

# **A PROTOTYPE THEORY OF CONSUMER EXPENSE MISPREDICTION**

Chuck Howard  
PhD Candidate in Marketing  
Sauder School of Business, University of British Columbia  
chuck.howard@sauder.ubc.ca  
Job Market Paper

with

David J. Hardisty  
Assistant Professor of Marketing  
Sauder School of Business, University of British Columbia  
david.hardisty@sauder.ubc.ca

Abigail B. Sussman  
Associate Professor of Marketing  
The University of Chicago Booth School of Business  
abigail.sussman@chicagobooth.edu

## 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, 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 personal finance apps 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 and 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 show that prompting consumers to consider reasons why their expenses might be different than usual increases 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 process explanation for why the expense prediction bias occurs, and it allows past findings in this area to be understood as a whole for the first time. 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 contributes to the broader literature on consumer misprediction in non-financial domains. 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 and *how* distributional skew impacts predictions accuracy. Similarly, we contribute to research on the use of prototype attributes in judgment and decision making—which has implied that the average

represented by prototype attributes is the *mean*—by providing evidence that these attributes represent *modal* 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). 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 a series of studies that test these hypotheses in the field and the lab. 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 perhaps most compelling because it has the potential to offer a parsimonious explanation for several results in the expense misprediction literature. For example, previous work in this area has demonstrated that consumers' expense predictions do not adequately incorporate either their distribution of actual past expenses or their distribution of possible future expenses (Peetz and Buehler 2012; Peetz et al. 2015). This kind of *distribution neglect* is easily explained by the prototype proposition because prototype attributes represent a relatively thin slice of a consumer's expense distribution. Relatedly, it has been shown that consumers behave as if their atypical or exceptional expenses will not reoccur (Sussman and Alter 2012). This can be explained by the prototype proposition because the "thin slice" that prototype attributes represent does not include these 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 can be explained by 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).

To explore the prototype proposition and help illuminate the cognitive processes that underlie the expense prediction bias we began by running a think-aloud protocol study. Fifty-five undergraduate commerce students at a large Canadian university were recruited to take part in a study about consumer financial decision making. Each participant was taken to a private room where they received written instructions to say aloud every thought that came to mind as they formulated their expense prediction for the next week. Participants' thoughts were recorded and later transcribed and coded by research assistants. Specifically, we had the research assistants independently code each transcription for references to typical spending, future oriented spending (i.e., spending specific to the next week), and adjustments for unexpected spending. We then had them code which of these thoughts appeared first in each transcript.

Table 1 presents the proportion of participants who referenced typical spending, future-oriented spending, and adjustments for unexpected spending, as well as examples of each. A significantly higher proportion of participants referenced typical spending than future-oriented spending (Mean difference = 29.09%, 95% CI = [11.88%, 44.12%],  $X_{(1)} = 10.80$ ,  $p = .001$ ). Typical spending also came to mind first for a strong majority of participants ( $z = 2.56$ , 95% CI = [53.29%, 79.32%],  $p = .010$ ). Finally, only half of the participants made a conscious adjustment for unexpected expenses. Taken together, these results provide preliminary support for the prototype proposition by demonstrating that prototype attributes are both easily accessible and a foundational component of expense predictions. We next outline in greater detail why prototypical prediction can be problematic.

**Table 1**  
**Results of the Think-Aloud Protocol Study**

Classification	Proportion	First Thought	Examples
Typical	83.64%	67.27%	" <b>Typically</b> I buy groceries every week. Thats about \$50 dollars or so." "On <b>average</b> , I would say I spend around \$10 per day on food and drinks." "So on <b>average</b> I spend \$500 per month." "On <b>average</b> I buy mainly a coffee or a lunch on campus so let's say I spend like up to \$10-ish maybe." "On <b>average</b> , I spend around \$45 to \$60 with food." " <b>Normally</b> I will spend uh approximately \$20 per day for food." "So groceries are <b>normally</b> \$100." "So I <b>usually</b> buy my lunch, cuz I have to." "On Friday I <b>usually</b> get gas so that's <b>usually</b> thirty dollars a week." "I <b>usually</b> for grocery shopping I spend, umm, around \$30 a week."
Future-Oriented	54.55%	32.73%	"Huh, I'm traveling <b>next week</b> too, traveling is...I'll say \$400, yeah." "Um <b>the next week, the next seven days</b> . Let's see...so this weekend I'm going...where am I going this weekend? So I'm going to ****'s party so that's probably around let's say \$15 of things." " <b>This Sunday</b> , I might go to the mall to get new work clothes for my co-op, so that might be dress shoes, that might be maybe \$120." "Are there any birthdays <b>coming up</b> ? Oh wait my brother's birthday. He is an expensive person so that's going to be about \$300."
Adjustment	50.91%	0.00%	"I might have <b>other expenses</b> , little things, stuff from vending machines, probably adds up to about \$10 over a week." "I'll put about \$20 for like <b>miscellaneous items</b> ." "And just for <b>miscellaneous</b> items I would put another \$10." "And then, shopping... <b>miscellaneous</b> , we'll just budget \$50 for that."

### *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 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. It is also consistent with 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 consumers should be fairly adept at predicting ordinary, recurring expenses, which are modal by definition.

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). However, 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 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).

### *HYPOTHESES*

One corollary of the prototype proposition is that predictions will largely neglect atypical expenses. In contrast, 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. Therefore, our first hypothesis is:

**H1:** On average, consumers predict their future expenses will be more typical than their past expenses.

Our 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. This follows from the fact that expenses are positively skewed: if prototype attributes represent modal spending, then a more prototypical prediction will generally be a lower prediction, because most of the distribution is higher than the mode. Using perceived typicality of future expenses as a measure of reliance on prototype attributes, we hypothesize that:

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

A clear implication of H1 and H2 is that consumers will under-predict their future expenses as compared to their recalled 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:

**H3a:** Consumers predict lower expenses for the future as compared to the expenses they recall for the past.

Our next hypothesis is 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 the results of the think aloud study suggest that predictions are not 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). We therefore hypothesize that:

**H3b:** Consumers under-predict their expenses for the target week or month as compared to the expenses they actually incur during that week or month.

Another corollary of the prototype proposition is 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 a typical week or month 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 “different” should therefore make predictions closer to the mean. We thus hypothesize that:

**H4:** Prompting consumers to consider three reasons why their expenses might be different from a typical week increases expense prediction accuracy.

Finally, 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  $\text{mode} < \text{mean}$  versus when it is normally distributed with  $\text{mode} = \text{mean}$ . In study 4 we experimentally manipulate the distribution of expenses in the lab, changing the mode while keeping other features of the distribution like mean and variance constant. In this context we hypothesize that:

**H5:** Consumer expense predictions will be lower (and farther from the mean) after observing a positively skewed distribution of expenses versus a normal distribution.

### *OVERVIEW OF STUDIES*

To test our hypotheses we conducted a series of four experiments. In study 1 we examine H1–H3b in a repeated-measures longitudinal field study with members of a financial cooperative, and test the efficacy of the atypical intervention (vs. control) in the last week of the study (H4). Study 1 also provides evidence suggesting that the effects documented by H1–H3b are not the result of a savings goal, trait optimism, or other individual differences that could offer alternative explanations for our results. In study 2 we establish that prototype attributes influence monthly as well as weekly expense predictions. In study 3—conducted with a nationally representative sample of adult Americans—we directly replicate the findings of study 1 with respect to H1–H3a and provide evidence that our intervention is effective because it makes atypical expenses more accessible when consumers construct their predictions (H4). In

study 4, 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 (H5).

As highlighted in our theoretical development, the distribution of expenses displays significant positive skew. This is a defining characteristic of the data, and an important element in our theory; it also presents serious challenges for responsible inferential analysis in studies 1 and 3, where we compare predicted expenses to reported expenses across experimental conditions. 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 1 and 3. 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. Our data, syntax, and study materials are available here: [hyperlink to be inserted here].

### *STUDY 1: A LONGITUDINAL STUDY OF EPB IN THE FIELD*

The first goal of study 1 was to test the prototype proposition and its associated hypotheses (H1-H2). The second goal of study 1 was to contribute a more comprehensive understanding of expense prediction bias (H3) by observing its magnitude and persistence within a highly engaged sample across multiple points in time. The third goal of study 1 was to examine the effectiveness of the atypical intervention (H4) in the field. To accomplish these goals we partnered with a financial cooperative to run a five-week longitudinal field study with its members.

#### *Method*

*Participants and Procedure.* 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 pilot studies. Each participant completed six surveys over the course of five weeks, as illustrated by each time period marked in figure 1.

**Figure 1**

#### **Data Collection Schedule in Study 1**



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 10<sup>th</sup>, 2017. Ninety-three people completed the survey before it was deactivated at 11:59pm on Monday, September 11<sup>th</sup>, 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.

All surveys were 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 next week as follows: “Please take some time to estimate your total expenses for the **next week**. By “total expenses” we mean everything you will pay for during the next week. (Page Break). Please enter your estimated total expenses for the **next week**.” Participants were then asked to indicate how typical they expected their expenses to be (“How different or similar do you think your expenses will be for the next week [month], relative to a typical week [month]?” 1 = very different, 7 = very similar), and how confident they were in their prediction accuracy (“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). We then had them repeat this prediction exercise for their monthly expenses.

The remaining five surveys began by asking participants to log into their online bank account and report their expenses for the past week, then predict their expenses for the next week. Both expense reports and predictions were followed by the same measures of perceived typicality and confidence used in survey 1. In the second-to-last survey (i.e., at T4), half of the

sample was randomly assigned to receive the atypical intervention, making the final week of the study a 2 (condition: control vs. atypical)  $\times$  2 (expenses: predicted vs. actual) between-within design.<sup>1</sup> In the atypical condition participants received the following instructions before making their prediction: “Please take some time to consider why your expenses for the **next week** might be different from a typical week. In the spaces provided below, please type 3 reasons why your expenses for next week might be different from a typical week.” Participants in the control condition received the same prediction instructions as in the previous surveys.

Across the six surveys we also measured the following individual differences: savings goals (Peetz and Buehler 2009), trait optimism (Scheier et al. 1994), short-term financial propensity to plan (Lynch et al. 2010), numeracy (Schwartz et al. 1997), spendthrift-tightwad tendencies (Rick, Cryder, and Lowenstein 2008), openness to experience (John, Donahue, and Kentle, 1991), temporal discounting (Kirby & Maraković 1996), and cyclical versus linear time orientation (Tam and Dholakia 2014). Additionally, we collected exploratory measures related to behaviors such as budgeting, borrowing, and spending. None of these measures were found to be related to the expense prediction bias, so we present these results in Web Appendix C.

### *Results*

*Skew.* As can be seen in table 2, the average consumer in study 1 had strong positive skew in their distribution of spending over the five weeks of the study. It is also notable that 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

---

<sup>1</sup> Participant dropout over the last week of the study was minimal ( $n = 4$ ) and did not differ by condition.

of predicted expenses are substantially lower than the mean and median of past expenses, which supports the expense prediction bias hypotheses (H3a and H3b).

**Table 2**

**Distributional Parameter Estimates for the Average Consumer in Study 1**

	<b>Median</b>	<b>Mean</b>	<b>SD</b>	<b>Skew</b>
<b>Reported Expenses</b>	\$500.00	\$665.89	407.84	1.03
<b>Predicted Expenses</b>	\$450.00	\$600.00	316.04	0.84

*Perceived Typicality.* To test our hypothesis that people predict their expenses will be more typical in the future than in the past (H1), 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 2 and table 3, participants predicted their expenses would be more typical in the next (vs. past) week at all four points in time. Figure 2 also shows that our atypical intervention succeeded in neutralizing this tendency at T4. In sum, these results provide strong support for H1.

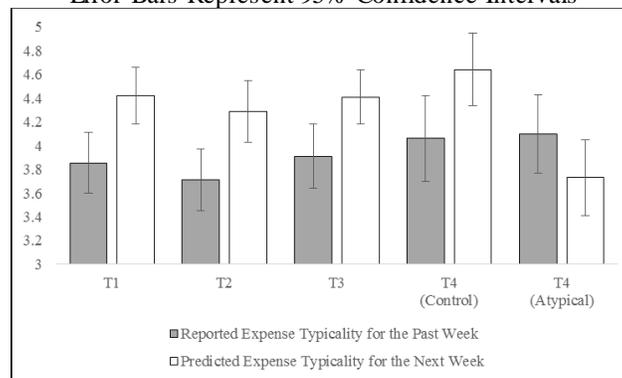
To test our hypothesis that perceived typicality of future expenses is negatively correlated with expense predictions (H2), we analyzed the correlation between perceived typicality of future expenses and weekly expense predictions for each week of the study, as well as for the month. Perceived typicality 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 ( $r(185) = -.03, p = .74$ ). We next regressed predictions onto perceived typicality of

future expenses and participant income 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. This analysis produced only one meaningful change in these results: The relationship between perceived typicality and prediction at T1 became significant ( $p = .013$ ). In aggregate, these findings provide support for H2.

**Figure 2**

**Mean Reported Expense Typicality for the Past Week vs. Mean Predicted Expense Typicality for the Next Week for Each Week of Study 1**

Error Bars Represent 95% Confidence Intervals



**Table 3**

**T-tests Comparing Reported vs. Predicted Typicality in Study 1**

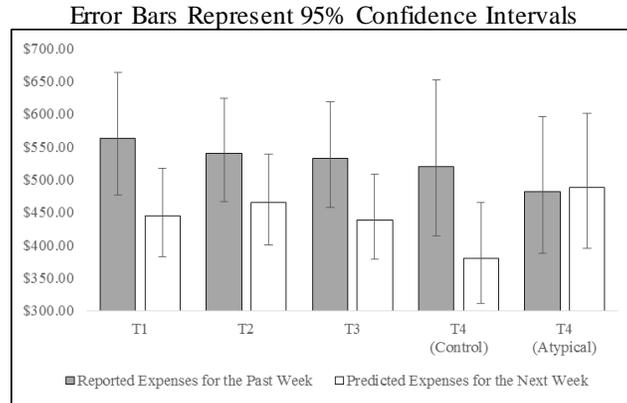
	T1	T2	T3	T4 (Control)	T4 (Atypical)
t	3.45	4.10	3.25	2.88	-1.78
df	174	173	177	89	88
p	0.001	< 0.001	0.001	0.005	0.080

*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 (H3a), 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 3 and table 4, predicted expenses were significantly lower than reported expenses at each stage of the study until the atypical intervention was deployed at T4. This provides strong support for H3a.

**Figure 3**

**Mean Expenses Incurred in the Past Week vs. Mean Predicted Expenses for the Next Week for Each Week of Study 1**



**Table 4**

**T-tests Comparing Reported Past-Week vs. Predicted Next-Week Expenses in Study 1**

	T1	T2	T3	T4 (Control)	T4 (Atypical)
t	3.45	2.40	3.25	3.88	-0.16
df	174	173	177	89	88
p	0.001	0.018	0.001	0.000	0.870

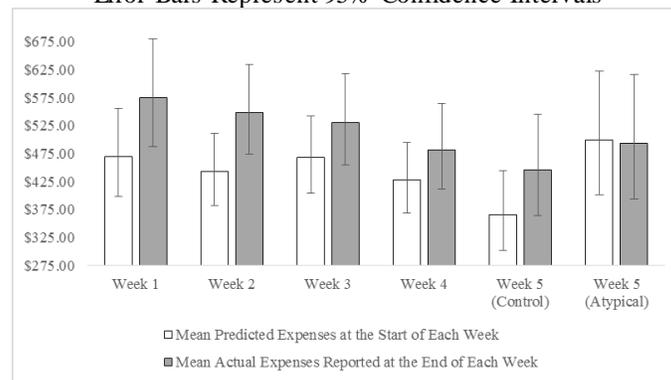
*Expense Prediction Bias (Actual – Predicted Expenses).* As illustrated in figure 4 and table 5, predicted expenses were significantly lower than incurred expenses in each week of the study, except during week 5 in the atypical condition, in which our intervention completely neutralized expense prediction 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 sacrifice correlational accuracy ( $r_{\text{control}(90)} = .81, r_{\text{treatment}(89)} = .80, z = .19, p = .85$ ). These findings provide strong support for H3b and H4.

**Figure 4**

**Mean Predicted Expenses vs. Mean Incurred Expenses for Each Week of Study 1**

Error Bars Represent 95% Confidence Intervals



**Table 5**

**T-tests Comparing Predicted Expenses vs. Incurred Expenses in Study 1**

	Week 1	Week 2	Week 3	Week 4	Week 5 (Control)	Week 5 (Atypical)
t	3.71	4.35	2.86	2.73	3.19	-0.19
df	179	179	180	182	91	90
p	< .001	< .001	0.005	0.007	0.002	0.85

*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 incurred in weeks 1 through 4 of the study. Results showed that predicted expenses for the target

month ( $M_{\text{pred}} = \$2276.74$ , 95% CI = [2031.64, 2551.15]) were \$416.77 (15.5%) lower than expenses actually incurred ( $M_{\text{actual}} = \$2693.51$ , 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.

*Intervention content analysis.* To gain more insight into the cognitive process underlying expense prediction we had two research assistants who were blind to our hypotheses independently code the content of participants' responses to the atypical intervention. Specifically, we had them code each response as either a reason expenses would be higher, a reason expenses would be lower, a reason expenses would not actually be different, or a reason that was too ambiguous to be coded as higher, lower or not different. Interrater reliability was 96.1% and disagreements were resolved through discussion between the raters.

The results of this coding exercise are presented in Table 6. The results in the top row ("Overall") are striking because they are highly consistent with a positively skewed distribution of expenses: most reasons given were reasons why expenses would be higher than typical, but a fair percentage of responses represent reasons why expenses would be lower. The results in the third column of the table ("Not Different") are also notable: the percentage of responses indicating that expenses would not be different grew by almost a factor of five between reasons 1 and 3 ( $X_{(1)} = 7.79$ ,  $p = .005$ ). This finding complements the results of the think aloud study (which showed the prototypical expenses are relatively accessible) by providing evidence that atypical expenses are relatively inaccessible.

**Table 6****Summary of Results for the Study 1 Intervention Content Analysis**

	<u>Higher</u>	<u>Lower</u>	<u>Not Different</u>	<u>Ambiguous</u>
<b>Overall</b>	73.5%	12.9%	9.3%	4.3%
<b>Reason 1</b>	83.9%	8.6%	3.2%	4.3%
<b>Reason 2</b>	68.8%	15.1%	9.7%	6.5%
<b>Reason 3</b>	67.7%	15.1%	15.1%	2.2%
<b>Examples</b>	"Buying some new duds" "I've started a keto diet and some of the extra food might be expensive." "Possible legal bill"	"No commute expenses" "Doggy daycare paid until December" "Recently stocked up on certain grocery items"	"No reason" "N/A" "None"	"Different routine" "Beginning of the month" "Jury duty"

*Prediction Confidence.* We measured prediction confidence in this study because it has been shown that lower prediction confidence leads people to adjust their expense predictions upward (Ulkumen et al. 2008). Therefore, it could be the case that the atypical intervention is effective merely because it is decreasing prediction confidence. To test this possibility we first performed an independent samples t-test with condition as the independent variable and prediction confidence as the dependent variable. This analysis revealed that the atypical intervention did decrease prediction confidence ( $M = 4.69, SD = 1.13, M = 5.04, SD = 1.10, t(181) = -2.13, p = .035$ ). However, the effect of condition (atypical versus control) on expense predictions remained significant when controlling for prediction confidence ( $F(1, 180) = 4.20, p = .042$ ), demonstrating that the effect of the atypical intervention is operating above and beyond conditional differences in confidence. Moreover, prediction confidence did not mediate the effect of condition on expense predictions (indirect effect = .01,  $SE = (.03), 95\% CI = [-.05, .06]$ ), indicating that prediction confidence was not directly related to expense predictions in this study.

*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 1 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 ( $EPB_{\text{outsideview}} = -\$116.05$ , different from zero,  $t(91) = -2.92, p = .004$ ), whereas participants in the atypical condition displayed remarkably high mean prediction accuracy ( $EPB_{\text{atypical}} = -\$6.65$ , 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 ( $r_{\text{outsideview}}(90) = .69, r_{\text{atypical}}(89) = .80, z = 1.65, p = .050$ ).

The second way to compare the effectiveness of the outside view and atypical interventions 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 ( $EPB_{\text{outsideview}} = -\$48.59$ , not different from zero,  $t(89) = -1.08, p = .28$ ), and their correlational accuracy would have been significantly worse ( $r_{\text{outsideview}}(89) = .65, r_{\text{atypical}}(89) = .80, z = 2.13, p = .017$ ). These analyses are exploratory and by no means conclusive, but they do suggest 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 1 offer longitudinal evidence from the field that provides compelling support for the prototype theory. First, consumers predicted that their expenses would be more typical (H1) and lower (H3a) 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 (H3b). Fourth, it was shown that the atypical intervention is capable of virtually eliminating the expense prediction bias in a real-world setting (H4). It was also shown that the atypical intervention outperforms an outside-view decision-rule that incorporates the full distribution of (recent) expenses. Finally, the null results of study 1 (reported in web appendix C) suggest that the effects documented by H1-H3b are the result of a general cognitive process like prototypical prediction rather than individual differences such as motivation to meet a savings goal or trait optimism. We return to this point in study 4.

### *STUDY 2: (A) TYPICALITY IN WEEKLY VS. MONTHLY PREDICTIONS*

The primary goal of study 2 was to more cleanly 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 2 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 1 demonstrates that an expense prediction bias exists for both time periods.

### *Method*

*Participants and Procedure.* We recruited 601 participants (48.6% female,  $M_{\text{age}} = 37.93$ ) 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)  $\times$  2 (intervention condition: control vs. atypical) between-subjects design that utilized the same prediction prompts as in study 1, modified where necessary for monthly predictions. Participants also reported perceived typicality and prediction confidence as in study 1.

### *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)  $\times$  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 intervention condition by time period interaction ( $F(1, 597) = 2.09$ ,  $p = .15$ ).

*Monthly Expense Predictions.* The impact of the atypical intervention on monthly expense predictions was consistent with its impact on weekly expense predictions in the final week of study 1: 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 ( $M_{\text{atypical}} = 4.79$ , 95% = [4.56, 5.01];  $M_{\text{control}} = 5.39$ , 95% CI<sub>control</sub> = [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 the final week of study 1 was replicated in study 2: 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 ( $M_{\text{atypical}} = 5.20$ , 95% = [4.99, 5.40];  $M_{\text{control}} = 5.35$ , 95% CI<sub>control</sub> = [5.17, 5.54];  $t(300) = 1.13$ ,  $p = .26$ ).

### *Discussion*

The results of study 2 reveal that perceived typicality of future expenses differs 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 is 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 2 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 1—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.

### *STUDY 3: THE ACCESSIBILITY OF ATYPICAL EXPENSES*

The primary goal of study 3 was to better understand the process by which the atypical intervention improves prediction accuracy. Specifically, we examined the extent to which our intervention makes atypical expenses more accessible (versus control), and the relationship between accessibility and predictions. We also used study 3 to better understand the expense prediction process by including an experimental control condition that prompted participants to consider three reasons why their expenses would be *similar* to a typical week. Our expectation was that predictions and perceived typicality would not differ between the experimental control condition and the pure control condition, because if the prototype proposition is correct, then

predictions in the pure control condition will already be swayed by typical expenses. The third goal of study 3 was to increase the generalizability of our findings by using a nationally representative sample.

### *Method*

*Participants.* A nationally representative sample of 1,108 US residents completed study 3 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 ( $M_{\text{age}} = 49.59$ ; 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: pure control, experimental control (i.e., the “typical condition”), or intervention (i.e., the “atypical condition”). In the pure control condition, participants recalled and predicted their weekly expenses for the past and next week as in study 1. 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 (H1). 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 (H4). The order of prediction and recall was counterbalanced in all conditions.<sup>2</sup>

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.

Finally, participants completed the same measures of perceived typicality used in study 1, and five exploratory measures designed to let us explore the relationship between expense prediction bias (EPB), financial slack (Berman et al. 2016; Zauberma n 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.

---

<sup>2</sup> 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).

## Results

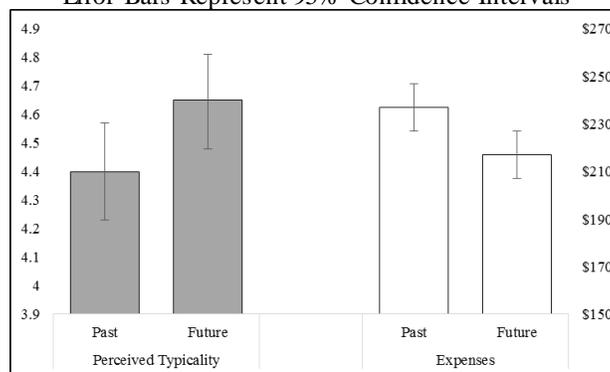
*Skew.* As in study 1, both recalled and predicted expenses in the control condition displayed a high degree of positive skew ( $\text{skew}_{\text{recall}} = 7.65$ ,  $\text{skew}_{\text{predict}} = 3.82$ ).

*Replicating Study 1.* As illustrated in figure 5, the results of study 1 were directly replicated in the control condition of study 3. Supporting H1, participants predicted their future expenses would be more typical than their past expenses ( $M_{\text{past week}} = 4.40$ , 95%  $CI_{\text{past week}} = [4.23, 4.57]$ ,  $M_{\text{next week}} = 4.65$ , 95%  $CI_{\text{next week}} = [4.48, 4.81]$ ,  $t(415) = -3.42$ ,  $p = .001$ ,  $d = .17$ ). Supporting H2, higher perceived typicality of future expenses was again associated with lower expense predictions ( $r(414) = -.21$ ,  $p < .001$ ), and this association remained significant after controlling for participant income ( $B = -.14$ ,  $SE = .03$ ,  $t(413) = -4.98$ ,  $p < .001$ ). Finally, in support of H3a, participants predicted their future expenses would be 8.6% lower than their past expenses ( $M_{\text{past week}} = \$237.46$ , 95%  $CI_{\text{past week}} = [\$217.02, \$262.43]$ ,  $M_{\text{next week}} = \$217.02$ , 95%  $CI_{\text{next week}} = [\$196.37, \$239.85]$ ,  $t(415) = 2.76$ ,  $p = .006$ ). We next expand our analyses to test for differences in perceived typicality and EPB across all three conditions.

**Figure 5**

### Mean Perceived Typicality of Past vs. Future Expenses and Mean Reported vs. Predicted Expenses in the Control Condition of Study 3

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

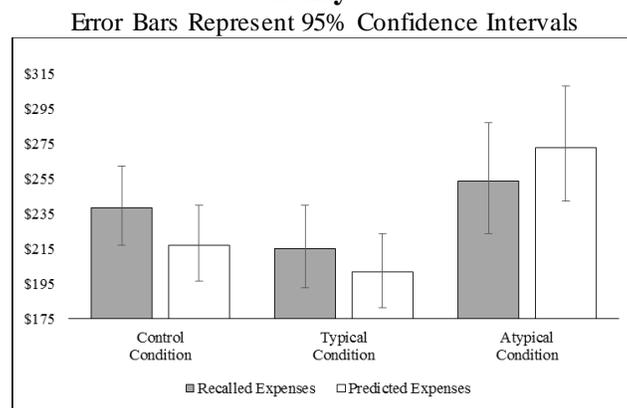
expenses as the dependent variable revealed a significant effect of condition on perceived typicality of future expenses ( $F(2, 1044) = 32.27, p < .001$ ). Planned contrasts further revealed that perceived typicality was virtually identical in the control and typical conditions ( $M_{\text{control}} = 4.64, 95\% \text{ CI}_{\text{control}} = [4.48, 4.81], M_{\text{typical}} = 4.65, 95\% \text{ CI}_{\text{typical}} = [4.48, 4.83], t(1044) = .03, p = .97$ ), but significantly lower in the atypical condition ( $M_{\text{atypical}} = 3.74, 95\% \text{ CI}_{\text{atypical}} = [3.56, 3.93], t(1044) = -8.02, p < .001, d = .55$ ).

*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 by our atypical intervention (H4), as illustrated in figure 6. A 3 (intervention: control vs. typical vs. atypical)  $\times$  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 ( $M_{\text{atypical}} = \$273.14, 95\% \text{ CI}_{\text{atypical}} = [\$242.26, \$307.97]$ ) were 25.9% (\$56.12) higher than in the control condition ( $M_{\text{control}} = \$217.02, 95\% \text{ CI}_{\text{control}} = [\$196.37, 239.85]; t(1045) = 2.91, p = .004$ ), and 35.0% (\$70.79) higher than in the typical condition ( $M_{\text{typical}} = \$202.35, 95\% \text{ CI}_{\text{typical}} = [\$181.27, \$223.63]; t(1045) = 3.67, p < .001$ ). 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 ( $M_{\text{atypical}} = \$254.68, 95\% \text{ CI}_{\text{atypical}} = [\$223.63, \$287.15]$ ) and control conditions ( $M_{\text{control}} = \$237.46, 95\% \text{ CI}_{\text{control}} = [\$217.02, \$262.43]; t(1045) = .79, p = .43$ ),

but they were somewhat lower in the typical condition ( $M_{\text{typical}} = \$214.86$ , 95%  $CI_{\text{typical}} = [\$192.48, \$239.85]$ ) than in the atypical condition ( $t(1045) = 2.01$ ,  $p = .045$ ).<sup>3</sup> 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 6**

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



*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 ( $M_{\text{atypical}} = 1.57$ ,  $SD_{\text{atypical}} = 1.60$ ) was significantly higher than in the control and typical conditions ( $M_{\text{control}} = 1.21$ ,  $SD_{\text{control}} = 1.48$ ,  $M_{\text{typical}} = 1.16$ ,  $SD_{\text{typical}} = 1.52$ ;

<sup>3</sup> 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.

$t(1045) = 6.80, p = .001$ ), 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 more accessible 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 show that our intervention succeeded in making atypical expenses more accessible, and that expense predictions were significantly higher in the intervention condition as well. To further investigate the relationship between the accessibility of atypical expenses and 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 ( $b = -.38, 95\% \text{ CI} = [.18, .59]; t(1046) = 3.66, p < .001$ ), and that this number was associated with higher expense predictions, even while controlling for condition ( $b = -.14, 95\% \text{ CI} = [.10, .18]; t(1045) = 6.67, p < .001$ ).<sup>4</sup>

### *Discussion*

Study 3 provides support for the prototype theory in three ways. First, it directly replicates the findings of study 1 with a nationally representative sample, demonstrating that consumers predict their future expenses will be both more typical (H1) and lower than their past

---

<sup>4</sup> 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]).

expenses (H3a), and that perceived typicality of future expenses is negatively correlated with predictions (H2). Second, it demonstrates that our intervention makes atypical expenses more accessible (versus control), and that the number of atypical expenses that come to mind is associated with higher expense predictions. Finally, 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.

#### *STUDY 4: MANIPULATING SKEW IN 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 4 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 (H5).

The design of study 4 also allows us to follow-up on the null results observed in study 1 with respect to the relationship between EPB and individual differences in motivation and optimism. If the expense prediction bias were driven primarily by motivation to spend less in the future or a general optimism that future expenses will be lower, then predictions should be lower than the mean in both distribution conditions. Indeed, it would be reasonable to expect that predictions in the normal condition could be even lower than in the positively skewed condition, because the left tail of the former is longer than the latter, which means it leaves more room for motivation and/or optimism to influence predictions. Therefore, support for H5 would not only provide further evidence that prototypical prediction is a primary driver of the bias, it would also suggest that motivation and optimism are not.

### *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,  $M_{\text{age}} = 36.9$ ). 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, Reinholz, 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 (H5), 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 ( $M_{avg} = \$194.90$ ,  $SD_{avg} = 15.59$ ,  $t(200) = .20$ ,  $p = .84$ ) or in the normal condition ( $M_{avg} = \$199.72$ ,  $SD_{avg} = 16.76$ ,  $t(199) = .40$ ,  $p = .69$ ).

### *Discussion*

Study 4 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. Notably, the results of study 4 also provide further evidence that motivation and optimism are not superordinate drivers of the bias.

### *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 that improves prediction accuracy by making atypical expenses easier to retrieve when consumers make their predictions. We next discuss the implications of our work for theory and practice, as well as directions for future research.

#### *Understanding the Expense Prediction Bias*

The prototype theory deepens our understanding of the expense prediction bias in several ways. First, it provides a novel and parsimonious explanation for why the bias occurs, revealing for the first time the interactive effect that prototype attributes and distributional skew have on prediction accuracy. Second, as highlighted in our theoretical development, the prototype theory helps explain several past findings in the expense misprediction literature, including distribution neglect (Peetz and Buehler 2012; Peetz et al. 2015), narrow bracketing of exceptional expenses

(Sussman and Alter 2012), and the inverse relationship between prediction confidence and adjustment (Ulkumen et al. 2008). The evidence supporting the prototype theory also shows that expense prediction bias is primarily a function of a basic cognitive process (i.e., prototypical prediction), rather than an individual difference like motivation to save (cf. Peetz and Buehler 2009). Similarly, our results demonstrate that the bias is not an instance of complete base-rate neglect (cf. Peetz and Buehler 2012). Instead, it can be said that consumers simply have access to the wrong base-rate (i.e., the mode, not the mean) in their minds. Finally, the prototype theory offers the first intervention that both informs theory and is practically useful in the field. We next expand on this point.

#### *Neutralizing the Expense Prediction Bias*

Mechanistically, the atypical intervention 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). However, there is 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 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 kind of distributional information is missing, and how distributional skew effects prediction accuracy. The atypical intervention also carries a practical advantage: It requires consideration of only a handful of reasons why expenses may be atypical (versus 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, although unpacking can increase mean prediction accuracy across a sample, it can also cause some people to make tangibly worse (i.e., less accurate) predictions than they would have otherwise. It is therefore notable that the atypical intervention not only improved mean prediction accuracy in study 1, it also maintained the same level of correlational accuracy as in the control condition.

#### *The “Average” Represented by Prototype Attributes*

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. An important direction for future research is to investigate when prototypes represent the mode, mean, median, or some combination.

### *Temporal Asymmetry*

Finally, the present research contributes to a nascent literature on temporal asymmetry, which hypothesizes that people think about 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 thinking about 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). Another direction for future research is to directly compare and contrast the strength of asymmetries in resources like time versus money.

### *Implications for Consumers, Financial Literacy Organizations and Firms*

An important contribution of the present research is that it provides a more comprehensive understanding of expense prediction bias as a phenomenon. For example, we present the first studies to identify the magnitude 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. The implications of our findings in this regard are clear: the magnitude of the bias (approximately \$100 per week or \$400 per month in study 1) is large enough to be economically meaningful for many consumers.

Thus, the prosocial benefit of our research is also clear—any consumer can make use of the atypical intervention to improve his or her expense prediction accuracy and make better informed decisions regarding their spending, borrowing, or saving behavior.

One promising channel through which the atypical intervention can be disseminated is financial literacy organizations. Currently, the modal approach of such organizations is to educate their stakeholders about debits and credits, interest rates, and so on. However, this approach is both time consuming and appears to have very limited impact (Fernandes, Lynch and Netemeyer 2014). In contrast, the atypical intervention does not need to be learned, per se, but it can be easily provided and used to effectively increase prediction accuracy.

Our findings also have 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 many behaviors that follow a skewed distribution may be subject to prototyping, 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. Indeed, yet another important direction for future research is to test the explanatory power of the prototype theory and the effectiveness of the atypical intervention beyond the context of expense misprediction.

## REFERENCES

- 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 Zauberaman (2016), "Expense neglect in forecasting personal finances," *Journal of Marketing Research*, 53 (4), 535-550.
- 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/>
- 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\\_financial-well-being-scale-technical-report.pdf](https://s3.amazonaws.com/files.consumerfinance.gov/f/documents/201705_cfpb_financial-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*.

- Fernandes, Daniel, John G. Lynch Jr, and Richard G. Netemeyer. "Financial literacy, financial education, and downstream financial behaviors." *Management Science* 60.8 (2014): 1861-1883.
- Gillespie, P. (2017), Intuit: Gig economy is 34% of US workforce. *CNN Money*.
- 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 (2010), "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 re-evaluation of the Life Orientation Test," *Journal of Personality and Social Psychology*, 67 (6), 1063.
- 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.
- 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.

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.

*WEB APPENDIX A: STATISTICAL ROBUSTNESS TESTS**STUDY 1*

The distribution of consumer expenses displays strong positive skew, which presents serious challenges for responsible inferential analysis in studies 1 and 3. We addressed this by excluding the data of participants whose reported expenses exceed their predicted expenses by more than a factor of 10 (or vice versa), then LN-transforming the distributions of reported and predicted expenses. The first advantage of this approach is that it allows us to analyze (and visualize) EPB as the difference between reported and predicted expenses. In contrast, if we adopted an approach like winsorization we could only present winsorized mean bias scores, because winsorizing reported and predicted expenses independently could erase or even reverse some bias scores.<sup>5</sup> Therefore, a clear drawback of winsorization is that presenting only bias scores means the reader has no way of knowing if the bias occurred (or not) because of the effect of our atypical intervention on predictions or because of variation in reported expenses between conditions. The second advantage of the approach we use in the main text is that once results are exponentiated into dollars they are very easy to interpret, regardless of a reader's background in statistics. In contrast, taking a non-parametric approach that analyzes the median difference between reported and predicted expenses yields results that can be counterintuitive for readers who are unfamiliar with these methods. For example, if the median of reported expenses is \$180

---

<sup>5</sup> To illustrate this point imagine a participant who reports expenses of \$800, predicts expenses of \$700, and therefore has a bias score of \$100. If the dollar value at the 95<sup>th</sup> percentile of reported and predicted expenses is equivalent and  $\leq$  \$700, then this participant's bias score will become zero. Consider also a scenario where the 95<sup>th</sup> percentile of reported expenses is \$690, and the 95<sup>th</sup> percentile of predicted expenses is \$700. In this case winsorizing reported and predicted expenses independently would reverse this participant's bias score from \$100 to -\$10.

and the median of predicted expenses is \$150, the median *difference* between reported and predicted expenses will quite likely *not* equal \$30.00. Issues such as this are further compounded in multi-condition studies where condition is a between-subjects variable and expenses (reported vs. predicted) are measured within-subject.

Despite the drawbacks of winsorization and non-parametric analysis we feel it is important to present them here so that readers can rest assured the analysis we present in the main text is consistent with these alternative approaches. In this appendix we also present the non-parametric correlations between perceived typicality and predicted expenses (H2), and we present the results of our central parametric analyses to demonstrate that these results do not change in any meaningful way when no data is excluded. Finally, in Web Appendix B, we summarize the impact that our transformation process has with respect to the magnitude of the bias, homogeneity of variance between conditions and time periods, and normality.

*Perceived Typicality.* To test our hypothesis that people predict their expenses will be more typical in the future than in the past (H1), we compared reported vs. predicted expense typicality at T1, T2, T3, and T4 using parametric paired samples t-tests. 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 7 and Table 7, participants predicted their expenses would be more typical in the next (vs. past) week at all four points in time. Figure 7 also illustrates that our atypical intervention reversed this tendency at T4. In sum, these results provide strong support for H1.

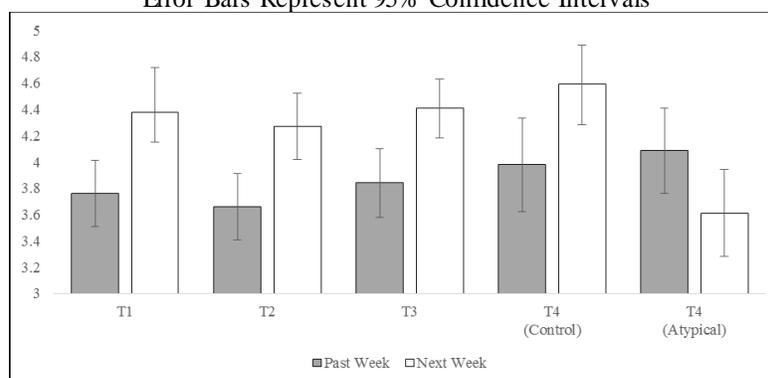
To test our hypothesis that perceived typicality of future expenses is negatively correlated with expense predictions (H2), we analyzed the non-parametric correlation (Spearman's rho) between perceived typicality of future expenses and weekly expense predictions for each week

of the study, as well as for the month. Perceived typicality of future expenses was negatively correlated with weekly expense predictions at T0 ( $r(185) = -.32, p < .001$ ), T2 ( $r(185) = -.23, p = .002$ ), and T4 ( $r(185) = -.22, p = .003$ ), as well as with monthly predictions ( $r(185) = -.15, p = .037$ ). The correlations at T1 and T3 were directionally consistent, though not significant (T1:  $r(185) = -.10, p = .17$ ; T3:  $r(185) = -.04, p = .64$ ).

**Figure 7**

**Mean Reported Expense Typicality for the Past Week vs. Mean Predicted Expense Typicality for the Next Week for each Week of Study 1**

Error Bars Represent 95% Confidence Intervals



**Table 7**

**Paired Samples T-tests Comparing Perceived Typicality of Reported Expenses for the Past Week vs. Perceived Typicality of Predicted Expenses for the Next Week in Study 1**

	T1	T2	T3	T4 (Control)	T4 (Atypical)
t-stat	3.79	4.37	3.74	2.98	2.27
p-value	<0.001	<0.001	<0.001	0.004	0.025

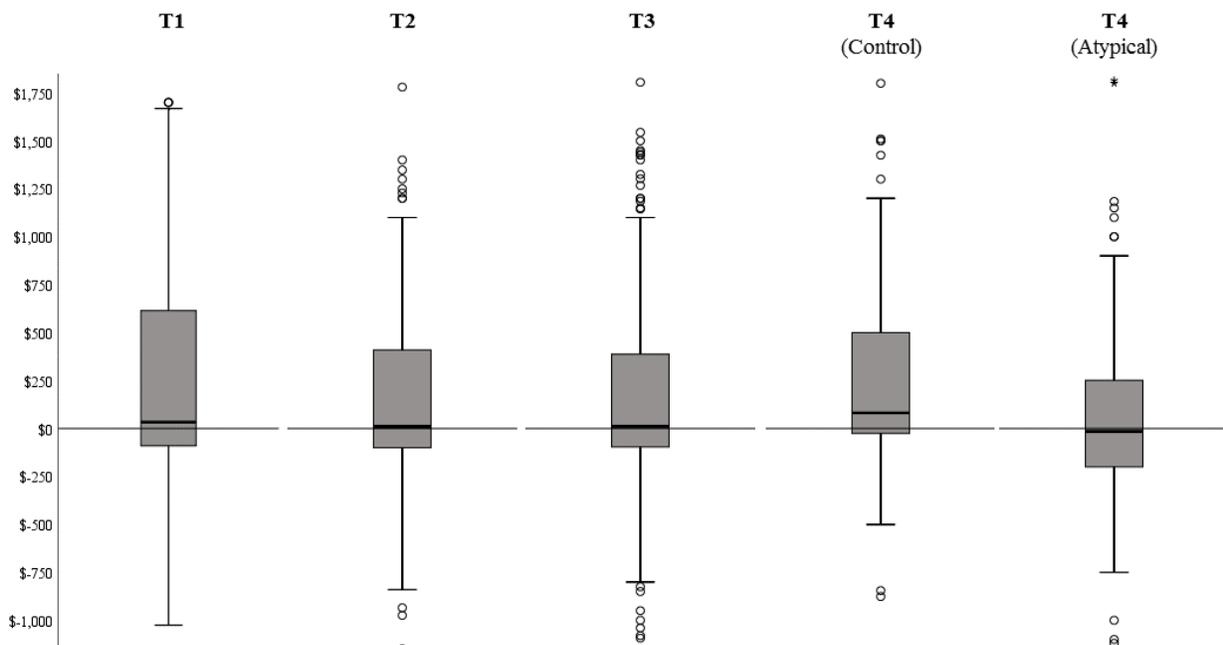
*Expense Prediction Bias (Recalled – Predicted Expenses)*. To test our hypothesis that consumers predict their future expenses will be lower than their past expenses (H3a), we compared reported expenses against predicted expenses at T1, T2, T3, and T4. That is, we tested whether participants predicted their expenses would be lower in week 2 than week 1, week 3 than week 2, and so on. As illustrated in Figure 8 and Table 8, both winsorized mean and raw

median bias scores were significantly different from zero, indicating that predicted expenses were significantly lower than reported expenses at each stage of the study until the atypical intervention was deployed.

**Figure 8**

**Distribution of Bias Scores (Reported Expenses for Past Week – Predicted Expenses for Next Week) at T1 – T4 in Study 1**

Bottom whisker = 1<sup>st</sup> percentile. Bottom of the box = 25<sup>th</sup> percentile. Midline = median. Top of the box = 75<sup>th</sup> percentile. Top whisker = 1.5x the height of the box. Circles = data points 1.5x–3x the height of the box. Stars = data points > 3x the height of the box.



**Table 8**

**Winsorized Mean and Raw Median Bias Scores (Recalled – Predicted Expenses) at T1 – T4 in Study 1**

	T1	T2	T3	T4 (control)	T4 (atypical)
Mean	\$316.41**	Mean \$96.78†	Mean \$226.91**	Mean \$290.21**	Mean -\$53.71
Median	\$33.00**	Median \$10.00*	Median \$10.79**	Median \$81.50**	Median -\$15.00
SD	823.41	SD 772.45	SD 710.87	SD 588.30	SD 658.81

Notes: †p < .10, \*p < .05, \*\*p < .01. Mean bias scores and SD's were calculated after winsorizing the distribution of bias scores at the 5th and 95th percentile. Winsorized mean bias scores were compared against zero using a parametric one-sample t-test. Raw median scores were compared against zero using a non-parametric one-sample Wilcoxon signed ranks test.

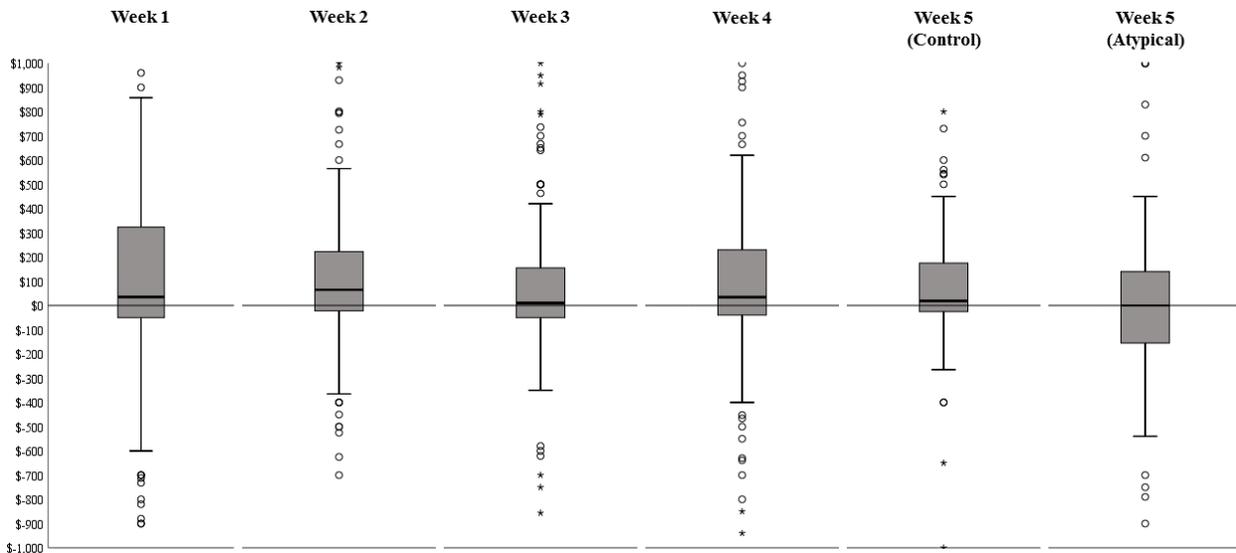
*Expense Prediction Bias (Actual – Predicted Expenses)*. As illustrated in Figure 9 and Table 9, there was a significant bias in each and every week of the study until week 5 when the atypical intervention effectively eliminated the bias, as measured by either the winsorized mean

or raw median. A pair of Mann-Whitney U tests further reveal that predictions were higher in the atypical condition than in the control (Median<sub>atypical</sub> = \$500.00, Median<sub>control</sub> = \$400.00, z = 1.95, p = .051), and that actual expenses did not differ between the two conditions (Median<sub>atypical</sub> = \$493.00, Median<sub>control</sub> = \$409.00, z = 1.17, p = .24). This confirms that the week 5 results are owed to the effect of our atypical intervention on predictions. It is also notable that our intervention slightly improved (non-parametric) correlational accuracy relative to the control condition (rho<sub>control</sub>(92) = .73, rho<sub>atypical</sub>(91) = .79).

**Figure 9**

**Bias Scores (Actual – Predicted Expenses) for each Week of Study 1**

Bottom whisker = 1<sup>st</sup> percentile. Bottom of the box = 25<sup>th</sup> percentile. Midline = median. Top of the box = 75<sup>th</sup> percentile. Top whisker = 1.5x the height of the box. Circles = data points 1.5x–3x the height of the box. Stars = data points > 3x the height of the box.



**Table 9**

**Winsorized Mean and Raw Median Bias Scores (Actual – Predicted Expenses) for each Week of Study 1**

	Week 1	Week 2	Week 3	Week 4	Week 5 (control)	Week 5 (atypical)
Mean	\$216.88**	\$171.25**	\$143.61**	\$154.93**	\$78.92*	-\$7.07
Median	\$35.00**	\$65.00**	\$10.00**	\$34.00**	\$19.00**	-\$0.00
SD	698.69	474.07	474.45	449.58	323.39	426.80
	Mean	Mean	Mean	Mean	Mean	Mean
	SD	SD	SD	SD	SD	SD

Notes: \*p < .10, \*\*p < .05, \*\*\*p < .01. Mean bias scores and SDs were calculated after winsorizing the distribution of bias scores at the 5th and 95th percentile. Winsorized mean bias scores were compared against zero using a parametric one-sample t-test. Raw median scores were compared against zero using a non-parametric one-sample Wilcoxon signed ranks test.

*Monthly Expense Prediction Accuracy (Actual – Predicted Monthly Expenses).* Both the winsorized mean and raw median monthly bias score are significantly different from zero, indicating that consumers under-predicted their monthly expenses ( $\text{Mean}_{\text{winsorized}} = \$751.77$ ,  $\text{Median}_{\text{raw}} = \$350.00$ ,  $\text{SD}_{\text{winsorized}} = 1675.62$ ; parametric one-sample t-test:  $t(186) = 6.14$ ,  $p < .001$ ; non-parametric one-sample Wilcoxon signed rank test:  $z = 5.10$ ,  $p < .001$ ). For further context: the median consumer in our sample under-predicted their monthly expenses by 13.3%. Furthermore, almost two-thirds of the sample (64.2%) under-predicted their monthly expenses ( $z = 3.88$ ,  $p < .001$ ).

### STUDY 3

*Replicating Study 1.* The results observed in study 1 were directly replicated in the control condition of study 3. Supporting H1, participants in the control condition of study 3 also predicted their expenses for the next week would be significantly more typical than their expenses for the past week ( $M_{\text{pastweek}} = 4.37$ ,  $\text{SD}_{\text{pastweek}} = 1.79$ ,  $M_{\text{nextweek}} = 4.61$ ,  $\text{SD}_{\text{nextweek}} = 1.72$ ; paired-samples t-test:  $t(436) = -3.31$ ,  $p = .001$ ,  $d = .14$ ). Supporting H2, the non-parametric correlation (Spearman's rho) between perceived typicality of future expenses and expense predictions was again significant and negative ( $\text{rho}(435) = -.19$ ,  $p < .001$ ). Finally, as detailed below, participants also predicted lower expenses for the next (vs. past) week in the control condition of study 3. We next expand our analysis to test for differences in perceived typicality and EPB across all three conditions.

*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 ( $F(2, 1105) = 34.60, p < .001$ ). Planned contrasts further revealed that perceived typicality was virtually identical in the control and typical conditions ( $M_{\text{control}} = 4.61, SD_{\text{control}} = 1.72, M_{\text{typical}} = 4.62, SD_{\text{typical}} = 1.64, t(1105) = .11, p = .91$ ), but significantly lower in the atypical condition ( $M_{\text{atypical}} = 3.69, SD_{\text{atypical}} = 1.66, t(1105) = -8.31, p < .001, d = .55$ ).

*Expense Prediction Bias (Recalled – Predicted Expenses)*. Table 10 summarizes the non-parametric median analyses performed on the study 3 expense data. First, by reading the control and typical columns vertically, it can be seen that the median of predicted expenses in these conditions is significantly lower than the median of recalled expenses. This provides clear support for H3a. Second, by reading the atypical column vertically it can be seen that the median of predicted expenses in this condition is significantly *higher* than the median of recalled expenses. In other words, this analysis suggests that our atypical intervention not only neutralized EPB in this study, it reversed it. This provides support for H4 because people tend to remember the past in slightly optimistic terms (Buehler, Griffin, and Peetz 2010), which suggests that over-prediction in the context of prediction vs. recall will translate into more accurate predictions in the context of prediction vs. reality (i.e., the expenses a consumer actually incurs during the target week). Finally, by reading horizontally across the top two rows of the table, it can be seen that the median of recalled expenses did not differ by condition, and that the median of predicted expenses did. Contrast analysis confirms that the median of predicted expenses did not differ between the control and typical conditions ( $p = .47$ ), and that the median of predicted expenses was significantly higher in the atypical condition than in the control and typical conditions ( $p = .002$ ).

**Table 10****Non-Parametric Median Tests for Recalled and Predicted Expenses in Study 3**

Wilcoxon Signed Ranks: Tests for median differences between recalled and predicted expenses within condition (i.e., within each column). Kruskal-Wallis H: Tests for median differences across conditions (i.e., across each row).

	Control	Typical	Atypical	Kruskal-Wallis H Test
<b>Median Recalled Expenses</b>	\$245.00	\$207.00	\$250.00	H = 1.57 <i>p</i> = .46
<b>Median Predicted Expenses</b>	\$200.00	\$200.00	\$300.00	H = 17.13 <i>p</i> < .001
<b>Wilcoxon Signed Ranks Test</b>	<i>z</i> = -2.43 <i>p</i> = .015	<i>z</i> = -3.29 <i>p</i> = .001	<i>z</i> = 2.39 <i>p</i> = .017	

Table 11 summarizes the parametric winsorized mean analyses performed on the study 3 expense data. First, by reading the control and typical columns vertically, it can be seen that the winsorized mean bias score in these conditions is significantly greater than zero, indicating that participants in these conditions predicted lower expenses for the next week than they recalled for the past week. This again provides clear support for H3a. Second, by reading the atypical column vertically it can be seen that the winsorized mean bias score in the atypical condition is significantly *less* than zero. So, as above, this analysis suggests that our atypical intervention not only neutralized EPB in this study, it reversed it. Finally, by reading horizontally across the bottom two rows of the table, it can be seen that the winsorized mean bias scores in the control and typical conditions did not significantly differ, and that the winsorized mean bias score in the atypical condition was significantly higher than in the control and typical conditions. (NB: The a one-way ANOVA with condition as the IV and winsorized bias scores as the DV produced a significant omnibus test ( $F(2, 1105) = 18.26, p < .001$ ).

**Table 11**  
**Winsorized Mean EPB Scores in each Condition of Study 3**

	Control (1)	Typical (2)	Atypical (3)
<b>Mean EPB</b>	\$27.93	\$37.54	-\$95.85
<b>(SD)</b>	(248.56)	(177.84)	(486.14)
<b>t-test</b>	t(436) = 2.35	t(346) = 3.93	t(323) = -3.55
<b>(diff from zero)</b>	p = .019	p < .001	p < .001
<b>Contrast</b>	t(1105) = .42		
<b>(1 vs. 2)</b>	p = .68		
<b>Contrast</b>	t(1105) = 6.04		
<b>(1 &amp; 2 vs. 3)</b>	p < .001		

Note: The distribution of EPB scores was winsorized at the 5th and 95th percentile of each condition. This table presents the resultant means and standard deviations with which statistical tests were performed.

*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 ( $F(2, 1105) = 6.88, p = .001$ ). Consistent with H4, planned contrasts further revealed that the number of expenses listed in the atypical condition ( $M_{\text{atypical}} = 1.54, SD_{\text{atypical}} = 1.59$ ) was significantly higher than in the control and typical conditions ( $M_{\text{control}} = 1.19, SD_{\text{control}} = 1.48, M_{\text{typical}} = 1.14, SD_{\text{typical}} = 1.52; t(1105) = 3.71, p < .001$ ), and the number of expenses listed in the control and typical conditions does not differ ( $t(1105) = .40, p = .69$ ). Furthermore, the non-parametric correlation between the number of expenses listed and predictions is positive and significant in all three conditions ( $r$ 's  $> .18, p$ 's  $< .01$ ). Finally, a one-way ANOVA with condition as the IV and average dollar amount of atypical expenses as the DV reveals no effect of condition ( $F(2, 619) = .88, p = .41$ ).

**WEB APPENDIX B:**  
**DETAILS OF THE DATA TRANSFORMATION PROCESS IN STUDIES 1 and 3**

**Table 12**

**Study 1 Distributional Analysis**

**Expense distributions w/ no exclusions and no transformation (n = 187)**

	<b>Control Condition (n = 94)</b>					<b>Atypical Condition (n = 93)</b>					<b>Levene's Test of HOV</b>	
	Median	Mean	SD	Skew	Normality Test	Median	Mean	SD	Skew	Normality Test	Levene Statistic	p-value
					p-value					p-value		
<b>Reported</b>	\$409.00	\$722.15	\$960.29	4.17	< .001	\$493.00	\$995.82	\$1,910.35	5.61	< .001	2.79	0.097
<b>Predicted</b>	\$400.00	\$620.88	\$710.94	2.20	< .001	\$500.00	\$867.88	\$1,326.50	5.94	< .001	2.20	0.139

**Expense distributions after excluding outliers (n = 183)**

	<b>Control Condition (n = 92)</b>					<b>Atypical Condition (n = 91)</b>					<b>Levene's Test of HOV</b>	
	Median	Mean	SD	Skew	Normality Test	Median	Mean	SD	Skew	Normality Test	Levene Statistic	p-value
					p-value					p-value		
<b>Reported</b>	\$431.50	\$733.22	\$967.57	4.14	< .001	\$493.00	\$852.70	\$1,232.91	5.05	< .001	0.47	0.495
<b>Predicted</b>	\$400.00	\$572.97	\$630.42	2.24	< .001	\$500.00	\$870.47	\$1,336.37	5.94	< .001	3.83	0.052

**Expense distributions after excluding outliers and applying LN transformation (n = 183)**

	<b>Control Condition (n = 92)</b>					<b>Atypical Condition (n = 91)</b>					<b>Levene's Test of HOV</b>	
	Median*	Mean*	SD	Skew	Normality Test	Median*	Mean*	SD	Skew	Normality Test	Levene Statistic	p-value
					p-value					p-value		
<b>Reported</b>	6.0668 (\$431.30)	6.0992 (\$445.50)	0.98	0.16	0.964	6.2005 (\$493.00)	6.2005 (\$493.00)	1.08	-0.35	0.149	0.06	0.810
<b>Predicted</b>	5.9915 (\$400.01)	5.9013 (\$365.51)	0.94	0.27	0.147	6.2146 (\$500.00)	6.2139 (\$500.00)	1.06	-0.05	0.793	1.03	0.310

\*Values in parentheses represent exponentiated medians and means.

Study 3 Distributional Analysis

Table 13

Expense distributions w/ no exclusions and no transformation (n = 1108)																	
Control Condition (n = 437)					Typical Condition (n = 347)					Atypical Condition (n = 324)			Levene's Test of HOV				
	Median	Mean	SD	Skew	Normality Test p-value	Median	Mean	SD	Skew	Normality Test p-value	Median	Mean	SD	Skew	Normality Test p-value	Levene Statistic	p-value
Recalled	\$245.00	\$446.14	\$880.19	6.85	< .001	\$207.00	\$363.41	\$424.98	2.79	< .001	\$250.00	\$479.02	\$841.59	6.08	< .001	4.47	0.012
Predicted	\$200.00	\$385.63	\$566.67	3.80	< .001	\$200.00	\$350.00	\$603.21	7.02	< .001	\$300.00	\$593.22	\$968.04	3.78	< .001	20.15	< .001
Expense distributions after excluding outliers (n = 1048)																	
Control Condition (n = 416)					Typical Condition (n = 332)					Atypical Condition (n = 300)			Levene's Test of HOV				
	Median	Mean	SD	Skew	Normality Test p-value	Median	Mean	SD	Skew	Normality Test p-value	Median	Mean	SD	Skew	Normality Test p-value	Levene Statistic	p-value
Recalled	\$247.50	\$426.33	\$793.74	7.65	< .001	\$200.00	\$342.99	\$382.21	2.92	< .001	\$250.00	\$459.93	\$701.04	5.25	< .001	4.93	0.007
Predicted	\$200.00	\$380.04	\$538.01	3.82	< .001	\$200.00	\$336.67	\$516.33	7.59	< .001	\$292.50	\$492.03	\$651.27	3.04	< .001	8.14	< .001
Expense distributions after excluding outliers and applying LN transformation (n = 1048)																	
Control Condition (n = 416)					Typical Condition (n = 332)					Atypical Condition (n = 300)			Levene's Test of HOV				
	Median*	Mean*	SD	Skew	Normality Test p-value	Median*	Mean*	SD	Skew	Normality Test p-value	Median*	Mean*	SD	Skew	Normality Test p-value	Levene Statistic	p-value
Recalled	5.5154 (\$248.49)	5.4742 (\$238.46)	1.02	0.17	0.002	5.3033 (\$201.00)	5.3713 (\$215.14)	1.00	-0.17	0.050	5.5255 (\$251.01)	5.5362 (\$253.71)	1.08	0.013	0.353	1.11	0.120
Predicted	5.3033 (\$201.00)	5.3803 (\$217.09)	1.03	0.24	0.005	5.3033 (\$201.00)	5.3056 (\$201.46)	1.00	0.02	0.185	5.6816 (\$293.42)	5.6094 (\$272.98)	1.10	0.000	0.160	1.64	0.194

\*Values in parentheses represent exponentialized medians and means.

*WEB APPENDIX C:  
ANALYSIS OF EXPLORATORY VARIABLES IN STUDY 1*

Study 1 explored the relationship between EPB and several theoretically and practically relevant individual differences that may be predictive of the bias. The first was the presence of a savings goal, because motivation to save has been tied to lower predictions (Peetz and Buehler 2009). The second was trait optimism (Scheier, Carver, and Bridges, 1994). We predicted that trait optimism would actually *not* be correlated with EPB because research on the planning fallacy has demonstrated that optimistic task completion time predictions are not the result of an optimistic disposition (Buehler, Griffin, and Ross, 1994). Nonetheless, because the relationship between trait optimism and expense predictions seems intuitively compelling and has not yet been explored, we felt there was value in testing it.

The third individual difference we measured was short-term financial propensity to plan (PTP; Lynch et al., 2010). One prediction regarding the relationship between short-term financial PTP and EPB is that consumers with a higher PTP will display lower EPB because they are more attuned to (or concerned with) future outcomes. However, it could be the case that greater plan-focus leads consumers to be less attentive to unplanned expenses and therefore more likely to under-predict their future expenses (cf. Buehler, Griffin, and Peetz, 2010). We therefore included the Lynch et al. (2010) measure of short-term financial PTP so that we could explore these competing hypotheses.

The fourth individual difference measure we included was the Rick, Cryder, and Lowenstein (2008) spendthrift-tightwad scale. There are three hypotheses for how this measure could be correlated with EPB. The first hypothesis is that tightwads may display *higher* EPB because anticipatory pain of paying causes them to predict even lower expenses than they

actually incur. The second hypothesis is that tightwads could display *lower* EPB if they are more sensitive to expenses and therefore more accurate. Finally, spendthrifts may display higher EPB because they lack pain of paying during purchase and therefore may be more likely to overspend vs. prediction.

The fifth individual difference measure we included was numeracy (Schwartz et al., 1997) because consumers who are unable to perform the mental calculations required to make an accurate expense prediction may display higher EPB. The sixth measure was linear vs. cyclical time orientation (adapted from Tam and Dholakia, 2014). Consumers with a stronger cyclical orientation may display lower EPB because they see life events as a series of recurring events (Tam and Dholakia, 2014) and therefore should be more easily able to incorporate past atypical expenses into their predictions for the future. The seventh individual difference we included was openness to experience (John, Donahue, and Kentle, 1991), because being more likely to consider a wider range of outcomes when making predictions could be associated with lower EPB. We also measured temporal discounting for both losses and gains (Kirby & Maraković 1996). Consumers with a relatively higher discount rate for losses (i.e., those with a stronger preference to pay more later vs. less now) may display higher EPB because they may want to postpone payment as much as possible, and our measure of EPB in this study was for the coming week. The same logic applied in reverse led us to believe that temporal discounting of gains may be negatively correlated with EPB.

In addition to the individual differences of theoretical interest above, we also included measures that let us explore the relationship between EPB and preferred form of payment (credit card, debit card, cash, other), budgeting behavior, perceived expense predictability, socio-economic status, gender, and education. We reasoned that heavier reliance on cards as a form of

payment might be correlated with EPB because using credit cards reduces pain of payment (Thomas, Desai, & Seenivasan, 2010), which could lead to spending more than predicted. Budgeting could lead to lower EPB by prompting consumers to consider a wide array of future outcomes, or higher EPB by causing them to focus on minimize predicted (but not actual) expenses in pursuit of a savings goal (Peetz and Buehler, 2009). We expected that higher perceived predictability could be correlated with higher EPB because they may be associated with higher prediction confidence which has been shown to lead to less accurate predictions (Ulkumen et al., 2008). Finally, we wanted to explore the relationship between EPB and socio-economic status (SES), and EPB and education, to determine if neutralizing the bias could be of particular benefit to vulnerable consumers (i.e., those who are low SES or who have less formal education).

As can be seen in table 14, our exploration of the relationship between EPB and the variables described above yielded mostly null results. One explanation for this could be a lack of power – our sample size was determined by a power analysis that included an estimate of the EPB effect size, not an estimate of the correlation between EPB and these measures. Of course, a second explanation is that there is no meaningful relationship between EPB and these measures. In our view, this possibility is perhaps most exciting, because it suggests that prototypical prediction is a general cognitive process that causes expense under-prediction regardless of a consumer's orientation with respect to the individual difference variables we measured. However, future research is required to test this conjecture.

**Table 14**

Correlation between Individual Difference Measures and EPB in Study 1											
	Savings Goal	Trait Optimism	Propensity to Plan	Numeracy	Spendthrift - Tightwad	Openness to Experience	Temporal Discounting (Losses)	Temporal Discounting (Gains)	Risk Aversion (Losses)	Risk Aversion (Gains)	Cyclical vs. Linear Time Orientation
Week 1 EPB	0.05	-0.03	0.03	0.06	0.05	-0.09	-0.08	0.02	-0.13	0.06	-0.06
Week 2 EPB	0.06	-0.04	0.01	-0.03	0.09	0.09	0.09	0.07	0.00	0.07	-0.01
Week 3 EPB	-0.06	-0.07	-0.03	0.09	0.00	-0.11	0.04	-0.03	0.00	-0.06	-0.04
Week 4 EPB	-0.06	-0.05	-0.15*	.06	-0.04	-0.11	0.06	0.06	-0.04	-0.02	0.03
Monthly EPB	-0.04	-0.09	-0.08	-0.02	-0.07	-0.10	-0.14	0.09	-0.01	-0.04	0.00

\* $p < .05$ 

Correlation between Individual Difference Measures and EPB in Study 1 (Continued)											
	Credit Card	Debit Card	Cash	Sets Budget Y/N	Time Spent Budgeting	Expense Predictability	LN(Available Resources)	SES	Annual Income	Gender (male=1; female=0)	Education
Week 1 EPB	0.06	-0.08	0.04	0.09	-0.20	0.07	0.03	0.00	0.10	0.12	0.04
Week 2 EPB	-0.02	0.01	0.05	0.11	-0.13	-0.03	-0.07	-0.02	-0.02	-0.12	-0.12
Week 3 EPB	0.05	-0.02	-0.09	-0.04	-0.05	-0.07	0.01	-0.09	0.08	-0.09	-0.05
Week 4 EPB	0.04	-0.01	-0.02	-0.13	-0.02	-0.03	-0.04	0.08	0.05	-0.01	-0.06
Monthly EPB	-0.01	0.05	-0.02	0.03	0.00	-0.04	0.00	0.03	0.08	0.13	-0.11

We also used study 1 to explore the relationship between EPB and potential correlates of the bias such as savings, debt, and subjective financial well-being. As illustrated in table 15, this exploration also yielded mostly null results. As noted above, this could be due to a lack of power or a truly null relationship. It could also be due to measurement error: it is possible that not all expenses were posted online when participants reported their expenses at the end of each week, so our measure of EPB might be somewhat conservative, which could be suppressing the link between EPB and these measures. It is also possible that it is persistence in monthly EPB that links under-prediction with these variables, which would mean we were unable to capture relationships using (predominately) weekly measures of EPB. Examining these possibilities falls well outside the scope of this article, but certainly warrants examination in future research.

**Table 15**

	<b>\$1,000 Allocation Decision Task</b>			<b>Savings</b>		<b>Debt</b>		<b>Subjective Financial Measures</b>			
	Save	Repay Debt	Spend	LN(Target Monthly Savings)	LN(Actual Monthly Savings)	Fair & Fast Loan Use (Y/N)	Borrows from Friends and Family	Financial Well-Being	Financial Security	Financial Slack	Spends more than expected
Week 1 EPB	-.12†	.13†	-.03	-.04	-.04	-.09	-.01	-.06	-.04	-.02	.06
Week 2 EPB	.05	-.09	-.16*	.15*	.13†	.02	.02	-.02	.07	.05	-.06
Week 3 EPB	.05	-.07	-.06	.00	.01	-.04	.05	-.08	-.13†	.07	.06
Week 4 EPB	0.03	-0.02	0.02	-0.05	-0.02	0.11	0.05	0.02	0.03	0.03	0.05
Monthly EPB	0.00	-0.07	0.00	0.02	0.07	.15*	0.11	-0.01	-0.01	0.07	0.06

*WEB APPENDIX D:  
ANALYSIS OF EXPLORATORY VARIABLES IN STUDY 3*

After participants completed the study 3 measures detailed in the main text of this article, we had them complete five measures designed to let us to explore the relationship between EPB, financial slack (Zauberman and Lynch, 2005; Berman et al. 2016), various measures of spending, and available resources. These measures and the results they yielded are detailed below.

*Financial Slack:* The financial slack measure was “Using the scale below, please indicate how much spare money you expect to have in the next week, compared to an average week in your life (1 = Very little spare money; 7 = A lot of spare money).” We hypothesized that predicted slack would be higher in the control and typical conditions (vs. the atypical condition) because EPB was higher in these conditions. To test this hypothesis we performed a 3(condition: control vs. typical vs. atypical) x 2(order: predict then recall vs. recall then predict) ANOVA with slack as the dependent variable. Neither the main effects nor the interaction were significant ( $p$ 's > .24).

*WTP for Dinner:* This measure asked “Imagine that a friend invites you to go out for a fancy dinner next week. You will each pay for your own food and drinks. How much money would you be willing to spend on dinner, including all your food, drinks, taxes, and tip? (1 = \$0-\$10; 11 = More than \$100).” We hypothesized that WTP would be higher in the control and typical conditions (vs. the atypical condition) because EPB was higher in these conditions. To test this hypothesis we performed a 3(condition: control vs. typical vs. atypical) x 2(order: predict then recall vs. recall then predict) ANOVA with WTP as the dependent variable. Neither the main effects nor the interaction were significant ( $p$ 's > .24).

*WTP for an Emergency Loan:* This measure asked “Imagine that next week you find yourself with an unexpected bill. For example, suppose that you use your vehicle to get to work, and it requires an expensive repair that is not covered by insurance. Until you get your car repaired, you will have to walk to work, which will add an extra 60 minutes onto your commute each way. Now imagine that to help cover the cost of fixing your vehicle you are able to take out a \$350 loan which will need to be repaid in 2 weeks along with the lender's fee. Using the scale below, please indicate the highest lender's fee you would be willing to pay to be able to borrow the \$350 (1 = \$0; 11 = More than \$50).” We hypothesized that WTP would be higher in the control and typical conditions (vs. the atypical condition) because EPB was higher in these conditions. To test this hypothesis we performed a 3(condition: control vs. typical vs. atypical) x 2(order: predict then recall vs. recall then predict) ANOVA with WTP as the dependent variable. Neither the main effects nor the interaction were significant ( $p$ 's > .20).

*Loan Repayment Confidence:* This measure asked “Assuming that you took the loan offered in the previous question, how confident are you that you would be able to pay back the loan (including the lender's fee) within 2 weeks? (1 = Extremely confident; 7 = Extremely unconfident).” We hypothesized that confidence would be higher in the control and typical conditions (vs. the atypical condition) because EPB was higher in these conditions. To test this hypothesis we performed a 3(condition: control vs. typical vs. atypical) x 2(order: predict then recall vs. recall then predict) ANOVA with confidence as the dependent variable. Neither the main effects nor the interaction were significant ( $p$ 's > .26).

*\$1,000 Allocation Task:* This measure asked “Imagine that you have just inherited \$1,000 that you weren't expecting. How much of the \$1,000 would you use for each of the following? (Please note that your total must equal \$1,000).” We hypothesized that less money would be

allocated to saving in the control and typical conditions (vs. the atypical condition) because EPB was higher in these conditions. To test this hypothesis we performed a 3(condition: control vs. typical vs. atypical) x 2(order: predict then recall vs. recall then predict) ANOVA with the sum allocated to saving as the dependent variable. Neither the main effects nor the interaction were significant ( $p$ 's > .23). Replacing the savings dependent variable with the sum of money allocated to spending or debt repayment yields the same pattern of results.

*Available Resources:* This measure asked "Imagine that you have to pay an unexpected bill immediately. For example, suppose that you require an expensive medical procedure that is not covered by insurance. Considering all possible resources available to you (including savings, borrowing, etc.), what is the maximum dollar amount that you could come up with on short notice? Please enter the amount below." This measure was included so that we could gain some insight as to whether or not EPB is associated with socio-economic status. There was no correlation between this measure and EPB ( $r(1023) = .04, p = .16$ ).