Narrative Expectations in Financial Forecasting

Samuel G. B. Johnson [1,2,3]
David Tuckett [2]

[1] University of Warwick, Department of Psychology
[2] University College London, Centre for the Study of Decision-Making Uncertainty
[3] University of Bath School of Management, Division of Marketing, Business, & Society

Author for Correspondence: David Tuckett or Sam Johnson

Email for Correspondence: d.tuckett@ucl.ac.uk or sam.g.b.johnson@warwick.ac.uk

Word Count: 9,886 (main text, including footnotes)

Last Revised: 26 February 2021 (final revision)

In press at *Journal of Behavioral Decision Making*

Acknowledgements: We thank Nigel Harvey, Tim Hill, Don Lancaster, Zoe Lee, and Aba Szollosi for comments. Experiment 1 was presented at the 40th Annual Conference of the Cognitive Science Society. We thank the reviewers and attendees for their suggestions. This work was supported by grants to the second author from the Institute for New Economic Thinking and from the ESRC-NIESR Rebuilding Macroeconomics network.

Declarations: The authors declare no conflicts of interest. Data are available on request from the authors.

Abstract

How do people form expectations about the future? We use amateur and expert investors' expectations about financial asset prices to study this question. Three experiments contrast the *rational expectations* assumption from neoclassical economics (investors forecast according to neoclassical financial theory) against two psychological theories of expectation-formation—*behaviorally-informed expectations* (investors understand empirical market anomalies and expect these anomalies to occur) and *narrative expectations* (investors use narrative thinking to predict future prices). Whereas neoclassical financial theory maintains that past public information cannot be used to predict future prices, participants used company performance information revealed *before* a base price quotation to project future price trends after that quotation (Experiment 1), contradicting rational expectations. Importantly, these projections were stronger when information concerned predictions about a company's future performance rather than actual data about its past performance, suggesting that people not only rely on financially irrelevant (but narratively relevant) information for making predictions, but erroneously impose temporal order on that information. These biased predictions had downstream consequences for asset allocation choices (Experiment 2) and these choices were driven in part by affective reactions to the company performance news (Experiment 3). There were some mild effects of expertise, but overall the effects of narrative appear to be consistent across all levels of expertise studied, including professional financial analysts. We conclude by discussing the prospects for a narrative theory of choice that provide new micro-foundational insights about economic behavior.

Keywords: Expectations, narratives, intuitive theories, behavioral decision-making, experimental finance

Introduction

Our expectations shape our choices. We purchase products we expect to enjoy; go to universities we expect to benefit our careers; marry who we expect to make us happy. In many domains of life, we aim to buy low and to sell high. This insight is a cornerstone of economics and helps to provide a bridge between our beliefs and our behavior (e.g., Friedman, 1957; Lachmann, 1943).

The best-known conception of expectations in economics is the *rational expectations* assumption (Lucas, 1972; Muth, 1961). This assumption simply states that the agents in an economic model are "rational" in the sense that they share the same assumptions as the modelers. For example, when predicting future stock prices, rational economic agents in a standard financial model would understand that stock prices take a random walk. This assumption has long been criticized directly (Davidson, 1982; Haltiwanger & Waldman, 1985; Lovell, 1986; Simon, 1979), while the broader behavioral economics revolution provides an indirect critique. But less work has experimentally contrasted theories of how people form expectations (cf. Adam, 2007; Copeland & Friedman, 1987; Harvey et al., 1994; Hommes, 2011; Plott & Sunder, 1982; Smith et al., 1988). Here, we test amateur and experienced investors' predictions about stock prices. We contrast rational expectations with two alternative, psychologically motivated theories of price expectations—expectations based on stock market anomalies (*behaviorally-informed expectations*) or based on narrative thinking (*narrative expectations*).

Making Sense of Financial Markets

Humans may have an intrinsic tendency to "truck, barter, and exchange" real goods and services (Smith, 1776; see Chen et al., 2006), but aspects of the experience trading financial assets is psychologically challenging (Tuckett, 2011). First, such assets appeared as recently as an evolutionary eyeblink. To the extent that we have adapted intuitions for trade, they would be adapted for barter, not sales of financial assets (Cosmides & Tooby, 1992; but see Johnson, Zhang, & Keil, 2020). Second, these assets are highly abstract, such as streams of future dividends or bundles of loans. Unlike traded consumer goods, financial assets often have no worth beyond what they can be traded for. Third, the value of a financial asset depends not on the value of the underlying asset as such, but on what *other* people believe this value is. As Keynes (1936) put it, the market is like a beauty contest wherein the goal is not assessing the beauty of the contestants, but predicting the other judges' scores. Fourth, traders in financial assets receive extremely noisy feedback given market volatility. Despite economists' insistence that stock prices follow a random walk and are essentially unpredictable, numerous manuals in "technical trading" fill the shelves of bookshops, promising to help investors to detect patterns in overwhelming noise. Finally, posing the greatest difficulty of all, financial decisions are often made under *Knightian* or *radical uncertainty* (Knight, 1921; Mises, 1949) with no principled way to assign probabilities to possible outcomes: What is the probability that a technical innovation can be accomplished on time, that consumers will have the taste for a new product, or that an economic downturn will tighten consumer spending?

We argue here that people circumvent these limited intuitions by using *narrative thinking* to understand financial assets, influencing forecasts of asset values and subsequent choices. Narratives are one way we satisfy our drive to understand the world (Bruner, 1990; Chater & Loewenstein, 2016). Story-telling is cross-culturally universal (Brown, 1991; Hogan, 2003), emerging early in child development (Applebee, 1978) and human history (Abbott, 2000). Stories appear to powerfully shape our cognition (Gottschall, 2012), pervading our memories (Bartlett, 1932; Mandler et al., 1980; Schank & Abelson, 1977) and imbuing our lives with meaning (McAdams, 1993).

Given our facility for storytelling, compared with the obscurity of finance, several thinkers have proposed that storytelling influences economic behavior. Taleb (2001) has argued that people are "fooled by randomness," committing a *narrative fallacy* in which they confabulate narrative explanations for random phenomena (Taleb, 2007). Nobel laureates Robert Shiller and George Akerlof have suggested that powerful stories capture the public's imagination in times of mania and panic, generating feedback loops that create bubbles and busts (Akerlof & Shiller 2009; Shiller 2000, 2017).

Empirical support for these proposals is incomplete, but existing evidence is suggestive. First, narrative thinking influences decision-making broadly. Juries are more swayed when the same witness testimony is arranged to tell a story (Pennington & Hastie, 1992). Similarly, consumers respond more strongly to information presented in a narrative rather than list form (Adaval & Wyer, 1998), form stronger connections to brands presented through narrative (Escalas, 2004), and adjust their attitudes and intentions when they get "lost" in a story (Van Laer et al., 2014).

Second, research on forecasting suggests that people use narratives to predict future events (Beach, 2020; Lawrence et al., 2006). For example, forecasters extrapolate from past trends and act on these extrapolations (De Bondt, 1993; Harvey et al., 1994; Hommes et al., 2008), perhaps because they expect existing causal forces to persist; indeed, later studies found that participants impose more sophisticated patterns on data rather than mere linear extrapolation (Johnson, Matiashvili, & Tuckett, 2019a). More directly, people incorporate causal information in their forecasts (Lim & O'Connor, 1996), particularly information about internal (rather than external) features of the firms (Johnson, Matiashvili, & Tuckett, 2019b). Forecasters often rely on "scenarios" of causally linked events to simplify predictions (Godet, 1982), with scenarios being particularly persuasive (Önkal et al., 2013) and explanations making users less likely to adjust forecasts (Gönül et al., 2009). Scenario thinking can open up forecasters to new possibilities, but can also lead to bias. In many domains, including economic prediction, people often make predictions that account for only the most likely scenario, rather than taking account of multiple possibilities (Johnson, Merchant, & Keil, 2020; Murphy & Ross 1994; but see Chen et al., 2014; Johnson et al., 2019). For example, when people believe that there is a 70% probability that the government will loosen fiscal policy (but a 30% probability against), they make forecasts as though there is a 100% probability of looser fiscal policy (Johnson & Hill 2017). Indeed, merely imagining a scenario to be true increases judgments that it *is* true (Koehler, 1991). In line with these findings, Beach (2020) has argued that conceptualizing scenarios as narratives makes sense of a variety of forecasting biases, although relatively little work has directly investigated forecasting through a narrative lens.

Finally, although not experimental, interviews with professional money managers (Tuckett, 2011, 2012) support the idea that professional investors rely largely on narratives to make decisions. Institutional investors face massive amounts of information and must filter out a minuscule fraction to inform their decision-making. These investors routinely make judgments not only about accounting data, but also about managers' abilities and intentions, the choices of governments, the outlook for the economy, and the whims of consumers. In situations of such profound uncertainty, what choice do investors have but to make their best guess as to what story best fits the facts?

Despite this suggestive evidence, few experiments have directly pitted narrative accounts of financial forecasting against other descriptive theories. Forecasting financial asset prices is different from many forecasting tasks more commonly studied in the literature—such as forecasting sales or earnings—in that stock prices are thought to be *fundamentally unpredictable*. That is, whereas even novices can outperform statistical forecasts in some domains (Lawrence et al., 1985), expertise not only does not seem to help but can even hurt in financial forecasting (Yates et al., 1991). In fact, there is little evidence that there *is* skill in predicting financial asset prices: money managers' performance varies randomly rather than systematically from year to year and few if any funds systematically outperform market returns (Jensen, 1968; Wermers, 2011). Although traders certainly believe there is some predictability in stock prices (otherwise they would not trade them!), many investors may well be aware of the exceptionally low signal-to-noise ratio in financial forecasting. Whether people use narratives in financial forecasting is therefore important to understand for finance research, given the profound differences between financial forecasting and other forecasting tasks. Moreover, few studies have examined how narratives shape forecasting in general, making financial forecasting a potentially useful case study. For example, Önkal et al. (2013) study the effects of *providing* scenarios to forecasters, whereas we are primarily interested in the cues people use to *impose* narratives to structure information. That is, we ask whether and how people convert information into narrative mental representations that facilitate forecasting (Szollosi & Newell, 2020). Studying narrative thinking may thus be useful for understanding the psychological mechanisms underlying scenario-based forecasting (Beach, 2020).

Three Theories of Expectations

Here, we study layperson investors' reactions to news about companies' performance. This is a rich domain for contrasting theories of financial decision-making, because news announcements have previously been studied in detail by financial economists and because plausible theories of investor behavior make sharply divergent predictions. The core question we ask is how investors predict (and act on the prediction of) companies' stock prices given either positive or negative news, and whether these predictions and choices differ in strength depending on whether the news concerns the company's past quarter performance or estimates of the company's next quarter performance. We contrast the hypothesis that investors form *narrative expectations* from the more orthodox hypotheses that investors form *rational expectations* or *behaviorally-informed expectations*.

Rational expectations. The *rational expectations* assumption is a centerpiece of neoclassical economics (Lucas, 1972; Muth, 1961). It holds that the agents in an economic theory form expectations of the future that are consistent with the theory in which they find themselves. In other words, agents in a neoclassical economic theory predict the future using neoclassical economic theory. Thus, to understand the predictions of the rational expectations account, we need to understand what neoclassical economics says about price movements following news announcements.

According to financial theory, stock prices are the market's best guess as to the security's stream of future dividends, discounted to reflect the fact that these dividend payments will occur in the future (Miller & Modigliani, 1961). Stock prices change as new information is revealed that is relevant to determining the company's future value. However, unless an investor has access to information that is not public, she can do no better than chance at predicting future price movements: that is, stock prices take a *random walk* (Fama, 1965). This follows from the logic of arbitrage. If future stock prices were predictable on the basis of publicly available news information, then a "smart money" arbitrageur would be able to capitalize on this predictability by buying or selling shares of the stock before the market moved. Because there are many traders attempting to predict the trajectory of the market, such arbitrage opportunities last for only a very short time—especially in a modern financial market with low transaction costs, near-instantaneous trading, and automated trading algorithms. Financial theorists have argued from this unpredictability that financial markets can be *efficient* in the sense that they incorporate all known information into security prices (Fama, 1970).

Thus, neoclassical theory predicts that positive or negative corporate news announcements will be followed rapidly by a shift in the company's share price, and that prices afterwards will follow a random walk from that new price. Therefore, if a share price is quoted *after* a news announcement (as in our experiments), investors with rational expectations would predict that share prices gradually increase over time at a rate reflecting the risk-adjusted opportunity cost of capital—that is, at roughly the historical rate for a stock of equivalent risk (Brealey et al., 2013). The nature of the announcement is irrelevant to future share prices because all publicly available information is *already* embedded in the share price. This is true whether the announcement is positive or negative relative to previous expectations, and whether it concerns actual past performance or predicted future performance.

Behaviorally-informed expectations. An individual investor really *would* be hard-pressed to make predictions or choices that improve over the predictions of the efficient markets hypothesis. Nonetheless, a variety of anomalies have been detected in stock price data, which, though modest in magnitude, constitute divergences from strictly efficient markets (Shefrin, 2002). Might people intuit these divergences and thereby make predictions that are actually *more* accurate than neoclassically rational expectations?

Empirically, stock prices do *not* follow a strict random walk after earnings announcements. Instead, investors appear to initially *underreact* to earnings announcements (Bernard, 1992; Chan et al., 1996). That is, if a security outperforms expectations, the rapid increase in share price (predicted by market efficiency) is followed by a continued upward *drift* in share prices in the short- to medium-term (Bernard, 1992; Bernard & Thomas, 1989, 1990). The converse is seen when a security underperforms expectations: the initial drop in share value is followed by an extended downward drift in share prices. Put differently, earnings announcements trigger a period of short-term *price momentum* (Cutler et al., 1991; Jegadeesh & Titman, 1993). Although these abnormal returns (relative to the market rate of return) are modest in magnitude, they are difficult to explain in a strict efficient markets framework.

This initial underreaction over short timeframes gives way over longer timeframes to *overreaction* (Chopra et al., 1992; De Bondt & Thaler, 1985; Stein, 1989). After a positive performance surprise, share prices will drift upward in the short- to medium-term, but will drift back *downwards* afterwards. Conversely, after a negative performance surprise, share prices will drift downward for a time, but drift back *upwards* afterwards. That is, security prices drift *too far* in this initial period, and adjust back to an equilibrium price afterwards so that the long-run return of the security is no different from the market overall. What comes up (out of equilibrium) must come down (back to equilibrium) and vice versa. That is, price momentum is followed by *reversion*.

Various models have been proposed to explain this pattern (Barberis et al., 1998; Daniel et al., 1998; Hong & Stein, 1999), but there is no consensus. For our purposes, we simply note that agents with *behaviorally-informed expectations* would anticipate this pattern. If predicting share prices, relative to a benchmark given *after* a performance surprise, a behaviorally-informed investor would predict short-term abnormal returns, over-and-above the market rate of return. (In our studies, we probe for this belief by asking for predictions at a 2-week interval after the announcement.) However, at a longer time interval, such an investor would predict that the prices should revert back toward the market rate of return (in our studies, at a 1-year interval). Although we certainly would not expect amateur investors to have learned about these patterns from the academic literature, it may be plausible that investors could intuit them. After all, investors *cause* them.

Narrative expectations. Although both of the above positions would be in keeping with existing financial theory, in one way or another, we predicted a different pattern because we hypothesize that people construct narratives to make sense of complex systems and guide behavior. *Conviction Narrative Theory* (CNT; Johnson, Bilovich, & Tuckett, 2020; Tuckett & Nikolic, 2017) is an account of choice under radical uncertainty. CNT defines narratives functionally as a mental representation that (i) explains available information, (ii) generates imagined futures, and (iii) motivates actions; thus narratives are a subset of a broader category of causal models or intuitive theories that can simultaneously accommodate past evidence and make future predictions. According to CNT, decision-makers faced with radical uncertainty marshal whatever evidence they can to generate a causal narrative to support their actions, which they extrapolate into the future, conditional on their potential choices. They then rely on their affective reactions to evaluate that possible future and choose between *narratives*; they are then motivated to approach the choice option imagined to bring about the desired outcome.

CNT differs in several ways from other views on offer. First, it goes beyond existing notions of causation in forecasting, such as scenario-based forecasting, by integrating the explanatory, imaginative, and motivational functions of narratives; thus, our approach is consonant with Beach's (2020) notion of grounding scenario-based forecasting in narrative thinking. Second, in appealing to recent cognitive science advances in explanatory reasoning (Lombrozo, 2016), CNT provides a link between the psychology of inference and decision-making. For example, causal knowledge is organized as causal mechanism schemata that permit mental simulation of event sequences (Hegarty, 2004; Johnson & Ahn, 2015)—thus, causal narratives naturally provide a link between past evidence and imagined futures. Likewise, we typically simulate a single possibility at a time (Evans, 2007; Johnson et al., 2020; Murphy & Ross, 1994), consistent with the commonsense intuition that stories follow a discrete sequence of events rather than existing in our minds as probability distributions. Finally, because narratives are a natural format for human communication (e.g., Smith et al., 2017), CNT may be promising for understanding how individual cognition embedded in a social context leads to "viral" beliefs (Shiller, 2017).

CNT makes many predictions about forecasting, of which we focus on a subset in this article. To derive these predictions, we reflect on the signature properties of stories (Bruner, 1990; Graesser et al., 1994; Mandler & Johnson, 1977; Mar & Oatley, 2008; Rumelhart, 1975). Stories refer to goal-directed activities, which become emotionally valenced as goals are approached or thwarted. Stories are temporally oriented, referring to sequences of discrete events occurring in a particular order. Stories provide causal explanations and rely on a set of schematic patterns to make sense of information organized over time. And as noted above, stories occur in discrete sequences rather than probability distributions.

Here, we restrict ourselves to testing two of these predictions, concerning the *structure* of narratives, though we test several other predictions of CNT elsewhere (Batteux et al., 2020, 2021; Bilovich, Johnson, & Tuckett, 2020; Johnson et al., 2019a, 2019b; Nyman et al., 2018; see Johnson et al., 2020 for a review), including predictions about narrative *content*. Here, we examine the consequences of narratives being *goal-oriented* and *temporally-oriented*.

First, stories (like investments) are goal-oriented. Their protagonists want to achieve certain objectives and developments in the narrative either facilitate or thwart these objectives. Thus, stories take on an emotional valence as goals become closer or more distant. Just as memories are often organized around narratives, memories often take on the emotional tinge of the associated narrative (Bartlett, 1932; Bower, 1981). If people use narratives to generate predictions, then they should use the valence of information to inform their predictions. According to CNT, narratives that generate approach emotions should be associated with more positive predictions and actions, while narratives that generate avoidance emotions should be associated with more negative predictions and avoidance behaviors (see Bilovich et al., 2020 on how approach vs. avoidance emotions influence decision confidence). This prediction is structural in the sense that narratives organized around goal approach should lead to more positive predictions compared to narratives in which goals are thwarted. Different attributions for goal approach (e.g., luck versus skill) may moderate this effect, but we test this possibility elsewhere (Johnson et al., 2019b).

Second, stories are temporally oriented (Mar & Oatley, 2008). They have a beginning, middle, and end, and causality flows in a single direction. If we can be informed directly about the future, that is a better clue to how the story ends compared to what has already happened in the past. Indeed, the future seems to be more psychologically "real" than the past. Future actions are seen as more intentional and, if unethical, more morally wrong (Caruso 2010; Burns et al., 2011). People ask for greater compensation for future harms (Caruso et al., 2008), future events evoke stronger affective reactions (Van Boven & Ashworth 2007), and the future feels "closer" than the past (Caruso et al. 2013). According to CNT, forecasts are narrative simulations (see also Beach & Mitchell, 1987). Such simulations can use information about the past, but, if *direct* information about the future is instead available, such information would be more readily incorporated into simulations as its implications for the future have already been "pre-processed." We would therefore expect future-oriented information to be weighed more heavily than past-oriented information, even if equally (ir)relevant.

The narrative expectations hypothesis therefore makes two distinct predictions based on narrative structure. First, both positive and negative trends should be projected into the future at all time horizons. Thus, a positive (negative) performance surprise should lead to predicted abnormal returns above (below) the market return both at short and long time horizons. This is distinct from neoclassically rational expectations, which would not use *past* performance surprises to predict future returns, as well as from behaviorally-informed expectations, which would predict trend continuation in the short-term but reversion to the market return in the long-term. Second, the effect of valence should be stronger when the news concerns predicted *future* rather than actual *past* performance. That is, the predicted abnormal returns induced by performance surprises should be amplified (more positive or more negative), following future-oriented news. This prediction cannot be motivated by neoclassically rational expectations, and given the lack of evidence for such an effect in the behavioral finance literature, it also appears to be inconsistent with behaviorally-informed expectations.

Summary of Experiments

We distinguish these theories of investor expectations across three studies. Experiment 1 measures expectations directly, following corporate performance surprises, measuring predicted stock prices at shorter and longer intervals. The rational, behavioral, and narrative expectations accounts make contrasting predictions about the effects of news valence and time-reference on these predictions. Experiment 2 tests whether these predictions translate into asset allocation choices that are biased by the standards of financial theory, with potentially negative implications for the real-world returns of amateur investors. Experiment 3 tests whether these choices occur because positive performance surprises generate approach emotions and negative performance surprises generate avoidance emotions, with particularly strong emotional resonance for future- rather than past-oriented news. In the Supplementary Materials, we examine evidence concerning expertise effects, both within the main studies and an additional sample with greater expertise.

Experiment 1

Participants in Experiment 1 learned and made judgments about the stock prices of realistic, but fictitious, companies. For each company, participants learned that an hour *previous* to the most recent price quotation, an announcement was made by analysts concerning the company's performance in either the *past* or *future* quarter, which was either *positive* or *negative*. Participants were then asked to predict the future trajectory of the price, at intervals of one-day, two-weeks, and one-year.

As described earlier, neoclassical financial theory predicts that the price should take a random walk after the initial adjustment following the announcement. Thus, if people have neoclassically rational expectations, they would predict that the price should rise at a rate consistent with other securities, adjusted for the riskiness of the asset (i.e., at the opportunity cost of capital). This is true both at short and long time intervals because future news is unknown at the time of prediction. Critically, there should be no difference between future price predictions for positive or for negative performance surprises (assuming the predictions are made after a short period is allowed for the information to be priced in), nor for surprises about past versus future performance.

If participants are more sophisticated and rely on a mental model concordant with behavioral finance theories, then they may predict modest post-announcement drift in the short-term (i.e., a more rapid price increase following a positive rather than negative surprise) followed by a reversal of this trend in the longer term. Thus, we would expect the divergence in price between positive and negative surprises to decrease and be eliminated in the long-run (i.e., our one-year horizon). As we are not aware of any econometric work documenting divergences in price momentum between surprises in past versus future performance, we do not believe that a behaviorally-inclined participant would differentiate between these conditions.

In contrast, because we hypothesize that people use narrative thinking to make predictions, we anticipated that participants would differentiate between positive and negative surprises, and between surprises in past and future performance.

Method

Participants. We recruited 225 American participants from the online crowdsourcing platform Amazon Mechanical Turk. This target sample size was set a priori for all experiments and achieves 90% power for within-subjects effects $d > 0.20$. For Experiment 1, 40 participants were excluded from analysis due to inattentiveness (see below). Participants were prevented from participating in multiple experiments reported in this article.

Relative to student samples, the demographics of Mechanical Turk are more appropriate for experiments in economic decision-making because the participants come from a wider range of age, education, and socioeconomic backgrounds. For Experiment 1, the sample ranged in age from 19 to 71 ($M = 37.8$, $SD = 11.4$) and in education from "did not complete high school" to "graduate degree" (median = "some college"). About half (49%) of participants held some financial assets (such as stocks, bonds, or mutual funds) and about half (53%) had taken at least one finance course. About 14% of participants majored in a business field, such as finance, management, accounting, or economics. Thus, although Mechanical Turk participants are generally not expert investors, they reasonably represent the investing experience of the American public, as about 52% of Americans hold stocks (McCarthy, 2016) and about 19% of American bachelor's degrees are awarded in business fields (National Center for Education Statistics, 2016).

Clearly, this population is not nearly as experienced as professional traders. However, as many as half of our participants belong to the category of low-information investors known as "noise traders" in the finance literature (Shleifer & Summers, 1990). Financial models turn greatly on the assumed behavior of these investors (e.g., Shleifer & Vishny, 1997), so it is important to characterize these investors' actual beliefs and behaviors. That said, Parts C and D in the Supplementary Materials test for expertise effects, both within our primary sample and in a sample of genuine experts.

Procedure. Each participant completed four items, each pertaining to a different fictitious company—Remlon Software Corporation (RWQ), Wilfinger Industries (WNV), Paravoz Exploration

(PVZ), and Excellerate Construction (XOL). Each company appeared in one of the four experimental conditions (past/positive, future/positive, past/negative, and future/negative), with the assignment of company to condition counterbalanced using a Latin square.

For each company, participants first read background information about the company and its current price. For example, one item read:

> Remlon Software Corporation (stock symbol RWQ) is a Dallas-based company that designs and markets business software to medium- and large-size firms.
>
> Here is the most recent price quotation for shares in RWQ stock: $56.00.

Then, participants were asked to make baseline predictions about the price trajectory of the shares ("Given that RWQ shares currently trade at $56, please estimate what you think the share price will be on the following dates") at time horizons of "tomorrow," "in two weeks," and "in in one year." Ratings were made on a sliding scale centered at the current price, and ranging from 50% less than the current price (e.g., $28 for RWQ) up to 50% more than the current price ($84 for RWQ). This measure was taken to understand participants' expectations about the price trajectory of each stock in the absence of performance data and to provide a comparison to news-induced price predictions.

On the next screen, participants read a piece of news from financial analysts concerning the security, which instantiated our experimental manipulations of valence (positive or negative) and time (past-oriented or future-oriented information). Critically, in both conditions, the news information was said to have come out an hour *before* the price quotation. Thus, market would have already incorporated this news into its valuations.

In the past condition, this information described past performance relative to average (with the bracketed text varying across the positive and negative conditions):

> *About an hour prior to the most recent price quotation ($56) for Remlon's stock (RWQ), the following piece of news was revealed:*
>
> Although average sales growth is expected for the next quarter, analysts determined that Remlon experienced [above-average / below-average] levels of sales growth over the past quarter.

Conversely, in the future condition, the information described expected future performance relative to average:

> *About an hour prior to the most recent price quotation ($56) for Remlon's stock (RWQ), the following piece of news was revealed:*
>
> Although average sales growth was observed for the past quarter, analysts anticipate that Remlon will experience [above-average / below-average] levels of sales growth over the next quarter.

The performance measures varied across the companies, and included sales growth (as above), as well as innovation, discoveries of mineral deposits, and new contracts. The full text of the instructions and items is reported in Part A of the Supplementary Materials.

Below this information, participants were asked to make a new prediction: "Given that RWQ shares currently trade at $56, please estimate again what you think the share price will be on the following dates." The time horizons and scale were the same as the baseline prediction.

After the main task, participants answered a set of recognition memory check questions (concerning the industries of the companies), to monitor attentiveness. Any participants incorrectly answering more than 30% of these questions were excluded from data analysis (*N*=40). However, Part B of the Supplementary Materials reports a version of the analyses including all participants.

Finally, after answering demographic questions including measures of financial expertise, participants were debriefed, explaining the purpose of the study and that the companies were fictitious.

Results

For statistical analyses, we converted participants' price estimates into percentage changes relative to the price quotation given in the problem, as shown in Table 1. Overall, the results largely confirm the predictions of the narrative account. Participants predicted much more bullish price changes after a positive surprise, relative to baseline, and much more bearish price changes after a negative surprise. For the positive surprises, these predicted changes were larger in light of future-oriented than for past-oriented performance information. Data are available at https://osf.io/hy3w2/.

The analyses below rely on simple comparisons between key cells for ease of presentation. In Part B of the Supplementary Materials, we report hierarchical regression models for all experiments. These models include robustness checks on sample exclusion criteria (repeating key analyses on both the full sample and a subset that excludes outliers) and using different specifications, such as including baseline forecasts as a covariate and (for the 2-week and 1-year intervals) including a lagged forecast variable (i.e., the 1-day and 2-week forecast, respectively). For the valence effect, results are generally robust to these analytical choices. For the time-reference effect, results are less robust, with predicted effects showing up in some (but typically not all) specifications in all three experiments.

Time Horizon	Baseline	Positive Surprise		Negative Surprise	
		Past	Future	Past	Future
1-day	1.7%	5.5%	6.5%	–2.7%	–2.9%
	(2.6%)	(7.8%)	(7.7%)	(6.7%)	(7.5%)
2-weeks	4.3%	9.3%	11.2%	–4.6%	–4.4%
	(4.9%)	(9.6%)	(9.0%)	(7.9%)	(8.6%)

1-year	8.7%	14.7%	17.5%	–5.7%	–6.1%
	(9.9%)	(15.4%)	(15.4%)	(12.8%)	(14.1%)
Number of Observations = 185					

Table 1. Results of Experiment 1.

Note. Entries are predicted changes from current value, as percentages. Possible scores range from –50% to +50%. The baseline column gives the mean of the baseline estimates made across the four within-subjects conditions, since these estimates were made prior to the manipulation. SDs in parentheses.

Baseline predictions. At the baseline, prior to reading any news, participants expected a moderate price increase over 1-day (+1.7%), 2-week (+4.3%), and 1-year (+8.7%) time horizons. Although the 1-day and 2-week predictions are optimistic, the 1-year prediction is consistent with historical returns (e.g., about a 10% nominal increase annually for the S&P 500). The variance in predictions increased at longer time intervals, in both the baseline and experimental conditions. This may reflect the greater uncertainty at long horizons about either specific firms or general market conditions. As shown in Part B of the Supplementary Materials, baseline forecasts are strongly predictive of forecasts in all conditions. However, models that include and exclude this variable tend to result in similar estimates of the valence and time-reference effects.

Valence of news. Because the news information given was from *before* the most recent price quotation, predictions about future prices should not depend on whether the news was positive or negative. Yet, Table 1 shows that predictions markedly differed depending on the news valence.

Looking at the positive surprise items collapsed across time conditions, participants predicted increases of +6.0% at a 1-day, +10.3% at a 2-week, and +16.1% at a 1-year timeframe. These predictions were significantly more positive than the baseline predictions [ts>8.9, ps<.001, ds>0.61], in violation of market efficiency. Strikingly, the divergences between the baseline and the positive surprise predictions were largest at *longer* time intervals. That is, the performance surprise led to a predicted premium of +4.3% at 1-day and +6.0% at 2-weeks, with the latter premium significantly larger [$t(184)=5.28$, $p<.001$, $d=0.25$], with a yet larger premium of +7.4% at 1 year [$t(184)=2.92$, $p=.004$, $d=0.15$]. In other words, the alleged predictive signal associated with the news announcement actually grew *larger* rather than smaller over longer time frames. Thus, participants predicted strong price momentum, with investors underreacting to news—a belief at least qualitatively consistent with empirical studies of asset prices. However, whereas in reality these trends reverse in the longer run, participants predicted an *ever-increasing* effect of positive news.

The story was similar for negative surprises, but even more dramatic (in line with other asymmetries between positive and negative events; e.g., Baumeister et al., 2001). Collapsing across time conditions, participants predicted *decreases* of –2.8% at 1-day, –4.5% at 2-weeks, and –5.9% at 1-year timeframes. Needless to say, these predictions diverged sharply from the baseline [*t*s>9.2, *p*s<.001, *d*s>0.93] as well as from the positive surprise condition [*t*s>10.4, *p*s<.001, *d*s>1.3]. And once again, the predicted shortfall relative to baseline increased at longer time horizons, with a shortfall of –4.5% at 1 day versus –8.8% at 2 weeks [$t(184)=9.50$, $p<.001$, $d=0.54$] and an even larger shortfall of –14.6% at 1 year [$t(184)=8.68$, $p<.001$, $d=0.48$]. Again, participants predicted both short- and long-term momentum, rather than long-term reversion as has been found empirically.

Time-reference of news. Though not as strong as the effect of valence, participants often took account of the time-reference of news inconsistently with financial theory. Predictions tended to be more extreme (i.e., positive in light of positive news and negative in light of negative news) given future-oriented information compared to past-oriented information. Collapsing across valence, future-oriented predictions were 0.6% more extreme at a 1-day horizon [$t(184)=1.84$, $p=.066$, $d=0.14$ vs. 0% in a one-sample test], 0.8% more extreme at a 2-week horizon [$t(184)=2.29$, $p=.023$, $d=0.17$], and 1.6% more extreme at a 1-year time horizon [$t(184)=2.50$, $p=.013$, $d=0.18$]. Thus, overall our prediction was supported that future-oriented information would be weighted more heavily than past-oriented information due to the inherently temporal nature of narrative thinking.

However, these effects were not symmetric across valences, but were instead driven by the positive valence conditions. For positive news, there was a substantial effect of time-reference at all horizons (1.0%, 1.9%, and 2.8%), whereas there was no significant effect at any horizon for the negative valence items (0.2%, –0.2%, and 0.4%). It is unclear what accounts for this asymmetry, which was not observed in subsequent experiments. One possibility is that participants were hesitant to predict more negative price changes than –6% in light of information that is only moderately negative, especially given that the stock market was quite bullish at the time of the experiment (March 2017). That is, our manipulation may have run into a tacit floor. If this is the case, then more extreme negative events could potentially lead to a time-reference asymmetry. Rather than pursuing this approach, however, subsequent experiments turn instead to alternate dependent measures.

Parts C and D of the Supplementary Materials examine the effects of expertise. Such effects appear to be modest. Within the range of expertise in our experiments, we find that investing *experience* seems to modestly decrease (but not eliminate) the effects, whereas self-reported investing *knowledge* seems, if anything, to exacerbate them (Part C), with these findings reasonably consistent across Experiments 1–3 but nonetheless exploratory. We also conducted a near-exact replication of Experiment 1 on a sample of individuals with greater expertise—professional financial analysts, PhD students in economics, and Masters students in finance (Part D). That study revealed nearly identical findings, suggesting that even the intuitions of experts such as investment professionals may differ little from those of our non-expert participants. Thus, although market experience may well attenuate some behavioral biases (e.g., List, 2003), there is little evidence that the narrative effects we see here are eliminated by expertise.

Discussion

These results support the idea that people rely on narratives when predicting the price trajectories of financial assets. Whereas participants with neoclassically rational expectations would predict increases in asset prices at the market rate of return, our participants sharply differentiated between positive and negative performance surprises, predicting dramatically superior growth in light of a positive rather than negative piece of news. This was the case even though the predicted price changes were made relative to the price *after* the news announcement. Instead, news information appears to trigger narratives in investors' minds. Since narratives are temporally extended, they can be used to make predictions about the future. The effect of news valence was not symmetric relative to the baseline predictions, with negative news exerting a larger effect than positive news. This is in line with well-documented negativity biases in many domains (Baumeister et al., 2001; Kahneman & Tversky, 1979), although negativity biases are not always seen in forecasting tasks (Fildes et al., 2019).

In addition, there was some evidence that participants differentiated between news concerning the past versus the future—a finding that appears at odds with both neoclassically rational and behaviorally-informed expectations. Positive surprises about past performance were seen as less positive than surprises about expected future performance, although the corresponding effect for negative surprises did not reach significance in this experiment. If people think about financial assets like economists—who recognize that it is *expectations* about the future that matter, which are quickly priced in to asset prices, whether new information concerns the past or the future—then the temporal direction of performance surprises should not matter. But if people use news information as raw material for constructing narratives about the company, then information about the future would indeed be more diagnostic about the company's future than information about the past.

Could these results be reconciled with neoclassical financial theory on the basis of participants' inferences about risk? According to financial models such as the Capital Asset Pricing Model (CAPM; Fama & French 2004; Lintner 1965; Sharpe 1964) and its successors, investors prefer, for a given rate of return, securities with lower variance around that expectation. That is, investors are risk-averse. According to this logic, investors will require a larger expected return to invest in a riskier security, and participants' tendency to predict higher returns for some securities would be consistent with mainstream theory if due to inferences about risk.

However, this explanation is not workable. For the risk-inference account to hold, people would need to believe that securities with *positive* performance surprises are riskier (having greater variance) than those with *negative* performance surprises. Further, the magnitude of the difference between the positive and negative surprises (of greater than 20% at a 1-year horizon) is empirically implausible as a risk premium. It is more plausible that participants would believe future information to be more risk-inducing than past information (justifying the higher expected return for positive future compared to past surprises). However, the risk account would also predict that future *negative* performance surprises should lead to stronger future returns compared to past negative surprises. The means generally went in the opposite direction (albeit non-significantly), and we will see a significant effect in the opposite direction in Experiment 2. Thus, inferences about risk are unlikely to account for participants' divergent predictions based on the valence and time-reference of news information.

The results also conflict with behaviorally-informed expectations. Investors with such expectations would predict short-term price momentum, followed by longer-term reversals. Our participants diverged from this pattern in three ways. First, their short-term momentum was overzealous compared with the econometric findings (Bernard & Thomas 1989, 1990). Second, rather than reverting back toward the market return in the longer-term, participants' predictions were precisely the opposite, diverging increasingly at longer horizons. Finally, we are not aware of any behavioral work that would predict a difference in predictions or choices based on the time reference of company news, so it is unclear how the behavioral account would explain the time-reference effect.

We also can consider several alternative explanations. First, perhaps most participants simply are not aware of the idea that known news is incorporated into current prices (explaining the valence asymmetry) and this ignorance is more common for future-oriented news (explaining the temporal asymmetry). One prediction made by this account is that the effects should disappear for those participants who are especially knowledgeable. However, as noted above and explored in Parts C and D of the Supplementary Materials, this was not the case.

Second, several researchers have raised concerns about demonstrations of irrationality that require participants to interpret and accept statements made by the experimenters (Hilton, 1995). For example, framing effects sometimes depend on inferences made about the implicit recommendations of the speaker (Sher & McKenzie, 2006), while the conjunction fallacy appears to depend in part on inferences about ambiguous terms such as "probability" (Fiedler, 1988; Hertwig & Gigerenzer, 1999). In our studies, the predictions of rational expectations particularly depend on participants' accepting that the news was publicly announced one hour prior to the price quotation—perhaps they don't. Related, participants might experience demand characteristics, feeling the need to differentiate between conditions based on a desire to use the given information.

While we cannot rule these out as contributing factors, they cannot be the whole story. These explanations seem to predict a difference in the first period, but do not make clear predictions about what specific pattern we will see in later periods, and in particular do not distinctly predict the narrative expectations pattern (differences between conditions get larger at later periods) versus behaviorally-informed expectations pattern (differences get smaller at later periods). Moreover, it is unclear how these factors would explain the time-reference effect. At the same time, we are sympathetic to aspects of these accounts, particularly the idea that participants make intelligent inferences about speakers' intentions. Indeed, such conversational inferences may contribute to the adoption of narratives.

Finally, perhaps the time-reference effect is due to confounding past/future with objective/subjective information, since the future, unlike the past, is inherently unknowable. Although we acknowledge that this confound exists, it actually seems to push in the opposite direction of our results—people should rely *less* on subjective than on objective information, yet people tend to rely *more* on future- rather than past-oriented information.

Experiment 2

Experiment 2 sought to extend the findings to a new measure—choice. After once again learning about positive or negative surprises in past or future performance, participants rated the likelihood that they would include the security in a portfolio they were constructing. Since accurate predictions of price growth represent profit opportunities, participants should prefer to hold securities that they expect to increase in value. Based on Experiment 1, we therefore predicted (a) a preference to hold securities with a positive rather than negative performance surprise, and (b) a more extreme preference when the surprise concerned expected future performance rather than actual past performance.

Method

We recruited 225 Americans from Mechanical Turk, excluding 51 using the same criterion as Experiment 1. However, the results are similar if these participants are included in the analyses.

The method was similar to Experiment 1. Participants made judgments about four companies, which faced either positive or negative surprises concerning their past or future performance. Relative to Experiment 1, we introduced two changes. First, participants were asked to make a portfolio allocation choice rather than a prediction about future value. Participants were asked to "Suppose that you are creating a portfolio of securities. Given that RWQ shares currently trade at $56, please rate the probability that you would include RWQ shares in your portfolio." These ratings were made on a scale from 0% to 100%. Second, where participants in Experiment 1 made predictions both before and after reading the news information, allocation choices were only made once in Experiment 2, after reading the news information. This change was made to avoid potential demand characteristics associated with asking for two ratings, which could be a possible concern about Experiment 1.

Negative Surprise		Positive Surprise	
Past	**Future**	**Past**	**Future**
29.7% (22.4%)	23.1% (18.0%)	62.3% (21.7%)	66.5% (22.9%)
Number of Observations = 174			

Table 2. Results of Experiment 2.

Note. Entries are probability judgments that the security would be invested in the participant's portfolio. Possible scores range from 0% to 100%. SDs in parentheses.

Results and Discussion

As shown in Table 2, the results replicate both the valence and the temporal asymmetries found in Experiment 1.

Looking first at valence, we collapse across the past and future conditions within each valence condition, reporting on average a 64.4% chance of including a security in their portfolio if it had experienced a positive performance surprise, and a 26.4% chance if it had experienced a negative performance surprise. These judgments, obviously, differed significantly from one another [$t(173)=20.35$, $p<.001$, $d=2.02$]. Thus, the effect of valence translates into portfolio allocation choices.

This effect was moderated by the time reference of the news. Like Experiment 1, allocation choices were 5.4% more extreme (collapsing across valence) when the news was future- rather than past-oriented [$t(173)=4.50$, $p<.001$, $d=0.34$ vs. 0% in a one-sample test]. Unlike Experiment 1, however, this effect was driven by both the positive and negative conditions. When the news was positive, participants were 4.2% *more* likely to invest if the news was future- rather than past-oriented [$t(173)=2.65$, $p=.009$, $d=0.20$] and when the news was negative, participants were 6.6% *less* likely to invest if the news was future- rather than past-oriented [$t(173)=4.45$, $p<.001$, $d=0.34$].

Experiment 3

What psychological mechanisms underpin the relationship between narrative thinking and choice? On the one hand, Experiment 1 demonstrated that predictions are driven in part by narrative thinking, and it is no surprise that participants' beliefs about the future translate into patterns of choices. In addition to this cognitive process, however, we note that both narratives and choices are often tinged with emotion. The most prototypical examples of narratives in our culture are novels and films which are enjoyable precisely because they evoke emotions. A related literature on explanatory reasoning has argued that people accept explanations in part because they "feel" satisfying (Gopnik, 1998; Johnson, 2017; Lipton, 2004), even in such abstract domains as mathematics (Johnson & Steinerberger, 2019). On the choice side, we are likelier to make choices that maximize approach emotions and minimize avoidance emotions (Carver, 2006). Emotions are critical metacognitive cues that help us to mediate between cognition and action, aiding us in planning and typically guiding us toward adaptive choices (Oatley & Johnson-Laird, 1986). However, they can also become untethered from more rational cognitive appraisals and lead to mistakes (Loewenstein et al., 2001).

More specifically, we propose that people hold emotion-tinged attitudes toward specific objects, such as financial securities, and that these emotions are influenced by the role of that object in the person's narrative (Bilovich et al., 2020; Johnson et al., 2020). If the narrative predicts a positive outcome for a security, this generates approach emotions, which lead people to purchase the security. Conversely, if the narrative predicts a negative outcome for the security, this generates avoidance emotions, leading people to sell the security.

Experiment 3 tests the idea that emotional processes mediate the relationship between performance surprise information and choices, for both the effect of valence and of time reference. That is, we expected that people would feel more positive, approach emotions in light of positive surprises, leading to choices to invest in those securities, and to feel more negative, avoidance emotions in light of negative surprises, leading to choices to avoid those securities. Further, we expected that this affective gap would be larger when the surprises were future- rather than past-oriented, and that this larger affective gap would lead to a larger gap in choice.

Method

We recruited 225 Americans from Mechanical Turk, excluding 45 using the same criterion as previous experiments.

The method was similar to Experiment 2, with two changes. First, we introduced a measure of participants' emotions for each item. On the same screen as the news information, participants were asked to "Suppose you held shares in Remlon Software Corporation (RWQ). How would the above information make you feel? Please check all that apply." Participants then checked items from a list of 20 avoidance emotions (e.g., "distrustful," "threatened," "worried") and 20 approach emotions (e.g., "confident", "passionate," "satisfied"), which were listed in a new random order for each item. Second, the choice measure was moved to a separate page to avoid explicitly reminding the participants which emotions they checked.

Results

Participants were likelier to include securities in their portfolios after a positive rather than negative surprise, as shown in Table 3, and once again this trend was exaggerated when the surprises concerned future rather than past performance. Going beyond Experiment 2, however, we find that these effects of valence and time-reference are mediated by emotions.

Given that our hypotheses are described in terms of mediation, our statistical analyses are organized around mediation tests. Because the manipulations were within-subjects, we followed the procedure of Judd, Kenny, and McClelland (2001), as implemented in the MEMORE macro for SPSS (Montoya & Hayes, 2017) using the percentile bootstrap method for computing confidence intervals.[1]

[1] This procedure is a path-analytic formulation of Judd et al.'s (2001) procedure for within-subjects mediation tests. This procedure estimates the effects in the following way. The total effect c of the independent variable (X) on the outcome variable (Y) is estimated by conducting a paired t-test comparing the two levels of Y (equivalently, regressing the difference between levels of Y on an intercept term). The effect a of X on the mediator (M) is estimated by conducting a paired t-test comparing the two levels of M (equivalently, regressing the difference between levels of M on an intercept term). The effect b of M on Y is estimated by regressing Y on M. The direct (unmediated) effect c' of X on Y is estimated by regressing Y on X and adjusting for the effect of the mediator. The indirect effect of X on Y via M is equal to ab, or equivalently $c - c'$. The distribution of ab is estimated using bootstrapping, allowing for confidence intervals and significance tests. We estimated p-values

	Negative Surprise		Positive Surprise	
	Past	**Future**	**Past**	**Future**
Avoidance Emotions	5.6 (3.9)	6.0 (4.1)	0.8 (5.2)	0.7 (5.3)
Approach Emotions	0.3 (1.3)	0.2 (0.6)	5.2 (3.9)	5.3 (4.0)
Choices	30.8% (20.1%)	29.6% (21.0%)	66.1% (19.5%)	68.3% (19.7%)
Number of Observations = 180				

Table 3. Results of Experiment 3.

Note. The first two rows indicate the mean number of negative and positive emotion words (each out of 20) checked in each condition. The third row shows probability judgments that the security would be invested in the participant's portfolio (from 0% to 100%). SDs in parentheses.

Emotion as a mediator of the valence effect. We first examine the effect of valence on emotion and choice. Because the design was within-subjects, the mediator and outcome variables each have two levels (i.e., the positive and negative surprise conditions). The outcome variables were the mean propensity to include the security in the participant's portfolio, collapsing across the past and future conditions, separately for positive and negative surprises. The emotion mediator variable was a *net emotion* score (approach minus avoidance emotions), collapsing again across time reference condition, separately for positive and negative surprises. The mediator variables thus could potentially range from –20 to +20, with higher numbers indicating a preponderance of approach over avoidance emotions.

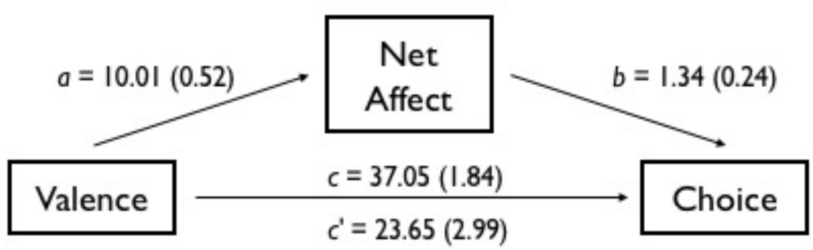

Figure 1. Mediation of valence effect on choice by affect.

Note. The *c* coefficient reflects the total effect of valence on choice, measured by the difference in choice between positive and negative surprises. The

---

for the tests of indirect effects by calculating the proportion of bootstrap samples for which the indirect effect estimate was not positive, and doubling this proportion (to create a two-tailed test).

indirect effect is equal to the product of the *a* and *b* coefficients, while the direct effect *c′* is the remaining effect of valence on choice after accounting for the indirect effect. Coefficients are unstandardized (SEs in parentheses).

As shown in Figure 1, there was a significant total effect of valence on choice, 95% CI[33.42,40.69], $p<.001$. That is, replicating Experiment 2, people were likelier to include assets in their portfolio following a positive rather than negative surprise. The bootstrapping procedure revealed that this effect has both a mediated component via emotion, 95% CI[8.38,18.90], $p<.001$, and an unmediated, direct component, 95% CI[17.75,29.54], $p<.001$. Since there were significant indirect (mediated) and direct (unmediated) paths, we conclude that the effect of valence on choice is partially mediated by emotion. Given that the mediation was partial, other (perhaps cognitive) mechanisms are also likely to be at play in explaining this effect.

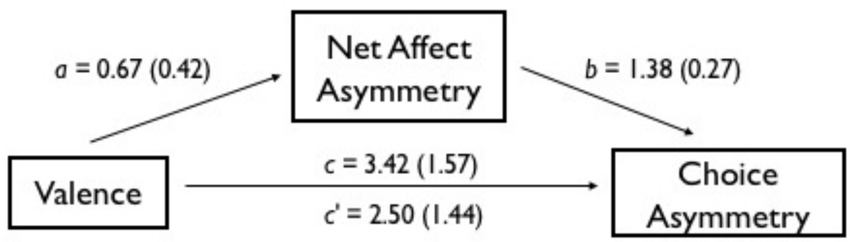

Figure 2. Mediation of time-reference effect on choice by affect.

Note. The *c* coefficient reflects the total effect of valence on time-reference-induced differences in choice (future minus past), measured by the difference in time-reference difference scores between positive and negative surprises. The indirect effect is equal to the product of the *a* and *b* coefficients, while the direct effect *c′* is the remaining effect of valence on choice after accounting for the indirect effect. Coefficients are unstandardized (SEs in parentheses).

Emotion as a mediator of the time-reference effect. We next examine the effect of time reference on emotion and choice. For these analyses, the mediator and outcome variables were the *differences* between net emotion and choice across the two time-reference conditions (future minus past), separately for positive and negative surprises. Thus, these scores should be positive for the positive surprises, because future-oriented positive news (relative to past-oriented positive news) should lead to a stronger preponderance of approach over avoidance emotions and a greater choice propensity, whereas future-oriented *negative* news should lead to the opposite pattern. The mediation analysis allows us to test whether these time-reference-induced differences in emotion lead to the time-reference-induced differences in choice.

Figure 2 reveals a significant total effect of valence on the time-reference asymmetry (i.e., future – past difference scores) in choice, 95% CI[0.32,6.53], $p=.031$. This replicates the effect of time-reference on choice in Experiment 2. Given that the time-reference effect was of smaller magnitude than the valence effect, however, the model was less well-powered to distinguish between the indirect (mediated) and direct (unmediated) effects of valence on time-reference. Both the indirect and direct effects were marginal, 95% CI[–0.26,2.14], p=.108 and 95% CI[–0.34,5.34], p=.085, respectively. Given the significant total effect and marginal partial effects, we suspect that, as with the effect of valence, both the direct and indirect paths are in operation.

Discussion

These results broadly support our predictions. First, we directly replicated the effects of valence and time-reference on choice that were found in Experiment 2. Second, we demonstrated these effects on yet another dependent measure—the preponderance of approach over avoidance emotions. Finally, we showed these effects on emotion partially mediate the downstream effects on choice. Thus, these effects appear to have both cognitive and affective components, consistent with the idea that people construct narratives about financial securities (a largely cognitive process) and use those narratives to inform their choices (a process likely to be tinged with emotion).

General Discussion

People are natural story-tellers. Do these narrative instincts help people to make sense of financial data and to make economic choices? Three studies suggest an affirmative answer. Contradicting the predictions of a rational expectations theory, people predict much higher increases in share prices after positive rather than negative news (Experiment 1). Even though in reality such trends are small and reverse over time, participants actually believed that these trends would grow larger at longer time intervals. These differences were larger when information concerned predictions about the future rather than facts about the past, suggesting that people impose temporal order on news information. These effects of news valence and time-reference had downstream consequences for portfolio allocation choices (Experiment 2) and were driven partly by affect (Experiment 3). Parts C and D pf of the Supplementary Materials revealed that any effects of expertise are modest.

These studies are not, of course, without limitations. First, performance was not financially incentivized, suggesting some caution about generalizing these results to real-world behavior. Ameliorating the issue somewhat, we would point to evidence that incentives often do surprisingly little to eliminate decision-making biases, sometimes even exacerbating biases and rarely eliminating them (Camerer & Hogarth, 1999). Second, the information given to participants was impoverished relative to many real-world contexts, which typically provide, among other things, time-series information. This is useful for dissecting theory, but it is possible that these effects differ in richer information environments. Replicating these studies in a real-time, interactive trading environment with financial incentives would be a valuable contribution to experimental finance. Future research might also examine the interactive effect of time-series and news data, which have usually been studied separately in experiments.

Neoclassical microfoundations for economic behavior, such as rational expectations, have difficulty accounting for our results. Further, they appear to fall outside the scope of existing behavioral decision theories, such as prospect theory (Kahneman & Tversky, 1979). Although prospect theory and its extensions capture much about human behavior in contexts where possibilities are enumerable and their probabilities are known (such as gambles), they have less to say about situations of Knightian uncertainty in which such probabilities are elusive. These results support the idea that in such situations, people use narratives as their primary tool for making sense of information and making choices leading to action. This position is known as *conviction narrative theory* (CNT; Chong & Tuckett, 2014; Johnson et al., 2020; Tuckett & Nikolic, 2017).

According to CNT, individuals faced with Knightian uncertainty marshal available information to form a narrative—a causally and temporally structured mental representation that explains this information, generates predictions about the future, and motivates action. To construct these explanatory narratives, people draw upon prior beliefs and lay theories (e.g., Furnham, 1988), causal reasoning abilities (e.g., Lombrozo, 2016; Sloman, 2005), and trusted sources (e.g., Hovland & Weiss, 1951; Mills, 2013; Sperber et al., 2010). Because narratives are causally and temporally extended, they can be projected forward to imagine future events (Beach & Mitchell, 1987). And because narratives are affectively rich, they can generate approach and avoidance motivations that allow an individual to build sufficient conviction to maintain a sustained decision over time. Interview studies of professional money managers support these ideas (Tuckett, 2011), and we believe these processes capture the phenomenology of choice under Knightian uncertainty. Moreover, text-mining studies find that the preponderance of positive over negative emotion words in sources such as central bank documents can predict macroeconomic variables such as industrial production and GDP (Nyman et al., 2018).

However, both qualitative and econometric methods are less useful for establishing rigorous causal evidence about the mental processes that underlie these effects, limiting their use for testing theories of microfoundations. The current studies add to our ongoing efforts to experimentally test these microfoundations. When predicting the future value of a stock under uncertain states of the world, investors tend to focus on a single possible state and act as though it *is* certain—choosing a narrative and sticking with it (Johnson & Hill, 2017). Investors are sensitive to the explanations offered by managers and analysts for changes in share prices and earnings, suggesting that these explanations can offer the raw material for making narrative projections (Johnson et al., 2018b). Other work has begun to examine how people evaluate competing narratives. For example, rather than naively extrapolating past price changes into the future when forming price expectations, people use sophisticated techniques (albeit erroneous, from the perspective of financial theory) to match past price patterns to future predictions (Johnson et al., 2018a). Our work on the interplay between narratives, confidence, and emotions suggests that narratives direct attention to particular information, which influences our confidence and in turn our emotions (Bilovich et al., 2020). Finally, social influence plays a crucial role, as people use seemingly irrelevant cues, such as an expert's moral and political values, to assess which financial advisor to trust (Johnson, Rodrigues, & Tuckett, 2020). This may be one important link in understanding why particular stories "go viral" in particular social groups (Shiller, 2017). The current work adds yet further evidence to this growing empirical case in favor of CNT.

Beyond this mounting empirical case, we might consider the theoretical merits and demerits of CNT. One merit, compared with the rational and behaviorally-informed expectations accounts as well as with existing notions about causation in the forecasting literature, is that CNT accounts for how the relevant mental models got into people's heads (Szollosi & Newell, 2020). As a psychological theory, it provides greater detail about mental representations and processes and provides a stronger bridge to other known facts about the mind. On the other hand, its psychological realism can also be a weakness, when taken as an economic model, in that one might reasonably critique CNT for being insufficiently constrained. This is of course a common critique of behavioral approaches by more classically-oriented economists—mainstream economics provides a strikingly unified approach, whereas behavioral approaches are often haphazard, documenting particular anomalies in a comparatively unsystematic fashion. Indeed, we accept that CNT is unlikely to achieve the precision of neoclassical models. Yet, we think that CNT's links with basic psychological mechanisms strike a good balance between systematic theory and integrity to the real world. We expect that further advances, both theoretical and empirical, will allow our understanding of forecasting and decision-making to advance on both fronts.

References

Abbott, H. P. (2000). The evolutionary origins of the storied mind: Modeling the prehistory of narrative. *Narrative*, *8*, 247–256.

Adam, K. (2007). Experimental evidence on the persistence of output and inflation. *Economic Journal*, *117*(520), 603–636.

Adaval, R., & Wyer, R. S. (1998). The role of narratives in consumer information processing. *Journal of Consumer Psychology*, *7*, 207–245.

Akerlof, G. A., & Shiller, R. J. (2009). *Animal spirits.* Princeton, NJ: Princeton University Press.

Applebee, A. N. (1978). *The child's concept of story.* Chicago, IL: University of Chicago Press.

Barberis, N., Shleifer, A., & Vishny, R. (1997). A model of investor sentiment. *Journal of Financial Economics*, *49*, 307–344.

Bartlett, F. C. (1932). *Remembering.* Cambridge, UK: Cambridge University Press.

Batteux, E., Bilovich, A., Johnson, S. G. B., & Tuckett, D. (2020). *Impressed by numbers: The extent to which novice investors favor precise numerical information in a context of uncertainty.* Available at SSRN: https://papers.ssrn.com/abstract_id=3595409.

Batteux, E., Bilovich, A., Johnson, S. G. B., & Tuckett, D. (2021). *The negative consequences of failing to communicate uncertainties during a pandemic: the case of COVID-19 vaccines*. Preprint.

Beach, L. R. (2020). Scenarios as narratives. *Futures & Foresight Science.* Early view.

Beach, L. R., & Mitchell, T. R. (1987). Image theory: Principles, goals, and plans. *Acta Psychologica*, *66*, 201–220.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, *57*, 289–300.

Bernard, V. (1992). Stock price reactions to earnings announcements: A summary of recent anomalous evidence and possible explanations. In R. Thaler (Ed.), *Advances in behavioral finance* (pp. 303–340). New York, NY: Russell Sage Foundation.

Bernard, V., & Thomas, J. (1989). Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research*, *27*, 1–36.

Bernard, V., & Thomas, J. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, *13*, 305–340.

Bilovich, A., Johnson, S. G. B., & Tuckett, D. (2020). *Constructing confidence: The role of affect and cognition in decisions under radical uncertainty.* Available at PsyArXiv: https://psyarxiv.com/tacpn/.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.

Brealey, R., Myers, S., & Allen, F. (2013). *Principles of corporate finance* (11th Ed.). New York, NY: McGraw-Hill.

Brown, D. (1991). *Human universals.* New York, NY: McGraw-Hill.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Burns, Z. C., Caruso, E. M., & Bartels, D. M. (2012). Predicting premeditation: Future behavior is seen as more intentional than past behavior. *Journal of Experimental Psychology: General*, *141*, 227–232.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*, 7–42.

Caruso, E. M. (2010). When the future feels worse than the past: A temporal inconsistency in moral judgment. *Journal of Experimental Psychology: General*, *139*, 610–624.

Caruso, E. M., Gilbert, D. T., & Wilson, T. D. (2008). A wrinkle in time: Asymmetric valuation of past and future events. *Psychological Science*, *19*, 796–801.

Carver, C. S. (2006). Approach, avoidance, and self-regulation of affect and action. *Motivation & Emotion*, *30*, 105–110.

Chan, L. K. C., Jegadeesh, N.,& Lakonishok, J. (1996). Momentum strategies. *Journal of Finance*, *51*, 1681–1713.

Chater, N., & Loewenstein, G. (2016). The under-appreciated drive for sense-making. *Journal of Economic Behavior & Organization*, *126*, 137–154.

Chen, M. K., Lakshminarayanan, V., & Santos L. R. (2006). How basic are behavioral biases? Evidence from capuchin monkey trading behavior. *Journal of Political Economy*, *114*, 517–537.

Chen, S. Y., Ross, B. H., & Murphy, G. L. (2014). Implicit and explicit processes in category-based induction: Is induction best when we don't think? *Journal of Experimental Psychology: General*, *143*, 227–246.

Chong, K., & Tuckett, D. (2014). Constructing conviction through action and narrative: How money managers manage uncertainty and the consequence for financial market functioning. *Socio-Economic Review*, *13*, 309–330.

Chopra, N., Lakonishok, J., & Ritter, J. R. (1992). Measuring abnormal performance: Do stocks overreact?

Copeland T. E., & Friedman D. (1987). The effect of sequential information arrival on asset prices: An experimental study. *Journal of Finance*, *42*, 763–797.

Cosmides L., Tooby J. (1992) Cognitive adaptations for social exchange. In J. H. Barkow (Ed.), ed. *The adapted mind* (pp. 163–228). New York, NY: Oxford University Press.

Cutler, D. M., Poterba, J. M., & Summers, L. H. (1990). Speculative dynamics. *Review of Economic Studies*, *58*, 529–546.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). A theory of overconfidence, self-attribution, and security market under- and over-reactions. *Journal of Finance*, *53*, 1839–1886.

Davidson, P. (1982). Rational expectations: A fallacious foundation for studying crucial decision-making processes. *Journal of Post Keynesian Economics*, *5*, 182–198.

De Bondt, W. F. M. (1993). Betting on trends: Intuitive forecasts of financial risk and return. *International Journal of Forecasting*, *9*, 355–371.

De Bondt, W. F. M., & Thaler, R. H. (1985). Does the stock market overreact? *Journal of Finance*, *40*, 793–808.

Escalas, J. E. (2004). Narrative processing: Building consumer connections to brands. *Journal of Consumer Psychology*, *14*, 168–180.

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. New York, NY: Psychology Press.

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*, 34–105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*, 383–417.

Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, *18*, 25–46.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research*, *50*, 123–129.

Fildes, R., Goodwin, P., & Önkal, D. (2019). Use and misuse of information in supply chain forecasting of promotion effects. *International Journal of Forecasting*, *35*, 144–156.

Friedman, M. (1957). *A theory of the consumption function*. Princeton, NJ: Princeton University Press.

Furnham, A. (1989). *Lay theories*. Amsterdam, Netherlands: Elsevier.

Gönül, S., Önkal, D., & Goodwin, P. (2009). Expectations, use and judgmental adjustment of external financial and economic forecasts: An empirical investigation. *Journal of Forecasting*, *28*, 19–37.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395.

Gopnik, A. (1998). Explanation as orgasm. *Minds & Machines*, *8*, 101–118.

Gottschall, J. (2012). *The storytelling animal*. New York, NY: Houghton Mifflin.

Haltiwanger, J., & Waldman, M. (1985). Rational expectations and the limits of rationality: An analysis of heterogeneity. *American Economic Review*, *75*(3), 326–340.

Harvey, N., Bolger, F., & McClelland, A. (1994). On the nature of expectations. *British Journal of Psychology*, *85*, 203–229.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*, 280–285.

Hertwig, R., & Gigerenzer, G. (1999). The 'conjunction fallacy' revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, *12*, 275–305.

Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, *118*, 248–271.

Hogan, P. C. (2003). *The mind and its stories*. Cambridge, UK: Cambridge University Press.

Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control*, *35*(1), 1–24.

Hommes, C., Sonnemans, J., Tuinstra, J., & van de Velden, H. (2008). Expectations and bubbles in asset pricing experiments. *Journal of Economic Behavior & Organization*, *67*, 116–133.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, *54*, 2143–2184.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, *15*, 635–650.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, *48*, 65–91.

Jensen, M. (1968). The performance of mutual funds in the period 1945–64. *Journal of Finance*, *23*, 389–416.

Johnson, S. G. B. (2017). *Cognition as sense-making*. Unpublished doctoral dissertation.

Johnson, S. G. B., & Ahn, W. (2015). Causal networks or causal islands? The representation of mechanisms and the transitivity of causal judgment. *Cognitive Science, 39*, 1468–1503.

Johnson, S. G. B., Bilovich, A., & Tuckett, D. (2020). *Conviction narrative theory: A theory of choice under radical uncertainty*. Available at PsyArXiv: https://psyarxiv.com/urc96/.

Johnson, S. G. B., & Hill, F. (2017). Belief digitization in economic prediction. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 2314–2319).

Johnson, S. G. B., Matiashvili, T., & Tuckett, D. (2019a). *Expectations based on past price patterns: An experimental study*. Working paper.

Johnson, S. G. B., Matiashvili, T., & Tuckett, D. (2019b). *How investors use explanations of price and earnings changes to predict future value.* Working paper.

Johnson, S. G. B., Merchant, T., & Keil, F.C. (in press). Belief digitization: Do we treat uncertainty as probabilities or as bits? *Journal of Experimental Psychology: General.*

Johnson, S. G. B., Murphy, G. L., Rodrigues, M., & Keil, F. C. (2019). Predictions from uncertain moral character. In A.K. Goel, C.M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 506–512). Austin, TX: Cognitive Science Society.

Johnson, S. G. B., Rodrigues, M., & Tuckett, D. (2020). Moral tribalism and its discontents: How intuitive theories of ethics shape consumers' deference to experts. *Journal of Behavioral Decision Making*. Advance online publication.

Johnson, S. G. B., & Steinerberger, S. (2019). Intuitions about mathematical beauty: A case study in the aesthetic experience of ideas. *Cognition*, *189*, 242–259.

Johnson, S. G. B., Zhang, J., & Keil, F. C. (2020). *Win–win denial: The psychological underpinnings of zero-sum thinking.* Available at PsyArXiv: https://psyarxiv.com/efs5y.

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subjects designs. *Psychological Methods*, *6*, 115–134.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–292.

Keynes, J. M. (1936). *The general theory of employment, interest, and money*. London, UK: Macmillan.

Knight, F. H. (1921). *Risk, uncertainty, and profit*. Boston, MA: Hart, Schaffner, & Marx.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin, 110*, 499–519.

Lachmann, L. M. (1943). The role of expectations in economics as a social science. *Economica*, *10*, 12–23.

Lawrence, M. J., Edmundson, R. H., & O'Connor, M. J. (1985). An examination of the accuracy of judgmental extrapolation of time series. *International Journal of Forecasting*, *1*, 25–35.

Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, *22*, 493–518.

Lim, J. S., & O'Connor, M. (1996). Judgmental forecasting with time series and causal information. *International Journal of Forecasting*, *12*, 139–153.

Lintner, J. (1965). The valuation of risk assets and the selection of investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, *47*, 13–37.

Lipton, P. (2004). *Inference to the best explanation* (2nd Ed.). New York, NY: Routledge.

List, J. A. (2003). Does market experience eliminate market anomalies? *Quarterly Journal of Economics*, *118*, 41–71.

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

Lovell, M. C. (1986). Tests of the rational expectations hypothesis. *American Economic Review*, *76*(1), 110–124.

Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences*, *20*, 748–759.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, *4*, 103–124.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*, 111–151.

Mandler, J. M., Scribner, S., Cole, M., & DeForest, M. (1980). Cross-cultural invariance in story recall. *Child Development*, *51*, 19–26.

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*, 173–192.

McAdams, D. P. (1993). *The stories we live by.* New York, NY: William Morrow & Company.

McCarthy, J. (2016, April 20). Just over half of Americans own stocks, matching record low. *Gallup*. Retrieved from http://www.gallup.com.

Miller, M. H., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *The Journal of Business*, *34*, 411–433.

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology*, *49*, 404–418.

Mises, L. v. (1949). *Human action*. New Haven, CT: Yale University Press.

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*, 6–27.

Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorizations. *Cognitive Psychology*, *27*, 148–193.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, *29*, 315–35.

National Center for Education Statistics (2016). Bachelor's degrees conferred by postsecondary institutions, by field of study. *Digest of Education Statistics*. Retrieved from http://nces.ed.gov.

Nyman, R., Kapadia, S., Tuckett, D., Gregory, D., Ormerod, P., & Smith, R. (2018). News and narratives in financial systems: Exploiting big data for systemic risk assessment. *Bank of England Working Paper 704*.
Oatley, K., & Johnson-Laird, P. N. (1986). Toward a cognitive theory of emotions. *Cognition and Emotion*, *1,* 29–50.
Önkal, D., Sayim, K. Z., & Gönül, M. S. (2013). Scenarios as channels of forecast advice. *Technological Forecasting & Social Change*, *80*, 772–788.
Pennington, N., & Hastie, R. (1992). Explaining the evidence: Testing the Story Model for juror decision making. *Journal of Personality and Social Psychology*, *62*, 189–206.
Plott, C. R., & Sunder, S. (1982). Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of Political Economy*, *90*(4), 663–698.
Rumelhart, D. E. (1975). Notes on a scheme for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding*. New York, NY: Academic Press.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*, 425–442.
Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.
Sher, S., & McKenzie, C. R. (2006). Information leakage from logically equivalent frames. *Cognition*, *101*, 467–494.
Shefrin, H. (2002). *Beyond greed and fear* (2nd Ed.). New York, NY: Oxford University Press.
Shiller, R. J. (2000). *Irrational exuberance*. Princeton, NJ: Princeton University Press.
Shiller, R. J. (2017). Narrative economics. *American Economic Review*, *107*, 967–1004.
Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *Journal of Economic Perspectives*, *4*, 19–33.
Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *Journal of Finance*, *52*, 35–55.
Simon, H. A. (1979). Rational decision making in business organizations. *American Economic Review*, *69*(4), 493–513.
Sloman, S. (2005). *Causal models*. New York, NY: Oxford University Press.
Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*. London, UK: Penguin.
Smith, D. et al. (2017). Cooperation and the evolution of hunter–gatherer storytelling. *Nature Communications*, *8*, 1853.
Smith, V., Suchanek, G. L., & Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica*, *56*(5), 1119–1151.
Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H, Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*, 359–393.
Stein, J. C. (1989). Overreactions in the options market. *Journal of Finance*, *44*, 1011–1023.
Szollosi, A., & Newell, B. (2020). People as intuitive scientists: Reconsidering statistical explanations of decision making. *Trends in Cognitive Sciences*, *24*, 1008–1018.
Taleb, N. N. (2001). *Fooled by randomness*. New York, NY: Random House.
Taleb, N. N. (2007). *The black swan*. New York, NY: Random House.
Tuckett, D. (2011). *Minding the markets*. London, UK: Palgrave Macmillan.
Tuckett, D. (2012). Financial markets are markets in stories: Some possible advantages of using interviews to supplement existing economic data sources. *Journal of Economic Dynamics and Control*, *36*, 1077–1087.
Tuckett, D., & Nikolic, M. (2017). The role of conviction and narrative in decision making under radical uncertainty. *Theory & Psychology*, *27*, 501–523.
Van Boven, L., & Ashworth, L. (2007). Looking forward, looking back: Anticipation is more evocative than retrospection. *Journal of Experimental Psychology: General*, *136*, 289–300.

Van Laer, T., De Ruyter, K., Visconti, L. M., & Wetzels, M. (2014). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research*, *40*, 797–817.

Wermers, R. (2011). Performance measurement of mutual funds, hedge funds and institutional accounts. *Annual Review of Financial Economics*, *3*, 537–574.

Yates, J. F., McDaniel, L. S., & Brown, E. S. (1991). Probabilistic forecasts of stock prices and earnings: The hazards of nascent expertise. *Organizational Behavior and Human Decision Processes*, *49*, 60–79.

SUPPLEMENTARY MATERIALS

Narrative Expectations in Financial Forecasting

Samuel G. B. Johnson & David Tuckett

Part A: Materials for Experiment 1

Instructions

On the following pages, you will read some information about the stock market and answer some questions. There will be some additional questions about this information at the end of the study, so please read it carefully.

There will be information about four different stocks to read about, and for each stock, this information will be displayed over two separate pages with questions.

Please click on the arrow to continue.

Item 1

Remlon Software Corporation (stock symbol RWQ) is a Dallas-based company that designs and markets business software to medium- and large-size firms.

Here is the most recent price quotation for shares in RWQ stock: $56.00

Given that RWQ shares currently trade at $56, please estimate what you think the share price will be on the following dates: [*Tomorrow / In two weeks / In one year*] [*Scale from $28.00 to $84.00*]

– Page break –

About an hour prior to the most recent price quotation ($56) for Remlon's stock (RWQ), the following piece of news was revealed:

[Negative/Past] Although average sales growth is expected for the next quarter, analysts determined that Remlon experienced <u>below-average levels of sales growth over the past quarter</u>.

[Negative/Future] Although average sales growth was observed for the past quarter, analysts anticipate that Remlon will experience <u>below-average levels of sales growth over the next quarter</u>.

[Positive/Past] Although average sales growth is expected for the next quarter, analysts determined that Remlon experienced <u>above-average levels of sales growth over the past quarter</u>.

[Positive/Future] Although average sales growth was observed for the past quarter, analysts anticipate that Remlon will experience <u>above-average levels of sales growth over the next quarter</u>.

Given that RWQ shares currently trade at $56, please estimate again what you think the share price will be on the following dates: [*Tomorrow / In two weeks / In one year*] [*Scale from $28.00 to $84.00*]

Item 2

Wilfinger Industries (stock symbol WNV) is a Charlotte-based company that produces telecommunications equipment.

Here are the most recent price quotations for shares in WNV stock: $32.00

Given that WNV shares currently trade at $32, please estimate what you think the share price will be on the following dates: [*Tomorrow / In two weeks / In one year*] [*Scale from $16.00 to $48.00*]

– Page break –

About an hour prior to the most recent price quotation ($32) for Wilfinger's stock (WNV), the following piece of news was revealed:

[Negative/Past] Although average levels of innovation are expected for the next quarter, analysts determined that Wilfinger experienced <u>below-average levels of innovation over the past quarter</u>.

[Negative/Future] Although average levels of innovation were observed for the past quarter, analysts anticipate that Wilfinger will experience <u>below-average levels of innovation over the next quarter</u>.

[Positive/Past] Although average levels of innovation are expected for the next quarter, analysts determined that Wilfinger experienced <u>above-average levels of innovation over the past quarter</u>.

[Positive/Future] Although average levels of innovation were observed for the past quarter, analysts anticipate that Wilfinger will experience <u>above-average levels of innovation over the next quarter</u>.

Given that WNV shares currently trade at $32, please estimate again what you think the share price will be on the following dates: [*Tomorrow / In two weeks / In one year*] [*Scale from $16.00 to $48.00*]

Item 3

Paravoz Exploration (stock symbol PVZ) is a Denver-based company that explores for mineral deposits.

Here are the most recent price quotations for shares in PVZ stock: $64.00

Given that PVZ shares currently trade at $64, please estimate what you think the share price will be on the following dates: [*Tomorrow / In two weeks / In one year*] [*Scale from $32.00 to $96.00*]

– Page break –

About an hour prior to the most recent price quotation ($64) for Paravoz's stock (PVZ), the following piece of news was revealed:

[Negative/Past] Although typical mineral discoveries are expected for the next quarter, analysts determined that Paravoz discovered <u>less-than-typical levels of new mineral deposits over the past quarter</u>.

[Negative/Future] Although typical mineral discoveries were observed for the past quarter, analysts anticipate that Paravoz will discover <u>less-than-typical levels of new mineral deposits over the next quarter</u>.

[Positive/Past] Although typical mineral discoveries are expected for the next quarter, analysts determined that Paravoz discovered <u>more-than-typical levels of new mineral deposits over the past quarter</u>.

[Positive/Future] Although typical mineral discoveries were observed for the past quarter, analysts anticipate that Paravoz will discover <u>more-than-typical levels of new mineral deposits over the next quarter</u>.

Item 4

Excellerate Construction (stock symbol XOL) is a Chicago-based company that offers contracting services for large commercial real estate projects.

Here are the most recent price quotations for shares in XOL stock: $40.00

Given that XOL shares currently trade at $40, please estimate what you think the share price will be on the following dates: [*Tomorrow / In two weeks / In one year*] [*Scale from $20.00 to $60.00*]

– Page break –

About an hour prior to the most recent price quotation ($40) for Excellerate's stock (XOL), the following piece of news was revealed:

[Negative/Past] Although a typical number of new contract signings are expected for the next quarter, analysts determined that Excellerate signed a <u>less-than-typical number of new contracts over the past quarter</u>.

[Negative/Future] Although a typical number of new contract signings were observed for the past quarter, analysts anticipate that Excellerate will sign a <u>less-than-typical number of new contracts over the next quarter</u>.

[Positive/Past] Although a typical number of new contract signings are expected for the next quarter, analysts determined that Excellerate signed a <u>more-than-typical number of new contracts over the past quarter</u>.

[Positive/Future] Although a typical number of new contract signings were observed for the past quarter, analysts anticipate that Excellerate will sign a <u>more-than-typical number of new contracts over the next quarter</u>.

Counterbalancing

These materials were counterbalanced using a Latin square, such that participants saw one of the four combinations of items and conditions:

	Item			
	1	**2**	**3**	**4**
Group 1	Neg/Past	Pos/Fut	Pos/Past	Neg/Fut
Group 2	Neg/Fut	Neg/Past	Pos/Fut	Pos/Past
Group 3	Pos/Past	Neg/Fut	Neg/Past	Pos/Fut
Group 4	Pos/Fut	Pos/Past	Neg/Fut	Neg/Past

Part B: Mixed Effects Models (Robustness Checks)

We fit a series of mixed effects models to supplement the analyses in the main text. These had two main purposes.

First, they allowed us to model both participants and items as random effects. Although items differed in several respects (e.g., industry, initial price, and conditions [sales growth, new contracts, etc.]), these were not systematically varied and we cannot tease apart their separate effects in our models due to confounding. We can, however, treat item as a random effect to permit generalization across these aspects of the stimuli.

Second, we used these models to conduct robustness checks. We repeat all models using three different subsamples—excluding only inattentive individuals (as in the main text), excluding inattentive individuals plus outliers, and the full sample. These generally give similar results. In addition, for Study 1, we test models that use baseline forecasts as a covariate (i.e., testing forecast change over baseline rather than absolute forecasts) and with prior time-lagged forecasts as covariates (i.e., testing forecast change over the previous period).

In Part C, we fit a set of analogous models to quantify effects of expertise, by including participants' responses to objective questions about financial assets, finance courses, and college major, as well as self-reported judgments of financial knowledge and financial experience.

Experiment 1

We report a series of mixed effects regressions for Experiment 1. In all models, the unit of observation is the forecast on each trial of the experiment. Random intercepts are fit for each participant and item.

Table B1 summarizes model fits for estimating the effect of news valence condition. In Model 1, we enter a fixed factor for the independent variable, *Valence Condition* (contrast-coded; –1 = negative news, 1 = positive news). In Model 2, we add the baseline prediction (made before receiving the news), to test the added effect of news over-and-above the initial forecast. That is, for modeling 1-day forecasts, we add the *1-day Baseline*, for modeling 2-week forecasts, we add the *2-week Baseline*, and for modeling 1-year forecasts, we add the *1-year Baseline*. In Model 3, we add the previous period's forecast (made after receiving the news) when modeling 2-week forecasts (entering the *1-Day Forecast*) or 1-year forecasts (entering the *2-Week Forecast*). Finally, Models 4 and 5 re-fit Model 1 on datasets using alternative exclusion criteria, as a further robustness check. In Model 4, in addition to excluding

Sample	**Ex. Inattentive Participants (*N* = 185)**			**Ex. Outliers (*N* = 139)**	**Full (*N* = 225)**
Model	**1**	**2**	**3**	**4**	**5**
			DV: 1-Day Forecasts		
Valence	.044 (.003)***	.044 (.003)***		.032 (.002)***	.041 (.003)***
Baseline Forecast		.186 (.057)**			
Intercept	.016 (.003)***	.013 (.003)***		.009 (.002)*	.017 (.003)***
R^2	.261	.271		.429	.205
			DV: 2-Week Forecasts		
Valence	.074 (.003)***	.074 (.003)***	.033 (.002)***	.058 (.002)***	.071 (.003)***
Baseline Forecast		.213 (.050)***			
1-Day Forecast			.924 (.027)***		
Intercept	.029 (.003)**	.020 (.004)**	.014 (.004)*	.023 (.003)**	.029 (.003)***
R^2	.415	.431	.772	.553	.336
			DV: 1-Year Forecasts		
Valence	.110 (.005)***	.110 (.005)***	.022 (.004)***	.095 (.004)***	.106 (.005)***
Baseline Forecast		.427 (.039)***			
2-Week Forecast			1.19 (.038)***		
Intercept	.051 (.006)***	.014 (.006)*	.017 (.005)***	.048 (.005)***	.055 (.006)***
R^2	.367	.454	.698	.515	.318

Table B1. Effect of valence on forecasts in Experiment 1.

Note. Entries are unstandardized coefficients (and SEs) for each model. The R^2 values refer to the marginal variance explained by the fixed effects. Sample sizes (*N*) refer to the number of participants; thus, each model contains 4*N* (trial-level) observations.

Sample	Ex. Inattentive Participants (N = 185)			Ex. Outliers (N = 139)	Full (N = 225)
Model	**1**	**2**	**3**	**4**	**5**
			Positive Valence Condition		
			DV: 1-Day Forecasts		
Time-Reference	.005 (.003)°	.004 (.003)		.003 (.002)	.007 (.003)*
Baseline Forecast		.226 (.065)***			
Intercept	.060 (.005)***	.057 (.005)***		.041 (.003)***	.057 (.005)***
R^2	.004	.026		.005	.006
			DV: 2-Week Forecasts		
Time-Reference	.009 (.003)**	.007 (.003)**	.005 (.002)*	.007 (.002)**	.009 (.003)**
Baseline Forecast		.043 (.060)***			
1-Day Forecast			.881 (.038)***		
Intercept	.103 (.007)***	.084 (.008)***	.050 (.005)***	.080 (.006)***	.099 (.006)***
R^2	.010	.115	.600	.013	.008
			DV: 1-Year Forecasts		
Time-Reference	.014 (.004)**	.012 (.004)**	.004 (.003)	.013 (.004)**	.014 (.004)**
Baseline Forecast		.055 (.005)***			
2-Week Forecast			1.12 (.006)***		
Intercept	.161 (.012)***	.114 (.012)***	.047 (.010)***	.143 (.009)***	.161 (.010)***
R^2	.008	.222	.471	.016	.008
			Negative Valence Condition		
			DV: 1-Day Forecasts		
Time-Reference	–.000 (.002)	–.001 (.002)		–.001 (.001)	–.000 (.003)
Baseline Forecast		–.091 (.066)			
Intercept	–.028 (.005)***	–.026 (.005)***		–.023 (.002)***	–.024 (.004)***
R^2	.000	.003		.000	.000
			DV: 2-Week Forecasts		
Time-Reference	.001 (.003)	.001 (.003)	.002 (.002)	.000 (.002)	.002 (.003)
Baseline Forecast		–.017 (.058)			
1-Day Forecast			.874 (.040)***		
Intercept	–.046 (.008)**	–.045 (.009)**	–.021 (.006)*	–.035 (.004)***	–.041 (.006)***
R^2	.000	.000	.571	.000	.000
			DV: 1-Year Forecasts		
Time-Reference	–.002 (.005)	–.002 (.005)	–.003 (.004)	–.007 (.003)°	–.000 (.005)
Baseline Forecast		.268 (.054)***			
2-Week Forecast			1.22 (.057)***		
Intercept	–.059 (.012)**	–.083 (.012)***	–.004 (.006)	–.046 (.008)***	–.051 (.010)***
R^2	.000	.062	.565	.006	.000

Table B2. Effect of time-reference on forecasts in Experiment 1.

Note. Entries are unstandardized coefficients (and SEs). R^2 values refer to the marginal variance explained by fixed effects. Sample sizes (N) refer to the number of participants; thus, each model contains 2N (trial-level) observations.

inattentive participants, any participant was excluded whose response on any trial was an outlier within that condition, as assessed by Tukey's boxplot method. In Model 5, the full sample was used with no exclusions. As shown in Table B1, the effect of valence condition on forecasts was large and statistically robust at all timescales and across all model specifications.

Table B2 summarizes model fits for the effect of time-reference condition on forecasts. These models are constructed in the same way as those in Table B1, with two exceptions. First, time-reference condition (contrast-coded; –1 = past; 1 = future) was used as the independent variable instead of valence. Second, we fit separate models for the positive and negative valence conditions, as time-reference was predicted to have different effects across conditions (a positive effect for positive valence and a negative effect for negative valence), and it is possible for one or the other or both of these predictions to hold . These models were analogous to those in Table B1, except that time-reference rather than valence condition was the key independent variable.

As shown in Table B2, the results reported in the main text are generally consistent with the alternative specifications and exclusion criteria tested here. For positively valenced news (Table B2), future time-reference led to significantly increased forecasts relative to past time-reference, with this effect reaching at least marginal significance in 11 out of 14 models. Consistent with the main text, this effect was larger and more robust at longer time-intervals. For negatively valenced news (Table B3), the effects were consistently weak, rarely reached even marginal statistical significance, and were sometimes in the wrong direction (the signs should be negative). Thus, reflecting the analyses reported in the main text, the time-reference effect occurred only for positive news in Experiment 1.

Experiment 2

We fit a similar series of models for testing the effects on choices in Experiment 2, except models with baseline forecasts or lags were not included since we only measured choices at one time point in Experiment 2. Like Experiment 1, these models included random intercepts for participants and items. We fit all models with the same sets of exclusion criteria used above—excluding inattentive participants only (Model 1, as in the main text), excluding outliers as well (Model 2), and the full sample with no exclusions (Model 3).

Table B3 shows model fits for the valence effect. Reflecting the findings in the main text, there was a very large effect of valence regardless of the exclusion criterion used.

Sample	**Ex. Inattentive Participants (*N* = 174)**	**Ex. Outliers (*N* = 124)**	**Full (*N* = 225)**
Model	**1**	**2**	**3**
Valence	19.01 (0.70)***	18.59 (0.68)***	18.64 (0.62)***
Intercept	45.42 (1.16)	48.65 (0.88)	46.01 (0.99)
R^2	.440	.562	.424

Table B3. Effect of valence on choices in Experiment 2.

Note. Entries are unstandardized coefficients (and SEs) for each model. The R^2 values refer to the marginal variance explained by the fixed effects. Sample sizes (*N*) refer to the number of participants; thus, each model contains 4*N* (trial-level) observations.

Table B4 shows model fits for the time-reference effect, separately for the positive and negative valence conditions, analogously. As in the main text, there is a positive effect of future-oriented (vs. past-oriented) news when the news was positive, and a negative effect of future-oriented news when the news was negative. These effects were statistically robust in both the positive and negative valence conditions across all exclusion criteria tested.

Sample	Ex. Inattentive Participants (N = 185)	Ex. Outliers (N = 139)	Full (N = 225)
Model	**1**	**2**	**3**
		Positive Valence Condition	
Time-Reference	2.02 (0.78)*	2.10 (0.81)*	1.64 (0.68)*
Intercept	64.42 (1.75)	67.19 (1.39)	64.65 (1.46)
R^2	.008	.016	.006
		Negative Valence Condition	
Time-Reference	–3.38 (0.74)***	–2.58 (0.72)***	–3.02 (0.67)***
Intercept	26.38 (1.64)	30.03 (1.37)	27.38 (1.42)
R^2	.027	.024	.020

Table B4. Effect of time-reference on choices in Experiment 2.
Note. Entries are unstandardized coefficients (and SEs) for each model. The R^2 values refer to the marginal variance explained by the fixed effects. Sample sizes (N) refer to the number of participants; thus, each model contains 2N (trial-level) observations.

Experiment 3

For Experiment 3, we fit a series of models for testing the effects on choices and on net emotions (the number of positive emotion words minus the number of negative emotion words), analogous to those used to analyze Experiment 2. These models included random intercepts for participants and items, and were fit with the same sets of exclusion criteria as the models above—excluding inattentive participants only (Model 1, as in the main text), excluding outliers as well (Model 2), and the full sample with no exclusions (Model 3).

As shown in Table B5, the effects of valence on both choices and net emotions were very large across all exclusion criteria, consistent with the highly statistically significant mediation results reported in the main text.

Table B6 reports model fits for the time-reference effect, separately for the positive and negative valence condition, on choices and net emotions. All results are directionally consistent with the key predictions, but vary in robustness. The effect of time-reference on choices is only statistically significant in the positive valence condition, and only reaches the .01 significance level when outliers are excluded from the analysis. On the other hand, the effect of time-reference on net emotions is only statistically significant in the negative valence condition, although this effect is consistent across different exclusion criteria. These results are broadly consistent with the mediation results reported in the main text, which test (i) whether the effect of valence on choices is larger in the future than in the past time-reference condition and (ii) whether these effects are related to differences in the impact of valence on net emotions across time-reference conditions. Since the effect of time-reference on both choices and emotions were significant in either the positive or negative valence condition (and directionally similar in both cases), this is compatible with the mediation results reported in the main text. The modest effect sizes and moderate levels of statistical significance correspond to the difficulty of the main text mediation model in separating out the direct pathway from the indirect pathway via emotions, which were both marginally significant despite the significant total effect on choice.

Sample	Ex. Inattentive Participants (N = 185)	Ex. Outliers (N = 139)	Full (N = 225)
Model	**1**	**2**	**3**
	DV: Choices		
Valence	18.52 (0.63)***	17.62 (0.66)***	17.11 (0.59)***
Intercept	48.70 (1.31)	51.46 (0.91)	48.39 (1.07)
R^2	.460	.560	.394
	DV: Net Emotions (Positive – Negative)		
Valence	5.02 (0.17)***	4.93 (0.19)***	4.82 (0.16)***
Intercept	–0.56 (0.47)	–0.45 (0.41)	–0.41 (0.43)
R^2	.550	.578	.506

Table B5. Effect of valence on choices and emotions in Experiment 3.
Note. Entries are unstandardized coefficients (and SEs) for each model. The R^2 values refer to the marginal variance explained by the fixed effects. Sample sizes (N) refer to the number of participants; thus, each model contains 4N (trial-level) observations.

Sample	Ex. Inattentive Participants (N = 185)	Ex. Outliers (N = 139)	Full (N = 225)
Model	**1**	**2**	**3**
	Positive Valence Condition		
	DV: Choices		
Time-Reference	1.11 (0.61)°	1.93 (0.65)**	0.75 (0.55)
Intercept	67.20 (1.82)	69.08 (1.00)	65.50 (1.75)
R^2	.003	.023	.001
	DV: Net Emotions (Positive – Negative)		
Time-Reference	0.08 (0.17)	0.14 (0.16)	0.08 (0.15)
Intercept	4.45 (0.51)	4.47 (0.52)	4.41 (0.48)
R^2	.000	.001	.000
	Negative Valence Condition		
	DV: Choices		
Time-Reference	–0.61 (0.56)	–0.48 (0.64)	–0.84 (0.52)
Intercept	30.17 (1.48)	33.84 (1.57)	31.29 (1.34)
R^2	.000	.000	.002
	DV: Net Emotions (Positive – Negative)		
Time-Reference	–0.26 (0.12)*	–0.26 (0.13)*	–0.28 (0.11)*
Intercept	–5.57 (0.44)	–5.37 (0.47)	–5.23 (0.47)
R^2	.004	.004	.003

Table B6. Effect of time-reference on choices and emotions in Experiment 3.
Note. Entries are unstandardized coefficients (and SEs). R^2 values refer to the marginal variance explained by fixed effects. Sample sizes (N) refer to the number of participants; thus, each model contains 2N (trial-level) observations.

Part C: Effects of Expertise in Experiments 1–3

Across all experiments, we collected several measures of financial expertise. These included objective questions about financial assets, courses, and major, and self-reported measures of financial knowledge, from 0 ("Very little knowledge") to 10 ("Extensive knowledge"), and investment experience, from 0 ("Very little experience") to 10 ("Extensive experience"). As mentioned in the main text, participants were relatively experienced compared to traditional (e.g., undergraduate) samples. For instance, in Experiment 1, 49% of participants held some financial assets, 53% had taken at least one finance course, and 14% had majored in a business field. Responses on the self-report scales were somewhat lower than the scale midpoint (M = 4.04 for knowledge and M = 3.05 for experience), with wide variance as befitting the range of answers to our objective questions (SD = 2.51 for knowledge and SD = 2.55 for experience).

Here, we test how these expertise measures interact with our primary effects of interest (of valence and time-reference). To do so, we fit a series of multilevel models using a similar approach to Part B. The unit of observation was each trial of each experiment, and random intercepts were fit for each participant and item.

Table C1 shows the results of models that include valence condition (contrast-coded as in Part B), the self-report expertise measures (experience and knowledge; mean-centered and scaled), the objective expertise measures (courses, assets, and major; dummy-coded), and the interaction between valence condition and each of the five expertise measures. Multicollinearity diagnostics were acceptable for all models (VIFs < 5) in Tables C1 and C2.

Experiment	**Experiment 1 (*N* = 179)**			**Experiment 2 (*N* = 171)**	**Experiment 3 (*N* = 174)**	
DV	**1-Day Forecasts**	**2-Week Forecasts**	**1-Year Forecasts**	**Choices**	**Choices**	**Net Emotions**
Valence	.046 (.006)***	.078 (.007)***	.118 (.011)***	20.47 (1.53)***	17.39 (1.36)***	4.45 (0.36)***
Experience	–.007 (.005)	–.006 (.006)	.009 (.011)	–1.17 (1.66)	0.95 (1.76)	0.50 (0.28)º
Knowledge	.001 (.005)	.000 (.006)	–.021 (.010)*	0.68 (1.68)	0.44 (1.64)	–0.00 (0.26)
Courses	.010 (.006)	.003 (.007)	.006 (.013)	1.04 (2.46)	0.93 (2.54)	–0.17 (0.41)
Assets	.001 (.007)	–.005 (.008)	.001 (.014)	–3.77 (2.66)	–0.88 (2.63)	–0.35 (0.42)
Major	.001 (.008)	–.010 (.010)	.000 (.017)	–0.34 (2.84)	–0.95 (2.83)	–0.14 (0.45)
Valence × Experience	–.014 (.005)**	–.018 (.006)**	–.019 (.010)º	–0.77 (1.07)	–2.93 (1.05)**	–0.55 (0.28)*
Valence × Knowledge	.015 (.005)**	.020 (.006)***	.022 (.009)*	0.40 (1.08)	1.57 (0.98)	0.33 (0.26)
Valence × Courses	–.020 (.006)**	–.015 (.007)*	–.013 (.012)	0.20 (1.59)	4.14 (1.53)**	1.11 (0.41)**
Valence × Assets	.019 (.007)**	.008 (.008)	–.001 (.010)	–1.08 (1.72)	–0.11 (1.58)	0.45 (0.42)
Valence × Major	–.007 (.008)	.000 (.010)	–.005 (.015)	–4.50 (1.84)*	–3.23 (1.70)º	–0.92 (0.45)*
Intercept	.010 (.006)º	.031 (.007)***	.048 (.012)***	47.01 (2.36)	48.77 (2.41)	–0.24 (0.59)
R^2	.287	.429	.375	.445	.472	.565

Table C1. Effects of Valence × Expertise in Experiments 1–3.

Note. Entries are unstandardized coefficients (and SEs) for each model. Valence was contrast-coded as in Appendix B. Experience and knowledge were centered and scaled, such that 0 represents the average response on these variables. Courses, assets, and major were dummy-coded, such that 0 represents the absence of these characteristics and 1 indicates their presence. The R^2 values refer to the marginal variance explained by the fixed effects. Sample sizes (N) refer to the number of participants; thus, each model contains 4N (trial-level) observations. Sample sizes differ slightly from the primary analyses due to missing data in the expertise variables.

Experiment	Experiment 1 (N = 179)			Experiment 2 (N = 171)	Experiment 3 (N = 174)	
DV	**1-Day Forecasts**	**2-Week Forecasts**	**1-Year Forecasts**	**Choices**	**Choices**	**Net Emotions**
			Positive Valence Condition			
Time-Reference	.001 (.006)	.012 (.006)º	.010 (.009)	2.53 (1.64)	0.74 (1.30)	0.39 (0.38)
Experience	-.021 (.009)*	–.023 (.012)*	–.009 (.020)	–1.93 (2.29)	–1.98 (2.26)	–0.05 (0.51)
Knowledge	.016 (.008)º	.020 (.011)º	.001 (.018)	1.07 (2.31)	1.99 (2.11)	0.32 (0.47)
Courses	–.010 (.011)	–.012 (.014)	–.008 (.024)	1.23 (3.40)	5.08 (3.26)	0.95 (0.73)
Assets	.019 (.012)	.003 (.015)	.001 (.026)	–4.86 (3.66)	–1.08 (3.38)	0.06 (0.76)
Major	.007 (.015)	–.010 (.019)	–.061 (.031)	–4.85 (3.92)	–4.09 (3.63)	–1.02 (0.81)
Time-Reference × Experience	–.008 (.005)	.005 (.005)	.009 (.008)	–2.98 (1.16)*	–2.84 (1.09)**	–0.56 (0.29)º
Time-Reference × Knowledge	–.010 (.005)*	–.006 (.005)	–.015 (.008)º	0.34 (1.17)	2.31 (0.95)*	0.51 (0.28)º
Time-Reference × Courses	.002 (.007)	–.002 (.006)	–.006 (.010)	–0.72 (1.72)	–0.82 (1.47)	–0.52 (0.43)
Time-Reference × Assets	–.000 (.007)	–.008 (.007)	.009 (.011)	–1.77 (1.85)	2.35 (1.51)	0.18 (0.44)
Time-Reference × Major	.018 (.009)*	.015 (.009)º	.010 (.013)	4.54 (1.98)*	–2.78 (1.63)º	–0.61 (0.47)
Intercept	.057 (.010)***	.109 (.013)***	.166 (.023)***	67.50 (3.29)	66.18 (3.17)	4.21 (0.80)
R^2	.043	.042	.018	.062	.035	.030
			Negative Valence Condition			
Time-Reference	–.006 (.005)	–.001 (.006)	–.010 (.010)	–2.17 (1.58)	–1.03 (1.23)	–0.36 (0.27)
Experience	.007 (.009)	.013 (.010)	.028 (.016)º	–0.39 (2.09)	3.86 (2.41)	1.06 (0.50)*
Knowledge	–.015 (.008)º	–.019 (.009)*	–.042 (.015)**	0.27 (2.11)	–1.16 (2.25)	–0.33 (0.46)
Courses	.030 (.011)**	.018 (.012)	.019 (.019)	0.76 (3.10)	–3.09 (3.48)	–1.23 (0.72)º
Assets	–.018 (.012)	–.013 (.013)	.002 (.021)	–2.76 (3.34)	–0.64 (3.60)	–0.73 (0.75)
Major	.008 (.014)	–.011 (.016)	.004 (.025)	4.16 (3.57)	2.22 (3.88)	0.72 (0.80)
Time-Reference × Experience	.005 (.005)	.005 (.005)	.004 (.009)	1.44 (1.11)	1.01 (0.96)	0.24 (0.21)
Time-Reference × Knowledge	–.007 (.004)º	–.007 (.005)	–.010 (.008)	–1.71 (1.12)	–1.30 (0.90)	–0.08 (0.20)
Time-Reference × Courses	.003 (.006)	.000 (.007)	.017 (.011)	0.05 (1.65)	2.02 (1.39)	0.17 (0.31)
Time-Reference × Assets	.004 (.006)	.003 (.007)	.001 (.012)	–0.79 (1.78)	–1.20 (1.44)	–0.11 (0.32)
Time-Reference × Major	.011 (.007)	.010 (.009)	–.006 (.015)	–2.25 (1.90)	–0.36 (1.55)	0.28 (0.34)
Intercept	–.036 (.010)***	–.047 (.013)**	–.070 (.002)**	26.61 (3.05)	31.23 (3.17)	–4.75 (0.73)***
R^2	.062	.033	.038	.040	.025	.041

Table C2. Effects of Time-Reference × Expertise in Experiments 1–3.

Note. Entries are unstandardized coefficients (and SEs) for each model. Time-reference was contrast-coded as in Appendix B. Experience and knowledge were centered and scaled, such that 0 represents the average response on these variables. Courses, assets, and major were dummy-coded, such that 0 represents the absence of these characteristics and 1 indicates their presence. The R^2 values refer to the marginal variance explained by the fixed effects. Sample sizes (N) refer to the number of participants; thus, each model contains 2N (trial-level) observations. Sample sizes differ slightly from the primary analyses due to missing data in the expertise variables.

The coefficients on valence were consistently highly significant across all DVs in all experiments, as one would expect from the models reported in Part B. Since interaction terms were included, these coefficients reflect the estimated effect of valence condition for a participant who was average on self-reported experience and knowledge (since these variables were mean-centered) and who did not take finance courses, own financial assets, or major in business (since these variables were dummy-coded). The simple effects of the expertise variables were not consistent and rarely reached significance.

Of primary interest here were interactions between valence condition and the expertise variables, which reflect the potential moderating effect of expertise. These effects were not consistent across measures of expertise, although some did appear to be robust across DVs. In particular, investors higher in self-reported experience consistently showed a smaller effect of valence, with this effect reaching significance for 5 out of 6 DVs; owning financial assets, however, showed no significant effect or consistent direction. Conversely, people higher in self-reported *knowledge* tended to show larger effects of valence, with this effect reaching significance for the three DVs in Experiment 1; the results for finance courses and major, however, were inconsistent across studies. Overall, however, all of these effects were small relative to the effect of valence condition, indicating that any effects of expertise (within the range of our sample) are not sufficient to neutralize the effect.

Table C2 reports the corresponding analyses for the time-valence effect, examining the effects of expertise separately for the positive valence condition (where the predicted effect of time-valence is positive, and therefore an attenuating effect of expertise would be negative) and negative valence condition (where the predicted effect of time-valence is negative, and therefore an attenuating effect of expertise would be positive). The specifications were otherwise similar to those in Table C1.

As in Table C1, we note that the coefficients reported for time-reference differ from those reported in Part B, as the presence of the interaction term changes their interpretation such that they reflect the estimated simple effect of valence condition for a participant who is average on self-reported experience and knowledge and who lacks financial assets, finance courses, or a business major. Most of these coefficients were not statistically significant, although all were in the predicted direction. This difference from Part B suggests that the results there were driven largely by subsets of the sample who showed larger effects.

The moderation results for time-valence condition were fairly comparable to those for valence condition, although not as robust. Self-reported financial experience tended to make time-reference effects smaller (i.e., negative coefficients for the positive condition and positive coefficients for the negative condition) in 10 out of 12 models, although this effect only reached statistical significance in 2 cases (both measures of choices in the positive valence condition). Self-reported financial knowledge, in contrast, tended to make time-valence effects *larger* in 9 out of 12 models, although again this effect did not reach significance in most models (and even reached significance in the opposite direction in one model). The moderating effects of the objective measures were inconsistent, as in Table C1.

These results must be considered exploratory. Taken together, they provide some evidence for diverging effects of two dimensions of expertise. Whereas investing experience seems to make the effect of valence smaller (and hence closer to normative models from economic theory), financial *knowledge* (or at least perceived financial knowledge) seems, if anything, to exacerbate the effect of valence. Similar effects were seen for the time-reference effect, although less consistently. Overall, the valence effect is strong even among participants high in expertise, as its magnitude swamps the moderating effects of expertise, whereas the time-reference effect may be driven to a large degree by the low-experience, high-knowledge investors in our sample.

Arguably, these results are consistent with some prior studies of expertise in behavioral economics and social psychology. Some economic anomalies, such as the endowment effect, seem to be eliminated with market experience (List, 2003), which is consistent with the view sometimes expressed by neoclassical economists in response to behavioral economics studies, that experienced and incentivized economic agents are unlikely to show systematic biases. On the other hand, the opposite effect of self-reported knowledge may reflect the conventional wisdom that "a little knowledge is a dangerous thing," as total novices are aware that they are novices, but a modest amount of learning can lead to Dunning–Kruger effects in which slightly-less-novice novices greatly overestimate their abilities (Sanchez & Dunning, 2018). However, given that the results were not particularly consistent and that we did not predict these findings in advance, such conclusions should be taken as speculative.

Part D: Experiment S1

In Part C, we found modest effects of expertise. However, the knowledge of laypeople is likely to be limited, even for those who are relatively high in investing experience. Would we find similar results among true experts, such as professional financial analysts or PhD students in economics? We tested this question in Experiment S1.

Method

We recruited 21 participants with financial expertise, who participated in several studies of financial decision-making. These participants were recruited from an email newsletter sent to members of the UK branch of the Chartered Financial Analysts (N = 5), from a seminar given to economics PhD students (N = 8), and from Masters students in finance and accounting at an elite UK university (N = 8). We excluded 5 from analysis due to incomplete data (N = 3) or inattentiveness (N = 2). The method was otherwise identical to Experiment 1.

Results and Discussion

As shown in Table D1, the results mainly replicated previous studies. At baseline, prior to receiving any news, participants predicted moderate price increases at 1-day (+3.0%), 2-week (+3.9%), and 1-year (+6.9%) time horizons. Like laypeople, the 1-day and 2-week predictions were unrealistically optimistic, but the 1-year prediction is in line with historical returns.

Also like laypeople, these predictions shifted greatly depending on the valence of news announcements. Collapsing across time-reference conditions, participants predicted increases of +5.6%, +8.4%, and +10.0% at 1-day, 2-week, and 1-year timeframes, but price *decreases* of –1.5%, –3.4%, and –2.9% at 1-day, 2-week, and 1-year timeframes. This led to a significant difference between the positive and negative news conditions at 1-day [$t(15) = 2.73$, $p = .015$, $d = 1.12$], 2-week [$t(15) = 3.27$, $p = .005$, $d = 1.40$], and 1-year [$t(15) = 3.45$, $p = .004$, $d = 1.28$] horizons. Like laypeople, this premium for positive rather than negative news is not expected to close even at a 1-year time horizon (a 12.9% difference), where the premium is similar to the 2-week time horizon (a 11.8% difference). This is inconsistent with both neoclassical economics and empirically derived behavioral expectations.

Time Horizon	Baseline	Positive Surprise		Negative Surprise	
		Past	Future	Past	Future
1-day	3.0% (3.8%)	6.5% (7.6%)	4.7% (9.7%)	–0.9% (6.1%)	–2.2% (6.3%)
2-weeks	3.9% (6.6%)	9.2% (10.4%)	7.5% (11.1%)	–2.5% (9.4%)	–4.3% (8.2%)
1-year	6.9% (8.1%)	10.2% (8.6%)	9.7% (16.6%)	+0.1% (11.5%)	–5.9% (11.3%)

Table D1. Results of Experiment S1.

Note. Entries are predicted changes from current value, as percentages. Possible scores range from –50% to +50%. The baseline column gives the mean of the baseline estimates made across the four within-subjects conditions, since these estimates were made prior to the manipulation. SDs in parentheses.

Experts also appear to have taken account of time-reference, though less statistically robustly. In Study 1, laypeople predicted more extreme outcomes for future- than past-oriented news when the news was positive, and showed no effect when the news was negative (though this asymmetry was not seen in Experiments 2 or 3). Experts in Experiment S1 actually showed the *opposite* asymmetry. For *negative* events, their predictions were more extreme in response to future- rather than past-oriented news, with this effect reaching significance at a 1-year horizon [$t(15) = 2.21$, $p = .043$, $d = 0.55$]. For positive events, on the other hand, predictions did not significantly differ at any time horizon [*t*s < 0.8, *p*s > .47] and tended to be in the opposite direction. This less robust result for experts could be because expert predictions line up more closely with neoclassical economics. However, given the relatively small sample and the significant results for the valence asymmetry, it seems just as likely that this result is due to lack of statistical power for detecting this more subtle effect.

Together, these findings with true experts are consistent with our findings from previous experiments, where the effects, particularly of valence, persisted over-and-above any effects of expertise. Experts' sensitivity to news valence was numerically somewhat weaker compared to that of laypeople, but of sufficient magnitude that it was easily detectible even in a fairly small sample. There was some evidence for an effect of time-reference as well, but this was less robust. Thus, although true experts may be somewhat less susceptible to these effects compared to laypeople or amateur investors, they do appear to rely on similar psychological processes.