**A PROTOTYPE THEORY OF CONSUMER EXPENSE MISPREDICTION**

**Abstract**

The present research theorizes that consumer expense predictions are shaped by prototype attributes that come to mind with relative ease when predictions are being constructed. These attributes represent average spending, where “average” is akin to the mode of a consumer’s expense distribution. This leads to an *expense prediction bias* in which consumers *under*-predict their expenses because the distribution of consumer expenses is positively skewed with mode < mean. Accordingly, it is proposed that prompting consumers to consider reasons why their expenses might be different than usual will increase prediction accuracy by making atypical expenses easier to retrieve. Five studies, including a longitudinal field study with members of a financial cooperative, provide support for this account of the bias.

*Keywords:*expense prediction bias, prototypes, prototype attributes, temporal asymmetry, consumer misprediction, consumer financial decision making.

Consumers tend to under-predict their future expenses, and this error can be costly. For example, over 25% of Americans with a 401(k) savings account withdraw funds early (i.e., before retirement), often to cover unexpected expenses (Fellowes and Willemin 2013). These early withdrawals cost consumers approximately $7 billion a year in penalties. Similarly, each year almost 2 million Americans use a payday loan to cover an unexpected expense (Pew 2012). The APR on these loans frequently exceeds 400% (Consumer Federation of America 2018). Many consumers also hold the expectation that they will be able to pay off their credit card balance each month (Yang et al. 2007). Yet American consumers collectively hold over one trillion dollars in credit card debt and pay associated interest costs (Federal Reserve Bank of New York 2018).

These examples suggest that increasing expense prediction accuracy can help consumers spend, save, and/or borrow money in a more efficient manner. An accurate assessment of future expenses can help consumers better allocate funds between their checking and 401(k) accounts to avoid penalties for early withdrawal. And if consumers had a clearer idea of how much they would spend in the future, they might choose to spend less in the present to avoid the costs associated with borrowing or using credit to cover expenses down the road. The prosocial value of helping consumers avoid these costs is evident. The rush by venture capital firms to fund FinTech start-ups offering app-based products that help consumers manage their expenses (CB Insights 2018) indicates there is also firm value in improving expense prediction accuracy.

Echoing the examples offered above, academic research also suggests that consumers tend to under-predict their future expenses (e.g., Ulkumen et al. 2008; Peetz and Buehler 2009), a phenomenon we label the *expense prediction bias*. The goal of the present research is to identify a key psychological driver of this bias, then leverage that theoretical insight to develop, test, and validate a simple cognitive tool that improves consumers’ expense prediction accuracy. To do so, we first theorize that consumer expense predictions are shaped by prototype attributes that represent modal expenses. We then propose that “prototypical prediction” causes consumers to under-predict their expenses because, generally speaking, the distribution of consumer expenses is positively skewed with mode < mean. Finally, we propose that prompting consumers to consider reasons why their expenses might be different than usual will increase prediction accuracy by making atypical expenses easier to retrieve.

By developing and testing a prototype theory of consumer expense misprediction, the present research makes the following contributions. First, the prototype theory offers a parsimonious explanation for why the expense prediction bias occurs, and it allows for the first unified interpretation of past findings in this area. Second, the present research introduces the first practical, effective, and field-tested intervention capable of neutralizing the expense prediction bias. Third, the present research provides the first step toward a comprehensive understanding of the bias itself. For example, this research is the first to identify the magnitude, prevalence, and persistence of the bias in non-student samples. It is also the first work to study the bias longitudinally and in the field, and to measure monthly expense predictions against actual expenses incurred during the target month.

The present research also makes two significant contributions to the broader literature on consumer misprediction. First, we highlight the important role that skew in the distribution of relevant outcomes can have with respect to prediction accuracy. Second, where past research demonstrates that predictions do not sufficiently incorporate distributional information (e.g., Buehler, Griffin, and Ross 1994; Kruger and Evans 2004), the present research elucidates *what* distributional information is neglected. Similarly, we contribute to research on the use of prototype attributes in judgment and decision making—which has largely remained agnostic about the measure of central tendency that prototype attributes represent—by providing evidence that these attributes represent modal (rather than mean) outcomes in the context of expense prediction. Finally, the present research contributes to a nascent literature that demonstrates a temporal asymmetry in which people mentally represent the future in more prototypical terms than the past (Kane, Van Boven, and McGraw 2012; Van Boven, Kane, and McGraw 2008). By comparing the nature of predicted versus recalled expenses, we extend this work to the domain of money.

In the following sections we present our prototype theory of expense prediction bias and lay out our hypotheses. We then present five studies that test these hypotheses in the lab and the field. To conclude, we discuss the theoretical and practical implications of this work, along with directions for future research.

**PREDICTIONS AND PROTOTYPE ATTRIBUTES**

The first theoretical proposition of the present research is that expense predictions are shaped by *prototype attributes*. Consistent with past research, prototype attributes are defined as representations of an average (Kahneman 2003; Kahneman and Frederick 2002). So, for example, prototype attributes can be conceptualized as the response to internal queries like “what expenses do I *typically* incur each [week/month]?” or “how much do I *typically* spend each [week/month]?” In contrast, the *target attribute* in expense prediction is the response to some variation of the query “how much will I *actually* spend next [week/month]?”

The proposition that expense predictions are shaped by prototype attributes (referred to hereafter as the “prototype proposition”) is derived from the observation that the target attribute in expense prediction is relatively low in accessibility, meaning it does not come to mind without deliberation or effort. To illustrate this point, consider that a comprehensive expense prediction first requires anticipating future expenses, then estimating the amount of each expense, then adding these amounts together. On the other hand, prototype attributes that represent accurate impressions of an average can be formed with relative ease, and they are highly accessible (Ariely 2001; Kahneman 2003; Kahneman and Frederick 2002; Rosch and Mervis 1975). Therefore, the prototype proposition is broadly consistent with the finding that accessibility often determines the content of judgments and decisions (e.g., Johnson, Häubl, and Keinan 2007; Kahneman 2003; Kahneman and Frederick 2002; Tversky and Kahneman 1974).

The prototype proposition is also supported by the existing body of work on consumer expense prediction. Previous work in this area has demonstrated that consumers’ expense predictions do not adequately incorporate either the distribution of their actual past expenses or the distribution of their possible future expenses (Peetz and Buehler 2012; Peetz et al. 2015). This kind of *distribution neglect* suggests that expense predictions are constructed from a thin slice of a consumer’s expense distribution rather than consideration of the distribution in its entirety. Past work in this area has also demonstrated that consumers incorporate ordinary, recurring expenses into their predictions, but not exceptional, infrequent expenses (Peetz and Buehler 2013; Sussman and Alter 2012). This implies that the “thin slice” from which expense predictions are constructed is the part of the expense distribution represented by prototype attributes, because ordinary, recurring expenses are representative of a consumer’s typical or average expenses. Finally, it has been shown that consumers under-predict their future expenses in part because they are overconfident in their prediction accuracy, and as a result they do not sufficiently adjust their initial predictions upward (Ulkumen et al. 2008). This supports the prototype proposition because confidence is indicative of cognitive ease (Alter and Oppenheimer 2009), and prototypes are “easy on the mind” (Winkielman et al. 2006).

Furthermore, the prototype proposition is consistent with research on misprediction in other domains. For example, research on affective forecasting has found that people overestimate the impact of future emotional events (Gilbert et al. 1998; Wilson et al. 2000). This supports the prototype proposition because the prototype of emotional events tends to be highly affective (Gilbert et al. 2004; Schkade and Kahneman 1998; Wilson et al. 2000), so it is natural that prototypical prediction would cause over-prediction in this context. Research on the planning fallacy has documented that people systematically under-predict their task completion times (e.g., how long it will take to file their taxes, or finish their holiday shopping), even when equipped with the knowledge that similar tasks have taken longer than planned in the past (Buehler, Griffin, and MacDonald 1997; Buehler, Griffin, and Ross 1994; Kahneman and Tversky 1979). One explanation for this behavior is that people tend to adopt a non-distributional “inside view” in which they focus on plan-based scenarios while generating predictions (Buehler, Griffin, and Ross 2010; Kahneman and Tversky 1979). A complementary perspective is that people base their predictions on prototype attributes that neglect atypical outcomes. For example, a consumer’s prediction that it will take 3 hours to finish their holiday shopping may be based on typical instances of such a task that include going to the mall, finding presents, and even waiting in line to pay, but not less typical instances like discovering the Lego set that little Timmy really wanted is sold out and needs to be ordered from another store, or getting rear-ended in the parking lot. In this case, prediction is not necessarily predicated on a plan to finish shopping quickly or by a certain time, it is simply a function of what is (and isn’t) included in the prototype attributes upon which the prediction is based. This perspective is notably consistent with work demonstrating that prospective mental representations of people, places, and events to which plans are not likely to be attached are more prototypical than retrospective representations of the same people, places, and events (Kane, Van Boven, and McGraw 2012).

Taken together, the evidence offered above provides compelling reasons to believe that expense predictions are shaped by prototype attributes that represent a consumer’s typical or average expenses. The following section defines “average” in greater detail and outlines why prototypical predictions can be problematic.

**PROTOTYPE ATTRIBUTES AND “AVERAGE” EXPENSES**

The second theoretical proposition of the present research is that the “average” represented by prototype attributes is akin to the mode of a consumer’s expense distribution rather than the mean. This proposition is directly supported by the finding that consumers are fairly accurate when predicting their ordinary, recurring expenses (Sussman and Alter 2012), because if predictions are shaped by prototype attributes that represent modal expenses, then it follows that consumers should be fairly adept at predicting ordinary, recurring expenses, which are modal by definition. This proposition is further supported by the basic psychological finding that impressions of prototypes are formed through repeated exposure to a stimulus (Ariely 2001), because modal expenses are those that are repeated most frequently. There are also strong reasons to believe that prototype attributes do *not* represent mean expenses. For example, prototype formation requires a certain degree of homogeneity within the considered set (Kahneman 2003). But calculating mean expenses requires incorporating heterogeneous outcomes (i.e., atypical expenses from the tails of the distribution), and past research suggests that consumers do not do this when constructing expense predictions (Peetz and Buehler 2013; Sussman and Alter 2012). It is also the case that “extensional” variables (i.e., sums) are relatively inaccessible (Kahneman and Frederick 2002; Tversky and Koehler 1994), which makes them poor candidates for prototype attributes. This suggests that prototype attributes do not represent mean expenses because calculating a mean necessitates summation.

**PROTOTYPICAL PREDICTIONS AND THE EXPENSE PREDICTION BIAS**

To understand why prototypical prediction leads to an expense prediction bias in which consumers under-predict their expenses, one need only look at the distribution of consumer expenses. If consumer expenses were normally distributed with mode = mean, then prototypical prediction would not be expected to cause an expense prediction bias, because the frequency and magnitude of over- and under-prediction would balance out between-subjects in a given sample at a given point in time, as well as within-subject over a sufficient length of time. However, our data make it clear that consumer expenses are positively skewed, with mode < mean. This is fairly intuitive from a mathematical perspective given that expenses are bounded by zero on the left side of the distribution, but they are free to run as high as a consumer’s credit will allow on the right. It is also fairly intuitive from a psychological perspective given the apparent ease with which consumers can overspend their budget (Sussman and Alter 2012) and the apparent difficulty they have spending less than their budget (Peetz and Buehler 2009).

In tandem, the theoretical propositions outlined above and the observation that consumer expenses are positively skewed lead to the following hypotheses regarding expense predictions. First, if predictions are shaped by prototype attributes, and prototype attributes represent modal expenses, then it follows that expense predictions will be lower when the distribution of expenses is positively skewed with mode < mean versus when it is normally distributed with mode = mean. Note that the distribution of expenses can be experimentally manipulated in a lab setting, changing the mode while keeping other features of the distribution like mean and variance constant. In this context it is hypothesized that:

**H1:** Consumer expense predictions will be lower (and farther from the mean) after observing a positively skewed distribution of expenses versus a normal distribution.

The second hypothesis is that, generally speaking, the more heavily a consumer relies on prototype attributes when predicting his or her expenses, the lower their expense prediction will be. Following the logic laid out above: If typical expenses are modal expenses, and expenses are positively skewed (with mode < mean), then it follows that if expenses are expected to be more typical, they will also be expected to be lower. Formally:

**H2:** Perceived typicality of future expenses is negatively correlated with expense predictions.

The third hypothesis of the present research is that prototypical predictions cause a *temporal asymmetry* in which consumers predict their future expenses will be more typical than they recall their past expenses being. This hypothesis follows from the observation that retrospection is grounded in reality (Johnson and Raye 1981; Kane, Van Boven, and McGraw 2012; Van Boven, Kane, and McGraw 2008), which implies that expense *recall* will include both typical and atypical expenses. In contrast, the prototype proposition suggests that *predictions* will largely neglect atypical expenses. This hypothesis is consistent with research documenting a temporal asymmetry in which prospective mental representations of people, places, and events are more prototypical (i.e., less detailed) than retrospective mental representations (Kane, Van Boven, and McGraw 2012). It is also consistent with work showing there is considerable neural differentiation during prospective (vs. retrospective) event construction (Addis, Wong, and Schacter 2007). It is therefore hypothesized that:

**H3:** Consumers, on average, predict their future expenses will be more typical than their past expenses.

A clear implication of H2 and H3 is that consumers will under-predict their future expenses as compared to their past expenses. In other words, if people think their expenses will be more typical in the future, and higher perceived typicality is associated with lower expense predictions, then it follows that:

**H4a:** Consumers predict lower expenses for the future as compared to the expenses they recall for the past.

Finally, it is hypothesized that consumers will under-predict their expenses as compared to the expenses they actually incur during the target week or month. This follows logically from the proposition that expense predictions are shaped by prototype attributes that represent modal expenses, and the observation that expenses are positively skewed with mode < mean. Although we do not expect that all consumers’ predictions are completely prototypical, research suggests that upward adjustment from a prototypical prediction will not be sufficient to prevent under-prediction (Epley and Gilovich 2006; Tversky and Kahneman 1974; Ulkumen et al. 2008). It is therefore hypothesized that:

**H4b:** Consumers under-predict their expenses for the target week or month as compared to the expenses they actually incur during that week or month.

**IMPROVING EXPENSE PREDICTION ACCURACY**

If the expense prediction bias is caused in part by prototypical prediction, then it follows that making atypical expenses more accessible when consumers are constructing their predictions will increase prediction accuracy. We reasoned that having people list three reasons why their expenses might be *different* than usual would serve as a simple cognitive tool that accomplishes this goal. The logic underlying this “atypical” intervention follows that of H2 above—although “different” could represent outcomes that would lead to lower predictions, most of the distribution of expenses that is different from the mode is also higher than the mode, and should therefore make predictions closer to the mean.

Mechanistically, this tool bears some resemblance to the “unpacking” intervention derived from support theory, in which people are asked to unpack their prediction into its component parts (e.g., individual expenses) to elicit greater consideration of the distribution of possible future outcomes (Kruger and Evans 2004; Peetz et al. 2015; Tversky and Koehler 1994). There is, however, an important distinction between our “atypical” intervention and unpacking: Where the unpacking intervention prompts people to consider the full distribution of possible outcomes, the atypical intervention prompts them to consider outcomes only in the right tail of the distribution. This is important from a theoretical perspective because the unpacking intervention says only that distributional information is missing from predictions. In contrast, the atypical intervention deepens our understanding of *what* distributional information is missing. The atypical intervention also carries a practical advantage: It requires consideration of only a handful of reasons why expenses may be atypical (vs. trying to unpack all possible expenses), which makes it easier to employ. This is noteworthy given that many expense predictions are made spontaneously (Peetz et al. 2016), which suggests that a simpler tool will be more widely used in practice. Finally, it is worth noting that some evidence supports the caveat that although the unpacking intervention improves mean prediction accuracy, it can sacrifice correlational accuracy (Kruger and Evans 2004; Peetz et al. 2015). In other words, while unpacking can increase predictions to better match the true mean of actual outcomes on average, it can also substantially increase variance, meaning the correlation between predicted and actual outcomes can be reduced by the intervention. We speculate that one reason this is true is that consumers who invest the effort that unpacking requires tend to make higher predictions, while those who invest very little effort make a lower prediction than they would have otherwise. We therefore posit that the simplicity of the atypical intervention will help it avoid this fate.

The logic underlying the atypical intervention also mirrors research demonstrating that “defocalizing” can improve affective forecasts. For example, Wilson et al. (2000) found that people induced to think about their post-event daily routines before they predicted how that event would make them feel were subsequently less likely to over-predict their affective reaction to the event. Because affective prototypes tend to be somewhat extreme (Gilbert et al. 2004; Schkade and Kahneman 1998; Wilson et al. 2000), a reasonable interpretation of this finding is that having people consider non-prototypical information lessened their reliance on prototype attributes and in doing so improved their prediction accuracy. The same logic applies to consumer expense predictions: While people naturally consider typical (i.e., modal) expenses, and therefore under-predict, an intervention that prompts them to consider reasons why their expenses might be different than usual should increase prediction accuracy by making atypical expenses more accessible. We therefore hypothesize that:

**H5a:** Prompting consumers to consider three reasons why their expenses might be different from a typical week increases expense prediction accuracy.

**H5b:** The effect of the atypical intervention (vs. control) on expense predictions is mediated by the number of atypical expenses that come to mind.

**OVERVIEW OF STUDIES**

To test our hypotheses we conducted a series of five experiments. In study 1, we manipulate the distribution of expenses to test the hypothesis that consumer expense predictions will be lower (and farther from the mean) after observing a positively skewed distribution of expenses versus a normal distribution (H1). In study 2, we examine the extent to which perceived typicality of future expenses is associated with expense predictions (H2), and whether consumers predict more typical and/or lower expenses for the future as compared to the past (H3 and H4a). In study 3—conducted with a nationally representative sample of adult Americans—we directly replicate study 2, test the effectiveness of the atypical intervention (H5a), and provide process evidence that demonstrates why the intervention is effective (H5b). In study 4, we examine H2–H4a in a repeated-measures longitudinal field study with members of a financial cooperative, compare weekly expense predictions against the expenses consumers incur during each week of the study (H4b), and test the efficacy of the atypical intervention (vs. control) in the last week of the study (H5a). Study 4 also measures the accuracy of monthly expense predictions. In study 5, we test the possibility that monthly expense predictions are less prototypical than weekly predictions and examine whether the atypical intervention is capable of increasing monthly predictions as well as weekly predictions.

**STUDY 1: MANIPULATING THE DISTRIBUTION OF EXPENSES**

 One way to test the veracity of the prototype theory is to experimentally manipulate the mode of the distribution of past expenses and observe subsequent predictions. If our theory is correct, then a positively skewed distribution with mode < mean should lead to a lower prediction than a normal distribution with mode = mean, holding the mean constant. We used study 1 to test this possibility by presenting participants with 52 weekly expense amounts drawn in random order from either a positively skewed distribution or a normal distribution. We then asked participants to predict their expenses for the next week, imagining that the expenses they had seen were an accurate representation of their weekly expenses over the past year. For clarity, it warrants mention that our expectation was *not* that predictions in the positive-skew condition would exactly equal the mode, because the probability that at least some participants will incorporate distributional information into their prediction is high, particularly when distributional information is made as salient as it is in this paradigm (Kahneman 2003; Kahneman and Frederick 2002). Rather, our expectation was that predictions in the positive-skew condition would be lower (and farther from the mean) than in the normal condition (H1). The preregistration, study materials, and data for this study can be found here: [insert hyperlink].

Method

*Participants and Procedure.* Four hundred and one American residents were recruited from Amazon Mechanical Turk to participate in an online study about consumer expenses (49.6% female, *M*age = 36.9). On the first page of the study participants read the following instructions: “On the next page you will see 52 values presented in quick succession. We want you to imagine that these values represent your weekly spending over the course of one year.” On the second page of the study participants were presented with a series of 52 weekly expense amounts, one at a time, every 1.2 seconds. This paradigm was adapted from André, Reinholtz, and de Langhe (2017) and chosen to parallel research on perception (Ariely 2001), because prototype formation is a perceptual process (Kahneman 2003). The 52 weekly expense amounts were drawn in random order from either a positively skewed distribution (Min = $170, Mode = $180, Mean = $200, Max = $400, SD = 38.81, skew = 3.13) or a normal distribution (Min = $85, Mode = Mean = $200, Max = $315, SD = 38.41, skew = 0.00). Mean and range were held constant across conditions, and variance was controlled as tightly as possible to ensure we were not also manipulating uncertainty. After viewing either the positive skew or normal distribution, participants were asked to predict their expenses by responding to the following question: “If the expenses you just saw were an accurate representation of your actual weekly spending over the last year, how much do you estimate you will spend in the next week?” Participants were also asked to estimate the average amount of the weekly expenses they had seen so that we could explore whether predictions were lower than the perceived average (i.e., optimistic), higher than the perceived average (i.e., pessimistic), or a direct representation of the perceived average.

Results

As hypothesized (H1), predicted expenses were significantly lower in the positive-skew condition (*M* = $195.14, SD = 18.95) than in the normal condition (*M* = $200.32, SD = 21.44), as confirmed by an independent samples t-test (*t*(399) = 2.98, *p* = .003, *d* = .26). Furthermore, predicted expenses in the positive-skew condition were significantly lower than the $200 mean of the distribution of past spending (*t*(197) = -3.63, *p* < .001), but predicted expenses in the normal condition were almost identical to the $200 mean of past spending (*t*(198) = .21, *p* = .83). Finally, average and predicted expenses did not differ significantly in the positive-skew condition (*M*avg = $194.90, SDavg = 15.59, *t*(200) = .20, *p* = .84) or in the normal condition (*M*avg = $199.72, SDavg = 16.76, *t*(199) = .40, *p* = .69).

Discussion

 Study 1 provides support for the prototype theory in two ways. First, it demonstrates that, ceteris paribus, expense predictions are lower and farther away from the mean when the distribution of expenses is positively skewed versus normal. Second, it provides evidence that the prototype (i.e., average) that people form when a distribution is positively skewed over-weights modal outcomes and does not accurately represent the mean. We view these results as a fairly conservative estimate of the impact of distributional skew on predictions, given that participants were presented with the full distribution of outcomes—including atypical outcomes—right before predicting, and in reality atypical outcomes are not likely even to come to mind. The limitations of study 1 are that the distribution of expenses is hypothetical, and participants are asked to make only one prediction. We address these limitations in studies 2–5 by soliciting predictions based on consumers’ knowledge of their own expenses, and measuring prediction accuracy in a repeated-measures longitudinal field study (study 4).

 **STUDY 2: PERCEIVED TYPICALITY OF PAST VS. FUTURE EXPENSES**

In study 2 we test our prototype theory by comparing the perceived typicality of predicted future expenses and recalled past expenses. Although memory is not perfect, recall does approximate reality (Kane, Van Boven, and McGraw 2012). This leads to the expectation that recalled past expenses will reflect the full distribution of expenses, including atypical expenses. However, the prototype proposition leads to the expectation that expense predictions will tend to emphasize typical expenses and neglect atypical expenses. Thus, the first goal of study 2 is to test the hypothesis that consumers predict their future expenses will be more typical than their past expenses (H3).

The second goal of study 2 is to examine the relationship between perceived typicality of future expenses and expense predictions. If typical expenses are modal expenses, and expenses are positively skewed, then it follows that if expenses are expected to be more typical, they will also be expected to be lower. We therefore use study 2 to test the hypothesis that perceived typicality of future expenses is negatively correlated with expense predictions (H2).

Finally, study 2 also tests the hypothesis that consumers predict lower expenses for the next week as compared to the expenses they recall for the past week (H4a), which follows logically from the two preceding hypotheses. We chose weekly expenses (as opposed to monthly or yearly expenses) as our unit of analysis for two reasons: 1) we wanted participants to be able to recall their past expenses while their memories were fresh, and 2) survey data (*n* = 1,514) collected from our participant pool prior to running this study suggested that people commonly think about their expenses in weekly terms (38.6% of survey respondents indicated thinking about their finances on a weekly basis vs. 27.1% who indicated thinking about their finances on a biweekly basis, 28.5% who indicated thinking about their finances on a monthly basis, and 5.7% who indicated “other”). In studies 4 and 5 we extend our investigation to include monthly expense predictions as well.

As highlighted in our theoretical development, our consumer expense data display strong positive skew, which will become increasingly clear starting with study 2. This is a defining characteristic of the data, and an important element in our theory; it also presents serious challenges for responsible inferential analysis. To address this situation, we present the unadulterated distributional skew of expenses (i.e., the distribution of expenses in the control condition with no exclusions or transformation) at the outset of studies 2–4. Then, for inferential analysis, we exclude the data of outlier participants whose reported expenses exceed their predictions by more than a factor of 10 (or vice versa), and LN-transform the distributions of reported and predicted expenses. For ease of interpretation, we then exponentiate our descriptive results and present them in dollar terms. Notably, this procedure ameliorates concerns related to homogeneity of variance assumption violations and the influence of outliers, but it does not change the pattern of results observed in the raw data. To illustrate this, we present a pair of statistical robustness tests (winsorization and non-parametric analysis) in web appendix A, and we detail the impact that the transformation process has on the expense data in web appendix B. The data, syntax, and study materials from studies 2–4 can be found here: [insert hyperlink].

Method

*Participants*. We recruited 499 US residents via Amazon Mechanical Turk to participate in a short consumer expense survey (*M*age = 33.51; 41.3% female). The reported expenses of 14 participants exceeded their predictions by more than a factor of 10 (or vice versa), leaving us with an effective sample size of 485 (*M*age = 33.67; 41.6% female).

*Procedure.* Participants were first asked to report their expenses for the past week and then, on the following page, to predict their expenses for the next week. Specifically, participants read the following instructions:

Please take some time to estimate your expenses for the past [next] week (i.e., the past [next] 7 days).

Please enter your total estimated expenses (in dollars) for the past [next] week. Your estimate should account for all the expenses you incurred [will incur] except monthly expenses like rent that happen[ed] to be due in the past [next] week.

We asked participants to exclude monthly expenses like rent from their estimates to reduce the possibility that any observed bias could be due to variation in the timing of these expenses. We next measured perceived typicality of expenses by asking, “How *different* or *similar* do you think your expenses were [will be] for the **past** [**next**] week, relative to a typical week?” (1 = Very different; 7 = Very similar). Finally, participants were asked to report basic demographic information.

Results

 *Skew.* Both recalled and predicted expenses displayed a high degree of positive skew (skewrecall = 4.47, skewpredict = 3.69). Naturally, this could be caused by a relatively small number of wealthy participants having extremely high expenses that drag the right tail of the distribution outward. Therefore, we also regressed expenses onto reported income (measured on a 1–16 scale in $10K increments) then measured skew in the distribution of residuals. Conceptually, this measures skew after controlling for income, which is, unsurprisingly, a significant predictor of both recalled and predicted expenses (*p*’s < .001). Notably, the distribution of residuals still displayed a high degree of positive skew (residual skewrecall = 4.03, residual skewpredict = 3.62), suggesting that the data are skewed because of naturally occurring expense variability and not between-subjects income variation. The longitudinal nature of study 4 allows us to revisit this topic using a within-subject measure of skew.

*Perceived Typicality.* As illustrated in the left side of figure 1, participants predicted that their future expenses would be more typical than their past expenses, (*M*pastweek = 4.71, 95% CIpastweek = [4.56, 4.86], *M*nextweek = 5.03, 95% CInextweek = [4.89, 5.16], *t*(484) = 4.22, *p* < .001, *d* = .20). Furthermore, correlational analysis showed that higher perceived typicality of future expenses was associated with lower expense predictions (*r*(483) = -.17, *p* < .01). Finally, we regressed predictions onto perceived typicality of future expenses and participant income as a robustness check designed to rule out the possibility that the correlation between predictions and typicality was due to wealthier participants with higher expenses also having more atypical expenses. Importantly, the negative relationship between predicted expenses and typicality remained significant after controlling for participant income (B = -.17, SE = .03, *t*(482) = -3.89, *p* < .001).

*Expense Prediction Bias (Recalled – Predicted Expenses)*. As illustrated in the right side of figure 1, predicted expenses for the next week were 10.80% ($19.55) lower than reported expenses for the past week, as confirmed by a paired t-test (*M*pastweek = $180.98, 95% CIpastweek = [$165.41, $198.03], *M*nextweek = $161.43, 95% CInextweek = [$147.57, $176.60], *t*(484) = 3.89, *p* < .001).

**Figure 1**

**Mean Perceived Typicality of Past vs. Future Expenses and Mean Reported vs. Predicted Expenses in Study 1**

Error Bars Represent 95% Confidence Intervals



*Ancillary Analysis*. We also investigated the distributional properties of reported versus predicted expenses. Of the 485 participants in our sample for this study, 18.6% (*n* = 90) predicted that their expenses for the next week would be the same as their expenses for the past week. Of the remaining 395 participants, 59.5% predicted that their future expenses would be lower than their past expenses versus 40.5% who predicted that their future expenses would be higher. Thus, among participants who predicted that their expenses would fluctuate, significantly more predicted that their expenses would fall rather than rise in the future (z= 3.78, 95% CI = [54.48%, 64.38%], *p* < .001).

Discussion

The results of study 2 paint a picture of consumer expense prediction that supports the prototype theory of expense prediction bias in three ways: Consumers predict their next week’s expenses will be more typical (H3) and lower (H4a) than their past week’s expenses, and perceived typicality of future expenses is negatively correlated with expense predictions (H2). It is also notable that predicted expenses displayed less skew than recalled expenses. This supports the prototype theory because prototypical prediction—or, equivalently, the neglect of atypical outcomes when predicting—should lead to a shorter right tail in the distribution of predicted (vs. recalled) expenses. In study 3, we directly replicate these results in a nationally representative sample of Americans and introduce our “atypical” intervention.

**STUDY 3: TESTING AN “ATYPICAL” INTERVENTION**

The primary purpose of study 3 is to test the effectiveness of our atypical intervention and provide insight into how it works. If the prototype theory is correct and expense predictions are prototypical, then prompting consumers to consider reasons why their expenses might be different than usual should increase prediction accuracy (H5a) by making atypical expenses more accessible (H5b). Therefore, if the intervention works as hypothesized, it will not only offer a simple way to improve consumers’ expense predictions, it will provide further support for the prototype theory. The control condition of study 3 also provides an opportunity to directly replicate the results of study 2.

Method

*Participants.* A nationally representative sample of 1,108 US residents completed study 2 via Time-Sharing Experiments for the Social Sciences. The recalled expenses of 60 participants exceeded their predictions by more than a factor of 10 (or vice versa), leaving us with an effective sample size of 1,048 (*M*age = 49.59; 53.0% female; 72.8% Caucasian, 9.4% Black, 10.7% Hispanic, 7.2% Other; Mode level of education = Bachelor’s degree; Median household annual income = $50–59,999).

*Procedure.* Participants were randomly assigned to one of three conditions: control, typical, or atypical. In the control condition, they recalled and predicted their weekly expenses for the past and next week, as in study 2. Participants in the typical condition also recalled and predicted their weekly expenses, but they received the following instructions before making their prediction: “Now consider why your expenses for next week might be *similar* to that of any other week. In the spaces provided below, please type 3 reasons why your expenses for next week might be *similar* to that of any other week.” We hypothesized that this would *not* significantly impact perceived typicality of future expenses or expense prediction amount (vs. control), because if the prototype proposition is correct, then predictions in the control condition should already be swayed by the illusion of a more typical future (H3). The atypical condition paralleled the typical condition but instructed: “Now consider why your expenses for next week might be *different* from that of any other week. In the spaces provided below, please type 3 reasons why your expenses for next week might be *different* from that of any other week.” We hypothesized that this would increase predicted expenses to the level of recalled past expenses by making atypical expenses more accessible (H5a). The order of prediction and recall was counterbalanced in all conditions.[[1]](#footnote-1)

We next presented participants with an atypical expense-listing task that asked, “Is there anything you believe you will spend money on in the next week that you did NOT spend money on during the past week?” and “Is there anything that you spent money on during the past week that you believe you will NOT spend money on in the next week?” Participants were then given the opportunity to list a description and corresponding dollar amount for up to five such expenses. Our principal expectation was that the atypical intervention would make it easier to retrieve atypical expenses for the next week, which would result in a higher number of expenses being listed in the atypical condition as compared to the control and typical conditions. Furthermore, we expected that the number of atypical expenses listed for the next week would mediate the relationship between experimental condition and predicted expenses (H5b).

Finally, participants completed the same measures of perceived typicality used in study 2, and five exploratory measures designed to let us explore the relationship between expense prediction bias (EPB), financial slack (Berman et al. 2016; Zauberman and Lynch 2005), various measures of spending (e.g., willingness to pay for an optional expense like a fancy dinner out with friends), and available resources. These exploratory measures yielded null results that are discussed in web appendix C.

Results

 *Skew.* Both recalled and predicted expenses in the control condition once again displayed a high degree of positive skew (skewrecall = 7.65, skewpredict = 3.82), as did the distribution of residuals generated by regressing expenses onto income (residual skewrecall = 7.58, residual skewpredict = 3.70).

*Replicating Study 2.* As illustrated in figure 2, the results observed in study 2 were directly replicated in the control condition of study 3, in which participants predicted their future expenses would be more typical than their past expenses (*M*pastweek = 4.40, 95% CIpastweek = [4.23, 4.57], *M*nextweek = 4.65, 95% CInextweek = [4.48, 4.81], *t*(415) = -3.42, *p* = .001, *d* = .17), supporting H3, and 8.6% lower than their past expenses (*M*pastweek = $237.46, 95% CIpastweek = [$217.02, $262.43], *M*nextweek = $217.02, 95% CInextweek = [$196.37, $239.85], *t*(415) = 2.76, *p* = .006), supporting H4a. Furthermore, higher perceived typicality of future expenses was again associated with lower expense predictions (*r*(414) = -.21, *p* < .001), supporting H2, and this association remained significant after controlling for participant income (B = -.14, SE = .03, *t*(413) = -4.98, *p* < .001). We next expand our analyses to test for differences in perceived typicality and EPB across all three conditions.

**Figure 2**

**Mean Perceived Typicality of Past vs. Future Expenses and Mean Reported vs. Predicted Expenses in the Control Condition of Study 2**

Error Bars Represent 95% Confidence Intervals

****

*Perceived Typicality of Future Expenses*. A one-way ANOVA with intervention condition (control vs. typical vs. atypical) as the independent variable and perceived typicality of future expenses as the dependent variable revealed a significant effect of condition on perceived typicality of future expenses (*F*(2, 1044) = 32.27, *p* < .001). Planned contrasts further revealed that perceived typicality was virtually identical in the control and typical conditions (*M*control = 4.64, 95% CIcontrol = [4.48, 4.81], *M*typical = 4.65, 95% CItypical = [4.48, 4.83], *t*(1044) = .03, *p* = .97), but significantly lower in the atypical condition (*M*atypical = 3.74, 95% CIatypical = [3.56, 3.93], *t*(1044) = -8.02, *p* < .001, *d* = .55).

*Expense Prediction Bias (Recalled – Predicted Expenses)*. Predicted expenses were 8.6% ($20.44) lower than recalled expenses in the control condition (*t*(415) = 2.76, *p* = .006) and 6.4% ($13.68) lower than recalled expenses in the typical condition (*t*(331) = 2.07, *p* = .039), but predicted expenses did not differ from recalled expenses in the atypical condition (*t*(299) = -1.49, *p* = .14). In other words, EPB was neutralized (and even slightly reversed) by our atypical intervention, supporting H5a, as illustrated in figure 3. A 3 (intervention: control vs. typical vs. atypical) × 2 (time period: past week vs. next week) between-within ANOVA with expenses as the dependent variable confirmed a significant main effect of intervention condition (*F*(2, 1045) = 4.64, *p* = .010), no main effect of time period (*F*(1, 1046) = 1.69, *p* = .19), and a significant intervention-by-time-period interaction (*F*(2, 1045) = 5.22, *p* = .006). Planned contrasts confirmed that predicted expenses in the atypical condition (*M*atypical = $273.14, 95% CIatypical = [$242.26, $307.97]) were 25.9% ($56.12) higher than in the control condition (*M*control = $217.02, 95% CIcontrol  = [$196.37, 239.85]; *t*(1045) = 2.91, *p* = .004), and 35.0% ($70.79) higher than in the typical condition (*M*typical = $202.35, 95% CItypical = [$181.27, $223.63]; *t*(1045) = 3.67, *p* < .001). Predictions did not differ between the control and typical conditions (*t*(1045) = .98, *p* = .33). Planned contrasts also revealed that recalled expenses did not differ between the atypical (*M*atypical = $254.68, 95% CIatypical = [$223.63, $287.15]) and control conditions (*M*control = $237.46, 95% CIcontrol = [$217.02, $262.43]; *t*(1045) = .79, *p* = .43), but they were somewhat lower in the typical condition (*M*typical = $214.86, 95% CItypical = [$192.48, $239.85]) than in the atypical condition (*t*(1045) = 2.01, *p* = .045).[[2]](#footnote-2) Notably, this makes our test of EPB between these two conditions quite conservative because lower (higher) recalled expenses decreases (increases) the size of the bias. However, despite lower recalled expenses in the typical condition and higher recalled expenses in the atypical condition, we observe a significant bias in the former but not in the latter.

**Figure 3**

**Mean Recalled vs. Predicted Expenses in the Control vs. Typical vs. Atypical Conditions in Study 3**

Error Bars Represent 95% Confidence Intervals

****

*Atypical Expense-Listing Task.* A one-way ANOVA with condition (control vs. typical vs. atypical) as the IV and number of atypical expenses (i.e., expenses predicted to occur in the next week that didn’t occur in the past week) as the DV revealed a significant effect of condition (*F*(2, 1045) = 6.80, *p* < .011). Planned contrasts further indicated that the number of expenses listed in the atypical condition (*M*atypical  = 1.57, SDatypical = 1.60) was significantly higher than in the control and typical conditions (*M*control = 1.21, SDcontrol = 1.48, *M*typical = 1.16, SDtypical = 1.52; *t*(1045) = 6.80, *p* = .001), consistent with H5b; and the number of expenses listed in the control and typical conditions did not differ (*t*(1045) = .34, *p* = .74). This provides evidence that our intervention makes atypical expenses easier to retrieve during prediction than they would be otherwise. The same ANOVA with average dollar amount of atypical expenses as the DV revealed no effect of condition (*F*(2, 598) = .74, *p* = .47).

*Mediation Analysis.* The results above confirm that our atypical manipulation succeeded in making atypical expenses easier to retrieve, and that expense predictions were significantly higher in the atypical condition as well. To further investigate the relationship between the number of atypical expenses that participants listed and expense predictions, we tested a mediation model with condition (atypical = 1 vs. control and typical = 0) as the independent variable, expense prediction as the dependent variable, and the number of atypical expenses listed as the mediating variable. The indirect effect of condition on expense prediction via number of atypical expenses was significant (indirect effect = .05, SE = .02, 95% CI = [.02, .09]). Specifically, the model confirms that the atypical intervention succeeded in increasing the number of atypical expenses listed (b = -.38, 95% CI = [.18, .59]; *t*(1046) = 3.66, *p* < .001), and that this number was associated with higher expense predictions, even while controlling for condition (b = -.14, 95% CI = [.10, .18]; *t*(1045) = 6.67, *p* < .001), supporting H5b.[[3]](#footnote-3)

Discussion

 Study 3 provides support for the prototype theory in three ways. First, it directly replicates the findings of study 2 with a nationally representative sample, demonstrating that consumers predict their future expenses will be both more typical (H3) and lower than their past expenses (H4a), and that perceived typicality of future expenses is negatively correlated with predictions (H2). Second, it demonstrates that the atypical intervention can increase predicted expenses to the level of recalled expenses by making atypical expenses easier to retrieve (H5a and H5b). This not only supports the prototype theory by providing evidence that atypical expenses are relatively inaccessible when consumers generate their predictions, it also offers a simple tool that can be easily applied to help consumers improve their prediction accuracy. Finally, while being careful not to over-claim the importance of a null result, we believe the fact that perceived typicality and expense predictions do not differ between the control and typical conditions supports the prototype theory by suggesting that predictions in these two conditions were equally prototypical.

The operationalization of EPB in studies 2 and 3 (recalled – predicted expenses) is useful for both theory building and hypothesis testing, but its natural limitation is that it doesn’t reveal true prediction accuracy. To address this, study 4 operationalizes EPB as the difference between predicted expenses and the expenses each participant actually incurs during the target week or month, and measures prediction accuracy within-subject using a repeated-measures longitudinal design. This allows us to more accurately measure the magnitude, prevalence, and persistence of the bias, and also gives us a within-person measure of expense-distribution skew.

**STUDY 4: A LONGITUDINAL STUDY OF EPB IN THE FIELD**

The purpose of study 4 was to test our hypotheses in the field. To do so, we partnered with a financial cooperative to run a five-week longitudinal field study with 187 of its members. This study let us observe the magnitude, prevalence, and persistence of the bias within a highly engaged sample across multiple points in time. Study 4 also allowed us to test and validate the effectiveness of the atypical intervention for increasing predicted expenses to the level of actual expenses.

Method

*Participants.* Participants for this study were members of a midsized financial cooperative (~60,000 members) recruited through an online panel of members (~5,000) that the cooperative maintains in order to conduct market research. We targeted a sample size of 200 based on the effect sizes observed in our previous studies. Each participant completed six surveys over the course of five weeks, as illustrated by each time period marked in figure 4.

**Figure 4**

**Data Collection Schedule in Study 4**



Because we had no prior experience sampling from this population, data collection took place in two waves. In wave 1, we sent a survey at time zero (T0) to 400 randomly selected members at noon on Sunday, September 10th, 2017. Ninety-three people completed the survey before it was deactivated at 11:59pm on Monday, September 11th, 2017. We then monitored attrition for two weeks before calculating that the T0 survey should be sent to another 800 randomly selected members (the maximum number allowed by our field partner) in the second wave of data collection so that we could recruit as close to 200 total participants as possible. In wave 2, 219 members completed the T0 survey. At the end of both waves of data collection we had collected complete data from 187 participants (61 from wave 1 and 126 from wave 2, *M*age = 51.12, 57.8% female). Compensation for each participant included a personalized spending report (provided at the end of the study) that served as an incentive to predict and report expenses as accurately as possible. Participants also received a $10 Amazon gift certificate for each completed survey.

*Procedure.* Each of the six surveys in this five-week field study was emailed to participants at noon on a Sunday and required completion before 11:59 pm the next day. The first survey asked participants to predict their expenses for the coming week and month, and indicate how similar or different they expected their expenses to be relative to a typical week and month. The remaining five surveys began by asking participants to log into their online bank account and report their expenses for the previous week, then predict their expenses for the upcoming week. Both expense reports and predictions were followed by the same measure of typicality used in survey 1 (“How different or similar do you think your expenses were [will be] for the **past** [**next**] week, relative to a typical week?” 1 = very different, 7 = very similar). In the second-to-last survey, half of the sample was randomly assigned to receive our atypical intervention, making the final week of the study a 2 (condition: control vs. atypical) × 2 (expenses: predicted vs. actual) between-within design.[[4]](#footnote-4) [[5]](#footnote-5) Each survey also included measures of prediction confidence (“How sure or confident are you that your estimate of your total expenses for the next week is accurate?” 1 = very unsure, 7 = very sure), individual difference measures such as optimism (Scheier et al. 1994) and numeracy (Schwartz et al. 1997), and some exploratory questions about spending and borrowing behavior. These measures yielded mostly null results and are summarized in web appendix D.

Results

 *Skew.* The expense data in studies 2 and 3 demonstrate that reported expenses are significantly more skewed than predicted expenses across participants at a single point in time, and that this is true even when controlling for income. Because study 4 measures each participant’s expenses for multiple weeks in a row, it offers a unique opportunity to follow up on this result by evaluating expense skew on a within-subject basis. To do this, we first calculated the mean, median, standard deviation, and skewness of each participant’s reported expenses over the course of the study. We then aggregated and averaged these values to estimate distributional parameters for the average consumer in our sample. Finally, we repeated this exercise for predicted expenses. As can be seen in table 1, the distribution of predicted expenses displays less skew and variance than the distribution of recalled expenses. This supports the prototype theory because prototypical prediction should lead to a tighter distribution if atypical expenses are being neglected. Moreover, the mean and median of predicted expenses are significantly lower than the mean and median of past expenses, which supports the expense prediction bias hypotheses (H4a and H4b).

**Table 1: Distributional Parameter Estimates for the Average Consumer in Study 4**



*Perceived Typicality.*To test our hypothesis that people predict their expenses will be more typical in the future than in the past (H3), we compared reported and predicted expense typicality at T1, T2, T3, and T4. In other words, we tested whether or not participants predicted their expenses would be more typical in week 2 than week 1, week 3 than week 2, and so on. As illustrated in figure 5 and table 2, participants predicted their expenses would be more typical in the next (vs. past) week at all four points in time. Figure 5 also shows that our atypical intervention succeeded in neutralizing (and even slightly reversing) this tendency at T4. In sum, these results provide strong support for H3.

To test our hypothesis that perceived typicality of future expenses is negatively correlated with expense predictions (H2), we analyzed the correlation between perceived typicality of future expenses and weekly expense predictions for each week of the study, as well as for the month. Perceived typicality of future expenses was negatively correlated with weekly expense predictions at T0 (*r*(185) = -.30, *p* < .001), T2 (*r*(185) = -.28, *p* < .001), and T4 (*r*(185) = -.25, *p* = .001). The correlation at T1 was marginally significant (*r*(185) = -.12, *p* = .09), as was the correlation between perceived typicality of monthly expenses and expense predictions for the month (*r*(185) = -.12, *p* = .09). The correlation at T3 was directionally consistent, though not significant (*r*(185) = -.03, *p* = .74). Regressing predictions onto perceived typicality and participant income produced only one meaningful change in these results: The relationship between perceived typicality and prediction at T1 became significant (*p* = .013). In aggregate, this replicates the findings of studies 2 and 3 with respect to H2.

**Figure 5**

**Mean Reported Expense Typicality for the Past Week vs. Mean Predicted Expense Typicality for the Next Week for Each Week of Study 4**

Error Bars Represent 95% Confidence Intervals



**Table 2**

**T-tests Comparing Reported vs. Predicted Typicality in Study 4**



 *Expense Prediction Bias (Recalled – Predicted Expenses).* To test our hypothesis that consumers tend to predict their future expenses will be lower than their past expenses (H4a), we compared reported expenses against predicted expenses at T1, T2, T3, and T4. That is, we tested whether participants predicted their expenses would be lower in week 2 than week 1, week 3 than week 2, and so on. As illustrated in figure 6 and table 3, mean predicted expenses were significantly lower than mean reported expenses at each stage of the study until the atypical intervention was deployed. This longitudinally replicates the findings of studies 2 and 3 and provides strong support for H4a.

**Figure 7**

**Mean Expenses Incurred in the Past Week vs. Mean Predicted Expenses for the Next Week for Each Week of Study 4**

Error Bars Represent 95% Confidence Intervals



**Table 3**

**T-tests Comparing Reported Past-Week vs. Predicted Next-Week Expenses in Study 4**



 *Expense Prediction Bias (Actual – Predicted Expenses).* As illustrated in figure 7 and table 4, mean predicted expenses were significantly lower than mean incurred expenses in each week of the study, except during week 5 in the atypical condition, in which our intervention completely neutralized mean expense prediction bias. A 2 (condition: control vs. atypical) × 2 (expenses: predicted vs. actual) between-within ANOVA confirmed a significant condition-by-expenses interaction (*F*(1, 181) = 5.08, *p* = .025). Planned contrasts further confirmed that predicted expenses were 36.7% higher in the atypical (vs. control) condition (*F*(1, 181) = 4.48, *p* = .036), and that actual expenses did not differ by condition (*F*(1, 181) = .44, *p* = .51).

 In dollar terms, EPB in the control condition was $79.99 (different from zero, t(91) = 3.19, *p* = .002) versus –$6.65 in the atypical condition (not different from zero, *t*(90) = -.20, *p* = .85). It is also notable that our intervention did not require sacrificing correlational accuracy (*r*control(90) = .81, *r*treatment(89) = .80, z = .19, *p* = .85). These findings provide support for H4b and H5a.

**Figure 7**

**Mean Predicted Expenses vs. Mean Incurred Expenses for Each Week of Study 4**

Error Bars Represent 95% Confidence Intervals



**Table 4**

**T-tests Comparing Predicted Expenses vs. Incurred Expenses in Study 4**



*Monthly Expense Prediction Accuracy.* We were also able to test the accuracy of participants’ expense predictions for the next month by comparing them to the expenses they actually incurred in weeks 1 through 4 of the study. Results showed that predicted expenses for the target month (*M*pred = $2276.74, 95% CI = [2031.64, 2551.15]) were $416.77 (15.5%) lower than expenses actually incurred (*M*actual = $2693.51, 95% CI = [2376.06, 3053.37]; *t*(184) = 3.85, *p* < .001). To put the size of this monthly EPB in perspective, consider that 46% of Americans report they do not have enough money to cover a $400 emergency expense (US Federal Reserve 2016). To the best of our knowledge this provides the first evidence that consumers under-predict their monthly as well as weekly expenses.

*The Atypical Intervention vs. an Outside View*. One decision rule that consumers could use to make their expense predictions is to simply project forward their past expenses. Indeed, taking this kind of “outside view” and basing predictions on past behavior is well established as a successful intervention in the literature on the planning fallacy (Buehler, Griffin, and Ross 1994; Kahneman and Tversky 1979). Notably, the structure of study 4 allows us to compare the effectiveness of the outside view against the effectiveness of the atypical intervention in two ways. First, we can average each control-condition participant’s expenses over the first four weeks of the study, replace their prediction at the start of week 5 with that value, and compare their new “prediction” accuracy against prediction accuracy in the atypical condition. This exercise reveals that if participants in the control condition had taken an outside view they would have significantly *over*-predicted their expenses (EPBoutsideview = -$116.05, different from zero, *t*(91) = -2.92, *p* = .004), whereas participants in the atypical condition displayed remarkably high mean prediction accuracy (EPBatypical = –$6.65, not different from zero, *t*(90) = -.20, *p* = .85). Furthermore, taking an outside view would have led to lower correlational accuracy in the control versus atypical condition (*r*outsideview(90) = .69, *r*atypical(89) = .80, z = 1.65, *p* = .050). The second way to compare the effectiveness of the outside view and atypical intervention is to compare prediction accuracy in the atypical condition against what would have happened if those participants had taken an outside view. In this case, the accuracy of participants in the atypical condition would have been directionally worse (EPBoutsideview = -$48.59, not different from zero, *t*(89) = -1.08, *p* = .28), and their correlational accuracy would have been significantly worse (*r*outsideview(89) = .65, *r*atypical(89) = .80, z = 2.13, *p* = .017). These analyses are exploratory and by no means conclusive, but they do suggest the exciting possibility that drawing people’s attention to atypical expenses is not only sufficient to increase prediction accuracy, it may actually be superior to drawing their attention to the more complete distribution of past expenses.

Discussion

 The results of study 4 offer longitudinal field evidence that provides compelling support for our prototype theory outside the lab. First, consumers predicted that their expenses would be more typical (H3) and lower (H4a) than their past expenses in each and every week of the study. Second, perceived typicality was shown to be inversely related to expense predictions in almost every time period of the study (H2). Third, consumers persistently under-predicted their weekly expenses as well as under-predicted their expenses for the month (H4b). Fourth, it was shown that the atypical intervention is capable of virtually eliminating the expense prediction bias in a real-world setting (H5a). It was also shown that the atypical intervention outperforms an outside-view decision-rule that incorporates the full distribution of (recent) expenses. Finally, the results of study 4 demonstrate that the magnitude of the expense prediction bias—approximately $100/week or $400/month—is economically significant.

**STUDY 5: (A)TYPICALITY IN WEEKLY VS. MONTHLY PREDICTIONS**

The primary goal of study 5 was to determine if perceived typicality of future expenses differs between weekly and monthly predictions. This is important from a theoretical perspective because it is possible that prototype attributes do not influence predictions for longer periods of time. For example, a more distant prediction horizon may be associated with more uncertainty, which could cause consumers to build an error term into their prediction that accounts for atypical expenses (Ulkumen et al. 2008). If true, we would expect to observe lower perceived typicality for monthly (vs. weekly) predictions in the absence of an intervention. It would also be reasonable to expect a weaker effect of the atypical intervention on perceived typicality for monthly (vs. weekly) predictions, because if monthly predictions are less influenced by prototype attributes there should be less room for perceived typicality to be shifted downward by the intervention. Study 5 tests these possibilities, and examines if the atypical intervention is capable of increasing monthly as well as weekly expense predictions. This is a matter of practical importance given that study 4 demonstrates that an expense prediction bias exists for both time periods.

Study 5 does not allow us to apply the same within-subject exclusion criterion used in studies 2–4 because it measures only predictions. We therefore opted to winsorize predictions at the 5th and 95th percentile instead. This is detailed in the study preregistration, which is included with the study materials and data file here: [insert hyperlink].

Method

 *Participants and Procedure.* We recruited 601 participants (48.6% female, *M*age = 37.93) from Amazon Mechanical Turk to take part in a consumer expense survey. Participants were randomly assigned to predict their expenses in a 2 (prediction time frame: week vs. month) × 2 (intervention condition: control vs. atypical) between-subjects design that utilized the same prediction prompts as in study 4, modified where necessary for monthly predictions. Participants again reported perceived typicality and prediction confidence.

Results

*Weekly vs. Monthly Prediction Typicality.* Perceived typicality did not differ as a function of time period, only as a function of intervention condition. Specifically, a 2 (prediction time period: week vs. month) × 2 (intervention condition: control vs. atypical) ANOVA with perceived typicality as the dependent variable revealed a main effect of intervention condition such that perceived typicality was higher in the control condition (*M* = 5.20, SD = 1.55) than in the atypical condition (*M* = 3.11, SD = 1.59; *F*(1, 597) = 263.73, *p* < .001, partial eta squared = .31). However, there was no main effect of time period (*F*(1, 597) = .72, *p* = .40) and no interaction of intervention condition by time period (*F*(1, 597) = 2.09, *p* = .15).

*Monthly Expense Predictions*. The impact of the atypical intervention on monthly expense predictions was consistent with its impact on weekly expense predictions in studies 3 and 4: Monthly expense predictions were 24.6% higher in the atypical condition (*M* = $1505.98, 95% CI = [$1314.88, $1724.86]) than in the control condition (*M* = $1208.34, 95% CI = [$1049.32, $1391.59]), as revealed by an independent-samples t-test (*t*(297) = 2.22, *p* = .027). Furthermore, perceived typicality was found to be significantly lower in the atypical condition (*M* = 2.97, 95% CI = [2.74, 3.21]) than in the control condition (*M* = 5.24, 95% CI = [4.99, 5.50]; *t*(297) = 12.91, *p* < .001). Prediction confidence did differ between the two conditions (*M*atypical = 4.79, 95% = [4.56, 5.01]; *M*control = 5.39, 95% CIcontrol = [5.19, 5.60]; *t*(297) = 3.94, *p* < .001), but the focal comparison between predicted expenses in the atypical and control conditions remained significant after controlling for confidence (*F*(1, 296) = 8.33, *p* = .004).

*Weekly Expense Predictions.* The pattern of results observed in studies 3 and 4 was replicated in study 5: Weekly expense predictions were 61.0% higher in the atypical condition (*M* = $335.26, 95% CI = [$290.03, $387.53]) than in the control condition (*M* = $208.20, 95% CI = [$181.22, $239.20]), as revealed by an independent-samples t-test (*t*(300) = 4.67, *p* < .001). Furthermore, perceived typicality was found to be significantly lower in the atypical condition (*M* = 3.27, 95% CI = [2.98, 3.55]) than in the control condition (*M* = 5.16, 95% CI = [4.93, 5.40]; *t*(300) = 10.16, *p* < .001). Finally, it was found that prediction confidence did not differ by condition (*M*atypical = 5.20, 95% = [4.99, 5.40]; *M*control = 5.35, 95% CIcontrol = [5.17, 5.54]; *t*(300) = 1.13, *p* = .26).

Discussion

 The results of study 5 revealed that perceived typicality of future expenses differed only as a function of intervention condition (control vs. atypical) and not as a function of prediction time period (week vs. month). Moreover, the impact of the atypical intervention on perceived typicality was roughly equal across time periods. In tandem, these findings suggest that the influence of prototype attributes on predictions—and the influence of the atypical intervention—is similar across these two common prediction time frames. The results of study 5 also demonstrate that prompting consumers to consider atypical expenses can increase monthly expense predictions as well as weekly expense predictions. Taken together with the results of study 4—which showed that consumers under-predict their monthly expenses as compared to their actual expenses during the target month—this suggests that the atypical intervention is capable of reducing monthly expense prediction bias as well as its weekly cousin.

**SUMMARY OF DESCRIPTIVE RESULTS ACROSS STUDIES**

A notable contribution of the present research is that it provides the first step toward a comprehensive understanding of the expense prediction bias itself. In other words, this is the first research to identify the magnitude, prevalence, and persistence of EPB in non-student samples, in a field study, and over time. Table 5 summarizes the descriptive results from the control conditions in studies 2–4, and several observations stand out. First, the magnitude of the bias varies substantially across studies. Although it is statistically significant in all studies, it is notably smaller when measured as the difference between *recalled* and predicted expenses than when measured as the difference between *actual* and predicted expenses. Each study samples from a different population, so we can’t say for sure if these differences are due to measurement method or sample characteristics, but systematically investigating measurement of the bias is a fruitful avenue for future research.

The prevalence of the bias is also notable. Across studies the percentage of participants who under-predicted (vs. over-predicted) their expenses is consistently greater than 50%. This suggests that EPB is a common phenomenon, not just driven by a handful of consumers who are particularly bad at expense prediction. Finally, the week-by-week results from study 4 demonstrate that the bias persists over time. In sum, these descriptive statistics provide evidence that the magnitude of EPB is economically significant, EPB is prevalent, and it is persistent. In other words, it is a phenomenon of substantive importance.

**Table 5**



**GENERAL DISCUSSION**

The present research develops and tests a prototype theory of consumer expense misprediction. The logic underlying the theory is that consumer expense predictions are shaped by prototype attributes that represent typical or modal expenses. Prototypical prediction then leads to an expense prediction bias in which consumers under-predict their expenses because the distribution of expenses is positively skewed with mode < mean. This led us to develop an “atypical” intervention designed to improve prediction accuracy by making atypical expenses easier to retrieve when consumers make their predictions.

Five studies provide support for the prototype theory and the effectiveness of the atypical intervention. Study 1 demonstrates that when the distribution of expenses is experimentally manipulated, a positively skewed distribution (with mode < mean) causes predictions that are lower and farther away from the mean than a normal distribution (with mode = mean), holding mean, range, and variance constant across conditions (H1). This provides direct support for the prototype theory because if prototypical predictions are shaped by modal outcomes, it follows that a lower mode should lead to lower predictions.

Study 2 begins by illustrating that the distribution of past expenses has stronger skew than the distribution of predicted expenses, even when controlling for income. This supports the prototype theory because prototypical predictions should have a shorter right tail if they neglect atypical outcomes. Study 2 then demonstrates that consumers predict their expenses for the next week will be more typical than the past week (H3). This supports the prototype theory because past expenses include atypical expenses, so it is natural to expect that prototypical prediction would cause this kind of temporal asymmetry. Study 2 also demonstrates that perceived typicality of future expenses is negatively correlated with expense predictions (H2), even when controlling for income. This supports the prototype theory because if typical expenses are modal expenses, and expenses are positively skewed (with mode < mean), then it follows that more "typical" predictions should be lower predictions. Finally, study 2 supports the existence of an expense prediction bias by demonstrating that participants predict lower expenses for the next week than the past week (H4a), an outcome that follows naturally from H2 and H3.

Study 3 begins by directly replicating the results of study 2 with a nationally representative sample of Americans. It then tests the effectiveness of the atypical intervention (listing three reasons why next week’s expenses might be different from a typical week before predicting) against a pure control condition and an experimental control condition that asks participants to list three reasons why their expenses might be similar to a typical week. The results demonstrate that expense prediction accuracy, defined in this study as the difference between predicted expenses for the next week and recalled expenses for the past week, is significantly higher in the atypical condition than in the control conditions (H5a). The results also show that participants were able to list a higher number of atypical expenses in the atypical condition (vs. controls), suggesting that the intervention increases prediction accuracy by making these expenses easier to retrieve. Finally, a mediation model with condition as the IV, number of atypical expenses listed as the mediator, and expense predictions as the DV demonstrates that the number of atypical expenses listed mediates the relationship between the atypical intervention (vs. controls) and predictions (H5b). Taken together, these results support the prototype theory by demonstrating that under-prediction can be ameliorated by bringing atypical expenses to mind.

Study 4 tests our hypotheses in a repeated-measures longitudinal field study that allows us to: a) examine the skew of expenses on a within-subject basis, b) test the persistence of the effects documented in studies 2 and 3 over time, and c) evaluate true expense prediction accuracy, defined as the difference between predicted expenses for a week or month and the expenses incurred during that week or month (H4b). On average, the within-subject distribution of recalled expenses has a higher mean, median, and standard deviation than the distribution of predicted expenses; it also has stronger positive skew. This bolsters the prototype theory by demonstrating that distribution neglect, which is a defining characteristic of prototypical prediction, is in fact a within-subject phenomenon. Study 4 also reveals that the effects predicted by H2–H4b are remarkably persistent over time, and that the atypical intervention is capable of virtually eliminating the expense prediction bias when it is defined as the difference between predicted and actual expenses.

Finally, study 5 demonstrates that the perceived typicality of predicted monthly expenses does not differ from the perceived typicality of predicted weekly expenses, and that the effect of the atypical intervention on perceived typicality is equivalent across these time frames. Taken together, these results suggest that prototype attributes influence monthly expense predictions in the same way that they influence weekly expense predictions. Study 5 also demonstrates that the atypical intervention increases monthly as well as weekly predictions. Given that study 4 documented a monthly expense prediction bias of over $400, this result is of practical importance because it suggests that the atypical intervention can improve consumers’ monthly expense prediction accuracy as well as their weekly prediction accuracy.

Contributions and Directions for Future Research

By developing and testing a prototype theory of consumer expense misprediction, the present research makes several contributions. First, the prototype theory deepens our understanding of the psychology of misprediction by offering a parsimonious explanation for why the expense prediction bias occurs. The prototype theory also offers a way to neutralize the bias, and the present research is the first to introduce a practical, effective, and field-tested intervention that achieves this goal. Furthermore, because the prototype theory offers a single explanation for distribution neglect in expense prediction (Peetz and Buehler 2012; Peetz et al. 2015), the underestimation of exceptional expenses in prediction (Sussman and Alter 2012), and the inverse relationship between expense prediction confidence and adjustment (Ulkumen et al. 2008), the prototype theory also offers a lens through which past findings in this area can be understood as a whole for the first time. Indeed, one exciting direction for future research is to extend the prototype theory by studying the direct links between these variables.

The present research also contributes an important first step toward a comprehensive understanding of the expense prediction bias itself. For example, we present the first studies to identify the magnitude, prevalence, and persistence of the bias in non-student samples. We are also the first to study the bias longitudinally and in the field, and to measure monthly expense predictions against actual expenses for the target month. This approach allows us to claim with some certainty that the expense prediction bias is meaningful in a real-world sense, and it is therefore a substantively important context in which marketing scholars can develop theory and test their hypotheses. A second direction for future research is to identify the consequences of the expense prediction bias and quantify the impact of neutralizing the bias.

In addition to its contributions to the literature on consumer expense misprediction, the present research makes an important contribution to the more general literature on consumer misprediction. Past work on phenomena such as the planning fallacy has tackled the problem of distribution neglect in prediction by prompting people to ground their predictions in the full distribution of either relevant past outcomes (Buehler, Griffin, and Ross 1994) or possible future outcomes (Kruger and Evans 2004; Peetz et al. 2015). Although these strategies have been shown to increase prediction accuracy, they don’t inform our understanding of what distributional information is being neglected, or how the shape of the distribution influences predictions. Therefore, the prototype theory and the atypical intervention advance knowledge in this regard by specifying what distributional information is and isn’t included in predictions. A third direction for future research is to generalize the prototype theory to other contexts such as the planning fallacy.

The present research also advances knowledge about the use of prototype attributes in judgment and decision making. Theorizing in this area has implied that the “average” represented by prototype attributes is often a simple mean. For example, Kahneman and Frederick (2002) conceptualize the peak-end rule—the phenomenon that global evaluations of a temporally extended experience can be predicted by averaging the peak and the end of the experience (Redelmeier and Kahneman 1996)—as an instance in which people are substituting a prototype attribute for an extensional one. Similarly, Kahneman (2003) offers an example of prototype attribute substitution in which the prototype is the mean of two sets of outcomes. It is therefore a notable contribution of the present research that, in the case of expense predictions at least, prototype attributes represent outcomes that follow the *mode* of a distribution. A fourth direction for future research is to investigate when prototypes represent the mode, mean, median, or some combination.

Finally, the present research contributes to a nascent literature on temporal asymmetry, which hypothesizes that people mentally represent the future in more prototypical terms than the past (Kane, Van Boven, and McGraw 2012; Van Boven, Kane, and McGraw 2009). By comparing perceived typicality of past versus future expenses, we extend this work to the domain of money. This provides a notably conservative test of the temporal asymmetry hypothesis because money is a relatively concrete and predictable resource (MacDonnell and White 2015; Zauberman and Lynch 2005), whereas the hypothetical people, places, and events that participants have been asked to mentally represent in other studies of temporal asymmetry are arguably much more ambiguous. Therefore, because prototypes are generalizations, it is reasonable to believe that people will rely on prototypes less when they are mentally representing a resource like money, which has very specific uses. Nonetheless, we find strong evidence that representations of future expenses (predictions) are more prototypical than representations of past expenses (recall). A fifth direction for future research is to compare and contrast the strength of asymmetries in resources like time versus money.

Practical Implications for Consumers and Firms

As outlined in the introduction to this article, there are many reasons to believe that improving expense prediction accuracy can improve consumers’ financial outcomes. Therefore, the prosocial benefit of the present research is clear—any consumer can make use of the atypical intervention to improve his or her expense prediction accuracy. Furthermore, the simplicity of the intervention means it can be easily implemented at scale by financial literacy organizations. Currently, the modal approach of such organizations is to educate their stakeholders about debits and credits, interest rates, and so on. We posit that a cognitive tool like the atypical intervention can serve as an important complement to these technical skills because it does not need to be learned, per se, but it can be easily provided and used to effectively increase prediction accuracy.

The present research also has practical implications for for-profit firms. For example, companies in the FinTech sector that are developing and managing budgeting apps can leverage our results to design their products in a way that helps users set more realistic budgets. Given that 63% of North Americans with a smartphone have at least one financial app on their phone (Barba 2018)—the key function of which is often budgeting—this could confer a substantial product advantage. Furthermore, because our theory is compatible with past research on misprediction in multiple domains, we believe that the atypical intervention can also be used to inform the design of products that aim to improve consumers’ predictions with respect to calories, exercise, time management, and a host of other variables that can positively impact consumers’ well-being.

**REFERENCES**

Addis, Donna Rose, Alana T. Wong, and Daniel L. Schacter (2007), "Remembering the past and

 imagining the future: Common and distinct neural substrates during event construction and elaboration," *Neuropsychologia,* 45 (7), 1363-1377.

Alter, Adam L. and Daniel M. Oppenheimer (2009), "Uniting the tribes of fluency to form a metacognitive nation," *Personality and Social Psychology Review,* 13 (3), 219-235.

André, Quentin, Nicholas Reinholtz, and Bart de Langhe (2017),"Variance Spillover in Intuitive Statistical Judgments," in *NA - Advances in Consumer Research*, 45, ed. Ayelet Gneezy, Vladas Griskevicius, and Patti Williams, Duluth, MN: Association for Consumer Research, 336-340.

Ariely, D. (2001), “Seeing sets: Representation by statistical properties,” *Psychological Science, 12,* 157–162.

Barba, R. (2018), <https://www.bankrate.com/personal-finance/smart-money/americans-and-financial-apps-survey-0218/>

Berman, Jonathan Z., An TK Tran, John G. Lynch Jr, and Gal Zauberman (2016), "Expense neglect in forecasting personal finances," *Journal of Marketing Research,* 53 (4), 535-550.

Buehler, Roger, Dale Griffin, and Heather MacDonald (1997), "The role of motivated reasoning in optimistic time predictions," *Personality and Social Psychology Bulletin,* 23 (3), 238-247.

Buehler, Roger, Dale Griffin, and Johanna Peetz (2010), "The planning fallacy: Cognitive, motivational, and social origins," *Advances in Experimental Social Psychology*, 43, Academic Press, 1-62.

Buehler, Roger, Dale Griffin, and Michael Ross (1994), "Exploring the ‘planning fallacy’: Why people underestimate their task completion times," *Journal of Personality and Social Psychology,* 67 (3), 366.

CB Insights (2018), Global FinTech Report Q1 2018. [https://www.cbinsights.com/research/report/FinTech-trends-q1-2018/](https://www.cbinsights.com/research/report/fintech-trends-q1-2018/)

Consumer Federation of America (2018), <https://paydayloaninfo.org/facts>

Consumer Financial Protection Bureau (2017), CFPB Financial Well-Being Scale: Scale Development Technical Report. Accessed April 14, 2018. [https://s3.amazonaws.com/files.consumerfinance.gov/f/documents/201705\_cfpb\_financia l-well-being-scale-technical-report.pdf](https://s3.amazonaws.com/files.consumerfinance.gov/f/documents/201705_cfpb_financia%09l-well-being-scale-technical-report.pdf)

Epley, Nicholas and Thomas Gilovich (2006), "The anchoring-and-adjustment heuristic: Why the adjustments are insufficient," *Psychological Science,* 17 (4), 311-318.

Federal Reserve Bank of New York (2018), Household Debt Continues Its Increase in the First Quarter of 2018. <https://www.newyorkfed.org/microeconomics/hhdc.html>.

Fellowes, Matthew and Willemin, Katy (2013), The Retirement Breach in Defined Contribution Plans. *HelloWallet Research Reports*.

Gilbert, Daniel T., Elizabeth C. Pinel, Timothy D. Wilson, Stephen J. Blumberg, and Thalia P. Wheatley (1998), “Immune neglect: A source of durability bias in affective forecasting," *Journal of Personality and Social Psychology,* 75 (3), 617.

Gilbert, Daniel, Carey Morewedge, Jane Risen, and Timothy Wilson (2004), “Looking forward to looking backward: The misprediction of regret,” *Psychological Science*, 15(5), 346-350.

Gillespie, P. (2017), Intuit: Gig economy is 34% of US workforce. *CNN Money*.

Hayes, Andrew F. and Kristopher J. Preacher (2014), "Statistical mediation analysis with a multicategorical independent variable," *British Journal of Mathematical and Statistical Psychology,* 67 (3), 451-470.

John, Oliver P., Eileen M. Donahue, and Robert L. Kentle (1991), "The big five inventory—versions 4a and 54."

Johnson, Marcia K. and Carol L. Raye (1981), "Reality monitoring," *Psychological Review,* 88 (1), 67.

Johnson, E. J., G. Häubl, and A. Keinan (2007), “Aspects of endowment: A query theory of value construction,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33* (3), 461.

Kahneman, Daniel (2003), “A perspective on judgment and choice: Mapping bounded rationality,” *American Psychologist*, *58* (9), 697.

Kahneman, Daniel and Shane Frederick (2002), "Representativeness revisited: Attribute substitution in intuitive judgment," *Heuristics and Biases: The Psychology of Intuitive Judgment,* 49, 81.

Kahneman, Daniel and Amos Tversky (1979), “Intuitive prediction: Biases and corrective procedures,” *TIMS Studies in Management Science, 12,* 313-327.

Kane, Joanne, Leaf Van Boven, and A. Peter McGraw (2012), "Prototypical prospection: Future events are more prototypically represented and simulated than past events," *European Journal of Social Psychology,* 42 (3), 354-362.

Kirby, Kris N. and Nino N. Maraković (1996), "Delay-discounting probabilistic rewards: Rates decrease as amounts increase," *Psychonomic Bulletin & Review,* 3 (1), 100-104.

Kruger, Justin and Matt Evans (2004), "If you don't want to be late, enumerate: Unpacking reduces the planning fallacy," *Journal of Experimental Social Psychology,* 40 (5), 586-598.

Lauriola, Marco, Irwin P. Levin, and Stephanie S. Hart (2007), "Common and distinct factors in decision making under ambiguity and risk: A psychometric study of individual differences," *Organizational Behavior and Human Decision Processes,* 104 (2), 130-149.

Lynch Jr., John G., Richard G. Netemeyer, Stephen A. Spiller, and Alessandra Zammit (2009), "A generalizable scale of propensity to plan: The long and the short of planning for time and for money," *Journal of Consumer Research,* 37 (1) 108-128.

MacDonnell, Rhiannon and Katherine White (2015), "How construals of money versus time impact consumer charitable giving," *Journal of Consumer Research,* 42 (4), 551-563.

Peetz, Johanna and Roger Buehler (2009), “Is there a budget fallacy? The role of savings goals in the prediction of personal spending,” *Personality and Social Psychology Bulletin*, *35* (12), 1579-1591.

Peetz, Johanna and Roger Buehler (2012), "When distance pays off: The role of construal level in spending predictions," *Journal of Experimental Social Psychology,* 48 (1), 395-398.

Peetz, Johanna and Roger Buehler (2013), “Different goals, different predictions: Accuracy and bias in financial planning for events and time periods,” *Journal of Applied Social Psychology*, *43* (5), 1079-1088.

Peetz, Johanna, Roger Buehler, Dereck J. Koehler, and Esther Moher (2015), “Bigger not better: Unpacking future expenses inflates spending predictions,” *Basic and Applied Social Psychology*, *37* (1), 19-30.

Peetz, Johanna, Melanie Simmons, Jingwen Chen, and Roger Buehler (2016), "Predictions on the go: Prevalence of spontaneous spending predictions," *Judgment and Decision Making,* 11 (1), 48.

Pew Charitable Trusts (2012), “Payday Lending in America: Who Borrows, Where They Borrow, and Why,” *Safe Small-Dollar Loans Research Project*.

Rick, Scott I., Cynthia E. Cryder, and George Loewenstein (2007), “Tightwads and spendthrifts,” *Journal of Consumer Research*, *34* (6), 767-782.

Rosch, E. and C. B. Mervis (1975), “Family resemblances: Studies in the internal structure of categories,” *Cognitive Psychology, 7,* 573–605.

Scheier, Michael F., Charles S. Carver, and Michael W. Bridges (1994), "Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test," *Journal of Personality and Social Psychology,* 67 (6), 1063.

Schkade, David A. and Daniel Kahneman (1998), "Does living in California make people happy? A focusing illusion in judgments of life satisfaction," *Psychological Science,* 9 (5), 340-346.

Schwartz, Lisa M., Steven Woloshin, William C. Black, and H. Gilbert Welch (1997), "The role of numeracy in understanding the benefit of screening mammography," *Annals of Internal Medicine,* 127 (11), 966-972.

Sussman, Abigail B. and Adam L. Alter (2012), “The exception is the rule: Underestimating and overspending on exceptional expenses,” *Journal of Consumer Research*, *39* (4), 800-814.

Tam, Leona and Utpal Dholakia (2014), "Saving in cycles: How to get people to save more money," *Psychological Science,* 25 (2), 531-537.

Thomas, Manoj, Kalpesh Kaushik Desai, and Satheeshkumar Seenivasan (2010), "How credit card payments increase unhealthy food purchases: Visceral regulation of vices," *Journal of Consumer Research,* 38 (1), 126-139.

Tversky, Amos and Derek J. Koehler (1994), "Support theory: A nonextensional representation of subjective probability," *Psychological Review,* 101 (4), 547.

Tversky, Amos and Daniel Kahneman (1974), "Judgment under uncertainty: Heuristics and biases," *Science,* 185 (4157), 1124-1131.

Ülkümen, Gulden, Manoj Thomas, and Vicki G. Morwitz (2008), “Will I spend more in 12 months or a year? The effect of ease of estimation and confidence on budget estimates,” *Journal of Consumer Research*, *35* (2), 245-256.

US Federal Reserve (2016), *Report on the Economic Well-Being of US Households in 2015*. <https://www.federalreserve.gov/2015-report-economic-well-being-us-households-201605.pdf>.

Van Boven, Leaf, Joanne Kane, and Peter A. McGraw (2008), “Temporally asymmetric constraints on mental simulation: Retrospection is more constrained than prospection,” ed. K. Markman, W. Klein, and J. Suhr, *The handbook of imagination and mental simulation* (pp. 131–149). New York: Psychology Press.

Wilson, Timothy D., Thalia Wheatley, Jonathan M. Meyers, Daniel T. Gilbert, and Danny Axsom (2000), "Focalism: A source of durability bias in affective forecasting," *Journal of Personality and Social Psychology,* 78 (5), 821.

Winkielman, Piotr, Jamin Halberstadt, Tedra Fazendeiro, and Steve Catty (2006), "Prototypes are attractive because they are easy on the mind," *Psychological Science,* 17 (9), 799-806.

Yang, Sha, Livia Markoczy, and Min Qi (2007), "Unrealistic optimism in consumer credit card adoption," *Journal of Economic Psychology,* 28 (2), 170-185.

Zauberman, Gal and John G. Lynch, Jr. (2005), "Resource slack and propensity to discount delayed investments of time versus money," *Journal of Experimental Psychology: General,* 134 (1), 23.

1. A set of 2 (order: predict first vs. recall first) × 3 (condition: control vs. typical vs. atypical) ANOVAs with predicted expenses, recalled expenses, and bias scores (recalled – predicted expenses) as the DVs revealed no order effect (*p*’s > .27). [↑](#footnote-ref-1)
2. As noted in footnote 1, the order of prediction and recall does not interact with condition, but adding order to the model does reveal directionally lower recall for participants in the typical condition who predicted first, and directionally higher recall for participants in the atypical condition who predicted first. Therefore, we believe the difference in recalled expenses between these two conditions is the result of the prediction manipulation in each condition spilling over into recall. [↑](#footnote-ref-2)
3. The same results are obtained when using only the atypical and pure control conditions as levels of the IV (indirect effect = .05, SE = .02, 95% CI = [.02, .09]) and when running a categorical mediation model that includes all three conditions (indirect effect of atypical dummy = .05, SE = .02, 95% = [.02, .08]; indirect effect of typical dummy = -.01, SE = .02, 95% CI = [-.04, .02]). [↑](#footnote-ref-3)
4. Participant dropout over the last week of the study was minimal (*n* = 4) and did not differ by condition. [↑](#footnote-ref-4)
5. The size of the sample provided by our field partner did not leave us with enough power to include the typical condition in this study. [↑](#footnote-ref-5)