

# Bigger Not Better: Unpacking Future Expenses Inflates Spending Predictions

Johanna Peetz

*Carleton University*

Roger Buehler

*Wilfrid Laurier University*

Derek J. Koehler

*University of Waterloo*

Ester Moher

*Children's Hospital of Eastern Ontario, University of Ottawa*

People often underestimate their future personal spending. Across four studies we examined an “unpacking” intervention to reduce this bias. Participants predicted spending for an upcoming week (Study 1), a weekend (Study 2a), a vacation (Study 2b), and for weeks versus self-nominated events (Study 3), and subsequently reported actual spending. In each case, unpacking the details of expected expenses increased spending predictions. In contexts where predictions tended to be too low (Study 1, 3), unpacking eliminated underestimation bias. However, in contexts where predictions were already unbiased, unpacking introduced an *overestimation* bias (Study 2, 3). Unpacking appears to make predictions bigger, not necessarily better.

In everyday life, people often try to estimate their future spending—when withdrawing money from the ATM, when contemplating a vacation or a shopping trip, or when planning next month’s budget—and mistakes can have serious consequences. For example, if people underestimate their weekly expenses, they might commit to events or purchases they cannot afford or take on excessive debt. The present research develops and tests a prediction strategy that involves breaking down or “unpacking” a future expense category into its constituent parts. We assess the effects of unpacking on spending predictions and explore contextual factors that may moderate the effectiveness of this strategy.

## BIAS IN SPENDING PREDICTIONS

Research suggests that people commonly underestimate how much money they will spend in the future (Peetz & Buehler, 2009, 2012; Sussman & Alter, 2012; Ülkümen, Thomas, & Morwitz, 2008; Yang, Markoczy, & Qi, 2007). For example, Ülkümen et al. (2008) asked university students to predict how much they would spend in the coming month, and subsequent reports indicated that they underestimated their actual spending by about 28%. Similarly, Peetz and Buehler (2009) asked university students to recall how much money they had spent in a previous week, to predict their spending for the upcoming target week, and later to report their actual spending during the target week (e.g., by tracking expenses with a daily diary). Participants predicted to spend about 30% less in the target week than they had spent previously; however, the amount they ended up

---

Correspondence should be sent to Johanna Peetz, Psychology Department, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada. E-mail: johanna\_peekz@carleton.ca

spending did not differ from the past, and thus they tended to underestimate spending substantially.

People's tendency to underestimate future spending may be moderated by contextual factors. Notably, recent research suggests that the bias is reduced or eliminated when people predict spending for discrete events (a birthday party, a specific shopping trip) rather than extended periods (the coming week or month) (Peetz & Buehler, 2012, 2013). Similarly, (unpublished) studies that examined students' spending predictions for their Christmas shopping found no evidence of bias (Peetz & Buehler, 2007; Spiller & Lynch, 2010). There is also evidence that people can predict very accurately their spending for regularly occurring, ordinary events during a week (Sussman & Alter, 2012). Thus, although people frequently underestimate their future spending, the bias appears to be less prevalent for prediction targets that are specific and concrete or ordinary.

### UNPACKING SPENDING PREDICTIONS

The tendency to underestimate future spending may be due, in part, to cognitive processes underlying prediction. To generate behavioral predictions, people typically create a mental representation of the target event, such as an imagined scenario of the event unfolding (Buehler, Griffin, & Peetz, 2010; Dunning, 2007; Epley & Dunning, 2000; Kahneman & Tversky, 1979). To predict their spending for an upcoming week, for instance, individuals may mentally simulate the upcoming week, imagining the expenses they will encounter as the week unfolds. The problem is that such scenarios typically do not provide a comprehensive representation but rather tend to be schematic and oversimplified. Predictors may focus on a limited set of salient or representative expenses but fail to anticipate many peripheral events that will require them to spend money.

We propose that people will generate larger spending estimates, and thus be less prone to an underestimation bias, if they are prompted to break down the overall

prediction target into smaller components. This proposal builds upon previous work examining effects of "unpacking" on judgment and prediction with its origins in support theory (Tversky & Koehler, 1994). According to support theory, the subjective probability of a multifaceted event increases when the event is "unpacked" into its constituent parts (Tversky & Koehler, 1994; also Fischhoff, Slovic, Lichtenstein, Read, & Combs, 1978; Rottenstreich & Tversky, 1997). For example, people conclude that it is more likely someone has died of "natural causes" when they are asked to estimate the likelihood of several specific, concrete natural causes (e.g., cancer, heart disease, kidney failure) than when they make a single overall estimate. By unpacking an event, possibilities that may have been initially overlooked are brought to mind, and those possibilities that have already been considered are made more salient (Tversky & Koehler, 1994).

Applied to behavioral prediction, an unpacking intervention prompts forecasters to break down a multifaceted future event into its smaller constituent parts before generating an overall prediction. Unpacking procedures have been advocated to address a variety of biases in self-relevant prediction and evaluation (Kruger & Evans, 2004; Savitsky, Van Boven, Epley, & Wight, 2005). Kruger and Evans (2004) showed that unpacking can curb people's tendency to underestimate task completion times (i.e., the planning fallacy; Buehler et al., 2010). In comparison to control participants, those prompted to unpack a target task (e.g., formatting a document) by listing all the individual subtasks required to carry it out (e.g., italicizing, punctuating, adding special characters) predicted the task would take longer, and thus were less likely to underestimate completion times (for a related effect, see Forsyth & Burt, 2008).

We extend the previous work on unpacking into the realm of financial prediction by testing an intervention that prompts individuals to unpack an overall spending category into smaller constituent parts. Before predicting their spending for an upcoming week, for instance, individuals are prompted to generate a detailed list of

TABLE 1  
Mean Predicted and Actual Spending by Unpacking Condition

	<i>Condition</i>	<i>Predicted Spending</i>	<i>Actual Spending</i>	<i>Effect Size (d)</i>	<i>Correlation (r)</i>
Study 1 Week	Control	83.14	130.71	.56	.77
	Unpacking	170.34	173.38	.01	.58
Study 2a Weekend	Control	59.52	73.10	.21	.80
	Unpacking	116.86	78.59	.49	.19 ns
Study 2b Vacation	Control	415.57	488.94	.23	.72
	Unpacking	623.63	516.75	.23	.76
Study 3 Week	Control	112.41	150.45	.48	.41
	Unpacking	178.54	179.63	.01	.71
Event	Control	48.11	50.97	.06	.94
	Unpacking	104.55	80.06	.30	.72

the specific expenses they will incur. We expect that people do not normally unpack their expenses in this detailed manner, and thus the intervention will prompt individuals to attend to nonfocal expense possibilities that would otherwise be neglected. Our primary hypothesis, then, is that unpacking will result in increased spending predictions. Given that people often underestimate spending, this could translate, at least sometimes, into more accurate predictions.

### IMPACT ON BIAS AND ACCURACY

To consider the impact of unpacking on prediction accuracy, we distinguish between two forms of accuracy, referred to as bias and discrimination (Epley & Dunning, 2006; see also Buehler et al., 2010; Dunning, 2007; Kruger & Evans, 2004). Discrimination or correlational accuracy refers to whether individuals scoring higher on prediction (relative to others in the sample) also score relatively high on actual behavior; this can be indexed here by the correlation between predicted and actual spending. Prediction bias refers to the extent to which the average prediction matches the average of actual behavior and can be operationalized here by comparing mean predicted spending to mean actual spending. Predictions low in bias are also said to be well calibrated.

We suggest that, by either of these standards, unpacking will not necessarily improve spending predictions. First, as Kruger and Evans (2004) noted, the theory underlying unpacking does not imply an increase in discrimination. Unpacking increases the accessibility of the subcomponents of a broader prediction target, thereby increasing overall estimates, but not necessarily making them less prone to error. In their studies of time prediction, unpacking tasks did not usually improve discrimination. Likewise, we do not anticipate that our unpacking intervention will improve the discrimination of spending estimates.

When it comes to bias, the unpacking intervention could be more effective. Kruger and Evans (2004) consistently found reductions in prediction bias—where participants had previously underestimated the time a project would take, so they were less biased after unpacking the steps involved in the project. However, it is still important to emphasize that the logic behind unpacking suggests only that it will increase predictions, which will not necessarily reduce bias. We suggest that the effect of unpacking on prediction bias may depend on contextual factors that determine whether people are naturally prone to bias in the first place. In contexts where people typically underestimate future spending, unpacking should attenuate this bias. But what effect will unpacking have in contexts where predictions are

normally unbiased? One possibility is that unpacking influences prediction in a selective or discriminative manner, increasing only those predictions that would otherwise have been too low. However, it may be that unpacking increases prediction indiscriminately, regardless of the initial level of bias. In this case, the intervention would inflate predictions even in contexts where predictors are naturally unbiased, creating an overestimation bias.

Notably, in their work on task completion prediction, Kruger and Evans (2004) were unable to resolve this issue because predictions were always biased to begin with. In the present research we vary, both across studies and within, a contextual factor that may determine whether spending predictions are initially biased—that is, whether the predictions target an extended period or a more concrete event (Peetz & Buehler, 2013). For extended periods, we expect that people will typically underestimate spending and that unpacking will curb this bias. For discrete events, in contrast, we expect that predictions will initially be unbiased and that unpacking may create an overestimation bias.

### OVERVIEW OF STUDIES

Four experiments examined the effects of an unpacking intervention on predictions of future spending. In each case, participants made a spending prediction and subsequently reported the amount they actually spent. Participants were prompted to create an “unpacked” list of specific, constituent expenses either *before* making their overall prediction (unpacking condition) or not until *after* they had made their overall prediction (control condition). We also varied whether the spending predictions concerned an extended period (i.e., weekly spending) or a more concrete event, to capture whether predictions were naturally prone to bias. In Study 1, participants predicted their spending for the upcoming week. In the next study participants predicted spending for more concrete targets—the upcoming weekend in Study 2A and a planned vacation in Study 2B—that were expected to be less prone to underestimation bias. In Study 3 we manipulated the type of prediction within a single study by assigning participants to predict spending for either the upcoming week or a specific, self-nominated event.

In each study we tested for prediction bias by comparing predicted spending with subsequent reports of actual spending, and examined the degree of correlation between predicted and actual spending. For an overview of predicted and actual spending across all studies, see Table 1. In addition to these quantitative tests, we examined the itemized lists generated during the unpacking exercise to determine how well these matched the items that were actually purchased.

## STUDY 1: UNPACKING BIASED PREDICTIONS

Participants predicted their overall spending for the upcoming week. They were also asked to create an itemized list of the specific expenses they expected to incur during the week either just before making the overall spending prediction (unpacking condition) or not until after making the prediction (control condition). Participants later reported their actual spending for the week. We expected that participants in the control condition would underestimate spending, in line with past research (Peetz & Buehler, 2009, 2012; Ülkümen et al., 2008). We hypothesized that unpacking expenses prior to prediction would result in overall larger spending predictions and consequently would reduce or eliminate the underestimation bias.

### Method

#### Participants

One hundred nineteen undergraduate students at a Canadian university were initially recruited. Fifty-three participants did not complete the second part of the study.<sup>1</sup> Two participants were excluded as outliers on predicted or actual spending ( $>3 SD$ ). The final sample consisted of 13 male and 51 female participants ( $M_{\text{age}} = 19.70$  years,  $SD = 2.57$ ). Participants were compensated with course credit. One person completed the second part but did not report actual spending (resulting in discrepancies between the total sample size and the reported degrees of freedom).

#### Procedure

The study was conducted online and had two parts. In the first part, participants were asked to indicate their age and gender and then were randomly assigned to either the unpacking ( $n = 29$ ) or control ( $n = 35$ ) condition. In the unpacking condition, participants were asked to generate an itemized list of all the individual expenses they would incur during the next week. Specifically they were asked,

Think about the activities you'll do during the week. What items/services will you buy? Now, please list as many details of the next week as you can think of. Include all expenses that you might incur during the week. Include expenses that you pay with cards (i.e., debit card, credit card) or cash. Use as many or as few lines as you need.

<sup>1</sup>In this and subsequent studies, attrition was not different between unpacking and control conditions.

Participants were provided with 10 lines on which to enter the items and corresponding lines on which to enter the price of each item. Then, they were asked to make an overall prediction of their expenses for the week:

Now, think about all your expenses for next week. Include all expenses associated with this week, expenses that you'll pay with cards (i.e., debit card, credit card) or cash. How much money will you spend in the next week (i.e., the next 7 days)?

Participants also rated their confidence about this prediction, on scale ranging from 1 (*not confident at all*) to 7 (*very confident*).<sup>2</sup> Participants in the control condition were not asked to generate the itemized list before making the overall spending prediction. However, to control for potential effects of unpacking on actual spending, the control participants generated the list at the end of the session. We recorded the number of items participants listed during the unpacking task.

In this study and in all subsequent studies, participants were not told to keep track of their spending, and were not led to expect a spending report after the week. They were simply asked to expect a follow-up session to the study. Participants were prompted by e-mail to complete the second part of the study after 7 days had passed. On average, participants completed the second part 7.6 days after the first part, and 90% completed the survey on the 7th or 8th day. Participants were asked, using wording consistent with the initial survey, to provide an overall estimate of how much they actually spent during the week and to rate their confidence in this overall estimate. Finally, participants provided an itemized list of their individual purchases and the price of each. We recorded the number of items participants reported purchasing.

A research assistant compared the lists of itemized expenses for the first and the second part and assessed the number of purchases that had been foreseen compared to the number of purchases that had not been foreseen. For example, if a person listed "supper" on their list of expected expenses and "supper at A&W" on their list of actual expenses, this was coded as a foreseen expense. Similarly, if a participant listed "bar tab" and then reported "alcohol" as actual expense, that was coded as a foreseen expense. However, if participants

<sup>2</sup>We assessed confidence in this study, and each subsequent study, primarily to determine whether the unpacking procedure might influence participants' intuitions about the accuracy of their estimates. Unpacking did not influence confidence in any study and confidence levels were generally high (means ranged from 4.71 to 5.38 on a 7-point scale across all studies). As we had no specific hypotheses concerning this measure, and it did not reveal any effects, it is not discussed further.

reported purchasing items that could not be matched to any of the listed anticipated expenses (e.g., “mittens,” “movie”), this was coded as an unforeseen expense. Finally, we computed the proportion of predicted expenses that were actually bought during the week by dividing the number of foreseen expenses by the total number of predicted expenses.

## Results

### Overall Spending

First, to test for prediction bias, we compared mean levels of predicted and actual spending using. Participants in the control condition predicted to spend less money ( $M = 83.14$ ,  $SD = 74.90$ ) than they subsequently reported spending ( $M = 130.71$ ,  $SD = 95.09$ ,  $d = .56$ ).<sup>3</sup> In other words, the control participants underestimated their expenses by about 34%, in line with previous evidence of bias for weekly spending predictions (Peez & Buehler, 2009, 2013). In contrast, in the unpacking condition, predicted spending ( $M = 170.34$ ,  $SD = 153.55$ ) and actual spending ( $M = 173.38$ ,  $SD = 161.95$ ) differed very little ( $d = .01$ ). Thus the unpacking procedure eliminated the underestimation bias found in previous research. Further analyses revealed that the unpacking manipulation resulted in increases in predicted spending ( $d = .72$ ) that were not matched by equal increases in actual spending ( $d = .32$ ).<sup>4</sup>

Correlational analyses revealed a strong correlation between predicted and actual spending ( $r = .64$ ), which was similarly strong in both the control condition ( $r = .77$ ) and the unpacking condition ( $r = .58$ ). Although the unpacking manipulation reduced bias in prediction, it did not result in improved discrimination.

### Specific Purchases

We also examined the correspondence between participants' expected purchases and the actual purchases they reported. On average, participants predicted to buy 3.84 items or services ( $SD = 1.77$ ) over the week, and later reported buying a similar number of items or services ( $M = 3.64$ ,  $SD = 1.92$ ),  $d = .11$ . Of the actual purchases, however, only 1.84 purchases had been foreseen by the participants ( $SD = 1.02$ ). In other words, only half of the predicted items on participants' itemized

lists generated at Time 1 were actually purchased. The percentage of accurately predicted items differed very little for participants in the unpacking condition ( $M = 50.86$ ,  $SD = 26.61$ ) and the control condition ( $M = 50.00$ ,  $SD = 30.35$ ,  $d = .03$ ).

## Discussion

Study 1 provided the first test of the effectiveness of an unpacking intervention on predictions of future spending. When left to their own devices, participants underestimated their spending for an upcoming period, as in past research (Peez & Buehler, 2009; Ülkümen et al., 2008). However, prompting participants to create an “unpacked” list of specific expenses before predicting their weekly spending served to eliminate the bias. Notably, it is not yet clear whether unpacking simply increases overall estimates, or has effects that are more discriminating. Because there was substantial underestimation bias to begin with, an intervention that simply inflates estimates can effectively reduce bias. Indeed some of the findings suggest that unpacking simply inflated predictions: The unpacking manipulation did not strengthen the correlation between predicted and actual spending, and lists of anticipated expenses did not match actual expenses very well.

## STUDY 2: UNPACKING UNBIASED PREDICTIONS

The purpose of the next two studies (Study 2a and 2b) was to examine whether the unpacking procedure increases spending predictions in a discriminant or indiscriminate manner. One way to address this question is to test whether unpacking increases spending predictions in contexts where there is no initial bias. Previous research suggests that, unlike predictions for weekly spending, there is not a systematic bias in spending predictions that target a specific day (Peez & Buehler, 2009) or more concrete events (Peez & Buehler, 2007, 2013; Spiller & Lynch, 2010). Thus we asked participants to estimate spending for events that were more concrete than a 1-week period: the coming weekend (Study 2a) and a planned vacation (Study 2b). A week is a unit of time measurement that recurs regularly and is structured similarly each time (e.g., go shopping once, go exercise twice, go to work five times). When thinking about one's next week, this unit of time might therefore appear as relatively abstract and commonplace—1 week looks similar to many other weeks in its structure and event content. In contrast, the weekend is less structured (i.e., different events might happen on different weekends) and it might be perceived as more out of the ordinary. When considering a weekend, specific events might

<sup>3</sup>Note that we do not report null hypothesis significance testing and instead focus on descriptive statistics and effect sizes, in line with recent suggestions about the logic of inferential statistics (Trafimow, 2003, 2014).

<sup>4</sup>Because the distributions of predicted and actual spending tended to be positively skewed, we also log-transformed these variables and recomputed the analyses. This transformation did not alter the pattern of unpacking effects in any study. For ease of interpretation we report analyses performed on the untransformed data.

come to mind that define this particular weekend against others. Similarly, a vacation may be seen as a unique experience and may be defined through concrete activities and specific events during this vacation. Thus, a week may evoke more abstract thoughts (e.g., of the week as one like many others), whereas weekends or vacations might evoke more concrete thoughts (e.g., of unique experiences).

## Study 2a: Weekend Expenses

### Method

**Participants.** Sixty undergraduate students at a Canadian university were recruited. Sixteen participants did not complete the second part of the study and were excluded from the analyses. One participant was an outlier in actual spending ( $>3$   $SD$ ) and was excluded. The final sample consisted of 13 male and 30 female participants ( $M_{\text{age}} = 20.14$  years,  $SD = 2.41$ ). Participants were compensated with partial course credit for their participation.

**Procedure.** The study was conducted online in two sessions. In the first session, participants indicated their age and gender and were randomly assigned to either the unpacking condition ( $n = 22$ ) or the control ( $n = 21$ ) condition. In the unpacking condition, participants were asked to generate an itemized list of all the individual expenses they would incur during the upcoming weekend (defined as Friday evening to Sunday evening). Instructions were adapted from Study 1. Again, participants were provided with 10 lines to list the specific expense items and prices. Then they were asked to make an overall spending prediction for the upcoming weekend and to rate their confidence in the prediction as in Study 1. Participants in the control condition were not asked to generate the itemized list before making their overall spending prediction but generated this list at the end of the session. Participants completed the first session only on Monday, Tuesday, Wednesday, or Thursday, to ensure that they were not making predictions *during* the weekend.

Participants were prompted per e-mail to complete the second session on Monday morning after the weekend. On average, participants completed the second session 7.19 days after the first session ( $SD = 2.31$  days), and 90% of them completed it on the Monday or Tuesday after the weekend. Participants gave an overall estimate of how much they had actually spent over the entire weekend and rated their confidence in this overall estimate. Finally, participants provided an itemized list of the individual purchases they had made and the price of each purchase. A research assistant compared the itemized lists from the first and the second sessions

and determined the number of purchases that were foreseen and the number of purchases that were foreseen.

## Results

**Overall spending.** To test for prediction bias, we compared mean levels of predicted and actual spending. In the control condition, as expected, there was no clear evidence of prediction bias. Predicted spending ( $M = 59.52$ ,  $SD = 52.39$ ) was only slightly lower than actual spending ( $M = 73.10$ ,  $SD = 70.93$ ,  $d = .21$ ). In contrast, in the unpacking condition, participants predicted to spend *more* ( $M = 116.82$ ,  $SD = 89.57$ ) than they actually spent ( $M = 78.59$ ,  $SD = 63.73$ ,  $d = .49$ ). Thus, participants in the unpacking condition demonstrated a slight *overestimation* bias. Further tests examined separately the impact of unpacking on predicted and actual spending. The unpacking manipulation resulted in higher spending predictions ( $d = .78$ ) but did not have an effect on actual spending ( $d = .08$ ). Thus in the present study, unpacking did not reduce underestimation bias (of which there was little evidence to begin with) but instead introduced a bias in the opposite direction.

Also correlational analyses again indicated that the unpacking manipulation did not improve discrimination. Indeed the correlation between predicted and actual spending was much stronger in the control condition ( $r = .80$ ) than in the unpacking condition ( $r = .19$ ). Although this decreased correlation was not hypothesized, it may reflect participants' inability to foresee accurately the specific expenses they would incur. In any case, there was again no evidence that unpacking led to improved discrimination.

**Specific purchases.** We again examined whether participants' expected purchases matched the actual purchases they reported. On average, participants predicted to buy 3.12 items or services ( $SD = 1.77$ ) over the weekend, and later reported a similar number of purchases ( $M = 3.64$ ,  $SD = 1.92$ ,  $d = .11$ ). Of the actual purchases, however, only 1.36 purchases had been foreseen by the participants ( $SD = 1.02$ ). The percentage of accurately predicted items was similar for participants in the unpacking condition ( $M = 46.19$ ,  $SD = 41.02$ ) and the control condition ( $M = 43.58$ ,  $SD = 30.11$ ,  $d = .07$ ).

## Study 2b: Vacation Expenses

### Method

**Participants.** One hundred eighty-five students at a German university were recruited through advertisements on a psychology mailing list. Participants were volunteers who were incentivized with a lottery entry

for participation in the first session and another entry for participation in the second session. Of the original sample, only 83 participants completed the second session (the second session was again completely voluntary and was collected on average about 2 months later, which may explain the high attrition rate). Of these 83, eight participants were excluded (one did not complete the unpacking manipulation, three did not end up taking a vacation, four were outliers [ $>3 SD$ ] in predicted or actual spending). The final sample consisted of 46 female, 28 male, and one gender unidentified participants ( $M_{\text{age}} = 24.64$  years,  $SD = 4.66$ ).

**Procedure.** Participants completed the initial session in May, at the beginning of summer break. In an online survey, they first completed demographic items and then were asked to nominate a vacation they had planned for the summer break that (a) would cost at least some money and (b) would take at least 2 days. Participants indicated the approximate time they would depart for the vacation and the approximate time of their return. The vacations started an average 37.13 days ( $SD = 13.89$  days) after the initial survey session.

Participants were assigned to either an unpacking condition ( $n = 40$ ) or a control condition ( $n = 35$ ).<sup>5</sup> Participants in the unpacking condition were asked, as in previous studies, to list the individual expenses for their vacation. In this study, participants were inconsistent in whether they identified categories of expenses (“flights,” “meals,” “souvenirs”) or the exact items (“breakfast,” “dinner”) in the unpacking lines, even though instructions did not differ from the previous studies. We therefore counted the number of lines completed as an indicator of the number of expenses participants were contemplating, but the different levels of expense construal made it impossible to meaningfully compare the lists. Then, participants made an overall prediction of how much money they would spend for the vacation and rated their confidence about this prediction. Participants in the control condition were not asked to generate the itemized list until after making the overall spending prediction.

Participants were e-mailed a link to the follow-up survey on the Monday after their vacation had ended. Participants who listed an approximate vacation time (e.g., “August”) or a range of possible times were contacted at the end of their listed time range (e.g., August

30). On average, participants completed the second session 53.61 days after the first session. In the second session, participants gave an overall estimate of how much they had actually spent for the vacation and rated their confidence about this overall estimate. Finally, participants provided an itemized list of the individual purchases they had made and the price of each purchase.

## Results

**Overall spending.** We again tested for bias in spending predictions using paired  $t$  tests. As expected, there was little evidence of bias in the control condition: Participants predicted to spend only a little less ( $M = 415.57$ ,  $SD = 245.34$ ) than they later reported spending ( $M = 488.94$ ,  $SD = 387.32$ ,  $d = .23$ ). In contrast, participants in the unpacking condition predicted to spend a little more ( $M = 623.63$ ,  $SD = 563.49$ ) than they later reported spending ( $M = 516.75$ ,  $SD = 364.12$ ,  $d = .23$ ). Further analyses examined separately the impact of unpacking on predicted and actual spending. Participants in the unpacking condition predicted to spend more than those in the control condition ( $d = .48$ ), but actual spending differed very little across the two conditions ( $d = .07$ ).

Correlational analyses revealed a strong correlation between predicted and actual spending ( $r = .68$ ), suggesting that predictions did discriminate well between participants who went on to spend more versus less money, relative to others in the sample. The correlation between predicted and actual spending was equally strong in the control condition ( $r = .72$ ) and unpacking condition ( $r = .76$ ). Thus again, unpacking yielded no improvement in correlational accuracy.

**Specific purchases.** Finally, we examined the number of purchases/purchase categories participants listed. On average, participants predicted to make 4.51 purchases ( $SD = 1.42$ ) over the vacation, and ended up actually making slightly fewer purchases ( $M = 4.07$ ,  $SD = 1.33$ ,  $d = .32$ ). Participants predicted to buy somewhat more items in the unpacking condition ( $M = 4.75$ ,  $SD = 1.58$ ) than in the control condition ( $M = 4.23$ ,  $SD = 1.16$ ,  $d = .38$ ), and participants also bought slightly more items in the unpacking condition ( $M = 4.23$ ,  $SD = 1.49$ ) than in the control condition ( $M = 3.88$ ,  $SD = 1.09$ ,  $d = .27$ ). As mentioned previously, in the present study we were unable to assess the degree of correspondence between the predicted purchases and the items that were actually purchased.

## Discussion

The findings of Studies 2a and 2b provide further evidence that the underestimation bias observed for

<sup>5</sup>A subset of participants also completed three items assessing their budgeting attitudes toward vacations (e.g., “I would wait to purchase a ticket if it means that the vacation will be a lot cheaper.”). Completing these items (vs. not completing them) did not have an effect on predicted or actual spending and did not covary with the unpacking condition. Therefore, we present the data aggregated across all participants.

extended periods (e.g., the coming week or month) does not necessarily extend to prediction targets that are more temporally compact and concrete (see also Peetz & Buehler, 2013; Spiller & Lynch, 2010). Control participants did not systematically underestimate, or overestimate, how much they would spend on the weekend or a planned vacation. Consequently the two studies were able to test the effect of unpacking in the absence of initial bias, and thus to shed further light on how the intervention influences spending predictions.

The findings suggest that unpacking inflates spending predictions but is insensitive to the level of initial bias. Because there was no bias to begin with, the unpacking procedure possibly produced an *overestimation* of actual spending. Thus rather than eliminating bias, the procedure simply introduced bias of another sort. Also in terms of discrimination, there was again no evidence that unpacking strengthened the (already strong) correlation between predicted and actual spending. Taken together, this pattern of findings suggests that unpacking increases spending predictions rather indiscriminately. Unpacking results in larger spending estimates, but whether this translates into a reduction in bias, or an increase in bias, appears to depend on whether there is bias in the first place.

Notably, although it appears that unpacking is insensitive to initial bias, the evidence so far is based on comparisons across studies that differed in many respects. For example, although the spending targets in Study 2 were briefer and more concrete than the prediction target in Study 1, they also differed in other ways. For example the planned vacations were more temporally distant, more expensive, and less familiar than regular weekly spending. Conceivably it was one of these factors, rather than the initial absence of bias, that resulted in overestimation. Also, more generally, cross-study comparisons need to be interpreted cautiously, as the effects may be attributable to idiosyncrasies in the sample, period, or experimental context. Thus we sought to replicate the pattern of effects in a final study that varied the target of prediction within a single study.

### STUDY 3: UNPACKING TIME PERIODS VERSUS EVENTS

The first three studies suggest that unpacking increases predictions that would otherwise be too low (thus attenuating the underestimation bias) as well as predictions that are already unbiased (thus introducing an overestimation bias). In the present study we sought to replicate this pattern of effects while controlling for factors (e.g., samples, period, experimental context, etc.) that may have contributed to the pattern seen across studies. We recruited participants from the same sample

and assigned them to make spending predictions for either a period (the coming week) or a very specific event *set* within that same week (i.e., a specific, anticipated event). Again, we expected that for weekly spending predictions, participants in the control condition would underestimate their spending and that the unpacking procedure would eliminate this bias. We also expected that for a specific event, participants in the control condition would make unbiased predictions and that unpacking their expenses would still increase their predictions to the point of creating an overestimation bias.

## Method

### *Participants*

Of an initial sample of 144 undergraduate students, 110 returned data of their actual spending (11 participants did not do the event follow-up and 23 participants missed more than two diary entries). Of these 110, another four participants were excluded (one had missing data and three were outliers  $>3 SD$  in predicted or actual spending). The final sample consisted of 106 participants (52% female) between 17 and 28 years ( $M$  age = 18.43 years,  $SD = 1.45$ ).

### *Procedure*

In the first session, participants were asked to predict their spending for either the next week (weekly spending condition;  $n = 41$ ) or a self-nominated event (event spending condition;  $n = 55$ ). Participants in the weekly spending condition were told to think about and predict their spending for the next 7 days, as in Study 1. Participants in the event spending condition first nominated and briefly described an event that would occur in the next 7 days that would involve spending at least some money. Nominated events included going out for dinner with friends ( $n = 17$ ), seeing movies or concerts ( $n = 10$ ), festivals, going to bars or parties ( $n = 8$ ), birthdays or other celebrations ( $n = 11$ ), going shopping ( $n = 5$ ), and traveling ( $n = 5$ ). Events were on average 4.17 days in the future ( $SD = 2.85$ ).

In addition, participants were randomly assigned to an unpacking or control condition,<sup>6</sup> using instructions identical to Studies 1 and 2, with one exception. In the weekly spending condition, participants were provided

<sup>6</sup>For exploratory reasons, in this final study we also included a modified unpacking condition, in which participants were asked to contemplate their specific purchases rather than listing out all of the purchases and expected prices. In terms of prediction bias, this modified unpacking procedure resulted in levels of bias that fell between the control and unpacking condition but differed very little from each, making it somewhat difficult to interpret the impact. Thus we have limited our presentation of results to the conditions used in the previous studies.



with one text box per day (i.e., seven boxes) and asked to list the expenses they expected for each day in the respective box, separated by commas. In the event spending condition, participants were provided with 10 text boxes and asked to list the expenses they expected, separated by commas. After their prediction, participants also rated their confidence about the prediction, as in Studies 1 and 2.

In the second session, actual spending was assessed. Participants in the event condition were contacted by e-mail the day after their nominated event, reminded of the event, and asked to complete an online follow-up questionnaire. In this questionnaire, participants listed their purchases for and during the target event and reported how much they had spent overall for the event. Participants completed the second session an average 6.24 days ( $SD = 3.02$ ) after the first session. Participants in the weekly condition completed a spending diary during the target week: At the end of each day they accessed an online survey to report their spending for that day. Participants did not receive daily reminders and, consequently, they sometimes missed entries. Those who missed only one entry ( $n = 14$ ) or two entries ( $n = 4$ ) were included in the sample, and the missing values were replaced by the mean of the remaining entries. Those who missed more than two entries were excluded (see Study 3 Participants section).<sup>7</sup> After their report, participants also rated their confidence about the actual spending report. The confidence ratings of the diary entries were averaged into a confidence index ( $\alpha = .79$ ).

A research assistant compared the lists of itemized expenses for the first and the second session (i.e., the follow-up or the diaries) and again assessed the number of purchases that had been foreseen compared to the number of purchases that had not been foreseen.

## Results

### Overall Spending

We again tested for bias in prediction by comparing the means of predicted and actual spending, beginning with the weekly spending condition. For weekly predictions, participants in the control condition predicted to spend substantially less ( $M = 112.41$ ,  $SD = 72.63$ ) than they actually spent ( $M = 150.45$ ,  $SD = 86.29$ ), thus revealing an underestimation bias ( $d = .48$ ). This underestimation bias was eliminated by the unpacking intervention. Participants in the unpacking condition predicted to spend an amount ( $M = 178.54$ ,

$SD = 122.19$ ) very close to what they actually spent ( $M = 179.63$ ,  $SD = 129.59$ ,  $d = .01$ ).

For event predictions, there was no initial bias in prediction: Participants in the control group predicted to spend about as much ( $M = 48.11$ ,  $SD = 45.51$ ) as they actually spent ( $M = 50.97$ ,  $SD = 48.03$ ,  $d = .06$ ). However, participants in the unpacking condition predicted to spend *more* ( $M = 104.55$ ,  $SD = 92.09$ ) than they actually spent ( $M = 80.06$ ,  $SD = 70.02$ ,  $d = .30$ ).

Further analyses examined the effect of unpacking separately on predicted and actual spending. For weekly spending, unpacking resulted in increased predictions ( $d = .66$ ), without an equivalent increase in actual spending ( $d = .26$ ), and thus the underestimation bias was eliminated. For event spending, participants in the unpacking condition predicted to spend more than participants in the control group ( $d = .78$ ) without an equivalent difference in actual spending ( $d = .48$ ). Thus regardless of the initial level of bias, unpacking led to increased spending predictions without a corresponding impact on actual spending. Whether the increased predictions eliminated bias, or created overestimation bias, depended on the initial, baseline levels of bias.

Also, there was again no evidence that unpacking strengthened the correlation between predicted and actual spending. For weekly spending, there was a moderate correlation between predicted and actual spending ( $r = .62$ ), and this correlation was a little stronger in the unpacking condition ( $r = .71$ ) than in the control condition ( $r = .41$ ). For event spending, there was a strong correlation between predicted and actual spending ( $r = .80$ ), and this correlation was a little stronger in the control condition ( $r = .94$ ) than in the unpacking condition ( $r = .72$ ).

### Specific Purchases

We again examined whether participants' expected purchases matched the actual purchases they reported. On average, participants predicted to buy 13.20 items or services ( $SD = 8.10$ ) over the week and ended up actually buying only 9.16 items or services ( $SD = 8.30$ ),  $d = .49$ . Of the actual purchases, only 5.29 purchases (59%) had been foreseen by the participants ( $SD = 5.78$ ). Similarly, for the event predictions, participants predicted to buy 3.85 items or services ( $SD = 1.63$ ) over the week and actually bought a similar number of items or services ( $M = 3.65$ ,  $SD = 1.97$ ,  $d = .11$ ). Of the actual purchases, 2.56 purchases (76%) had been foreseen by the participants ( $SD = 1.42$ ). In sum, participants were more accurate in foreseeing what kind of items they would buy when predicting a specific event than when predicting a week ( $d = .65$ ). However, the percentage of accurately foreseen items was similar across the unpacking and control conditions ( $d = .02$ ).

<sup>7</sup>We conducted all the analyses presented next with different exclusion criteria (excluding those who missed one or more diary entries, three or more entries, four or more entries) and the results do not change.

## Discussion

This study replicated the pattern of effects seen across previous studies, demonstrating that unpacking expenses increased spending predictions regardless of initial bias. Being prompted to consider the details of an upcoming spending instance led to overall higher spending estimates—regardless of initial bias—but did not generally result in more accurate predictions. Indeed, when predictions were unbiased to begin with (for predictions concerning specific events), unpacking introduced an overestimation bias. There was also, once again, no evidence that unpacking resulted in improved discrimination.

## GENERAL DISCUSSION

People frequently underestimate their future personal spending, and the present studies tested an intervention developed to eliminate this bias. Participants assigned to an unpacking condition were prompted to identify the constituent expenses for a spending category before making an overall spending prediction. On the basis of research in other domains (Kruger & Evans, 2004; Moher & Koehler, 2013; Tversky & Koehler, 1994), we proposed that this unpacking process would yield increased spending estimates and counter any tendency toward underestimation. Consistent with the hypothesis, all four experiments revealed that unpacking increased spending estimates, by a factor of between +51% and +117%, depending on the group examined. Such increases should help to offset an underestimation bias and, in fact, in contexts where individuals normally underestimated spending, the intervention eliminated this bias.

The studies further explored whether the unpacking effects were sensitive to initial levels of prediction bias. The logic behind unpacking suggests only that it will increase predictions and not that it will do so in a selective or discriminative manner (Kruger & Evans, 2004; Tversky & Koehler, 1994), and indeed we found that the unpacking effects were insensitive to initial bias. Unpacking *always* increased predictions, even in contexts where predictions were typically unbiased, leading to a slight overestimation bias in those cases.

To summarize this pattern of effects, we standardized predicted and reported spending in each study and analyzed the impact of unpacking across all studies. In tests of weekly spending predictions (Study 1 and Study 3), where there was an underestimation bias in the control condition, no bias remained in the unpacking condition ( $d = .03$ ). For more concrete spending targets (Study 2a, Study 2b, and Study 3), where there was no bias to begin with, unpacking produced a small *overestimation* bias ( $d = .14$ ). Also, as an index of prediction bias we

computed the signed difference between predicted and actual spending and tested the impact of unpacking on the difference scores in a meta-analysis across all studies. For weekly spending, the difference scores were smaller (i.e., less underestimation bias) in the unpacking condition ( $M_{\text{diff}} = -2.15$ ) than in the control condition ( $M_{\text{diff}} = -41.99$ ,  $d = .39$ ). For concrete spending targets, the difference scores were larger (i.e., greater overestimation bias) in the unpacking condition ( $M_{\text{diff}} = 67.32$ ) than in the control condition ( $M_{\text{diff}} = -33.11$ ,  $d = .45$ ). Taken together, then, the studies revealed clearly that unpacking did not reduce bias in future spending predictions.

In addition, this meta-analysis across studies showed no evidence that unpacking yields improvement in correlational accuracy. Predicted and actual spending were quite highly correlated in general. For weekly spending, correlations were  $r = .63$  and  $r = .63$  for control and unpacking condition, respectively. For the concrete events, correlations were  $r = .85$  and  $r = .84$  for control and unpacking condition, respectively. In no study was there a marked improvement in the unpacking condition compared to the control condition. Of interest, unpacking future expenses also did not improve participants' intuitions concerning the accuracy of their own predictions (i.e., prediction confidence; see Footnote 1). Taken together, the findings suggest that unpacking made predictions bigger but not necessarily better.

## Implications

This research contributes to an emerging body of work examining people's attempts to predict their future spending. The findings provide further evidence of a tendency to underestimate spending for extended time periods (Peetz & Buehler, 2009, 2012; Ülkümen et al., 2008) and suggest that this bias does not generalize to spending targets that are more temporally compact and concrete (see also Peetz & Buehler, 2013). Moreover, the studies provided the first empirical test of an approach to prediction that has proven effective in other domains, and thus extend our knowledge of factors that moderate spending estimates. The findings also extend the research literature on debiasing strategies in general (Larrick, 2004; Lilienfeld, Ammirati, & Landfield, 2009) and the unpacking strategy in particular (Kruger & Evans, 2004; Moher & Koehler, 2013; Savitsky et al., 2005). Our findings highlight the importance of assessing multiple forms of predictive accuracy (e.g., calibration and discrimination) across multiple judgmental contexts (e.g., contexts with preexisting bias and no preexisting bias) when evaluating the effectiveness of an intervention.

The findings have widespread practical implications, because people's spending estimates guide many choices

and decisions. Decisions ranging from everyday choices (where to buy lunch, how to spend the weekend) to major life decisions (e.g., when to retire, whether to have another child) require a consideration of future expenses. Thus errors in prediction—either underestimation or overestimation—could have serious personal, social, and economic consequences. Our findings imply that an unpacking intervention may be most advisable in contexts where underestimation is prevalent, or there are little costs associated with overestimating expenses. However, if bias is unlikely to begin with, or overestimation would cause serious problems, it may be better to adopt another approach.

### Limitations and Future Directions

Unpacking procedures can take many forms, and a notable feature of our manipulation was that individuals listed not only the specific items they expected to purchase but also the price of each item. Participants might not have accurate knowledge of prices (e.g., Dickson & Sawyer, 1990; Evanschitzky, Kenning, & Vogel, 2004). Also, conceivably, each of the estimated prices might be rounded up (Forsyth & Burt, 2008; Huttenlocher, Hedges, & Bradburn, 1990), and several rounded-up prices might inflate the composite estimate. Again, such accounts assume that the specific, unpacked content is incorporated directly into the composite prediction. The fact that many of the listed items were not actually bought (and others were bought instead) suggests instead that even perfect price estimation would not have improved the calibration of predicted spending. In any case, future research should test the generality of the effects using different variants of unpacking.

The studies were not designed to directly test the process by which unpacking might increase spending prediction, but we did gain some insight by examining the lists of specific expenses. Participants did not list a great number of items (e.g., about 13 for the coming week in Study 3) and were not very good at identifying the specific expenses they would incur (i.e., only about two thirds of the purchases were foreseen). For improved discrimination, it may be crucial that the content of the lists is comprehensive and accurate. Indeed, a meta-analysis across studies showed that the accuracy of the unpacking lists (the proportion of foreseen expenses on these lists) was linked to lower prediction bias ( $r = -.22$ ). It may also be that effects of unpacking on prediction are not tied closely to the specific content in the generated lists, but instead the process of unpacking creates a heightened awareness, at a more general level, that there are many expense possibilities that were not originally considered. Further research will be needed to understand the precise mechanism by which

unpacking influences prediction bias. For example, thought protocols recorded during the prediction process might shed light on how generating detailed expense lists change the prediction process.

Also, there are almost certainly other moderators and boundary conditions to the unpacking effects that could be examined in future research. For example, research in other domains suggests that unpacking will not influence predictions if there are not enough subcomponents to unpack (Kruger & Evans, 2004), or if the target behavior is occurring in the distant future (Moher & Koehler, 2013).

### CONCLUSIONS

Spending predictions are often unrealistically optimistic. Finding ways to improve the accuracy of spending predictions could be very beneficial and reduce the very real costs associated with misprediction. The present research suggests that prompting predictors to list the specific expenses they expect to incur—an unpacking intervention—can sometimes help to curb an underestimation bias. However, this same intervention can hurt predictions in contexts where people are less inclined to underestimate spending in the first place. Here unpacking leads individuals to overestimate their spending. It appears that unpacking makes predictions bigger but not necessarily better, and thus its effectiveness as a debiasing tool will depend on the context in which it is applied.

### ACKNOWLEDGMENTS

We thank Taha Attiah, Oscar Smallenbroek, and Tess Walsh for assistance with data coding.

### FUNDING

This work was supported by the Social Sciences and Humanities Research Council of Canada.

### REFERENCES

- Buehler, R., Griffin, D., & Peetz, J. (2010). Chapter one—The planning fallacy: Cognitive, motivational, and social origins. *Advances in Experimental Social Psychology*, *43*, 1–62. doi:10.1016/S0065-2601(10)43001-4
- Dickson, P. R., & Sawyer, A. G. (1990). The price knowledge and search of supermarket shoppers. *The Journal of Marketing*, *54*, 42–53. doi:10.2307/1251815

- Dunning, D. (2007). Prediction: The inside view. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology* (pp. 69–90). New York, NY: Guilford Press.
- Epley, N., & Dunning, D. (2000). Feeling “holier than thou”: Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology*, *79*(6), 861. doi:10.1037//0022-3514.79.6.861
- Epley, N., & Dunning, D. (2006). The mixed blessings of self-knowledge in behavioral prediction: Enhanced discrimination but exacerbated bias. *Personality and Social Psychology Bulletin*, *32*(5), 641–655. doi:10.1177/0146167205284007
- Evanschitzky, H., Kenning, P., & Vogel, V. (2004). Consumer price knowledge in the German retail market. *Journal of Product & Brand Management*, *13*, 390–405. doi:10.1108/10610420410560299
- Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences*, *9*, 127–152. doi:10.1007/bf00143739
- Forsyth, D. K., & Burt, C. D. B. (2008). Allocating time to future tasks: The effect of task segmentation on planning fallacy bias. *Memory & Cognition*, *36*, 791–798. doi:10.3758/mc.36.4.791
- Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 196–213. doi:10.1037//0278-7393.16.2.196
- Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *TIMS Studies in Management Science*, *12*, 313–327.
- Kruger, J., & Evans, M. (2004). If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Experimental Social Psychology*, *40*, 586–598. doi:10.1016/j.jesp.2003.11.001
- Larrick, R. P. (2004). *Debiasing*. Malden, MA: Blackwell.
- Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science*, *4*, 390–398. doi:10.1111/j.1745-6924.2009.01144.x
- Moher, E., & Koehler, D. J. (2013). *Tempering optimistic bias in temporal predictions: The distance-dependent unpacking effect*. Unpublished manuscript.
- Peetz, J., & Buehler, R. (2007, May). *Examining the “Budget Fallacy”*. Poster presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL.
- Peetz, J., & Buehler, R. (2009). Is there a budget fallacy? The role of savings goals in the prediction of personal spending. *Personality and Social Psychology Bulletin*, *35*, 1579–1591. doi:10.1177/0146167209345160
- Peetz, J., & Buehler, R. (2012). When distance pays off: The role of construal level in spending predictions. *Journal of Experimental Social Psychology*, *48*, 395–398. doi:10.1016/j.jesp.2011.07.016
- Peetz, J., & Buehler, R. (2013). Different goals, different predictions: Accuracy and bias in financial planning for events and time periods. *Journal of Applied Social Psychology*, *43*, 1079–1088. doi:10.1111/jasp.12072
- Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review*, *104*, 406–415. doi:10.1037//0033-295x.104.2.406
- Savitsky, K., Van Boven, L., Epley, N., & Wight, W. M. (2005). The unpacking effect in allocations of responsibility for group tasks. *Journal of Experimental Social Psychology*, *41*, 447–457. doi:10.1016/j.jesp.2004.08.008
- Spiller, S. A., & Lynch, J. G. (2010). *Individuals exhibit the planning fallacy for time but not for money*. Unpublished manuscript. Retrieved from <http://ssrn.com/abstract=1458380>
- Sussman, A. B., & Alter, A. L. (2012). The exception is the rule: Underestimating and overspending on exceptional expenses. *Journal of Consumer Research*, *39*, 800–814. doi:10.1086/665833
- Trafimow, D. (2003). Hypothesis testing and theory evaluation at the boundaries: surprising insights from Bayes's theorem. *Psychological Review*, *110*, 526–535. doi:10.1037/0033-295x.110.3.526
- Trafimow, D. (2014). Editorial. *Basic and Applied Social Psychology*, *36*, 1–2. doi:10.1080/01973533.2014.865505
- Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*, 547–567. doi:10.1037//0033-295x.101.4.547
- Ülkümen, G., Thomas, M., & Morwitz, V. G. (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. doi:10.1086/587627
- Yang, S., Markoczy, L., & Qi, M. (2007). Unrealistic optimism in consumer credit card adoption. *Journal of Economic Psychology*, *28*, 170–185. doi:10.1016/j.joep.2006.05.006