties, whereas sadness prompts rumination on past or ongoing negative events. Both can lead to cognitive overload and hamper the ability to identify the most relevant predictive cues.

mindset, expanding one's attention scope to include more creative or associative thought patterns.

In prediction tasks, this broadened cognitive focus can enhance the ability to spot key signals without becoming lost in irrelevant details. Empirical work shows that happiness also fosters reliance on heuristic processing and "general knowledge structures" (Bless et al., 1996, p. 666), enabling decision-makers to filter out extraneous information and draw on familiar cues or intuitive patterns and enabling individuals to focus on essential cues while avoiding unnecessary deliberation. This sense of efficacy and optimism can reduce indecision, allowing individuals to more confidently prioritize diagnostic information. As a result, happiness supports a flexible yet efficient processing style that can lead to more accurate forecasts.

*Hypothesis 4. Happiness during the process of making a prediction increases the accuracy of the prediction compared to control*

The implicit assumption in this study is that individuals who have a neutral emotional state will perform poorly compared to the other emotions. Neutral affect is typically conceptualized as mid-valence and low arousal on core affect dimensions (Russell, 1980), potentially lacking salient appraisals (e.g., certainty) and motivational tendencies (approach or avoid) that discrete emotions like anger, happiness, fear, or sadness provide (Lerner & Keltner, 2000). Because neutrality does not activate strong emotional signals, individuals in this state may fail to access the "black box" of implicit knowledge encoded in feelings that can guide JDM (Pham et al., 2012). In other words, neutrality may deprive forecasters of crucial heuristic shortcuts or motivational biases. Lacking clear affective guidance, neutral individuals might process predictive tasks more passively, thereby overlooking pivotal cues or patterns in information. Thus, without emotional cues to prioritize key information, neutrality can lead to lower prediction accuracy.

*Hypothesis 5. Neutral emotions during the process of making a prediction decrease the accuracy of the prediction compared to control*

Hypothesis 1 to 5 focuses on the idea that discrete emotions influence prediction accuracy through their distinct appraisals (i.e., certainty) and motivational tendencies

(approach vs. avoidance). Anger and happiness, characterized by high certainty and approach motivation, are expected to enhance accuracy by promoting heuristic processing. In contrast, fear and sadness, characterized by low certainty and avoidance motivation, are expected to hinder accuracy by promoting overly deliberative or ruminative processing. Finally, neutral states, which lack strong appraisals or motivations, are hypothesized to also encourage weak prediction performance.

### *Trust-in-Feelings (H6)*

The TIF manipulation has been shown by Pham et al. (2012) across multiple studies to increase the accuracy of predictions. However, the lack of a replication study raises critical questions about its efficacy and generalizability. This effect is tested in our experiment under the following hypothesis:

*Hypothesis 6. High TIF increases prediction accuracy compared to no TIF*

### *Emotion and Trust-in-Feelings Interactions (H7-H11)*

Building on Avnet et al. (2012), further reliance on feelings is expected to accentuate the effects of discrete emotions on prediction accuracy. By increasing reliance on emotional cues, high TIF is expected to enhance the adaptive potential (or maladaptive tendencies) of the different discrete emotions. Specifically, TIF is expected to strengthen the anger and happiness to foster heuristic processing and quick cue prioritization, while amplifying overanalysis seen in fear and sadness.

Thus, high TIF is hypothesized to improve prediction accuracy for anger and happiness (H1 and H4), but worsen it for fear, sadness, and neutral (H2, H3, H5). These dynamics are summarized in the following hypotheses:

*Hypothesis 7. High TIF positively moderates the relationship between anger and prediction accuracy, such that angrier individuals with high TIF will have a higher prediction accuracy*

*Hypothesis 8. High TIF negatively moderates the relationship between fear and prediction accuracy, such that more fearful individuals with high TIF will have a lower prediction accuracy*

*Hypothesis 9. High TIF negatively moderates the relationship between sadness and prediction accuracy, such that sadder individuals with high TIF will have a lower prediction accuracy*

*Hypothesis 10. High TIF positively moderates the relationship between happiness and prediction accuracy, such that happier individuals with high TIF will have a higher prediction accuracy*

*Hypothesis 11. High TIF negatively moderates the relationship between neutral emotions and prediction accuracy, such that neutral individuals with high TIF will have a lower prediction accuracy*

# Method

## Design and Sample

In total, 842[8] randomly selected participants completed the study through the Prolific online panel (see Appendix A for an overview of sample composition). Upon closer inspection, a total of 102 (approx. 12%) participants were disqualified from the subject pool due to failure of attention checks, incomplete responses, or clear indications of inaccurate input (data preparation procedure is detailed later). The hypotheses were tested using a random sample of 740 online participants through the Prolific online panel (see Table 1 for a detailed breakdown by group).

**Table 1:** *Detailed Breakdown of Sample*

| Group | No TIF (*n*) | High TIF (*n*) | Group Total |
|---|---|---|---|
| Control | 120 | 106 | 226 |
| Anger | 57 | 50 | 107 |
| Fear | 53 | 50 | 103 |
| Happiness | 59 | 45 | 104 |
| Sadness | 54 | 45 | 99 |
| Neutral | 52 | 49 | 101 |
| Total | 395 | 345 | **740** |

While the total usage sample was 740, we report hypothesis-specific *N* for each result. Prolific was chosen as the platform of choice as multiple scholars have pointed to

[8] The pre-screen 842 participant sample includes 241 from another paper.

the more favorable participant pre-screening and verification features on the platform, higher participant attention, and overall data quality (e.g., Albert & Smilek, 2023; Douglas, Ewell & Bauer, 2023). The participants received a remuneration of approximately 7 GBP (or approx. 9 USD at the time of the experiment)[9] for their time and were randomly assigned to one of 12 conditions of a 2 (high vs. no trust-in-feelings[10]) x 6 (happiness, anger, sadness, fear, neutral and control[11]) between-subjects design in a web-based survey format. Participants were pre-screened to confirm eligibility for the prediction questions. The random allocation of participants was performed by the survey platform, Qualtrics, to ensure that there is a relatively equal number of participants placed in each group. Participants were told that they would be taking part in four allegedly unrelated studies.

Prior to starting the four studies, each group of participants (except the two control groups[12]) were asked to self-report emotions. This was to gauge the baseline emotional state of the participant. Then, the participants were asked to watch a short neutral film clip (nature documentary excerpt) to minimize emotional carryover and to moderate any prior emotions (similar techniques used by e.g., Gabert-Quillen et al., 2014, Hewig et al., 2005). The first study manipulated participants' emotions through a film excerpt. After watching the film excerpt, participants were asked to continue to the second study, where their' momentary trust in their feelings were manipulated (another prompt for the control group), further explained in the following section. They are given seven minutes to complete their respective tasks (i.e., name 2 situations in which they were correct to trust their feelings to make a judgment or decision, or name 5 most recent Google searches for the no TIF group).

The participants were then introduced to a third filler task. The filler closely aligns with those used by Pham et al. (2012) and was meant to reduce demand effects from participants. The main task was administered in a supposedly unrelated fourth study, where

[9] The remuneration for the experiment varied depending on time investment as some groups received additional tasks after making the predictions; the value is a weighted average.
[10] As the low TIF was not different from the control group in Pham et al. (2012), a low TIF group was not created
[11] "Control" refers to the two no-induction conditions (TIF and no TIF) in which participants did not receive an emotion induction film (i.e., no attempt was made to alter their emotional state)
[12] The control groups did not self-report emotions

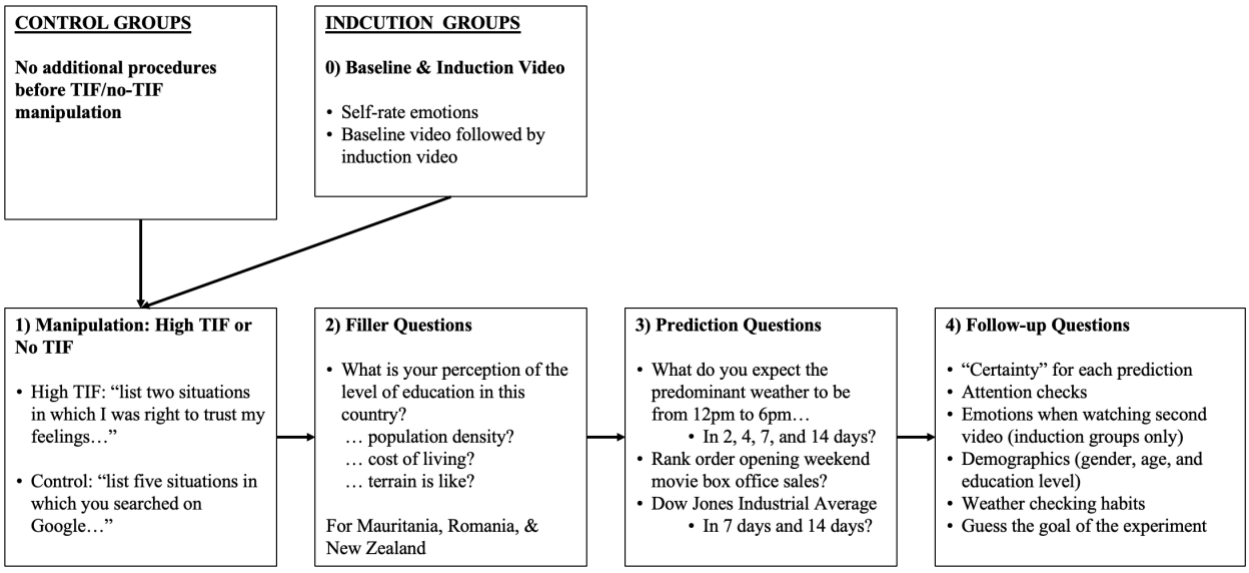
data for the dependent variable, forecasts (also detailed below), is collected[13]. At the completion of the fourth study, the participants were also asked to rate their confidence for each prediction on a 1 (very unsure) to 5 (almost certain) as a manipulation check of TIF. The participants were further asked to rate how involved they were in the prediction task by indicating their agreement with five statements (e.g., "I took this prediction task seriously." "I put effort into this task") (Pham et al., 2012, p. 464-465).

Additionally, the participants were requested, again, to self-report their emotions experienced from the second film excerpt (i.e., the film intended to induce specific discrete emotions). Following instructions first used by Philippot (1993) and Schaefer et al. (2010, p.8), participants were asked to report what they "actually felt when watching the film excerpt", with an explicit emphasis on "not what you believe people should feel". This procedure of retrospective evaluation of emotions is known to be a good predictor of the emotions felt during the induction (e.g., Fredrickson & Kahneman, 1993).

As the last step, the participants were given an attention check (Abbey & Meloy, 2017) question, invited to guess the goal of the study, and asked to answer basic demographic questions. The amount of time each participant took to make their predictions was recorded by the online survey program. At the end of the experiment, the participants were thanked and paid for their participation (see Figure 4 for an overview of experimental procedure, and Appendix B for screenshots of the survey administered in the study).

[13]The induction of a positive (or negative) emotion does not affect the level of trust in the emotion. The induction only affects the emotional information present, but not the validity of that information in the participant's mind (i.e., only the information from the participant's emotion provided is altered) (Avnet & Pham, 2005)

**Figure 4:** *Overview of Experiment Procedure*




## Measures

The induced emotions were the main independent variables. While prior research has shown that the film excerpts used to prime specific emotions are effective and alter the salient emotion to be the induced emotion, self-reported measures of one's emotions were still measured as a manipulation check. The moderator variable was one's trust-in-feelings and is a binary variable. Trust-in-feelings was also considered an independent variable when the direct effects were being measured (i.e., for H6). The dependent variable, prediction accuracy, was measured using the same approach as by Pham et al. 2012). The accuracy was determined by comparing the actual results of the prediction question against the predicted result of each participant (e.g., box office prediction vs. actual). Each participant's prediction results were coded as correct or incorrect. The specific calculation of prediction accuracy for each question is further explained in a later section.

### *Trust-in-Feelings (TIF) Manipulation*

After a brief explanation of the distinction between using feelings versus facts and logical reasons to make decisions, participants are asked to describe two "situations in which you trusted your feelings to make a judgment or decision, and it was the right thing to do" (Pham et al., 2012, p. 463), with an explicit requirement that only situations that did happen to be listed. Participants are to be given seven minutes to complete the task. The

participants could proceed to the next page sooner than seven minutes but were automatically sent to the next page once the seven-minute timer was up. The degree to which a participant had trust-in-feelings was not explicitly measured. Please see Appendix C.1 and Appendix C.2 for detailed instructions for both high TIF, and no TIF.

***Emotion Induction.***

Borrowing research from the field of psychology, there are five widely-recognized emotion induction techniques: visual stimuli (e.g., Gross & Levenson, 1995), listening to music (e.g., Krumhansl, 2002), autobiographical recall (e.g., Prkachin et al., 1999), situational procedure (e.g., creating a social situation) (e.g., Bushman & Baumeister, 1998), and mental imagery (i.e., creating vivid mental representations) (e.g., Mayer et al., 1995). Out of all the techniques, visual stimuli, in the form of the film excerpts (with sound) has been shown to generate effective induction results (for a review, refer to Siedlecka & Denson, 2019).

Research by Gilman et al. (2017) and Gabert-Quillen et al. (2014) point to large databases of emotion-eliciting film excerpts that have been shown to induce happiness, anger, sadness, or fear. Table 2 below lists a set of recognized film excerpts that have been reported by at least two groups of scholars to elicit strong inductions for a specific emotion (Gilman et al., 2017; Rottenberg et al., 2007; Schaefer et al., 2010; Wang et al., 2022). The initial minute-long baseline film excerpt, from Planet Earth (2006), was used for the attenuation of any pre-existing emotions prior to induction.

**Table 2:** *Description of Film-Excerpts for Induction*

| **Happiness** | **Anger** | **Sadness** | **Fear** | **Neutral**[14] |
|---|---|---|---|---|
| Wall-E (2008), 3:13<br><br>*Against the backdrop of outer space, two enchanting robots dance in harmony, captivating onlookers aboard a spaceship. The scene concludes as the enamored robots fly away together.* | Crash (2004), 6:33<br><br>*In a diner, a police officer has a stressful conversation on the phone. Later, the officer sexually assaults a black couple during a traffic stop. The scene ends with the violated wife re-entering the car.* | My Girl (1991), 2:37<br><br>*A funeral for a young boy unfolds. The boy's friend, struggling to comprehend his death, tearfully tries to offer him glasses. Overwhelmed, she runs out of the house, displaying a strong sense of loss.* | The Ring (2002), 2:49<br><br>*Tension builds as a man's TV mysteriously turns on, revealing a girl crawling out from the screen. The girl is trying to grab the man. The scene concludes with the TV screen fading to static, leaving an unsettling atmosphere.* | Blue Planet II (2017), 2:35<br><br>*A colorful display of tropical fish and coral swimming freely in the ocean.* |

Anger was induced using an excerpt from Crash (2004) lasting 6 minutes 33 seconds. Fear was induced using an excerpt from The Ring (2002) lasting 2 minutes 49 seconds. Happiness was induced using an excerpt from Wall-E (2008) lasting 3 minutes and 13 seconds. Sadness was induced using an excerpt from My Girl (1991) lasting 2 minutes 37 seconds. Neutrality was induced using an excerpt from Blue Planet II (2017) lasting 2 minutes and 35 seconds. The duration of these film excerpts ranged from 2 minutes 35 seconds to 6 minutes 33 seconds. Length was not shown to impact the efficacy and quality of emotion induction (e.g., Gabert-Quillen et al., 2014). Unlike the induction conditions, participants in the control condition did not view an induction clip.

To assess discrete emotional dimensions, a modified version of the Differential Emotions Scale questionnaire (DES; Izard et al., 1974) from McHugo et al. (1982) is used pre- and post-induction. The DES is a widely used and reliable self-reported scale of discrete emotions (Young et al., 2003) in emotion research. Participants rated on a 7-point Likert scale the intensity of feelings in ten emotion dimensions (sad, excited, angry, disgusted, happy, surprised, fearful, anxious, interested and calm) both prior to and after

[14]Nature documentary clips are commonly used as control stimuli in emotion-induction research (e.g., Gino & Schweitzer, 2008).

the emotion induction procedure. Because film excerpts are expected to co-elicit multiple affective states rather than a single "pure" emotion, we assessed the entire range of emotions (not only the five of interest) to test for co-elicitation and potential confounds (Schaefer et al., 2010; Siedlecka & Denson, 2019, p. 95).

***Prediction Tasks***

Three distinct prediction tasks were employed to assess forecast accuracy across different domains and time horizons, replicating Pham et al. (2012).

**Local Weather Forecasting (Prediction 1).** The first prediction task asked participants to forecast weather conditions in their respective locations using their postal codes as geographic identifiers. Participants made predictions for four different time horizons: 2 days, 4 days, 7 days, and 14 days ahead, specifically focusing on the weather between 12pm and 6pm[15]. For each prediction, participants selected from among seven meteorological categories: (a) sunny/fine/clear, (b) partially sunny with some clouds, (c) cloudy/overcast, (d) haze/fog, (e) thunderstorms, (f) rain, or (g) snow. Prediction accuracy was evaluated through binary classification: predictions matching the actual recorded weather conditions during the specified time window were coded as correct "1", while any deviations were coded as incorrect "0". To control for potential confounds, participants were also surveyed about their weather-checking behaviors.

**Opening Weekend Box Office (Prediction 2)**. The second task assessed participants' ability to predict relative box office performance. Participants were presented with three films slated for domestic U.S. release on April 14, 2023: Nefarious, Mafia Mamma, and Sweetwater. Each film was presented with standardized information including promotional artwork (advertising poster), plot synopsis (sourced from IMDB), genre classification (e.g., romance), and MPAA rating (e.g., 13+); see Appendix B, Section 21. Participants ranked the three films in predicted order of opening weekend box office revenue. Binary scoring was implemented: participants who correctly ordered all three films received a score of "1", while any deviation from the correct order was scored as "0".

[15] Pham et al. (2012) relied on the "predominant" daily weather as their source of truth; however, the authors could not specify exact source or platform used for their data. We chose the 12pm to 6pm time window given the availability of accurate weather data

The selected films were deliberately chosen to have comparable pre-release popularity, avoiding obvious market favourites that might skew predictions.

**Dow Jones Industrial Average Index (Prediction 3)**. The final prediction task focused on financial market forecasting, specifically targeting the Dow Jones Industrial Average (DJI). To ensure baseline finance and economics literacy, all participants were pre-screened (Appendix D) using a brief assessment of fundamental economic concepts, including an understanding of GDP and price-earnings ratios. Given the study's timing during a period of significant market uncertainty in early 2023 (IMF, 2023), participants were tasked with predicting the DJI closing values at two future points: 7 and 14 days from the study date. To ensure similar background knowledge, participants were provided with contextual information about current and historical DJI, following methodological precedents established by Pham et al. (2012) (see Appendix B, Section 23). Prediction accuracy was quantified by calculating the absolute difference between predicted and actual closing values for each target date.

## Outcome Measures and Scoring

### *Weather Accuracy Scoring*

During the study, participants were asked to predict the weather in their respective zip code. They were given seven options to choose from: 1) sunny/fine/clear, 2) partly sunny with some clouds, 3) rain, 4) thunderstorm, 5) cloudy/overcast, 6) snow, 7) haze/fog. Manual coding was required to convert the actual weather data from each zip code (obtained from timeanddate.com) to a corresponding option in the survey. For example, real weather data shown as “scattered clouds”, or “broken clouds” were then coded as “cloudy/overcast” to determine whether the respective weather prediction by the participant for a particular date was correct (see Appendix E for coding guide). Given the substantial number of data points that need to be hand coded, the newly coded dataset was spot checked twice (random selection of 25% twice). The error rate of the random spot checks were less than 0.1%, none of which made a difference to the initial manual coding. Predictions were coded as correct (1) if they matched the actual weather condition, or incorrect (0) if they did not match.

### *Box Office Accuracy Scoring*

The opening weekend box office performance of the three movies (Nefarious, Mafia Mamma, and Sweetwater), were checked through BoxOfficeMojo, a division of IMDB. BoxOfficeMojo receives relevant data directly from film studios and was leveraged as a data source in prior academic work (e.g., Mestyán et al., 2013). The ranking of the results was determined by ranking the gross sales of the movies from Friday, April 14 to Sunday, April 16, 2023, aligning with the opening weekend. Participants who correctly ranked all three movies were coded as correct "1"; any deviation from the correct ranking was coded as incorrect "0".

### *DJI Accuracy Scoring*

The value of the Dow Jones Industrial Average stock market index was determined through publicly available data on Yahoo! Finance, using the respective daily close value. Forecast accuracy was operationalized as the absolute difference between the actual close and the predicted close, where lower values indicate higher prediction accuracy.

## Sample Size and Power Analysis

To determine the minimum number of participants required for the experiment, we calculated ideal sample sizes for three types of analyses planned. Given that discrete emotions represent transient psychological states compared to stable trait-level constructs, we adopted slightly more conservative effect size estimates. The current investigation of emotional induction employed small-to-medium effect sizes ($f$ = .20, $w$ = .25, $f^2$ = .08). This conservative approach acknowledges that experimentally induced emotional states typically demonstrate more modest effects than stable individual differences (Lerner et al., 2015), while still detecting theoretically meaningful relationships.

### *ANOVA*

Using G*Power 3.1 (Faul et al., 2007) with an adjusted standard setting for behavioral sciences (statistical power = .80, small-to-medium effect size $f^2$ = .08, α = .05, $df$ = 10, number of groups = 12), the analysis indicated an ideal sample size of 416. This method was used to compute the direct effects of emotions and TIF for prediction questions with continuous outcomes (i.e., DJI predictions).

### *Chi-Square Tests*

G*Power 3.1 indicated a minimum total sample size of 206 using an adjusted standard statistical setting (statistical power = .80, small-to-medium effect size $w$ = .25, α = .05, df = 5). This method was used to examine the direct effects of emotions and TIF for prediction questions with binary outcomes (i.e., weather and opening weekend movie box office predictions).

### *Regression Analysis*

For logistic (weather and movie box office) and hierarchical (DJI) regression analyses, G*Power 3.1 indicated a minimum total sample size requirement of 167 using an adjusted standard statistical setting (statistical power = 0.8, small-to-medium effect size of $f^2$ = .08, α = .05, predictors = 11). We utilized the linear regression module in G*Power as a conservative approximation also for logistic regression sample size requirements. The regression methods were used to examine the precise effects of our independent variables on prediction accuracy, with a particular focus on emotion*TIF interactions (H7-H11).

Our final sample of 740 participants was adequately powered for main effect analyses (H1-H6) across tests. Group sizes ranged from 45 to 120 participants ($M$ = 61.67, $SD$ = 24.52) across 12 experimental conditions (6 emotions x 2 TIF levels). While all groups exceeded the recommended minimum of 30 participants (Van Smeden et al. 2016), the study was likely underpowered for detecting interaction effects (H7-H11). The high TIF emotion conditions ($n$ = 45-54) and no TIF emotion conditions ($n$ = 59-62) provided limited power for detecting emotion*TIF interactions (e.g., anger*TIF vs. anger*no TIF), which typically require substantially larger (i.e., at least 2x) samples than main effects (Simonsohn, 2015). Any interaction findings, null or otherwise, are interpreted cautiously, as we cannot confidently conclude whether TIF moderates emotion effects due to the present sample size. The control conditions had larger samples ($n$ = 120 no TIF, $n$ = 106 high TIF) and served as reference groups for many comparisons. Across all tests (direct and interaction effects), the sample sizes may vary slightly as there were a rare number of undocumented public weather data in some US zip codes.

## Pre-test Study

In preparation for the main experiment, a pre-test was conducted to emulate the experimental conditions and procedures. This pre-test, which replicated key aspects of the actual experiment, served as a valuable pilot study to assess the feasibility of the study design, and to identify any challenges or issues. Insights gained from the pre-test informed adjustments to the study design.

### *Participants*

A total of 91 (3 emotion conditions [anger, fear, neutral] x 2 TIF conditions [high, no]) participants from a Swiss University's Behavioral Lab student participant pool were part of this study. All participants were required to be fluent in English, as the film excerpts are in English. Approximately half of the participants were male (49 out of 91), and half were female (42 out of 91). The vast majority have a bachelor's degree (72 out of 91), while some have a master's degree (17 out of 91), Ph.D. (1 out of 91), or attended trade school (1 out of 91). The age of participants was predominately between 18 and 45 (34 between ages 18 and 25, 35 between ages 26 and 35, 3 between ages 36 and 45, and 1 between ages 46 and 55).

### *Procedures and Measures*

The pre-test was conducted online, and both the procedure and measures of the pre-test aligned with the procedure described previously. The only noteworthy difference in experimental design is that participants were randomly distributed to one of *only three* experimental conditions: fear, anger, and neutral. Moreover, some prediction questions (e.g., weather in the future) were omitted, as the main goal was to pre-test the survey mechanics and Qualtrics platform.

## Data Analytic Strategy

### *Data preparation*

Prior to conducting the primary analyses, the survey data underwent a thorough data cleaning and preparation process. First, the responses were screened for completeness, with partial responses being excluded. All responses less than 99% completion (the last 1% is clicking the "finish" button in Qualtrics, which participants may forget as they are asked

to re-visit the Prolific website for entering their completion code from Qualtrics) were disqualified. Second, the data were scrutinized for obvious outliers (e.g., took 60 minutes to complete the study), or non-sensical responses failing logical checks (e.g., 50000 for DJI prediction, virtually impossible to achieve). Third, all responses failing the simple attention check at the end (i.e., please select which animal was not shown in the first film) were discarded[16]. Finally, participants were excluded from the study if they had incomplete or truncated responses to the TIF manipulation. This decision stemmed from two considerations: a) if participants were not able to complete the manipulation within the generous timeframe of seven minutes, the likelihood of eliciting a high trust-in-feelings decreases substantially, and b) recognizing that a significant portion of the participants are motivated by financial incentives, participants may be concerned that failure to complete a survey section might lead to a rejection, potentially resulting in loss of income. This misinterpretation, in turn, might evoke undesirable emotions throughout the remainder of the study.

### *Analytic Tests Overview*

We implemented a multi-method analytic approach to test our hypotheses. For binary outcomes (weather and movie predictions), we employed chi-square tests and hierarchical binary logistic regression. For continuous outcomes (DJI predictions), we employed independent samples t-tests, two-way ANOVA, and hierarchical linear regression. Chi-square tests and t-tests provide descriptive comparisons and, for H6 specifically, maintained methodological consistency with Pham et al. (2012). The hierarchical regression models served as our primary hypothesis tests, enabling examination of both main effects and interaction effects while controlling for multiple effects simultaneously.

### *Test of Emotion Effects (H1-H5)*

To examine relationships between emotions (anger, fear, happiness, sadness, neutral) and prediction accuracy, we conducted chi-square tests for binary outcomes (weather predictions at 2, 4, 7 and 14 days; movie box office predictions). The control group served

[16] This attention check was not available for the control groups as they did not undergo emotion induction. However, the data from these participants were screened using all other techniques.

as the reference category for all comparisons. We report Pearson's chi-square statistics ($\chi^2$), degrees of freedom, *p*-values, effect sizes (Cramer's *V*), and percentage correct by emotion group. These tests provide descriptive comparisons of prediction accuracy across emotion conditions.

For continuous outcomes (DJI predictions at 7 and 14 days), we examined emotion main effects through hierarchical linear regression and two-way ANOVA. We report *F*-statistics, *p*-values, and partial eta-squared ($\eta^2p$) as effect size measures for ANOVA results.

***Test of Trust-in-Feelings (H6)***

To test whether high TIF increases prediction accuracy compared to no TIF, we conducted separate chi-square tests comparing TIF groups (high TIF vs. no TIF) for each binary outcome (weather predictions at 2, 4, 7, and 14 days; movie box office predictions), maintaining methodological consistency with Pham et al. (2012). We report Pearson's chi-square statistics ($\chi^2$), *p*-values, Cramer's *V*, and percentage correct by TIF group.

For continuous outcomes (DJI predictions), we conducted independent samples t-tests (equivalent to one-way ANOVA for two groups) comparing high TIF versus no TIF groups. We report *t*-statistics, *p*-values, and Cohen's *d* effect sizes. Levene's test assessed homogeneity of variance; if violated, we report results using Welch correction for unequal variances.

***Test of Interactions (H7-H11)***

We utilized hierarchical binary logistic regression for dichotomous outcomes (weather and movie box office) and hierarchical linear regression for continuous outcomes (DJI predictions) to investigate whether TIF moderates the relationship between discrete emotions and prediction accuracy.

For both regression approaches, emotions were dummy-coded (e.g., Aiken & West, 1991; Hayes, 2017) with the control group as the reference category (i.e., all dummy variables = 0). TIF was coded as 0 (no TIF) or 1 (high TIF). Interaction terms were created as product terms between each emotion dummy variable and TIF (e.g., Anger_Dummy*TIF)

Block 1 housed main effects (five emotion dummy variables and TIF) to establish direct relationships. Block 2 added five interaction terms (Anger*TIF, Fear*TIF, Happiness*TIF, Sadness*TIF, Neutral*TIF) to test moderation hypotheses (H7-H11). Model improvement from Block 1 to Block 2 was assessed through $R^2$ change ($\Delta R^2$) and chi-square change statistics for logistic regression, or $R^2$ change and $F$ statistics for linear regression.

For main effect hypotheses (H1-H6), we report Block 1 coefficients, which provided unconditional tests of emotion and TIF direct effects. These coefficients represent the average effect of each emotion across all participants, regardless of TIF condition. For interaction hypotheses (H7-H11), we report Block 2 coefficients and model fit statistics ($\Delta R^2$). Where Block 2 significantly improved model fit or where main effects notably changed from Block 1 to Block 2, we discuss these changes, as they provide insight into potential moderation patterns even when individual interactions may not reach statistical significance.

***Regression models***

**Logistic regression.** For binary outcomes, we report coefficient ($B$), odds ratios (OR), 95% confidence intervals, $p$-values, and Nagelkerke $R^2$. Block 1 coefficients test whether each emotion differs from control in prediction accuracy (H1-H5 via emotion dummy variables) and whether TIF increases accuracy overall (H6 via TIF coefficient). Block 2 interaction coefficients test whether TIF moderates each emotion's effect on accuracy (H7-H11). A significant positive interaction indicates high TIF amplifies that emotions' effect; a negative interaction indicates TIF attenuates the effect.

**Linear Regression.** For continuous outcomes (DJI predictions), we report standardized coefficients ($\beta$), $p$-values, variance inflation factors (VIF), $R^2$, and $R^2$ change ($\Delta R^2$). Model improvement was assessed through $F$ statistics. VIF values were used to assess multicollinearity.

**Two-way ANOVA.** To complement the regression analysis for continuous outcomes, we conducted two-way ANOVA with the emotion group (6 categories of emotions) and TIF (2 categories) as fixed factors. We report $F$-statistics for main effects, as well as each

emotion*TIF interaction, *p*-values, and partial eta-squared ($\eta^2p$). Levene's test assessed homogeneity of variance; we assumed normality given our sample size ($N$=740).

### *Coding and significance criteria*

All analyses used $\alpha = .05$ for statistical significance. Given the exploratory nature of examining multiple discrete emotions across prediction domains, we also note marginally significant effects ($p < .10$) but interpret them cautiously.

### *Post-hoc Tests and Multiple Comparisons*

It is important to note that a Bonferroni correction was not applied to all the analyses results. Researchers have shown that there is greater preference amongst researchers for avoiding the Bonferroni correction (e.g., Perneger, 1998), as it is unnecessarily underpowered and often commits Type 2 errors (e.g., Cabin & Mitchell, 2000; Nakagawa, 2004). Given the exploratory nature of this research and the theoretical motivations for each hypothesis, we cautiously prioritize substantive interpretation over strict control of family-wise error rates (Rothman, 1990).

## Results

### Pre-test Results

The pre-test illuminated key insights. First, the goal of the study was unguessable by any of the 91 participants. This affirms the robustness of the design for avoiding demand effects. Second, the induction films used for the selection emotions (i.e., anger, fear, and neutral) were able to effectively induce the desired emotional states. The anger treatment group reported feeling angrier ($M = 5.48$, $SD = 1.88$), compared to the fear treatment group ($M = 2.5$, $SD = 1.91$) and the neutral treatment group ($M = 1.3$, $SD = 0.70$), $F(2, 88) = 55.43$, $p < .001$. Similarly, the fear treatment group reported feeling more fearful ($M = 5.1$, $SD = 1.69$), compared to anger treatment group ($M = 3.68$, $SD = 1.49$), as well as neutral treatment group ($M = 1.2$, $SD = 0.48$), $F(2, 88) = 65.84$, $p < .001$. The neutral treatment group showed the least amount of emotional arousal *(*$M = 2.23$, $SD = 0.68$), computed as the average of all nine emotions, compared to the anger treatment group ($M = 4.25$, $SD = 0.66$), and the fear treatment group ($M = 3.83$, SD = 0.93). Third, the pre-test showed the necessity of adding an emotion to directly measure the state of neutrality. Calmness has

been shown to be an emotion to measure the effect neutral film clips (Gabert-Quillen et al., 2015) and was introduced to the self-reported emotions questionnaire as an additional emotion.

***Descriptive Statistics***

A total of 740 respondents were included as a part of this study. Of the respondents, 466 (62.97%) identified as males, 260 (35.14%) as females, 12 (1.62%) as non-binary, and 2 (0.27%) preferred not to say. Moreover, 83 (11.22%) respondents were within the ages of 18-25, 242 (32.70%) were within 26-35, 172 (23.24%) within 36-45, 125 (16.89%) within 46-55, and 118 (15.95%) aged 56 and above. 154 (20.81%) respondents held a high school as the highest education level, 384 (51.89%) respondents had a bachelor's degree, 137 (18.51%) respondents had a master's degree, 30 (4.05%) respondents had a Ph.D. or higher, 29 (3.92%) attended a trade school, and 6 (0.81%) preferred not to say.

Randomization checks confirmed that experimental conditions did not differ significantly on demographic variables (gender, age, education), weather checking behavior, or duration (all $p$s $> .05$), indicating successful random assignment. A small number of comparisons on baseline task engagement measures reached significance (4 of 26), but all had negligible effect sizes ($\eta^2 < .02$) and likely reflect Type I error given the number of tests conducted.

***Emotion induction***

Baseline emotional states were comparable across experimental groups prior to induction, indicating successful randomization. Following the film excerpts, the target emotions were successfully induced in each condition. For descriptive purposes, both TIF conditions were collapsed within each respective emotion group. Participants in anger condition reported elevated anger ($M = 5.13$), and notably, elevated disgust ($M = 5.47$) and anxiety ($M = 4.59$). Participants in the fear condition reported elevated fear ($M = 3.87$) and anxiety ($M = 4.06$), with substantially reduced happiness ($M = 2.24$) and calmness ($M = 2.74$). Participants in the happiness condition reported highly elevated happiness ($M = 4.81$) and calmness ($M = 5.01$), with low levels of negative emotions. Participants in the sadness condition reported elevated sadness ($M = 4.60$), with substantially reduced happiness ($M = 1.69$) and excitement ($M = 1.70$). Participants in the neutral condition reported elevated

calmness ($M$ = 5.55) and moderate happiness ($M$ = 4.33), with minimal levels of negative emotions across all dimensions. While the inductions successfully elevated target emotions relative to baseline and relative to other conditions, the self-reported emotion profiles revealed co-activation of related emotions, particularly anger co-occurring with disgust and anxiety, fear co-occurring with anxiety and happiness with neutral. This pattern of emotional co-activation is consistent with prior research (e.g., Schaefer et al., 2010, Lench et al. 2011) showing emotion inductions often produce blends of emotions. See Table 3 an overview of self-reported emotion ratings.

**Table 3:** *Self-Reported Pre and Post Emotion Induction Ratings*

| Group | | Interested | Sad | Excited | Angry | Disgusted | Happy | Surprised | Fearful | Anxious | Calm |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Anger | *Pre-Ind.* | 5.11 | 1.55 | 3.30 | **1.25** | 1.21 | 4.29 | 1.68 | 1.38 | 1.78 | 5.07 |
| | *Post-Ind.* | 4.83 | 4.21 | 2.21 | **5.13** | 5.47 | 1.51 | 3.40 | 3.48 | 4.59 | 2.21 |
| Fear | *Pre-Ind.* | 5.13 | 1.69 | 3.29 | 1.28 | 1.29 | 4.26 | 1.83 | **1.48** | 2.20 | 5.02 |
| | *Post-Ind.* | 4.69 | 1.76 | 3.12 | 1.43 | 2.36 | 2.24 | 3.78 | **3.87** | 4.06 | 2.74 |
| Happiness | *Pre-Ind.* | 4.88 | 1.63 | 3.35 | 1.38 | 1.26 | **4.43** | 1.82 | 1.43 | 1.83 | 5.11 |
| | *Post-Ind.* | 5.17 | 1.94 | 3.62 | 1.21 | 1.34 | **4.81** | 3.10 | 1.40 | 1.49 | 5.01 |
| Neutral | *Pre-Ind.* | 4.97 | 1.81 | 3.03 | 1.31 | 1.17 | 4.12 | 1.57 | 1.42 | 2.28 | **4.70** |
| | *Post-Ind.* | 4.93 | 1.37 | 2.63 | 1.07 | 1.05 | 4.33 | 2.42 | 1.17 | 1.47 | **5.55** |
| Sadness | *Pre-Ind.* | 4.98 | **1.41** | 3.13 | 1.17 | 1.07 | 4.30 | 1.63 | 1.32 | 1.86 | 4.79 |
| | *Post-Ind.* | 4.56 | **4.60** | 1.70 | 1.60 | 1.30 | 1.69 | 2.42 | 1.57 | 2.39 | 3.39 |

## Hypotheses Results

We tested our hypotheses using chi-square tests and hierarchical logistic regression for binary outcomes (weather and movie box office predictions), as well as independent samples t-test, two-way ANOVA and hierarchical linear regression for continuous outcomes (DJI predictions). See Table 4 for differences (percentage point for weather/movie box office predictions and absolute error for DJI predictions) relative to control, broken down by emotion group and TIF condition.

**Table 4:** *Emotion-Specific Differences from Control by TIF Condition*

| | | Difference Compared to Control (in pp) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Prediction** | **Condition** | **Anger** | **Fear** | **Sadness** | **Happiness** | **Neutral** | **Control (%)** | **Test Statistic** |
| *2-Day Weather* | | | | | | | | $\chi^2 = 13.13$* |
| | Overall | 17.60 | 10.90 | 10.30 | 13.20 | 5.80 | 23.40% | |
| | No TIF | 13.50 | 5.00 | 11.90 | 11.20 | 0.20 | 23.30% | |
| | High TIF | 22.20 | 17.20 | 8.20 | 15.90 | 12.00 | 23.60% | |
| *4-Day Weather* | | | | | | | | $\chi^2 = 9.33^{\dagger}$ |
| | Overall | 4.30 | 10.70 | 13.10 | 10.70 | 11.00 | 23.30% | |
| | No TIF | 5.80 | 17.00 | 10.80 | 16.80 | 12.00 | 18.80% | |
| | High TIF | 3.00 | 3.70 | 16.10 | 3.50 | 10.00 | 28.30% | |
| *7-Day Weather* | | | | | | | | $\chi^2 = 16.32$** |
| | Overall | 14.00 | 16.20 | 16.00 | 11.40 | 4.50 | 19.70% | |
| | No TIF | 15.40 | 19.90 | 14.30 | 12.50 | (4.30) | 19.70% | |
| | High TIF | 12.10 | 12.20 | 18.00 | 9.70 | 14.20 | 19.80% | |
| *14-Day Weather* | | | | | | | | $\chi^2 = 11.84$* |
| | Overall | 12.10 | 9.90 | (0.60) | 5.80 | 10.70 | 15.80% | |
| | No TIF | 7.70 | 5.30 | (11.80) | 4.80 | 9.50 | 15.50% | |
| | High TIF | 17.30 | 15.30 | 12.90 | 7.30 | 12.30 | 16.00% | |
| *Movie Box Office* | | | | | | | | $\chi^2 = 1.57$ |
| | Overall | 0.40 | (2.80) | 0.80 | (1.00) | 3.40 | 16.40% | |
| | No TIF | 8.60 | 0.70 | 9.70 | 2.80 | 2.90 | 12.50% | |
| | High TIF | (8.80) | (6.80) | (9.70) | (5.20) | 3.70 | 20.80% | |

| | | Difference Compared to Control (in points) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Prediction** | **Condition** | **Anger** | **Fear** | **Sadness** | **Happiness** | **Neutral** | **Control (pts)** | **Test Statistic** |
| *DJI 7-Day* | | | | | | | | $F = 0.707$ |
| | Overall | (19) | 8 | 246 | 42 | 70 | 813 | |
| | No TIF | 38 | (18) | 217 | (45) | 55 | 802 | |
| | High TIF | (84) | 37 | 280 | 154 | 86 | 826 | |
| *DJI 14-Day* | | | | | | | | $F = 0.437$ |
| | Overall | 168 | 111 | 149 | 18 | 220 | 1206 | |
| | No TIF | 273 | 105 | 206 | 34 | 332 | 1183 | |
| | High TIF | 47 | 119 | 80 | (6) | 103 | 1233 | |

*Note.* Values in the upper panel ("Difference Compared to Control [in pp]") are percentage-point differences in prediction accuracy relative to the control condition within the same TIF condition (control accuracy is shown in the "Control (%)" column). Positive values indicate higher accuracy than control. Values in the lower panel ("Difference Compared to Control [in points]") are point differences in absolute prediction error, computed as control minus condition. Positive values indicate lower prediction error than control (better performance). For example, proportion of correct predictions for 2-Day Weather/No-TIF is 36.8% (23.30% control + 13.50pp) $^{\dagger}p < .10$, $*p < .05$, $**p < .01$.

### *Direct Effects of Emotions on Prediction Accuracy (H1-H5)*

**Weather predictions.** Chi-square tests revealed significant associations between emotion groups and prediction accuracy for three of four weather prediction tasks. For 2-day weather predictions, emotion groups differed significantly in accuracy, $\chi^2(5, N = 724)$

= 13.13, $p$ = .022, Cramer's $V$ = .135. Participants in the anger condition achieved the highest accuracy (41.0%), followed by happiness (36.6%), fear (34.3%), sadness (33.7%), and neutral (29.2%), all exceeding the control group (23.4%).

For 4-day predictions, the association approached significance, $\chi^2(5, N = 732) = 9.33$, $p$ = .097, Cramer's $V$ = .113, with similar patterns: sadness (36.4%), neutral (34.3%), happiness (34.0%), fear (34.0%), and anger (27.6%) all numerically exceeded control (23.3%).

For 7-day predictions, emotion groups differed significantly, $\chi^2(5, N = 730) = 16.32$, $p$ = .006, Cramer's $V$ = .149. Accuracy rates are fear (35.9%), sadness (35.7%), anger (33.7%), happiness (31.1%), and neutral (24.2%), all exceeding control group (19.7%).

For 14-day predictions, a significant association emerged, $\chi^2(5, N = 726) = 11.84$, $p$ = .037, Cramer's $V$ = .128, with anger (27.9%), neutral (26.5%), fear (25.7%), and happiness (21.6%) outperforming control (15.8%), while sadness (15.2%) performed similarly to the control group.

To examine the unique contribution of each emotion while controlling for TIF, we conducted hierarchical logistic regression. We report Block 1 coefficients for main effects (H1-H5), as these provide unconditional tests of emotion across all participants.

***2-day weather predictions.*** Anger significantly predicted accuracy (OR = 2.284, 95% CI [1.387, 3.761], $p$ = .001), as did fear (OR = 1.709, 95% CI [1.022, 2.858], $p$ = .041) and happiness (OR = 1.917, 95% CI [1.149, 3.197], $p$ = .013). Sadness showed a marginally significant effect (OR = 1.673, 95% CI [0.992, 2.822], $p$ = .053). Neutral showed no effect (OR = 1.350, 95% CI [0.787, 2.315], $p$ = .276).

Notably, when interaction terms were added in Block 2, the model did not significantly improve ($\Delta\chi^2$ = 2.59, $p$ = .764, $\Delta R^2$ = .004), but the emotion main effects weakened considerably: anger (OR = 1.923, $p$ = .063), fear (OR = 1.301, $p$ = .484), happiness (OR = 1.735, $p$ = .119), and sadness (OR = 1.789, $p$ = .106) all lost statistical significance or approached marginal significance. This pattern suggests that interaction terms absorbed variance from main effects despite not being individually significant, potentially indicating subtle moderating effects that the study was underpowered to detect.

***4-day weather predictions.*** Fear significantly predicted accuracy (OR = 1.691, 95% CI [1.012, 2.826], $p = .045$), as did happiness (OR = 1.719, 95% CI [1.028, 2.874], $p = .039$), sadness (OR = 1.895, 95% CI [1.132, 3.172], $p = .015$), and neutral (OR = 1.724, 95% CI [1.026, 2.898], $p = .040$). Anger showed no significant effect (OR = 1.262, 95% CI [0.743, 2.144], $p = .389$).

When interaction terms were added in Block 2, the model did not significantly improve ($\Delta\chi^2 = 3.71$, $p = .591$, $\Delta R^2 = .007$). Most emotion effects remained significant or strengthened slightly in Block 2: fear (OR = 2.413, $p = .018$), happiness (OR = 2.386, $p = .016$), and neutral approached significance (OR = 1.919, $p = .088$), though sadness weakened and lost significance (OR = 1.818, $p = .116$).

***7-day weather predictions.*** Anger (OR = 2.067, 95% CI [1.224, 3.489], $p = .007$), fear (OR = 2.280, 95% CI [1.355, 3.836], $p = .002$), happiness (OR = 1.839, 95% CI [1.080, 3.131], $p = .025$), and sadness (OR = 2.263, 95% CI [1.333, 3.839], $p = .002$) all significantly predicted accuracy. Neutral showed no effect (OR = 1.302, 95% CI [0.739, 2.292], $p = .361$).

When interaction terms were added in Block 2, the model did not significantly improve ($\Delta\chi^2 = 5.61$, $p = .347$, $\Delta R^2 = .010$). Most effects remained significant in Block 2: anger (OR = 2.209, $p = .029$), fear (OR = 2.682, $p = .007$), sadness (OR = 2.102, $p = .046$), though happiness weakened to marginal significance (OR = 1.941, $p = .068$).

***14-day weather predictions.*** Anger (OR = 2.094, 95% CI [1.192, 3.676], $p = .010$), fear (OR = 1.863, 95% CI [1.047, 3.316], $p = .034$), and neutral (OR = 1.947, 95% CI [1.092, 3.471], $p = .024$) significantly predicted accuracy. Happiness (OR = 1.512, 95% CI [0.832, 2.746], $p = .175$) and sadness (OR = 0.964, 95% CI [0.498, 1.864], $p = .913$) showed no effects.

For 14-day predictions, Block 2 improvement was marginally significant ($\Delta\chi^2 = 10.10$, $p = .073$, $\Delta R^2 = .021$), though driven primarily by a significant sadness*TIF interaction. In Block 2, sadness showed a significant negative main effect (OR = 0.209, 95% CI [0.047, 0.938], $p = .041$), representing sadness's effect in the no TIF condition. This effect must be interpreted in conjunction with the significant sadness*TIF interaction

(discussed in the interaction effects section below). No other emotions showed significant effects in Block 2: anger (OR = 1.646, $p = .221$), fear (OR = 1.426, $p = .404$), happiness (OR = 1.390, $p = .425$), or neutral (OR = 1.815, $p = .146$).

**Movie box office predictions.** For movie box office predictions, no significant association emerged between emotion groups and accuracy, $\chi^2(5, N = 740) = 1.57$, $p = .905$, Cramer's $V = .046$. Accuracy rates across groups were nearly identical: neutral (19.8%), sadness (17.2%), anger (16.8%), control (16.4%), happiness (15.4%), and fear (13.6%).

Hierarchical logistic regression Block 1 confirmed no significant emotion effects: anger (OR = 1.033, 95% CI [0.557, 1.915], $p = .917$), fear (OR = 0.802, 95% CI [0.413, 1.560], $p = .516$), happiness (OR = 0.932, 95% CI [0.492, 1.765], $p = .828$), sadness (OR = 1.060, 95% CI [0.565, 1.991], $p = .856$), or neutral (OR = 1.260, 95% CI [0.689, 2.302], $p = .453$). Block 2 did not significantly improve fit ($\Delta\chi^2 = 7.76$, $p = .170$, $\Delta R^2 = .018$).

**DJI predictions.** Two-way ANOVA revealed no significant main effect of emotion group on DJI prediction error at 7 days, $F(5, 728) = 0.707$, $p = .618$, $\eta^2p = .005$, or at 14 days, $F(5, 728) = 0.437$, $p = .823$, $\eta^2p = .003$. Mean prediction errors ranged from 567 (sadness) to 832 (anger) for 7-day predictions, and from 986 (neutral) to 1,206 (control) for 14-day predictions, but these differences were not statistically significant. Levene's test indicated homogeneity of variance for both 7-day predictions ($p = .792$) and 14-day predictions ($p = .351$).

Hierarchical linear regression Block 1 similarly showed no significant emotion effects for either 7-day ($R^2 = .005$, $F(6, 733) = 0.580$, $p = .747$) or 14-day predictions ($R^2 = .005$, $F(6, 733) = 0.631$, $p = .705$). No individual emotion coefficients reached significance for either time horizon, though sadness showed a marginally significant negative effect on 7-day predictions ($\beta = -0.069$, $p = .092$), indicating lower prediction error. Block 2 did not improve fit for 7-day predictions ($\Delta R^2 = .001$, $\Delta F = 0.195$, $p = .964$) or 14-day predictions ($\Delta R^2 = .001$, $\Delta F = 0.160$, $p = .977$) and the full models remained nonsignificant.

### *Direct Effects of Trust-in-Feelings on Prediction Accuracy (H6)*

**Weather and Box Office Predictions.** Chi-square tests examined whether high TIF improved prediction accuracy compared to no TIF. For 2-day weather predictions, no significant difference emerged, $\chi^2(1, N = 724) = 1.862$, $p = .172$, Cramer's $V = .051$, with high TIF achieving 34.0% accuracy versus 29.3% for no TIF. For 4-day predictions, the difference approached significance, $\chi^2(1, N = 732) = 2.791$, $p = .095$, Cramer's $V = .062$, with high TIF at 33.2% versus 27.6% for no TIF. For 7-day predictions, no significant difference emerged, $\chi^2(1, N = 730) = 0.095$, $p = .758$, Cramer's $V = .011$ (high TIF at 28.9% vs. no TIF at 27.9%). However, for 14-day predictions, high TIF significantly improved accuracy, $\chi^2(1, N = 726) = 5.795$, $p = .016$, Cramer's $V = .089$, with 25.0% accuracy for high TIF versus 17.7% for no TIF. For movie predictions, no significant difference emerged, $\chi^2(1, N = 740) = 0.178$, $p = .674$, Cramer's $V = .015$ (high TIF at 17.1% vs. no TIF at 15.9%).

In hierarchical logistic regression Block 1, TIF showed no significant main effects for 2-day weather (OR = 1.267, 95% CI [0.922, 1.740], $p = .145$) or 7-day weather (OR = 1.062, 95% CI [0.766, 1.472], $p = .717$), and was only marginally significant for 4-day weather (OR = 1.322, 95% CI [0.961, 1.818], $p = .086$). TIF significantly predicted 14-day weather accuracy (OR = 1.566, 95% CI [1.091, 2.250], $p = .015$) and showed no effect on movie predictions (OR = 1.086, 95% CI [0.736, 1.603], $p = .678$).

**Dow Jones Index Predictions.** Independent samples t-tests revealed no significant effect of TIF on DJI prediction error. For 7-day predictions, high TIF ($M = 764.02$, $SD = 1353.50$) did not differ significantly from no TIF ($M = 768.72$, $SD = 1072.88$), $t(738) = 0.053$, $p = .958$, Cohen's $d = 0.004$, 95% CI for mean difference [-170.61, 180.01]). For 14-day predictions, the difference was also not significant, $t(738) = -1.205$, $p = .229$, Cohen's $d = -0.089$, 95% CI for mean difference [-348.57, 83.39], with high TIF ($M = 1184.78$, $SD = 1707.81$) showing numerically larger errors than no TIF ($M = 1052.19$, $SD = 1276.16$).

In Block 1 of hierarchical linear regression, TIF showed no significant effects for both 7-day predictions ($\beta = -0.003$, $p = .944$) and 14-day predictions ($\beta = 0.045$, $p = .221$). The results show no support for Pham et al.'s (2012) findings, as TIF effects emerged significantly in only one of seven prediction tasks.

### *Emotion*Trust-in-Feelings Interactions (H7-H11)*

**Weather Predictions.** To test whether TIF moderated emotion effects (H7-H11), we examined Block 2 of the hierarchical logistic regression, which added five interaction terms (Anger*TIF, Fear*TIF, Happiness*TIF, Sadness*TIF, Neutral*TIF) to the Block 1 main effects.

***2-day weather predictions.*** Block 2 did not significantly improve model fit ($\Delta\chi^2$ = 2.586, $p$ = .764, $\Delta R^2$ = .004). No individual interaction terms reached significance: anger*TIF (OR = 1.426, 95% CI [0.525, 3.873], $p$ = .487), fear*TIF (OR = 1.717, 95% CI [0.611, 4.827], $p$ = .305), happiness*TIF (OR = 1.221, 95% CI [0.438, 3.408], $p$ = .703), sadness*TIF (OR = 0.845, 95% CI [0.296, 2.413], $p$ = .753), or neutral*TIF (OR = 1.762, 95% CI [0.595, 5.217], $p$ = .306).

***4-day weather predictions.*** Block 2 did not significantly improve model fit ($\Delta\chi^2$ = 3.714, $p$ = .591, $\Delta R^2$ = .007). Anger*TIF (OR = 0.819, 95% CI [0.283, 2.371], $p$ = .713), fear*TIF (OR = 0.494, 95% CI [0.176, 1.384], $p$ = .180), happiness*TIF (OR = 0.495, 95% CI [0.175, 1.401], $p$ = .185), sadness*TIF (OR = 1.115, 95% CI [0.394, 3.152], $p$ = .838) and neutral*TIF (OR = 0.819, 95% CI [0.289, 2.323], $p$ = .708) all did not reach statistical significance.

***7-day weather predictions.*** Block 2 showed no improvement of model fit ($\Delta\chi^2$ = 5.605, $p$ = .347, $\Delta R^2$ = .010). The neutral*TIF interaction approached significance (OR = 2.811, 95% CI [0.873, 9.053], $p$ = .083), suggesting TIF may amplify neutral emotion's effect, though this pattern should be interpreted cautiously given the marginal significance. No other interaction terms reached significance: anger*TIF (OR = 0.859, 95% CI [0.300, 2.461], $p$ = .777), fear*TIF (OR = 0.710, 95% CI [0.250, 2.019], $p$ = .521), happiness*TIF (OR = 0.874, 95% CI [0.299, 2.559], $p$ = .806), or sadness*TIF (OR = 1.169, 95% CI [0.405, 3.373], $p$ = .772).

***14-day weather predictions.*** Block 2 improvement approached significance for 14-day weather predictions ($\Delta\chi^2$ = 10.097, $p$ = .073, $\Delta R^2$ = .021), driven primarily by a single interaction. The sadness*TIF interaction was highly significant (OR = 10.157, 95% CI [1.833, 56.284], $p$ = .008). To interpret this interaction, we examined the pattern of simple

effects. In Block 2, the sadness main effect (OR = 0.209, 95% CI [0.047, 0.938], $p$ = .041) represents sadness's effect in the no TIF condition, indicating that sadness significantly decreased accuracy compared to the control group when participants did not receive the TIF manipulation. However, the significant positive interaction indicates that TIF completely reversed this pattern. The combined effect of sadness with high TIF yields an estimated OR of approximately 2.12 (0.209*10.157) relative to control, transforming sadness from the most detrimental emotional state (79% reduction in odds without TIF) into a beneficial one (112% increase in odds with TIF) for long-horizon weather predictions. However, the wide confidence interval on the interaction term signals considerable instability in this estimate.

Anger*TIF (OR = 1.590, 95% CI [0.516, 4.901], $p$ = .419), fear*TIF (OR = 1.669, 95% CI [0.526, 5.299], $p$ = .385), happiness*TIF (OR = 1.141, 95% CI [0.346, 3.764], $p$ = .828) and neutral*TIF (OR = 1.136, 95% CI [0.359, 3.597], $p$ = .828) all did not reach or approach significance.

**Box Office Predictions.** Block 2 did not significantly improve model fit ($\Delta\chi^2$ = 7.759, p = .170, $\Delta R^2$ = .018). Two interactions reached or approached significance: Anger*TIF (OR = 0.279, 95% CI [0.077, 1.006], $p$ = .051) and Sadness*TIF (OR = 0.239, 95% CI [0.063, 0.909], $p$ = .036). However, given the overall null effects for movie predictions (no significant emotion main effects, no TIF main effect) and the lack of significant model improvement, these interactions should be interpreted with considerable caution. Type I error (Gelman & Loken, 2014) may be present in the context of multiple comparisons rather than meaningful moderation effects.

**Dow Jones Index Predictions.** Hierarchical linear regression revealed no significant interaction effects for DJI predictions. For 7-day predictions, Block 2 showed no improvement, $\Delta R^2$ = .001, $F(5, 728)$ = 0.195, $p$ = .964, with all interaction terms non-significant (all $p$s > .49). For 14-day predictions, Block 2 similarly showed no improvement, $\Delta R^2$ = .001, $\Delta F(5, 728)$ = 0.160, $p$ = .977, with all interaction terms non-significant (all $p$s > .10). All VIF values were below 3.3, indicating no multicollinearity concerns.

Two-way ANOVA confirmed these patterns, revealing no significant emotion group and TIF interactions for 7-day predictions, $F(5, 728) = 0.195$, $p = .964$, $\eta^2p = .001$, or 14-day predictions, $F(5, 728) = 0.160$, $p = .977$, $\eta^2p = .001$. Refer to Appendix F for Block 1 and Block 2 output for regressions and Two-Way ANOVA.

# Discussion

## Summary of Findings

Our study investigated whether discrete affective states (anger, fear, sadness, happiness, neutral) will influence prediction accuracy and whether trust-in-feelings (TIF) moderates these relationships. Drawing on appraisal-tendency framework (Lerner & Keltner, 2000, 2001) and approach-avoidance model (Carver & Harmon-Jones, 2009), supported by key insights from dual-process frameworks (Epstein, 1994; Kahneman, 2003), we hypothesized that approach-oriented emotions with high certainty appraisals (anger, happiness) would enhance accuracy through heuristic processing, while avoidance-oriented emotions with low certainty appraisals (fear, sadness) would impair accuracy through over-analysis and rumination. We further predicted that TIF would amplify these emotion-specific effects.

Our results diverge from core theoretical predictions in four ways. First, contrary to H2 and H3, fear and sadness significantly *improved* rather than impaired prediction accuracy for weather predictions, with fear showing effects across all four time horizons (ORs ranging from 1.69 to 2.28, all $p$s $< .05$) and sadness improving 4-day (OR = 1.895, $p = .015$) and 7-day predictions (OR = 2.263, $p = .002$). However, these emotion effects were domain-specific, emerging only for weather and largely absent for movie box office and DJI forecasts. Second, the neutral group outperformed the no-induction control for certain weather horizons (4-day: OR = 1.724, $p = .040$; 14-day: OR = 1.947, $p = .024$), suggesting that low-arousal inductions (and/or the act of receiving an induction stimulus) may be psychologically active rather than inert, complicating interpretations based solely on discrete, valenced emotions. Third, TIF showed minimal direct effects on accuracy, contradicting Pham et al.'s (2012) findings, with significant effects emerging only for 14-day weather predictions (OR = 1.566, 95% CI [1.091, 2.250], $p = .015$). Fourth,

emotion*TIF interactions were generally absent, though one isolated sadness*TIF interaction for 14-day weather ($p = .008$) warrants cautious attention and future replication. See Table 5 below for a summary of findings.

**Table 5:** *Summary of Findings*

| H# | Hypothesis | *n* | # of Sig. Results | Conclusion | Evidence |
|---|---|---|---|---|---|
| H1 | Anger increases accuracy | 333 | 3 of 7 | **Partial support** | Weather: 2-day ($p = .001$), 7-day ($p = .007$), 14-day ($p = .010$) |
| H2 | Fear decreases accuracy | 329 | 4 of 7 (**opposite**) | No support | Weather: 2-day ($p = .041$), 4-day ($p = .045$), 7-day ($p = .002$), 14-day ($p = .034$); opposite direction |
| H3 | Sadness decreases accuracy | 325 | 2 of 7 (**opposite**) | No support | Weather: 2-day ($p = .053^{\dagger}$), 4-day ($p = .015$), 7-day ($p = .002$); DJI 7-day ($p = .092$); opposite direction |
| H4 | Happiness increases accuracy | 330 | 3 of 7 | **Partial support** | Weather: 2-day (p = .013), 4-day (p = .039), 7-day ($p = .025$) |
| H5 | Neutral decreases accuracy | 327 | 2 of 7 (**opposite**) | No support | Weather: 4-day ($p = .040$), 14-day ($p = .024$); opposite direction |
| H6 | High TIF increases accuracy | 740 | 1 of 7 | **Partial support** | Weather: 4-day ($p = .086^{\dagger}$), 14-day ($p = .015$) |
| H7 | TIF positively moderates anger effect | 107 | 0 of 7 | No support | None significant |
| H8 | TIF negatively moderates fear effect | 103 | 0 of 7 | No support | None significant |
| H9 | TIF negatively moderates sadness effect | 99 | 1 of 7 (**opposite**) | No support | Weather 14-day ($p = .008$); opposite direction |
| H10 | TIF positively moderates happiness effect | 104 | 0 of 7 | No support | None significant |
| H11 | TIF negatively moderates neutral effect | 101 | 0 of 7 | No support | Weather 7-day ($p = .083^{\dagger}$); opposite direction |

$^{\dagger}p<.10$

## Theoretical Implications

### *Negative Emotions in Prediction*

Our theoretical framework, grounded in the appraisal-tendency framework and approach-avoidance model, generated clear directional predictions: fear and sadness, characterized by low-certainty appraisals and avoidance motivation, should impair prediction accuracy via over-analysis and rumination. Meanwhile, anger and happiness, characterized by high-certainty appraisals and approach motivation, should enhance prediction accuracy via efficient heuristic processing. This empirical pattern partially supports the latter prediction (i.e., anger and happiness) but runs counter to the former – fear and sadness significantly *improved* rather than impaired weather prediction accuracy.

This contradiction has meaningful theoretical implications. If emotions with opposing appraisal profiles (high-certainty anger vs. low-certainty fear; approach-oriented happiness vs. avoidance-oriented sadness) produce similar benefits, then the proposed mechanism, that differential engagement of heuristic versus systematic processing, cannot fully account for the observed effects. To understand what might account for these patterns, we turn to the specific emotion findings.

**Fear and Sadness.** Fear improved accuracy across all four time horizons: 2, 4, 7 and 14-day weather predictions. Sadness significantly improved 4 and 7-day weather predictions, with marginal effects for 2-day weather predictions. These findings directly contradict H2 and H3.

One interpretation consistent with our original theoretical framework is that the systematic processing promoted by fear and sadness may sometimes function adaptively in prediction contexts. While Gigerenzer and Gaissmaier (2011) demonstrated that heuristics outperform complex strategies when cue validity is high and environments are uncertain, they also noted that systematic processing can be beneficial when diagnostic cues are numerous and their relationships complex. Weather prediction, particularly at longer horizons, may represent such a context where careful attention to multiple indicators (recent trends, seasonal patterns, geographic factors) provides predictive value that quicker heuristic judgments might miss.

Alternatively, fear and sadness may support enhanced prediction accuracy through heightened attention and reduced overconfidence rather than through systematic processing per se. Fear's vigilance and sadness's reflective orientation (Bar-Haim et al., 2007; Karnaze & Levine, 2018) may have prompted participants to attend more carefully to available information and avoid premature conclusions on initial impressions. This interpretation aligns with research showing that negative affect reduces reliance on stereotypes and superficial cues (Bodenhausen et al., 1994), potentially enabling more accurate environmental assessment.

We emphasize that these interpretations are post-hoc. Our original predictions, grounded through the rumination and overanalysis lens, anticipated impairment rather than improvement. The reversal suggests either that our theoretical derivation may be somewhat flawed, that prediction tasks differ meaningfully from other judgment contexts where fear and sadness impair performance, or that design features (discussed later) obscure the true relationships.

**Anger and Happiness.** Anger improved accuracy for 2, 7 and 14-day weather predictions, though not for 4-day predictions, movie box office, or DJI forecasts. Happiness significantly improved accuracy for 2, 4, and 7-day weather predictions, but showed no effects for 14-day weather, movie, or DJI predictions.

These findings provide partial support for H1 and H4. However, comparing effect sizes across emotions undermines the theoretical rationale for these predictions. Anger and happiness showed ORs ranging from 1.72 to 2.28 for significant weather predictions; fear and sadness showed ORs ranging from 1.67 to 2.28 for the same outcomes. This similarly does not support the prediction that approach-oriented, high-certainty emotions would be distinctly beneficial compared to avoidance-oriented, low-certainty emotions. Rather, multiple discrete emotions improved weather prediction accuracy relative to no-induction control without clear differentiation by appraisal tendencies or motivational orientation – a pattern more consistent with a general engagement mechanism than with emotion-specific cognitive processing.

### *Neutral Condition*

We hypothesized (H5) that neutral states would impair accuracy due to lack of emotional signals to guide judgment. Surprisingly, the neutral group outperformed the no-induction control group for 4 and 14-day weather predictions.

This finding is theoretically important because it could complicate interpretation of all emotion effects in this study. The neutral film clip (Blue Planet II footage featuring tropical fish and coral) was selected to minimize emotional arousal; however, self-reported emotions revealed that participants experienced elevated calmness ($M$ = 5.56) and moderate happiness ($M$ = 4.33). This suggests that our "neutral" condition induced a low-arousal positive state rather than strict affective neutrality.

Our findings raise a critical interpretive question. If a condition designed to be mostly emotionally inert still improved prediction accuracy relative to the uninduced control, then the benefits observed for discrete emotions may not require the specific appraisal patterns or motivational tendencies that those emotions provide. The neutral condition's performance, driven by mild positive affect from marine life, is difficult to reconcile with frameworks that attribute predictive benefits to specific appraisal dimensions or motivational orientations. Instead, this pattern suggests that even minimal emotional activation may be sufficient to enhance prediction accuracy.

### *Engagement as an Alternative Mechanism*

The pattern of results across emotion conditions points toward a broader interpretation. Our no-induction control group did not watch any film, while all other groups watched film clips before making predictions. This creates an interpretive ambiguity that pervades our findings: differences between any induction condition and control may reflect a) the emotional content of the film, or b) the general attentional engagement (e.g., Hasson et al., 2008) from watching any film, or c) a combination of both. Several observations favour a potential engagement interpretation: all induced conditions improved weather accuracy relative to control despite theoretically opposing profiles, the neutral condition produced comparable effect sizes, and research suggests brief engaging stimuli can focus attention and reduce mind-wandering (Smallwood & Schooler, 2015).

If true, our findings would then relate to Seo and Barrett's (2007) work demonstrating that emotional intensity, rather than specific discrete emotions or valence, predicted superior investment performance. They argued that individuals who experience more intense emotions achieved better simulated returns, suggesting that emotional engagement broadly, not particular emotions, facilitated judgment quality. Although our investment-related predictions did not replicate this pattern, our weather prediction results are consistent with their interpretation, as anger, fear, sadness, and happiness all improved weather prediction accuracy relative to uninduced control, with comparable effect sizes across emotions that markedly differed in valence, appraisal profiles and motivational orientation. The neutral condition's (albeit slightly positively charged) strong performance further supports an engagement account – even low-arousal affective states induced by watching tropical fish outperformed the no-induction baseline for predictions.

This convergence suggests a reframing of how emotions may influence prediction accuracy. Rather than specific appraisal patterns or processing styles driving differential outcomes, what may matter most is whether individuals are emotionally engaged versus operating in a more passive, disengaged state. Emotional states, regardless the type or valence, may then signal task relevance, mobilize cognitive resources, and direct attention toward available information. Silvestrini and Gendolla (2019) suggest that emotional arousal can improve task performance generally by increasing effort and prioritizing cue processing. Moreover, this aligns with evidence that emotional arousal, regardless of valence, enhances cognitive performance by enhancing the availability of task-relevant information (Mather & Nesmith, 2008) and biasing attention toward goal-relevant information (Mather & Sutherland, 2011). In contrast, the absence of emotional engagement (i.e., our no-induction control) may produce more superficial automatic processing, characterized by higher task disengagement and attentional lapses, that misses relevant cues (Smallwood et al., 2004). Such interpretation also more directly aligns with affect-as-information accounts (Schwarz & Clore, 1983) suggesting that feelings serve as informational inputs signalling how to allocate cognitive effort.

However, important differences between our study and Seo and Barrett's (2007) warrant consideration. While both studies examined non-expert participants in simulation

contexts – the author studied investment club members in an internet-based stock trading simulation, whereas we examined Prolific participants making weather, movie and stock index predictions – a critical distinction remains in the origin of the emotional states. Seo and Barrett's (2007) participants' emotions arose organically from engagement with the trading task itself over 20 consecutive business days; these integral emotions developed through repeated interaction with the same domain and may have carried task-relevant diagnostic information accumulated across trading sessions. In contrast, our participants' emotions were incidentally induced by unrelated film clips immediately before making single-point predictions, emotions that, by definition, contained no diagnostic information about weather patterns, movie performance, or stock movements. The fact that we observed robust emotion effects for weather predictions, but null effects for DJI predictions (despite the latter being more comparable to Seo and Barrett's investment context) suggests that either emotional engagement operates differently for single-point versus iterative predictions, or that the benefits of emotional intensity may depend on emotions arising from sustained engagement with the prediction domain rather than from a brief external induction.

We cannot definitively adjudicate between emotion-specific and engagement-based explanations with the present results and design. Future research should include a film-matched control condition (e.g., visually engaging but affectively minimal content such as geometric patterns or household objects) to better isolate emotional content from general film exposure effects. Nonetheless, the convergence of evidence – comparable effect sizes across theoretically opposing emotions, the neutral condition's unexpected performance, and alignment with Seo and Barrett's (2007) intensity findings – makes emotional engagement the most parsimonious account of the present data.

### *Inability to Replicate Trust-in-Feelings Manipulation*

Our findings[17] largely fail to replicate Pham et al.'s (2012) "emotional oracle effect". Across all seven predictions tasks, TIF significantly improved accuracy only for 14-day weather predictions. TIF showed marginally significant effects for 4-day weather predictions. TIF showed no significant effects for 2 or 7-day weather, movie box office, or

[17] The present analysis utilizes the full sample ($N = 740$) to test TIF effects.

7-day DJI predictions. To the best of our knowledge, our study appears to be the first independent replication attempt of Pham et al.'s (2012) TIF manipulation, and the largely null findings raise important questions about the robustness and generalizability of TIF effects. Several explanations warrant consideration.

First, Pham et al. (2012) emphasized that TIF only improves predictions when forecasters possess sufficient domain knowledge, arguing that feelings aggregate implicit knowledge accumulated through experience. Despite pre-screening participants using virtually the same simple techniques as Pham et al. (2012), our participants may have lacked adequate expertise for the more complicated prediction domain (DJI and movies). Using this logic, encouraging reliance on feelings does little when the underlying feelings may not be grounded in sufficient domain knowledge.

Second, Avnet (personal communication, Dec 20, 2022) indicated that TIF effects may persist for six or more predictions. While plausible, no formal studies have examined the temporal dynamics of the TIF manipulation. Our fixed prediction task sequence (weather, movies, DJI) means that later predictions occurred further from the manipulation, potentially explaining why weather forecasts showed the most consistent TIF effects, while later tasks showed null effects.

Third, one might argue that our experimentally induced emotions represent the "wrong" feelings to trust. Pham et al. (2012) did not induce discrete emotions; their participants made predictions in naturally occurring affective states after receiving either a high TIF manipulation or a control task. Our design differed in that most participants ($n$ = 514) received discrete emotion inductions prior to forecasting. One might argue that experimentally induced emotions from unrelated film clips represent the "wrong" feelings to trust (i.e., feelings containing no task-relevant diagnostic information). However, this distinction between "naturally occurring" and "experimentally induced" incidental emotions may be somewhat artificial. Pham et al.'s (2012) participants were not in emotion-free states; they arrived with whatever incidental emotions they had accumulated that day, from traffic, social interactions, or media exposure, none of which are guaranteed to contain specific task-relevant diagnostic information about weather, elections, or stock markets. If incidental emotions interfere with TIF's efficacy, then Pham et al.'s positive findings are

difficult to explain, as their participants' uncontrolled affective states were equally incidental. Conversely, if TIF worked despite their participants' incidental emotions, it remains unclear why our experimentally induced states would be uniquely problematic. In any case, our design may still approximate real-world forecasting conditions, where decision-makers rarely make predictions in affectively neutral contexts.

Regardless of the explanation, the largely null TIF findings across a well-powered study ($N$ = 740) with multiple prediction domains suggest that TIF effects may be more fragile or context-dependent than originally proposed. The limited replication evidence highlights the need for systematic investigation for boundary conditions before TIF-based interventions are utilized.

**Emotion*TIF Interactions.** The only clearly significant interaction emerged for 14-day weather predictions: sadness*TIF interaction. However, this finding contradicts H9 in direction, not merely magnitude. H9 predicted that high TIF would *negatively* moderate sadness' effect – that is, trusting one's feelings while sad would amplify rumination and further impair accuracy. Instead, we found the opposite, where TIF *reversed* sadness' negative effect to positive.

To interpret this interaction, we examine the pattern of simple effects in Block 2. The sadness main effect (OR = 0.209, 95% CI [0.047, 0.938], $p$ = .041) represents sadness' effect in the no TIF condition, indicating that sadness significantly decreased accuracy compared to control when participants did not receive the TIF manipulation. However, the significant positive interaction indicates that TIF completely reversed this pattern. The combined effect of sadness with high TIF suggests that sad participants who trusted their feelings showed improved accuracy relative to control.

This pattern, if reliable, would suggest that the ruminative tendencies associated with sadness become adaptive when channelled through feeling-based judgment, perhaps by converting repetitive analysis into useful signal extraction. However, we urge extreme caution in interpreting this finding. First, the confidence interval is extremely wide (from 1.83 to 56.28), indicating an unstable estimate. Second, this is one significant interaction among 35 tested (7 outcomes * 5 emotions), and a significant sadness*TIF interaction for

movie predictions showed the opposite direction (OR = 0.239, $p$ = .036). The inconsistent pattern across domains suggests these may represent Type I errors (Gelman & Loken, 2014) in an underpowered analysis rather than a reliable phenomenon (Simonsohn, 2015). With 35 tests at $\alpha = .05$, about one to two false positives are expected by chance alone. We report this finding for transparency, but do not draw theoretical conclusions pending replication with larger sample sizes.

***Task-Dependent Results***

The striking domain specificity of our findings, somewhat robust effects for weather forecasting, but largely null results for movie box office and DJI predictions, demands careful interpretation. However, a critical design limitation prevents confident attribution of this pattern to domain characteristics, as our prediction tasks were administered in a fixed sequence (weather, movies, and then DJI) for all participants.

This fixed ordering creates an inextricable confound between domain and temporal proximity to the experimental manipulations. Our strongest effects emerged for weather predictions (administered first), followed by largely null effects for movie predictions and DJI predictions. This pattern is consistent with domain-specific effects but equally consistent with decay of emotion induction and/or TIF manipulation over time.

Research on film-induced emotions suggest that such states can be short-lived, potentially dissipating within several minutes (Verduyn et al., 2015). Given that participants completed filler tasks and multiple prediction sets, induced emotional states may have weakened substantially by the time they reached the DJI predictions. Moreover, no formal studies have examined how long TIF manipulation effects persist. If induced emotions decayed while TIF effects remained active (or vice versa), this temporal mismatch could explain both the inconsistent emotion*TIF interactions and the largely non-existent effects for later prediction tasks.

We cannot distinguish domain-specific effects from temporal proximity effects with the present design. The most parsimonious interpretation is that both factors likely contribute, whereby emotion and TIF effects may be strongest immediately after manipulation (favoring weather predictions) and weather predictions may be more

accesible to feeling-based judgment due to accessible cues and everyday experience. Future research should counterbalance prediction questions across participants to isolate these factors. Until such research is conducted, claims about domain specificity in our findings should be treated as tentative.

***Implications for Appraisal-Tendency and Approach-Avoidance Frameworks***

Our findings challenge direct applications of both appraisal-tendency framework (Lerner & Keltner, 2000) and approach-avoidance models (Carver & Harmon-Jones, 2009) to prediction contexts.

ATF predicts that emotions with different certainty and control appraisals should produce different cognitive processing styles, and consequently, different judgment outcomes. High-certainty emotions like anger should promote confident, heuristic-based judgments, while low-certainty emotions like fear should trigger more deliberative, systematic processing. Our results do not support this differentiation. Emotions with opposing appraisal profiles (e.g., 2-day weather: anger, OR = 2.284; $p = .001$ vs. fear, OR = 1.709, $p = .041$) produced comparable improvements in weather prediction accuracy with overlapping confidence intervals. Rather than the differentiated pattern predicted by ATF, we observed a general benefit of emotional induction relative to no induction.

Similarly, approach-avoidance models predict that approach-oriented emotions (anger, happiness), should facilitate task engagement, while avoidance emotions (fear, sadness) should promote withdrawal or excessive caution. If motivational orientation were the primary mechanism, we would expect systematic differences in prediction accuracy favouring approach emotions. Instead, approach-oriented anger (e.g., 7-day, OR = 2.067) and avoidance-oriented fear (7-day, OR = 2.280) produced effects of similar magnitude, and avoidance-oriented sadness (e.g., 7-day, OR = 2.263) had numerically stronger effects than approach-oriented happiness (7-day, OR = 1.839), with largely overlapping confidence intervals across the board. This pattern suggests that the action tendencies characterizing approach and avoidance motivation may be less relevant for prediction tasks than for the behavioural domains where these frameworks were originally validated.

One possible reconciliation is that approach-avoidance distinctions may manifest differently for cognitive versus behavioural outcomes. The "avoidance" component of fear and sadness may have expressed itself not as task disengagement but as heightened environmental scanning, a form of vigilance that could prove adaptive for weather prediction where subtle cues carry predictive value. However, this interpretation cannot be confirmed with the present data.

More broadly, the convergent null findings for both theoretical frameworks point toward a common conclusion where emotion-specific mechanisms may be less predictive of forecasting outcomes than domain characteristics, task sequence, or general engagement factors. The fact that emotions with opposing appraisal profiles and opposing motivational orientations produced similar benefits suggests that what distinguishes effective from ineffective prediction may operate at a different level of analysis than discrete emotional states. These findings do not invalidate either framework; rather, they suggest that prediction tasks may not engage the appraisal pathways or motivational systems in the manner originally theorized, or that our design features, particularly the film induction confound discussed above, obscured emotion-specific effects.

**Practical Implications**

Our findings, while requiring replication, suggest several tentative implications for forecasting practice. First, emotional states need not be suppressed or eliminated during prediction tasks. Traditional decision-making advice often emphasizes “cooling off” (e.g., Oechssler et al., 2015) or achieving emotional disengagement before important decisions. Our results suggest this recommendation may be misguided, at least for certain forecasting contexts. Participants experiencing discrete emotions, including supposedly negative emotions like fear and sadness, outperformed those in emotional uninduced states for weather predictions. If replicated and across more domains, these finding challenges the assumption that forecasters should strive for emotional detachment.

Second, our findings offer no support for encouraging people to "trust their feelings" as a general forecasting strategy. The TIF manipulation produced significant effects in only one of seven prediction tasks. Until boundary conditions are better understood, particularly regarding necessary domain expertise and temporal dynamics of the manipulation,

practitioners should not adopt such approaches based on Pham et al.'s (2012) findings alone. Notably, our largely null findings occurred under conditions that closely approximate everyday forecasting contexts, where incidental emotions are ubiquitous.

In general, organizations should consider whether their work environments or task designs inadvertently suppress potentially beneficial emotional engagement. Sterile presentation formats, overly routinized procedures, or disengaging contexts may induce the kind of passive processing our control condition exhibited. To the extent that emotional engagement promotes attention and careful consideration, workplaces that feel psychologically "flat" or unstimulating may inadvertently impair judgment quality (e.g., Rich et al., 2010). This does not mean organizations should induce fear or sadness, but it suggests that some degree of encouraged emotional investment in tasks, rather than emotionally muted execution, may be beneficial.

Nevertheless, we emphasize that these broader implications should be applied cautiously. The null results for movie box office and stock market predictions – whether due to task order, domain characteristics, or other factors – disqualify confident extension to these (and similar) domains. Moreover, our participants made low-stakes predictions with no consequences for accuracy. Whether emotion effects would persist, strengthen, or reverse under conditions of consequential decision-making remains an open question.

## Limitations and Future Research Direction

Several limitations constrain interpretations of our findings. We have discussed several in detail above; here, we consolidate them and identify additional methodological considerations.

### *Design Limitations*

The two design features that create interpretive ambiguity have been addressed in preceding discussion, namely, that our no-induction control group did not watch any film, confounding emotional content with general film exposure effects (see “Engagement as an Alternative Mechanism” section) and our fixed task sequence (weather, then movies, then DJI) prevents distinguishing domain-specific effects from temporal decay of manipulations (see “Task-dependent results” section).

**Emotion induction.** While our film-based induction procedure successfully elevated target emotions relative to baseline, the resulting emotional states were not discretely pure. Manipulation checks revealed meaningful co-activation of related emotions (e.g., anger co-occurred with elevated disgust, fear co-occurred with anxiety, and neutral with mild happiness). This pattern of emotional blends is consistent with prior research on film-based inductions (Lench et al., 2011; Schaefer et al., 2010) and reflects a fundamental challenge in emotion research: discrete emotions rarely occur in isolation. Consequently, we cannot confidently attribute prediction effects to single target emotions. As such, an anger effect seen may partly reflect the effects of co-activated disgust. Furthermore, our neutral condition was not affectively inert as participants reported elevated calmness ($M = 5.56$) and happiness ($M = 4.33$), indicating a low-arousal positive state rather than strict emotional neutrality. This complicates interpretation of the neutral condition as a true emotional baseline and may explain its unexpected outperformance of the no-induction control. Future studies should consider employing more targeted induction procedures (e.g., autobiographical recall with careful screening, standardized imagery scripts) and selecting specific validated neutral stimuli to produce minimal affective response in participants (e.g., Gasper, 2018; Siedlecka & Denson, 2019)

**Processing mechanisms.** Our theoretical framework centered on heuristic versus systematic processing as the mechanism linking emotions to prediction accuracy, yet we did not measure processing style directly. Without process measures, we cannot determine whether emotions shifted processing as theorized or whether the observed effects operated through different pathways entirely (e.g., attention, motivation, confidence). Future research employing process-tracing methods (e.g., think-aloud protocols; Someren et al., 1994) could more directly test whether emotion-induced differences in processing style mediate prediction outcomes.

### *Sample Size and Generalizability*

Our sample of 740 participants was adequately powered for main effects, but likely underpowered for interaction effects, with sub-sample sizes of 45-62 for interaction groups, which is substantially below the two to three times larger samples (compared to main

effects) typically required for detecting interactions (Simonsohn, 2015). The largely null interaction findings should therefore be interpreted cautiously.

Beyond sample size, our participants were recruited through Prolific, a platform where participants are often highly motivated to provide quality responses due to rejection policies (Peer et al., 2022) and income dependence. This motivation may have increased systematic processing and attenuated both emotion induction effects (through emotion regulation) and TIF effects (through override of feeling-based judgment). Additionally, our low-stakes prediction tasks differ substantially from consequential real-world forecasting. Whether emotion effects would persist, strengthen, or reverse under conditions of meaningful consequences remains an open question.

Moreover, while we adopted Pham et al.'s (2012) screening procedure for DJI predictions to ensure methodological consistency, simple questions about GDP and P/E ratios may not capture the deep, accumulated knowledge that makes feelings genuinely diagnostic. Knowing textbook definitions differs from having intuitions developed through years of market experience. This may explain why TIF effects emerged for weather, where participants likely possessed genuine everyday experience, but not more nuanced topics.

## Conclusion

Our study provides the first systematic examination of how discrete emotions – anger, fear, sadness, happiness, and neutral states – influence prediction accuracy and offers new evidence on the generalizability of Pham et al.'s (2012) trust-in-feelings (TIF) manipulation.

Two key findings emerged. First, contrary to theoretical predictions derived from appraisal-tendency and approach-avoidance frameworks, "negative" emotions did not impair forecasting. Fear and sadness improved weather prediction accuracy, with effect sizes comparable to anger and happiness. The similar benefits observed across emotions with opposing appraisal profiles and motivational orientations suggest that emotional engagement, rather than emotion-specific processing styles, may be what matters most for prediction accuracy.

Second, TIF effects were notably absent, emerging significantly in only one of seven prediction tasks. This limited replication raises important questions about the robustness of

the "emotional oracle effect." Critically, Pham et al.'s (2012) participants were not in emotion-free states either; they arrived with incidental emotions from daily life that were equally task-irrelevant. If TIF worked for them despite such emotions, it remains unclear why our findings diverged, and if TIF is fragile to incidental emotions, its practical utility for real-world forecasting, where such emotions are ubiquitous, becomes questionable.

Overall, our findings suggest that the conventional wisdom to “set emotions aside” when making decisions – forecasting included – may be misguided. But so too may be the advice to “trust your feelings”. Emotions appear to matter for prediction accuracy; however, how best to leverage them remains an open question.

# References


Abbey, J. D., & Meloy, M. G. (2017). Attention by design: Using attention checks to detect inattentive respondents and improve data quality. *Journal of Operations Management, 53-56*, 63-70.

Aboody, D., Barth, M. E., & Kasznik, R. (1999). Revaluations of fixed assets and future firm performance: Evidence from the UK. *Journal of Accounting and Economics, 26,* 149-178.

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.

Albert, D. A., & Smilek, D. (2023). Comparing attentional disengagement between Prolific and MTurk samples. *Scientific Reports, 13*(1), 20574.

Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences, 103*(24), 9369-9372.

Ambady, N., & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology, 83*(4), 947-961.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110*(3), 486-498.

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., Neumann, G. R., Ottaviani, M., Schelling, T. C., Shiller, R. J., Smith, V. L., Snowberg, E., Sunstein, C. R., Tetlock, P. C., Tetlock, P. E., ... Zitzewitz, E. (2008). The promise of prediction markets. *Science, 320*(5878), 877-878.

Avnet, T., & Pham, M. T. (2005). Should I trust my feelings or not? The metacognition of affect as information in judgment. *Advances in Consumer Research, 32*, 38-39.

Avnet, T., Pham, M. T., & Stephen, A. T. (2012). Consumers' trust in feelings as information. *Journal of Consumer Research, 39*(4), 720-735.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1-24.

Batra, R., & Stayman, D. M. (1990). The role of mood in advertising effectiveness. *Journal of Consumer Research, 17*(2), 203-214.

Bayazit, M., & Mannix, E. A. (2003). Should I stay or should I go? Predicting team members' intent to remain in the team. *Small Group Research, 34*(3), 290-321.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293-1295.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience, 19*(13), 5473-5481.

Benson, P. G., & Önkal, D. (1992). The effects of feedback and training on the performance of probability forecasters. *International Journal of Forecasting, 8*(4), 559-573.

Bless, H., Clore, G. L., Schwarz, N., Golisano, V., Rabe, C., & Wölk, M. (1996). Mood and the use of scripts: Does a happy mood really lead to mindlessness? *Journal of Personality and Social Psychology, 71*(4), 665-679.

Bodenhausen, G. V. (1993). Emotions, arousal, and stereotypic judgments: A heuristic model of affect and stereotyping. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition and stereotyping* (pp. 13-37). Elsevier.

Bodenhausen, G. V., Gabriel, S., & Lineberger, M. (2000). Sadness and susceptibility to judgmental bias: The case of anchoring. *Psychological Science, 11*(4), 320-323.

Bodenhausen, G. V., Sheppard, L. A., & 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.

Brust-Renck, P. G., Weldon, R. B., & Reyna, V. F. (2021). *Judgment and decision making*. In *Oxford Research Encyclopedia of Psychology*. Oxford University Press.

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology, 75*(1), 219-229.

Cabin, R. J., & Mitchell, R. J. (2000). To Bonferroni or not to Bonferroni: When and how are the questions. *Bulletin of the Ecological Society of America, 81*(3), 246-248.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115*(3), 401-423.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin, 135*(2), 183-204.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*(5), 752-766.

Chaiken, S., & Ledgerwood, A. (2012). A theory of heuristic and systematic information processing. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 1, pp. 246-266). SAGE Publications.

Clore, G. L., & Parrott, W. G. (1994). Cognitive feelings and metacognitive judgments. *European Journal of Social Psychology, 24*(1), 101–115.

Clore, G. L., & Tamir, M. (2002). Affect as embodied information. *Psychological Inquiry, 13*(1), 37-45.

Clore, G. L., Schwarz, N., & Conway, M. (1994). Affective causes and consequences of social information processing. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes; Applications* (2nd ed., pp. 323-417). Lawrence Erlbaum Associates.

Collins, G. S., & Moons, K. G. M. (2019). Reporting of artificial intelligence prediction models. *The Lancet, 393*(10181), 1577-1579.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. Putnam.

Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE, 18*(3), e0279720.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3-4), 169-200.

Ekman, P. (2007). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life* (2nd ed.). Holt.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*(8), 709-724.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology, 71*(2), 390-405.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255-278.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Finucane, M. L., Peters, E., & Slovic, P. (2003). Judgment and decision making: The dance of affect and reason. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 327-364). Cambridge University Press.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117*(1), 39-66.

Frankish, K. (2010). Dual-process and dual-system theories of reasoning. *Philosophy Compass, 5*(10), 914-926.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*(3), 218-226.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology, 65*(1), 45-55.

Gabert-Quillen, C. A., Bartolini, E. E., Abravanel, B. T., & Sanislow, C. A. (2015). Ratings for emotion film clips. *Behavior Research Methods, 47*(3), 773-787.

Gaissmaier, W., & Marewski, J. N. (2011). Forecasting elections with mere recognition from small, lousy samples: A comparison of collective recognition, wisdom of crowds, and representative polls. *Judgment and Decision Making, 6*(1), 73-88.

Gasper, K. (2018). Utilizing neutral affective states in research: Theory, assessment, and recommendations. *Emotion Review, 10*(3), 255-266.

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist, 102*(6), 460-465.

George, J. M., & Dane, E. (2016). Affect, emotion, and decision making. *Organizational Behavior and Human Decision Processes, 136*, 47-55.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*(1), 451-482.

Gilbert, D. T. (2006). *Stumbling on happiness*. Knopf.

Gilman, T. L., Shaheen, R., Nylocks, K. M., Halachoff, D., Chapman, J., Flynn, J. J., Matt, L. M., & Coifman, K. G. (2017). A film set for the elicitation of emotion in research: A comprehensive catalog derived from four decades of investigation. *Behavior Research Methods, 49*(6), 2061-2082.

Gino, F., & Schweitzer, M. E. (2008). Blinded by anger or feeling the love: How emotions influence advice taking. *Journal of Applied Psychology, 93*(5), 1165-1173.

Graver, M. R. (2007). *Stoicism and emotion*. University of Chicago Press.

Greifeneder, R., Bless, H., & Pham, M. T. (2011). When do people rely on affective and cognitive feelings in judgment? A review. *Personality and Social Psychology Review, 15*(2), 107-141.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition and Emotion, 9*(1), 87-108.

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). *Neurocinematics: The neuroscience of film*. Projections, 2(1), 1–26.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Hengstler, M., Holland, R. W., van Steenbergen, H., & van Knippenberg, A. (2014). The influence of approach–avoidance motivational orientation on conflict adaptation. *Cognitive, Affective, & Behavioral Neuroscience, 14*(2), 548-560.

Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., & Bartussek, D. (2005). A revised film set for the induction of basic emotions. *Cognition and Emotion, 19*(7), 1095-1109.

Hogarth, R. M. (2005). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. In T. Betsch & S. Haberstroh (Eds.), *The routines of decision making* (pp. 67-82). Lawrence Erlbaum Associates.

Huebner, B., Dwyer, S., & Hauser, M. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences, 13*(1), 1-6.

International Monetary Fund. (2023). *World economic outlook: A rocky recovery* (April 2023). https://www.imf.org/en/publications/weo/issues/2023/04/11/world-economic-outlook-april-2023

Isen, A. M. (2001). An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. *Journal of Consumer Psychology, 11*(2), 75-85.

Izard, C. E., Dougherty, F. E., Bloxom, B. M., & Kotsch, N. E. (1974). *The Differential Emotions Scale: A method of measuring the meaning of subjective experience of discrete emotions*. Vanderbilt University Press.

Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology, 24*(2), 187-192.

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology, 45*(1), 20-31.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697-720.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science, 39*(1), 17–31.

Karnaze, M. M., & Levine, L. J. (2018). Sadness, the architect of cognitive change. In H. C. Lench (Ed.), *The function of emotions: When and why emotions help us* (pp. 45-58). Springer.

Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science, 11*(2), 45-50.

Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316-337). Blackwell.

Larsen, J. T., McGraw, A. P., Mellers, B. A., & Cacioppo, J. T. (2004). The agony of victory and thrill of defeat: Mixed emotional reactions to disappointing wins and relieving losses. *Psychological Science, 15*(5), 325-330.

Lazarus, R. S. (1999). *Stress and emotion: A new synthesis*. Springer.

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin, 137*(5), 834-855.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion, 14*(4), 473-493.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology, 81*(1), 146-159.

Lerner, J. S., & Tiedens, L. Z. (2006). Portrait of the angry decision maker: How appraisal tendencies shape anger's influence on cognition. *Journal of Behavioral Decision Making, 19*(2), 115-137.

Lerner, J. S., Goldberg, J. H., & Tetlock, P. E. (1998). Sober second thought: The effects of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin, 24*(6), 563-574.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66*, 799-823.

Lieberman, M. D. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. R. Williams, & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes* (pp. 44-67). Cambridge University Press.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65*(3), 272-292.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267-286.

Lyubomirsky, S., & Nolen-Hoeksema, S. (1995). Effects of self-focused rumination on negative thinking and interpersonal problem solving. *Journal of Personality and Social Psychology, 69*(1), 176-190.

Mackie, D. M., & Worth, L. T. (1989). Processing deficits and the mediation of positive affect in persuasion. *Journal of Personality and Social Psychology, 57*(1), 27-40.

Mather, M., & Nesmith, K. (2008). Arousal-enhanced location memory for pictures. *Journal of Memory and Language, 58*(2), 449–464.

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on psychological science*, *6*(2), 114-133.

Mayer, J. D., Allen, J. P., & Beauregard, K. (1995). Mood inductions for four specific moods: A procedure employing guided imagery vignettes with music. *Journal of Mental Imagery, 19*(1-2), 151-159.

McHugo, G. J., Smith, C. A., & Lanzetta, J. T. (1982). The structure of self-reports of emotional responses to film segments. *Motivation and Emotion, 6*(4), 365-385.

Mellers, B., Ungar, L., Baron, J., Ramos, J., Gurcay, B., Fincher, K., Scott, S. E., Moore, D., Atanasov, P., & Swift, S. A. (2014). Psychological strategies for winning a geopolitical forecasting tournament. *Psychological Science, 25*(5), 1106–1115.

Mestyán, M., Yasseri, T., & Kertész, J. (2013). Early prediction of movie box office success based on Wikipedia activity big data. *PLoS ONE, 8*(8), e71226.

Mikels, J. A., Löckenhoff, C. E., Maglio, S. J., Goldstein, M. K., Garber, A., & Carstensen, L. L. (2010). Following your heart or your head: Focusing on emotions versus information differentially influences the decisions of younger and older adults. *Journal of Experimental Psychology: Applied, 16*(1), 87-95.

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology, 15*(6), 1044-1045.

Oechssler, J., Roider, A., & Schmitz, P. W. (2015). Cooling off in negotiations: Does it work? *Journal of Institutional and Theoretical Economics, 171*(4), 565-588.

Ortmann, A., Gigerenzer, G., Borges, B., & Goldstein, D. G. (2008). The recognition heuristic: A fast and frugal way to investment choice? *Handbook of Experimental Economics Results, 1*, 993-1003.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*(4), 1643-1662.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ, 316*(7139), 1236-1238.

Pham, M. T., Lee, L., & Stephen, A. T. (2012). Feeling the future: The emotional oracle effect. *Journal of Consumer Research, 39*(3), 461-477.

Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition and Emotion, 7*(2), 171-193.

Prkachin, K. M., Williams-Avery, R. M., Zwaal, C., & Mills, D. E. (1999). Cardiovascular changes during induced emotion: An application of Lang's theory of emotional imagery. *Journal of Psychosomatic Research, 47*(3), 255-267.

Reyna, V. F. (2004). How people make decisions that involve risk: A dual-processes approach. *Current Directions in Psychological Science, 13*(2), 60-66.

Rich, B. L., LePine, J. A., & Crawford, E. R. (2010). Job engagement: Antecedents and effects on job performance. *Academy of Management Journal, 53*(3), 617–635.

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology, 1*(1), 43-46.

Rottenberg, J., Ray, R. D., & Gross, J. J. (2007). Emotion elicitation using films. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 9-28). Oxford University Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion, 24*(7), 1153-1172.

Scherer, K. R., Wranik, T., Sangsue, J., Tran, V., & Scherer, U. (2004). Emotions in everyday life: Probability of occurrence, risk factors, appraisal and reaction patterns. *Social Science Information, 43*(4), 499-570.

Schimmack, U., Oishi, S., & Diener, E. (2002). Cultural influences on the relation between pleasant emotions and unpleasant emotions: Asian dialectic philosophies or individualism-collectivism? *Cognition and Emotion, 16*(6), 705-719.

Schooler, L. J., & Hertwig, R. (2005). *How forgetting aids heuristic inference*. *Psychological Review, 112*(3), 610–628.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 527-561). Guilford Press.

Schwarz, N., & Clore, G. L. (1988). How do I feel about it? The informative function of affective states. In K. Fiedler & J. P. Forgas (Eds.), *Affect, cognition, and social behavior* (pp. 44-62). Hogrefe.

Schwarz, N., & Clore, G. L. (2003). Mood as information: 20 years later. *Psychological Inquiry, 14*(3-4), 296-303.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*(2), 195-202.

Seneca, L. A. (1928). *Moral essays* (Vol. 1, J. W. Basore, Trans.). Harvard University Press. (Original work published ca. 41-65 CE)

Seo, M.-G., & Barrett, L. F. (2007). Being emotional during decision making—Good or bad? An empirical investigation. *Academy of Management Journal, 50*(4), 923-940.

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134*(2), 207-222.

Siedlecka, E., & Denson, T. F. (2019). Experimental methods for inducing basic emotions: A qualitative review. *Emotion Review, 11*(1), 87-97.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail—but some don't*. Penguin Press.

Silvestrini, N., & Gendolla, G. H. E. (2019). Affect and cognitive control: Insights from research on effort mobilization. *International Journal of Psychophysiology, 143*, 116-125.

Simonsohn, U. (2015). No-way interactions. *The Winnower*. https://thewinnower.com/papers/no-way-interactions

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2002). Rational actors or rational fools: Implications of the affect heuristic for behavioral economics. *The Journal of Socio-Economics, 31*(4), 329-342.

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology, 66*, 487-518.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and cognition*, *13*(4), 657-690.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48*(4), 813-838.

Snook, B., Taylor, P. J., & Bennell, C. (2004). Geographic profiling: The fast, frugal, and accurate way. *Applied Cognitive Psychology, 18*(1), 105-121.

Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical approach to modelling cognitive processes*. Academic Press.

Stephen, A. T., & Pham, M. T. (2008). On feelings as a heuristic for making offers in ultimatum negotiations. *Psychological Science, 19*(10), 1051-1058.

Tetlock, P. E., & Gardner, D. (2015). *Superforecasting: The art and science of prediction*. Crown Publishers/Random House.

Tetlock, P. E., Mellers, B. A., Rohrbaugh, N., & Chen, E. (2014). Forecasting tournaments: Tools for increasing transparency and improving the quality of debate. *Current Directions in Psychological Science, 23*(4), 290-295.

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology, 81*(6), 973-988.

Trampe, D., Quoidbach, J., & Taquet, M. (2015). Emotions in everyday life. *PLoS ONE, 10*(12), e0145450.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207-232.

Van Smeden, M., de Groot, J. A. H., Moons, K. G. M., Riley, R. D., & Reitsma, J. B. (2016). No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Medical Research Methodology, 16*, 163.

Verduyn, P., Delvaux, E., Van Coillie, H., Tuerlinckx, F., & Van Mechelen, I. (2009). Predicting the duration of emotional experience: Two experience sampling studies. *Emotion, 9*(1), 83-91.

Wang, Y., Li, Z., Liu, Y., Shi, D., & Xu, Y. (2022). A systematic review of emotion induction methods. *Frontiers in Psychology, 13*, 942271.

Wegener, D. T., & Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 141-208). Academic Press.

Whitmer, A. J., & Gotlib, I. H. (2013). An attentional scope model of rumination. *Psychological Bulletin, 139*(5), 1036-1238.

Wänke, M. (2013). Almost everything you always wanted to know about ease-of-retrieval effects. In C. Unkelbach & R. Greifeneder (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behaviour* (pp. 151-169). Psychology Press.

Young, S. G., Strom, P. A., & Green, J. D. (2003). Film-induced affect and impression formation: Direct and indirect effects. *Cognition and Emotion, 17*(4), 601-614.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151-175.

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist, 102*(6), 460-465.

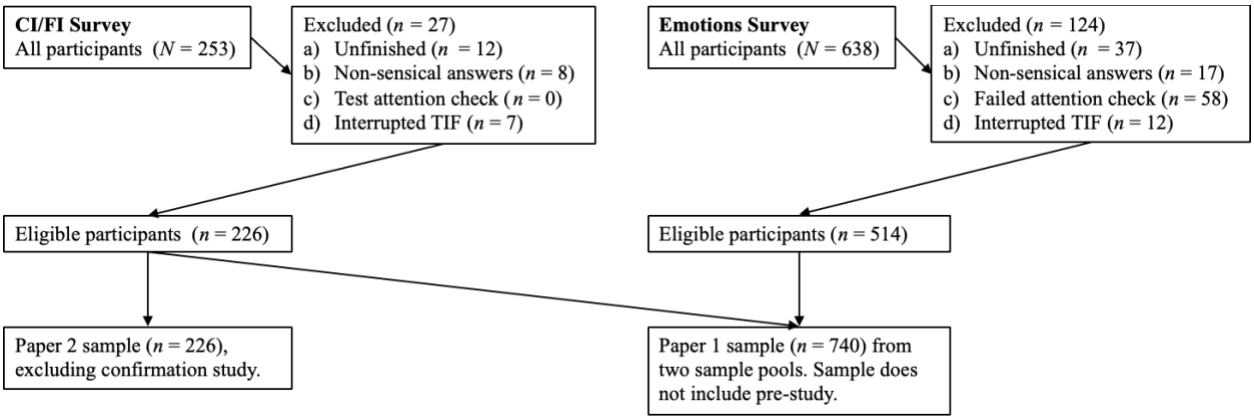
George, J. M., & Dane, E. (2016). Affect, emotion, and decision making. *Organizational Behavior and Human Decision Processes, 136*, 47-55.

# Appendix

## Appendix A

### Sample Composition

(*Paper 1 is current paper*)



# Appendix B

## Screenshots of Administered Participant Surveys

*Note.* The section numbers may not be sequential as they are part of a larger survey.

### Section 1

Thank you for participating in our research!

Please **DO NOT use a mobile phone** to participate in this study. You **CANNOT** use any external sources (e.g., Google) during the study. Please make sure you are answering this survey in a quiet environment. You must have access to sound, either through headphones or loud speakers.

**General Information**
The study is conducted only for scientific purposes, and takes approximately 25 minutes to complete. To receive your compensation, you must answer all questions provided. Kindly note that there are no right or wrong answers - all that we ask is for you to give the study your full attention. During the study, you may view films that may be distressing for you.

All collected data will be confidential. Participation in this study is anonymous. Results are published only in an aggregated format, and will not contain identifiable information to any specific participant. Please note that participation in this study is entirely voluntary, and you may discontinue at any time.

**Study Information**
The goal of the study is test your mental flexibility and memory. You will be asked to complete a range of different tasks. Should you have any questions, please do not hesitate to contact me through Prolific.

Do you agree to proceed with the study?

I agree

I do not agree

## Section 2

What is your Prolific ID?

*Please note that this response should auto-fill with the correct ID*

## Section 3

We would first like to have an understanding of how you **feel at this moment** (i.e., at this point in time and not the general mood of your day). There are no right or wrong answers, as we are only interested in your understanding of your feelings.

On the following page, rank the intensity in which you are currently experiencing for each of the emotions.

## Section 4

**At this moment**, I feel...

| | Not at all | | | Moderately | | | Very Intensely |
|---|---|---|---|---|---|---|---|
| Interested | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Sad | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Excited | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Angry | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Disgusted | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Happy | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Surprised | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Fearful | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Anxious | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Calm | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

## Section 5

For the first task, you will be asked to watch two separate videos. You will be asked about them later in the study.

## Section 6

Please continue to the first video.

## Section 7



## Section 8

Please continue to the second video.

## Section 9



## Section 10

For the second task, please read the instructions below carefully.

There are two basic ways of making decisions. Sometimes people make decisions based on facts and logical reasons. Sometimes, they make decisions based on their feelings, that is, “does it feel good or bad?”, “does it feel right or wrong?”.

For this task, we are interested in how much confidence individuals have in their feelings when making decisions.

**Instructions:**

We would like you to briefly describe different situations in which you trusted your feelings to make a judgment or a decision and it was the right thing to do.

In other words, describe specific situations that really happened to you in which you felt confident in making a judgment or a decision based on how you felt, and later found out that you were right in using your feelings for this judgment or decision. The situations you list could involve any type of decision, provided that they were real situations in which you felt pretty certain that you should go by what your feelings were telling you to do, and it turned out that your feelings were right. For example, one participant in a previous session of this study mentioned the following:

*“I had a very good feeling toward my new roommate from the moment I met him. Although objectively we were quite different, my feelings told me that we would be good friends. My feelings were correct! We became best friends after only a week. Trusting how I felt about my new roommate was definitely the right thing to do.*"

In the next **seven minutes**, please list **TWO (2)** similar situations that happened to you in which you felt quite confident in your feelings, and you were right in trusting them to make your decision (three or four sentences per situation will be enough).

**Please list below TWO (2) situations in which you were right in trusting your feelings...**

## Section 11

You will be automatically advanced to the next page after seven (7) minutes.

**1st situation in which I was right in trusting my feelings:**

**2nd situation in which I was right in trusting my feelings:**

## Section 12

We would like to assess people's ***perceptions*** of various countries that are not particularly well-known, but have an interest in expanding their tourism sectors. In this questionnaire, we will ask you about your general perceptions (mental images) about three relatively less-known countries. There are no right or wrong answers. We are only interested in your own perceptions of these countries. This information will be used to assess how these countries could better market themselves. Please answer on the scales provided below.

**Country 1: Mauritania**

What is your perception of the level of education in this country?

| Far below average | Below average | Average | Above average | Far above average |
|---|---|---|---|---|

What is your perception of the population density in this country?

| Very dense | Somewhat dense | Neither dense nor sparse | Somewhat sparse | Very sparse |
|---|---|---|---|---|

What is your perception of the cost of living in this country?

| Very low | Somewhat low | Neither high nor low | Somewhat high | Very high |
|---|---|---|---|---|

What is your perception of what the terrain is like in this country?

| Mostly mountainous | Somewhat mountainous | Neither mountainous nor flat | Somewhat flat | Mostly flat |
|---|---|---|---|---|

## Section 13

**Country 2: Romania**

What is your perception of the level of education in this country?

| | | | | |
|---|---|---|---|---|
| Far below average | Below average | Average | Above average | Far above average |

What is your perception of the population density in this country?

| | | | | |
|---|---|---|---|---|
| Very dense | Somewhat dense | Neither dense nor sparse | Somewhat sparse | Very sparse |

What is your perception of the cost of living in this country?

| | | | | |
|---|---|---|---|---|
| Very low | Somewhat low | Neither high nor low | Somewhat high | Very high |

What is your perception of what the terrain is like in this country?

| | | | | |
|---|---|---|---|---|
| Mostly mountainous | Somewhat mountainous | Neither mountainous nor flat | Somewhat flat | Mostly flat |

## Section 14

**Country 3: New Zealand**

What is your perception of the level of education in this country?

| | | | | |
|---|---|---|---|---|
| Far below average | Below average | Average | Above average | Far above average |

What is your perception of the population density in this country?

| | | | | |
|---|---|---|---|---|
| Very dense | Somewhat dense | Neither dense nor sparse | Somewhat sparse | Very sparse |

What is your perception of the cost of living in this country?

| | | | | |
|---|---|---|---|---|
| Very low | Somewhat low | Neither high nor low | Somewhat high | Very high |

What is your perception of what the terrain is like in this country?

| | | | | |
|---|---|---|---|---|
| Mostly mountainous | Somewhat mountainous | Neither mountainous nor flat | Somewhat flat | Mostly flat |

## Section 15

For the fourth task, we would like to understand your perspectives about potential events in the future.

## Section 16

What zip code are you currently residing in? (e.g., 10010)

What do you expect the **predominant** weather to be from **12pm to 6pm** at your zip code in **2 days** (Sunday, April 16)?

| Sunny/fine/clear | Partially sunny with some clouds | Cloudy/overcast | Haze/fog | Thunderstorm | Rain | Snow |
|---|---|---|---|---|---|---|

## Section 17

What do you expect the **predominant** weather to be from **12pm to 6pm** at your zip code in **4 days** (Tuesday, April 18)?

| Sunny/fine/clear | Partially sunny with some clouds | Cloudy/overcast | Haze/fog | Thunderstorm | Rain | Snow |
|---|---|---|---|---|---|---|

## Section 18

What do you expect the **predominant** weather to be from **12pm to 6pm** at your zip code in **7 days** (Friday, April 21)?

| Sunny/fine/clear | Partially sunny with some clouds | Cloudy/overcast | Haze/fog | Thunderstorm | Rain | Snow |
|---|---|---|---|---|---|---|

## Section 19

What do you expect the **predominant** weather to be from **12pm to 6pm** at your zip code in **14 days** (Friday, April 28)?

| Sunny/fine/clear | Partially sunny with some clouds | Cloudy/overcast | Haze/fog | Thunderstorm | Rain | Snow |
|---|---|---|---|---|---|---|

## Section 20

In the movie industry, the "opening weekend box office sales" is the amount of money generated during the first weekend (Friday to Sunday) when a movie is released in theatres.

On the next page, you will be provided basic information about three upcoming movies this weekend. Based on the information, please rank, in order of **highest to lowest**, what you expect the ranking of the weekend box office sales to be domestically (i.e., only within United States).

## Section 21

**Name:** Mafia Mamma
**MPAA:** R
**Main Cast:** Toni Collette, Monica Bellucci, Sophia Nomvete, Alessandro Bressanello
**Genre:** Action Comedy Crime
**Brief Synopsis:** An American mom inherits her grandfather's mafia empire in Italy. Guided by the firm's consigliere, she hilariously defies everyone's expectations as the new head of the family business.




**Name:** Nefarius
**MPAA:** R
**Main Cast:** Sean Patrick Flanery, Jordan Belfi, James Healy Jr., Eric Hanson
**Genre:** Horror Thriller
**Brief Synopsis:** On the day of his scheduled execution, a convicted serial killer gets a psychiatric evaluation during which he claims he is a demon, and further claims that before their time is over, the psychiatrist will commit three murders of his own.




**Name:** Sweetwater
**MPAA:** PG-13
**Main Cast:** Cary Elwes, Eric Roberts, Richard Dreyfuss, Jeremy Piven
**Genre:** Biography Drama Sport
**Brief Synopsis:** Hall of Famer Nat "Sweetwater" Clifton makes history as the first African American to sign an NBA contract, forever changing how the game of basketball is played.



## Section 22

**Name:** Nefarius
**MPAA:** R
**Main Cast:** Sean Patrick Flanery, Jordan Belfi, James Healy Jr., Eric Hanson
**Genre:** Horror Thriller
**Brief Synopsis:** On the day of his scheduled execution, a convicted serial killer gets a psychiatric evaluation during which he claims he is a demon, and further claims that before their time is over, the psychiatrist will commit three murders of his own.




**Name:** Mafia Mamma
**MPAA:** R
**Main Cast:** Toni Collette, Monica Bellucci, Sophia Nomvete, Alessandro Bressanello
**Genre:** Action Comedy Crime
**Brief Synopsis:** An American mom inherits her grandfather's mafia empire in Italy. Guided by the firm's consigliere, she hilariously defies everyone's expectations as the new head of the family business.




**Name:** Sweetwater
**MPAA:** PG-13
**Main Cast:** Cary Elwes, Eric Roberts, Richard Dreyfuss, Jeremy Piven
**Genre:** Biography Drama Sport
**Brief Synopsis:** Hall of Famer Nat "Sweetwater" Clifton makes history as the first African American to sign an NBA contract, forever changing how the game of basketball is played.




For the three movies, what do you expect the ranking of the opening weekend box office sales to be domestically?

To rank the movies, please drag the name of the movie to the corresponding box on the right.
(**Rank #1** = **highest** sales out of the three, **Rank #2** = **second highest** sales out of the three, and **Rank #3** = **lowest** sales out of the three).

| Items |
| --- |
| Nefarius |
| Sweetwater |
| Mafia Mamma |

| Rank #1 |
| --- |
| |

| Rank #2 |
| --- |
| |

| Rank #3 |
| --- |
| |

# Section 23

The “Dow Jones” Index (officially known as the Dow Jones Industrial Average Index), is one of the most famous and most widely followed stock market indices in the world. This index is a weighted-average of the stock prices of thirty of the largest publicly traded companies in the United States, such as Boeing, Disney, Microsoft and Goldman Sachs. The stock prices of these companies reflect the value that investors collectively perceive in these companies: their profitability, growth potential, etc. The level of the Dow Jones is thus widely regarded as a summary indicator of the performance of the stock market, investors' expectations of future performance of the stock market, and the strength and growth potential of the US economy.

Here are some indicative values of the Dow Jones Index (DJI) over the past years:

- On February 27, 1991, the day the first Gulf War ended and the economy was in a recession, the DJI closed at 2,889
- On January 14, 2000, at the height of the “Dot-Com Bubble”, the DJI closed at 11,723
- On March 13, 2000, just after the “Dot-Com Bubble” burst, the DJI closed at 9,947
- On May 6, 2010, after a computer-driven trading glitch, the DJI closed at 10,520
- On Jan 15, 2020, after a multi-year bull run and lessening trade disputes between China and US, the DJI closed at 29,030
- On March 12, 2020, after the US President declared a public health emergency (COVID), the DJI closed at 21,200
- On Nov 24, 2020, the awaited transition to the Biden administration began, the DJI closed at 30,046
- On Sept 13, 2022, after an inflation report of 8.3% year-over-year, DJI dropped 1,276 points and closed at 31,105
- On April 12, 2023 - this Wednesday, two days ago, the DJI closed at 33,647

**Where do you predict the Dow Jones index, at closing, will be on the following days in the future? Please enter numbers only (i.e, without ",").**

One week from today **(Friday, April 21)**

Two weeks from today **(Friday, April 28)**

## Section 28

How certain are you about each of your predictions?

| | Very unsure | Somewhat unsure | Neither sure nor unsure | Somewhat certain | Almost certain |
|---|---|---|---|---|---|
| Prediction 1a: Weather in 2 days | ○ | ○ | ○ | ○ | ○ |
| Prediction 1b: Weather in 4 days | ○ | ○ | ○ | ○ | ○ |
| Prediction 1c: Weather in 7 days | ○ | ○ | ○ | ○ | ○ |
| Prediction 1d: Weather in 14 days | ○ | ○ | ○ | ○ | ○ |
| Prediction 2: Weekend movie box office | ○ | ○ | ○ | ○ | ○ |
| Prediction 3a: Dow Jones Index in 7 days | ○ | ○ | ○ | ○ | ○ |
| Prediction 3b: Dow Jones Index in 14 days | ○ | ○ | ○ | ○ | ○ |

## Section 29

To what extent do you agree with the following statements?

| | Strongly disagree | Somewhat disagree | Neither agree nor disagree | Somewhat agree | Strongly agree |
|---|---|---|---|---|---|
| I found, on average, the prediction tasks difficult | ○ | ○ | ○ | ○ | ○ |
| I took the prediction tasks seriously | ○ | ○ | ○ | ○ | ○ |
| I was distracted during the prediction tasks | ○ | ○ | ○ | ○ | ○ |
| I was excited to make predictions | ○ | ○ | ○ | ○ | ○ |
| I put effort into the prediction tasks | ○ | ○ | ○ | ○ | ○ |

## Section 30

Have you checked what the weather is for this Sunday (2 days from now)?

- Yes
- No

## Section 31

Have you ever seen any of the two videos before?

| | No | Yes |
|---|---|---|
| First video clip (Wildlife) | ○ | ○ |
| Second video clip | ○ | ○ |

## Section 32

Thinking back to the first video (i.e., land animal montage) you saw, which animal was **not** shown during the video?

- Seal
- Tortoise
- Panda
- Camel
- Deer

## Section 33

When watching the second video, please tell us what feelings you felt. Specifically, please only report **what you actually felt when you were watching the video excerpt**, and not what you believe people should feel.

On the scales below, rank the intensity in which you experienced each emotion, if any.

| | Not at all | | | Moderately | | | Very Intensely |
|---|---|---|---|---|---|---|---|
| Interested | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Sad | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Excited | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Angry | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Disgusted | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Happy | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Surprised | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Fearful | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Anxious | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Calm | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

## Section 34

Having completed all tasks, what would you say is the overall goal of this study?

## Section 35

What gender do you identify as?

- Male
- Female
- Non-binary / third gender
- Prefer not to say

What is your age?

- 18-25
- 26-35
- 36-45
- 46-55
- 56+

What is the highest degree or level of education you have completed?

- High school
- Bachelor's degree
- Master's degree
- Ph.D. or higher
- Trade school
- Prefer not to say

# Appendix C.1

## High Trust-in-feelings (TIF) Manipulation Instructions

For the first task, please read the instructions below carefully.

There are two basic ways of making decisions. Sometimes people make decisions based on facts and logical reasons. Sometimes, they make decisions based on their feelings, that is, “does it feel good or bad?”, “does it feel right or wrong?”.

For this task, we are interested in how much confidence individuals have in their feelings when making decisions.

**Instructions:**

We would like you to briefly describe different situations in which you trusted your feelings to make a judgment or a decision and it was the right thing to do.

In other words, describe specific situations that really happened to you in which you felt confident in making a judgment or a decision based on how you felt, and later found out that you were right in using your feelings for this judgment or decision. The situations you list could involve any type of decision, provided that they were real situations in which you felt pretty certain that you should go by what your feelings were telling you to do, and it turned out that your feelings were right. For example, one participant in a previous session of this study mentioned the following:

*“I had a very good feeling toward my new roommate from the moment I met him. Although objectively we were quite different, my feelings told me that we would be good friends. My feelings were correct! We became best friends after only a week. Trusting how I felt about my new roommate was definitely the right thing to do.*"

In the next **seven minutes**, please list **TWO (2)** similar situations that happened to you in which you felt quite confident in your feelings, and you were right in trusting them to make your decision (three or four sentences per situation will be enough).

**Please list below TWO (2) situations in which you were right in trusting your feelings...**

## Appendix C.2

## No Trust-in-feelings (TIF) instructions

For the first task, please read the instructions below carefully.

The purpose of this task is to examine how online search technologies are used to facilitate people's lives. We would like you to briefly describe different situations in which you had to do a search on "Google" to find information that you needed and managed to find the information you were looking for.

In other words, describe specific situations that really happened to you in which you needed information about something and were able to use Google to search for this information. The situations you list could involve any type of situation, provided that they were real situations in which you needed some information and did a Google search to find it.

For example, one participant in a previous session of this study mentioned the following:

*"I needed to find a good Thai restaurant in my neighborhood because I had just moved to the area. I used Google to search for Thai restaurants near my home and found a number of places, which I then checked out further."*

In the next **seven minutes**, please list **FIVE (5)** similar situations that happened to you in which you needed information and you did a Google search to find this information (three or four sentences per situation will be enough).

**Please list FIVE (5) situations in which you searched on Google…**

## Appendix D

## Pre-Screening Questions for Participants

1. **Gross Domestic Product, or GDP, _____** (please select the most correct answer)
    a. Measures national income and output, including consumption, investment, and government spending.
    b. Measures national income and output, including consumption, and government spending.
    c. Measures national income and output, including consumption, investment, government spending, and net exports (i.e., exports minus imports).
    d. Measures national income and output, including consumption, government spending and inflation

2. **What is a P/E ratio?**
    a. The percentage of profits that a company reinvests in its operations
    b. The amount of debt a company has compared to its equity
    c. The percentage of the market capitalization represented by the company's net income
    d. The ratio of a company's stock price to its earnings per share (EPS)

3. **What is the typical relationship between changes in interest rates and stock market levels?** (please circle the most correct answer)
    a. Positive: if interest rates go up, stock prices also go up
    b. Negative: if interest rates go up, stock prices go down
    c. Same: Stock prices and interest rates move in the same direction, whether up or down
    d. There is no relationship

4. **Have you checked the weather in your area (zip code) for any time within the next two weeks?**

    Yes _ No _

# Appendix E

## Real Weather Data Mapped to Survey Categories

| | Sunny/fine/clear | Partly sunny with some clouds | Rain | Thunderstorm | Cloudy/Overcast | Snow | Haze/Fog |
|---|---|---|---|---|---|---|---|
| Source Conditions | sunny | partly sunny | light rain. Mostly cloudy | thunderstorms. Fog | overcast | light snow. Ice fog | fog |
| | | scattered clouds | light rain. Fog | thundershowers. Mostly cloudy | mostly cloudy | | |
| | | passing clouds | light rain. Cloudy | thunderstorms. Broken clouds | broken clouds | | |
| | | | light rain. Overcast | thunderstorms. Passing clouds | more clouds than sun | | |
| | | | light rain. Low clouds | thunderstorms. Mostly cloudy | low clouds | | |
| | | | rain. Low clouds | | cloudy | | |
| | | | rain. Overcast | | | | |
| | | | rain. Fog | | | | |
| | | | rain. More clouds than sun | | | | |
| | | | light rain. Partly sunny | | | | |

*Note*. Source is timeanddate.com, with data procured from a reputable third-party vendor.

# Appendix F

## Block 1 and 2 Results for Weather, Movie Box Office and DJI Predictions (Regressions and Two-Way ANOVA)

**2-Day Weather**

| Predictor Variable | | *B* | OR (Exp(B)) | 95% C.I. for OR Lower Bound | 95% C.I. for OR Upper Bound | *p* | Sig | Nagelkerke $R^2$ |
|---|---|---|---|---|---|---|---|---|
| *Block 1* | | | | | | | | 0.030 |
| | Anger | 0.826 | 2.284 | 1.387 | 3.761 | 0.001 | ** | |
| | Fear | 0.536 | 1.709 | 1.022 | 2.858 | 0.041 | * | |
| | Happiness | 0.651 | 1.917 | 1.149 | 3.197 | 0.013 | * | |
| | Sadness | 0.515 | 1.673 | 0.992 | 2.822 | 0.053 | † | |
| | Neutral | 0.300 | 1.350 | 0.787 | 2.315 | 0.276 | | |
| | TIF | 0.236 | 1.267 | 0.922 | 1.740 | 0.145 | | |
| *Block 2* | | | | | | | | 0.034 |
| | Anger | 0.654 | 1.923 | 0.965 | 3.831 | 0.063 | † | |
| | Fear | 0.263 | 1.301 | 0.623 | 2.718 | 0.484 | | |
| | Happiness | 0.551 | 1.735 | 0.869 | 3.465 | 0.119 | | |
| | Sadness | 0.582 | 1.789 | 0.884 | 3.622 | 0.106 | | |
| | Neutral | 0.014 | 1.014 | 0.466 | 2.206 | 0.972 | | |
| | TIF | 0.017 | 1.017 | 0.546 | 1.894 | 0.957 | | |
| | Anger*TIF | 0.355 | 1.426 | 0.525 | 3.873 | 0.487 | | |
| | Fear*TIF | 0.541 | 1.717 | 0.611 | 4.827 | 0.305 | | |
| | Happiness*TIF | 0.200 | 1.221 | 0.438 | 3.408 | 0.703 | | |
| | Sadness*TIF | -0.168 | 0.845 | 0.296 | 2.413 | 0.753 | | |
| | Neutral*TIF | 0.567 | 1.762 | 0.595 | 5.217 | 0.306 | | |

*Note*. $N$ = 724. Block 1 $\chi^2$ = 15.397, $p$ = .017; Block 2 $\Delta R^2$ = .004, $\Delta\chi^2$ = 2.586, $p$ = .764

**4-Day Weather**

| Predictor Variable | | *B* | OR (Exp(B)) | 95% C.I. for OR Lower Bound | 95% C.I. for OR Upper Bound | *p* | Sig | Nagelkerke $R^2$ |
|---|---|---|---|---|---|---|---|---|
| *Block 1* | | | | | | | | 0.024 |
| | Anger | 0.233 | 1.262 | 0.743 | 2.144 | 0.389 | | |
| | Fear | 0.526 | 1.691 | 1.012 | 2.826 | 0.045 | * | |
| | Happiness | 0.542 | 1.719 | 1.028 | 2.874 | 0.039 | * | |
| | Sadness | 0.639 | 1.895 | 1.132 | 3.172 | 0.015 | * | |
| | Neutral | 0.545 | 1.724 | 1.026 | 2.898 | 0.040 | * | |
| | TIF | 0.279 | 1.322 | 0.961 | 1.818 | 0.086 | † | |
| *Block 2* | | | | | | | | 0.031 |
| | Anger | 0.341 | 1.406 | 0.657 | 3.009 | 0.380 | | |
| | Fear | 0.881 | 2.413 | 1.165 | 4.998 | 0.018 | * | |
| | Happiness | 0.870 | 2.386 | 1.177 | 4.837 | 0.016 | * | |
| | Sadness | 0.598 | 1.818 | 0.862 | 3.833 | 0.116 | | |
| | Neutral | 0.652 | 1.919 | 0.907 | 4.061 | 0.088 | † | |
| | TIF | 0.533 | 1.705 | 0.910 | 3.192 | 0.096 | † | |
| | Anger*TIF | -0.200 | 0.819 | 0.283 | 2.371 | 0.713 | | |
| | Fear*TIF | -0.705 | 0.494 | 0.176 | 1.384 | 0.180 | | |
| | Happiness*TIF | -0.702 | 0.495 | 0.175 | 1.401 | 0.185 | | |
| | Sadness*TIF | 0.109 | 1.115 | 0.394 | 3.152 | 0.838 | | |
| | Neutral*TIF | -0.199 | 0.819 | 0.289 | 2.323 | 0.708 | | |

*Note*. $N$ = 732. Block 1 $\chi^2$ = 12.433, $p$ = .053; Block 2 $\Delta R^2$ = .007, $\Delta\chi^2$ = 3.714, $p$ = .591.

**7-Day Weather**

| Predictor Variable | | *B* | OR (Exp(B)) | 95% C.I. for OR Lower Bound | 95% C.I. for OR Upper Bound | *p* | Sig | Nagelkerke $R^2$ |
|---|---|---|---|---|---|---|---|---|
| *Block 1* | | | | | | | | 0.033 |
| | Anger | 0.726 | 2.067 | 1.224 | 3.489 | 0.007 | ** | |
| | Fear | 0.824 | 2.280 | 1.355 | 3.836 | 0.002 | ** | |
| | Happiness | 0.609 | 1.839 | 1.080 | 3.131 | 0.025 | * | |
| | Sadness | 0.817 | 2.263 | 1.333 | 3.839 | 0.002 | ** | |
| | Neutral | 0.264 | 1.302 | 0.739 | 2.292 | 0.361 | | |
| | TIF | 0.060 | 1.062 | 0.766 | 1.472 | 0.717 | | |
| *Block 2* | | | | | | | | 0.043 |
| | Anger | 0.793 | 2.209 | 1.086 | 4.492 | 0.029 | * | |
| | Fear | 0.987 | 2.682 | 1.312 | 5.481 | 0.007 | ** | |
| | Happiness | 0.663 | 1.941 | 0.953 | 3.954 | 0.068 | † | |
| | Sadness | 0.743 | 2.102 | 1.014 | 4.356 | 0.046 | * | |
| | Neutral | -0.297 | 0.743 | 0.308 | 1.792 | 0.509 | | |
| | TIF | 0.010 | 1.010 | 0.522 | 1.954 | 0.977 | | |
| | Anger*TIF | -0.152 | 0.859 | 0.300 | 2.461 | 0.777 | | |
| | Fear*TIF | -0.342 | 0.710 | 0.250 | 2.019 | 0.521 | | |
| | Happiness*TIF | -0.134 | 0.874 | 0.299 | 2.559 | 0.806 | | |
| | Sadness*TIF | 0.156 | 1.169 | 0.405 | 3.373 | 0.772 | | |
| | Neutral*TIF | 1.034 | 2.811 | 0.873 | 9.053 | 0.083 | † | |

Note. $N$ = 730. Block 1 $\chi^2$ = 16.784, $p$ = .010; Block 2 $\Delta R^2$ = .010, $\Delta\chi^2$ = 5.605, $p$ = .347.

**14-Day Weather**

| Predictor Variable | | *B* | OR (Exp(B)) | 95% C.I. for OR Lower Bound | 95% C.I. for OR Upper Bound | *p* | Sig | Nagelkerke $R^2$ |
|---|---|---|---|---|---|---|---|---|
| *Block 1* | | | | | | | | 0.038 |
| | Anger | 0.739 | 2.094 | 1.192 | 3.676 | 0.010 | * | |
| | Fear | 0.622 | 1.863 | 1.047 | 3.316 | 0.034 | * | |
| | Happiness | 0.413 | 1.512 | 0.832 | 2.746 | 0.175 | | |
| | Sadness | -0.037 | 0.964 | 0.498 | 1.864 | 0.913 | | |
| | Neutral | 0.666 | 1.947 | 1.092 | 3.471 | 0.024 | * | |
| | TIF | 0.449 | 1.566 | 1.091 | 2.250 | 0.015 | * | |
| *Block 2* | | | | | | | | 0.059 |
| | Anger | 0.498 | 1.646 | 0.741 | 3.657 | 0.221 | | |
| | Fear | 0.355 | 1.426 | 0.620 | 3.279 | 0.404 | | |
| | Happiness | 0.329 | 1.390 | 0.619 | 3.122 | 0.425 | | |
| | Sadness | -1.564 | 0.209 | 0.047 | 0.938 | 0.041 | * | |
| | Neutral | 0.596 | 1.815 | 0.812 | 4.056 | 0.146 | | |
| | TIF | 0.039 | 1.040 | 0.505 | 2.141 | 0.915 | | |
| | Anger*TIF | 0.464 | 1.590 | 0.516 | 4.901 | 0.419 | | |
| | Fear*TIF | 0.512 | 1.669 | 0.526 | 5.299 | 0.385 | | |
| | Happiness*TIF | 0.132 | 1.141 | 0.346 | 3.764 | 0.828 | | |
| | Sadness*TIF | 2.318 | 10.157 | 1.833 | 56.284 | 0.008 | ** | |
| | Neutral*TIF | 0.128 | 1.136 | 0.359 | 3.597 | 0.828 | | |

*Note*. $N$ = 726. Block 1 $\chi^2$ = 17.866, $p$ = .007; Block 2 $\Delta R^2$ = .021, $\Delta\chi^2$ = 10.097, $p$ = .073.

**Movie Box Office**

| Predictor Variable | | *B* | OR (Exp(B)) | 95% C.I. for OR Lower Bound | 95% C.I. for OR Upper Bound | *p* | Sig | Nagelkerke $R^2$ |
|---|---|---|---|---|---|---|---|---|
| *Block 1* | | | | | | | | 0.004 |
| | Anger | 0.033 | 1.033 | 0.557 | 1.915 | 0.917 | | |
| | Fear | -0.220 | 0.802 | 0.413 | 1.560 | 0.516 | | |
| | Happiness | -0.071 | 0.932 | 0.492 | 1.765 | 0.828 | | |
| | Sadness | 0.059 | 1.060 | 0.565 | 1.991 | 0.856 | | |
| | Neutral | 0.231 | 1.260 | 0.689 | 2.302 | 0.453 | | |
| | TIF | 0.082 | 1.086 | 0.736 | 1.603 | 0.678 | | |
| *Block 2* | | | | | | | | 0.022 |
| | Anger | 0.624 | 1.867 | 0.809 | 4.305 | 0.143 | | |
| | Fear | 0.063 | 1.065 | 0.407 | 2.787 | 0.898 | | |
| | Happiness | 0.231 | 1.260 | 0.516 | 3.076 | 0.612 | | |
| | Sadness | 0.693 | 2.000 | 0.864 | 4.629 | 0.105 | | |
| | Neutral | 0.241 | 1.273 | 0.503 | 3.218 | 0.610 | | |
| | TIF | 0.606 | 1.833 | 0.896 | 3.752 | 0.097 | | |
| | Anger*TIF | -1.277 | 0.279 | 0.077 | 1.006 | 0.051 | † | |
| | Fear*TIF | -0.539 | 0.584 | 0.153 | 2.218 | 0.429 | | |
| | Happiness*TIF | -0.583 | 0.558 | 0.154 | 2.030 | 0.376 | | |
| | Sadness*TIF | -1.433 | 0.239 | 0.063 | 0.909 | 0.036 | * | |
| | Neutral*TIF | -0.027 | 0.973 | 0.285 | 3.317 | 0.965 | | |

*Note*. $N$ = 740. Block 1 $\chi^2$ = 1.736, $p$ = .942; Block 2 $\Delta R^2$ = .018, $\Delta\chi^2$ = 7.759, $p$ = .170.

**7-Day DJI**

| Predictor Variable | *B* | SE | β | *t* | *p* | Sig | VIF |
|---|---|---|---|---|---|---|---|
| *Block 1* ($R^2$ = .005, $F$ = 0.580, $p$ = .747) | | | | | | | |
| Anger | 19.20 | 142.35 | 0.006 | 0.135 | 0.893 | | 1.26 |
| Fear | -8.08 | 144.22 | -0.002 | -0.056 | 0.955 | | 1.25 |
| Happiness | -42.32 | 143.77 | -0.012 | -0.294 | 0.769 | | 1.26 |
| Sadness | -246.46 | 146.21 | -0.069 | -1.686 | 0.092 | † | 1.25 |
| Neutral | -69.43 | 145.20 | -0.020 | -0.478 | 0.633 | | 1.25 |
| TIF | -6.25 | 89.44 | -0.003 | -0.070 | 0.944 | | 1.00 |
| *Block 2* ($R^2$ = .006, $\Delta R^2$ = .001, $F$ = 0.403, $p$ = .955, $\Delta F$ = 0.195, $p$ = .964) | | | | | | | |
| Anger | -37.70 | 195.67 | -0.011 | -0.193 | 0.847 | | 2.37 |
| Fear | 18.13 | 200.62 | 0.005 | 0.090 | 0.928 | | 2.41 |
| Happiness | 45.18 | 193.41 | 0.013 | 0.234 | 0.815 | | 2.26 |
| Sadness | -217.30 | 199.33 | -0.061 | -1.090 | 0.276 | | 2.30 |
| Neutral | -54.63 | 201.95 | -0.015 | -0.270 | 0.787 | | 2.40 |
| TIF | 24.26 | 162.14 | 0.010 | 0.150 | 0.881 | | 3.27 |
| Anger × TIF | 121.89 | 286.08 | 0.025 | 0.426 | 0.670 | | 2.58 |
| Fear × TIF | -55.02 | 289.48 | -0.011 | -0.190 | 0.849 | | 2.64 |
| Happiness × TIF | -199.64 | 290.25 | -0.039 | -0.688 | 0.492 | | 2.41 |
| Sadness × TIF | -63.19 | 294.23 | -0.012 | -0.215 | 0.830 | | 2.47 |
| Neutral × TIF | -31.53 | 291.44 | -0.006 | -0.108 | 0.914 | | 2.63 |

*Note.* $N$ = 740

**14-Day DJI**

| Predictor Variable | *B* | SE | β | *t* | *p* | Sig | VIF |
|---|---|---|---|---|---|---|---|
| *Block 1 ($R^2$ = .005, F = 0.631, p = .705)* | | | | | | | |
| Anger | -167.45 | 175.51 | -0.039 | -0.954 | 0.340 | | 1.26 |
| Fear | -113.32 | 177.82 | -0.026 | -0.637 | 0.524 | | 1.25 |
| Happiness | -13.28 | 177.27 | -0.003 | -0.075 | 0.940 | | 1.26 |
| Sadness | -147.24 | 180.27 | -0.034 | -0.817 | 0.414 | | 1.25 |
| Neutral | -222.48 | 179.03 | -0.051 | -1.243 | 0.214 | | 1.25 |
| TIF | 135.05 | 110.28 | 0.045 | 1.225 | 0.221 | | 1.00 |
| *Block 2 ($R^2$ = .006, $\Delta R^2$ = .001, F = 0.415, p = .950, $\Delta F$ = 0.160, p = .977)* | | | | | | | |
| Anger | -273.27 | 241.29 | -0.064 | -1.133 | 0.258 | | 2.37 |
| Fear | -105.49 | 247.39 | -0.024 | -0.426 | 0.670 | | 2.41 |
| Happiness | -33.68 | 238.50 | -0.008 | -0.141 | 0.888 | | 2.26 |
| Sadness | -205.51 | 245.79 | -0.047 | -0.836 | 0.403 | | 2.30 |
| Neutral | -332.06 | 249.03 | -0.076 | -1.333 | 0.183 | | 2.40 |
| TIF | 50.50 | 199.94 | 0.017 | 0.253 | 0.801 | | 3.27 |
| Anger × TIF | 226.13 | 352.77 | 0.038 | 0.641 | 0.522 | | 2.58 |
| Fear × TIF | -13.29 | 356.97 | -0.002 | -0.037 | 0.970 | | 2.64 |
| Happiness × TIF | 40.06 | 357.92 | 0.006 | 0.112 | 0.911 | | 2.41 |
| Sadness × TIF | 125.48 | 362.82 | 0.020 | 0.346 | 0.730 | | 2.47 |
| Neutral × TIF | 228.68 | 359.39 | 0.038 | 0.636 | 0.525 | | 2.63 |

*Note.* $N$ = 740

$^{\dagger}p < .10$, $*p < .05$, $**p < .01$.

**Two-Way ANOVA: Emotion Group * TIF on DJI Prediction Error**

| 7-Day DJI | SS | df | MS | *F* | *p* | $\eta^2p$ |
|---|---|---|---|---|---|---|
| Emotion Group | 5,230,399 | 5 | 1,046,080 | 0.707 | 0.618 | 0.005 |
| TIF | 31,416 | 1 | 31,416 | 0.021 | 0.884 | 0.000 |
| Emotion Group × TIF | 1,442,951 | 5 | 288,590 | 0.195 | 0.964 | 0.001 |
| Error | 1,077,159,121 | 728 | 1,479,614 | | | |

| 14-Day DJI | SS | df | MS | *F* | *p* | $\eta^2p$ |
|---|---|---|---|---|---|---|
| Emotion Group | 4,913,546 | 5 | 982,709 | 0.437 | 0.823 | 0.003 |
| TIF | 3,874,896 | 1 | 3,874,896 | 1.722 | 0.190 | 0.002 |
| Emotion Group × TIF | 1,798,281 | 5 | 359,656 | 0.160 | 0.977 | 0.001 |
| Error | 1,637,951,153 | 728 | 2,249,933 | | | |

*Note.* N = 740