2020

Using Our Past to Predict Our Future: Applying Reference Class Forecasting to Debias Individual Project Completion Predictions

Rebecca Friesdorf
Wilfrid Laurier University, frie3750@mylaurier.ca

USING OUR PAST TO PREDICT OUR FUTURE: APPLYING REFERENCE CLASS FORECASTING TO DEBIAS INDIVIDUAL PROJECT COMPLETION PREDICTIONS

by

Rebecca Friesdorf

Master of Arts, Wilfrid Laurier University, 2015

DISSERTATION

Submitted to the Department of Psychology

in partial fulfillment of the requirement for

Doctor of Philosophy in Psychology

Wilfrid Laurier University

Abstract

People often predict that they will finish projects sooner than they actually do, i.e., exhibit the planning fallacy (e.g., Buehler et al., 2010). This bias has important consequences for everyday life, including failure to meet deadlines, taking on too many projects, and increased stress. Several solutions have been proposed, including interventions which ask individuals to take an "outside view" (e.g., Kahneman & Lovallo, 1993), such as using information from past completion times to make predictions for a current project (e.g., Buehler et al., 1994). In this work, we take a novel approach to helping individuals use past project information: recalling past completion times in reference to predictions (i.e., "how much later did I finish compared to my original expectation?") as opposed to deadlines (i.e., how close to the deadline did I finish?") and reference class forecasting (RCF; Lovallo and Kahnemann, 2003).

In Study 1 and 2 (N = 322), we asked participants to report their planning fallacy beliefs, i.e., how many days after or before their predictions they believed they finished past projects. Although people on average reported finishing projects slightly later than predicted, awareness of this bias did not lead to less optimistic predictions for a current project. In Study 3-6 (N = 1,425), we instructed participants to recall relevant past project completion times using RCF, a technique that has been successful in reducing the planning fallacy for large-scale infrastructure projects (e.g., Flyvbjerg, 2008; Flyvbjerg et al., 2009), but has not been tested in individual, personal projects. Although our results were not completely consistent, we found evidence that RCF led to less optimistically biased completion predictions in three of the four studies.

Overall, our work suggests that RCF, especially with past completion times recalled in reference to predictions, is a promising strategy for helping people make more accurate completion predictions for their individual personal projects.

Acknowledgements

My deepest thanks go to my supervisor, Roger Buehler. Roger, thank you for being such a wonderful role model and collaborator. The time and effort that you invested in me, and in all your students, is truly admirable. I am extremely grateful for all your guidance throughout this process, the time you took to help me think through my sometimes crazy ideas, look at my data in different ways, and your general patience with my learning. Although I think we are both a little more pessimistic than average, I hope you know that I always walked away from our discussions feeling enriched, eager to learn more, and positive about taking action in the right direction. I must say I feel very lucky to have had a supervisor that wasn't phased by my seemingly relentless unrealistically optimistic task completion predicts, always running us up against tight deadlines, even though this bias is our very area of study. Overall, I feel greatly privileged to have been able to learn from you.

Thank you to my committee members, Christian Jordan and Anne Wilson. Your constructive feedback and insightful comments encouraged me to think more deeply about my research. I also want to thank Ian Newby-Clark and Greg Irving, who graciously volunteered to spend the time to serve on my committee and will challenge and push my thinking during the defence.

I am grateful for the many people that started out as colleagues in classes and research assistants in our lab and turned into close friends. I cherish all of the memories that we have made throughout graduate school and can't wait to see the impact that you will have on the world. On a personal note, my sincerest thanks go my family, whose unwavering support throughout this process has been humbling. To my partner Adam, your love and understanding at every step has been astonishing. I hope you know that you are my rock, and I hope you know that your support made this journey many many times better than it would have otherwise been.

Table of Contents

Abstract ii
Acknowledgements iii
Table of Contents iv
List of Figures vi
List of Tables vii
Introduction 1
The Planning Fallacy 1
Strategies to De-bias Predictions 3
Using Past Completion Times to Improve Predictions 3
Reference Class Forecasting 4
Application to Individual Projects 5
Past Project Completion Times 8
The Current Work 9
Study 1 11
Method 12
Results 15
Discussion 20
Study 2 21
Method 21
Results 25
Discussion 29
Study 3 31
Method 32
Results 38
Discussion 46
Study 4 47
Method 47
Results 48
Discussion 56
Study 5 57

Method 58
Results 62
Discussion 69
Study 6 70
Method 71
Results 72
Discussion 77
General Discussion 78
Limitations and Future Directions 83
Conclusions 87
References 88
Appendices 94
Figures 161
Tables 169

List of Figures

Figure 1.	Bar graph showing differences between completion predictions and completion times by condition for Study 3 (Part 2)……..	161
Figure 2.	Bar graph showing differences between completion predictions and completion times by condition for Study 4 (Part 2)……...	162
Figure 3.	Mediation models testing whether the effect of the RCF 1 procedure (vs. Control 1) leads to later completion predictions through past completion times or vice versa (Study 5)…….....	163
Figure 4.	Mediation models testing whether the effect of the RCF groups combined (vs. control groups) leads to later completion predictions through past completion times or vice versa (Study 5)……..	164
Figure 5.	Bar graph showing differences between completion predictions and completion times by condition for Study 5 (Part 2)……..	165
Figure 6.	Moderation results for RCF 1 vs. Control 1 and Control 2 – past completion times and predictions are more related in the RCF 1 than the control groups (Study 6)…….	166
Figure 7.	Moderation results for RCF 2 vs. Control 1 and Control 2 – past completion times and predictions are more related in the RCF 2 than the control groups (Study 6)…….	167
Figure 8.	Bar graph showing differences between completion predictions and completion times by condition for Study 6 (Part 2)……...	168

List of Tables

Table 1.	Planning Fallacy Beliefs Ratings by Condition (Study 1)	169
Table 2.	Mean Completion Predictions (days after study date) by Condition (Study 1)	170
Table 3.	Zero-Order Correlations (r) Between Planning Fallacy Beliefs and Predictions (Study 1)	170
Table 4.	Multiple Regression Analyses Testing for Beliefs by Beliefs Order Interactions on Predictions (Study 1)	172
Table 5.	Planning Fallacy Beliefs Ratings by Condition (Study 2)	173
Table 6.	Mean Completion Predictions (days before deadline) by Condition (Study 2)	174
Table 7.	Zero-Order Correlations (r) Between Planning Fallacy Beliefs and Predictions (Study 2)	175
Table 8.	Mediation Models Testing the Effect of the Article Manipulation on Completion Predictions through Planning Fallacy Beliefs (Study 2)	176
Table 9.	Moderated Mediation Models Testing the Effect of the Article Manipulation on Completion Predictions through Planning Fallacy Beliefs as Moderated by Target Type (Study 2)	177
Table 10.	Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample, Study 3)	178
Table 11.	Zero-Order Correlations Between Completion Predictions and Additional Measures (Part 1 sample, Study 3)	181
Table 12.	Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample A; Study 3)	182
Table 13.	Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample B; Study 3)	183
Table 14.	Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times by Condition (Part 2 sample A; Study 3)	184
Table 15.	Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample A; Study 3)	185
Table 16.	Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample; Study 4)	186
Table 17.	Zero-Order Correlations Between Completion Predictions and Additional Measures (Part 1 sample; Study 4)	189
Table 18.	Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample; Study 4)	190
Table 19.	Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times (Part 2 sample, Study 4)	191
Table 20.	Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample, Study 4)	192

Table 21.	Components of the Prediction Intervention Conditions in Study 5 and 6	193
Table 22.	Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample; Study 5)	194
Table 23.	Effects of the Reference Class Forecasting Manipulations on Past Project Completion Times (Part 1 sample; Study 5)	195
Table 24.	Multiple Regression Analyses Testing for Moderation Effects (Study 5)	196
Table 25.	Zero-Order Correlations Between Completion Predictions and Additional Measures (Part 1 sample; Study 5)	197
Table 26.	Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample; Study 5)	198
Table 27.	Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times (Part 2 sample, Study 5)	199
Table 28.	Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample, Study 5)	200
Table 29.	Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample; Study 6)	201
Table 30.	Effects of the Reference Class Forecasting Manipulations on Past Project Completion Times (Part 1 sample; Study 6)	202
Table 31.	Effects of the Reference Class Forecasting Manipulations on Predicted Project Completion Times (Part 1 sample; Study 6)	203
Table 32.	Multiple Regression Analyses Testing for Moderation Effects (Study 6)	204
Table 33.	Mediation Models Testing the Effect of the RCF procedure on Completion Predictions through Past Project Completion Beliefs (Study 6)	206
Table 34.	Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample; Study 6)	208
Table 35.	Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times (Part 2 sample, Study 6)	210
Table 36.	Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample, Study 6)	209
Table 37.	Overview of Main Study Results by Hypothesis Tested	211
Table 38.	Overview of the Practical Implications of the RCF Interventions for Predictions and Bias	211

Introduction

The ability to predict when tasks will be finished is important for individuals and organizations. People, firms, and governments make consequential decisions and binding commitments on the basis of time estimates. Underestimating how much time will be required to complete tasks or projects can be costly. Take the construction or software industry as examples; they often underestimate completion times by months or years, resulting in fines or even lawsuits leading to massive revenue loss. At the individual level, think of those times when you, your spouse, or your boss were too optimistic about how quickly an academic, personal, or work project could be completed. Such prediction errors often result in negative consequences such as overtime work, all-nighters, and failure to meet deadlines, with the potential for domino effects, where being late on one project leads to a cascade of falling behind on everything. These errors can also be accompanied by negative consequences for affect and well-being – lack of sleep, feelings of stress and fatigue, and negative self-evaluation (e.g., "Why can't I seem to keep up with the demands of life?"). In recognition of the importance of accurate completion time predictions, this dissertation explores one method for increasing their accuracy in the personal domain: nudging individuals to use completion information from similar past projects to guide their predictions for an upcoming project.

The Planning Fallacy

The tendency to underestimate the time it will take to complete tasks is well documented in the literature and known as the planning fallacy (Kahneman & Tversky, 1979). It is a form of optimistic bias in which people underestimate the time it will take to complete a task even though they are aware that similar tasks have typically taken longer than expected (for reviews see Buehler et al., 2010; Buehler & Griffin, 2015). The bias has been documented for a wide range of personal, academic, and work-related tasks (e.g., Buehler, Griffin, and Ross, 1994;

Griffin & Buehler, 1999; Kruger & Evans, 2004; Min & Arkes, 2012; Roy et al., 2005), as well as major infrastructure (e.g., Flyvbjerg, Holm, & Buhl, 2005) and software development projects (Halkjelsvik & Jørgensen, 2012; Jørgensen, 2004).

Kahneman and Tversky (1979) were the first to provide psychological explanations for the bias, theorizing that we can understand it by distinguishing between the "inside" and "outside" view of an event (Kahneman & Lovallo, 1993). When people adopt an inside view of an event they focus on the specific case at hand and its specific components. For example, when planning for a task, they focus on the specific characteristics of the task at hand. Unfortunately, our mental scenarios of events tend to be idealized, schematic, and oversimplified (Dunning, 2007; Liberman et al., 2007), thus leaving our plans prone to bias. In contrast, using an outside view means viewing a target event as one instance of a set of relevant comparison events, and making predictions about the event based on the distribution of similar events. Taking an outside view provides some protection against people's tendency to generate overly optimistic plans and idealized scenarios, and often yields more realistic predictions (Buehler et al., 1994; Kahneman & Lovallo, 1993). This is because the outside view does not require the forecaster to predict specific uncertain events that may affect the target project, but instead compares it to a distribution of known past outcomes. The outside approach requires explicit identification of a distribution of prior outcomes and regression of the target project toward that prior.

However, because the natural and intuitive way to think about a complex project is to focus on its specific and unique characteristics, as opposed to distributional information about similar past projects, individuals and organizations typically neglect the outside view in favor of the inside view (Buehler et al., 2010; Flyvbjerg et al., 2005; Kahneman & Lovallo, 1993).

Indeed, Kahnemann and Tversky (1979) concluded that “the prevalent tendency to underweigh or ignore distributional information is perhaps the major error of intuitive prediction".

Strategies to De-bias Predictions

Research has attempted to reduce the “error” inherent in the inside approach to prediction by helping forecasters to change their plan content and mental simulation. Strategies designed to alter the content of the inside view include decomposing the plan into smaller steps (Peetz et al., 2015), considering alternative scenarios for how a task may unfold (Newby-Clark et al., 2000), visualizing the plan using third-person imagery (Buehler et al., 2012), and generating the plan in reverse-chronological order (Wiese, Buehler, & Griffin, 2016). These strategies, which encourage greater awareness of potential interruptions, unanticipated events, declining motivation, and competing priorities, have had some success in attenuating the planning fallacy, but have rarely completely removed the bias. Finding ways to help individuals adopt an outside view to prediction remains a promising but understudied strategy to reducing the planning fallacy.

Using Past Completion Times to Improve Predictions

One of the few examples of past research instructing people to incorporate past project completion information into predictions for a current project was conducted by Buehler and colleagues (1994; Study 4). Undergraduate students were instructed to make completion predictions for school assignments, either without having reported their past experiences (control group), after having reported their past experiences with similar projects, i.e., how far before the deadlines they had typically finished (recall group), or after having reported their past experiences with similar projects and answering two more questions that required them to forge a connection between the past experiences and the target assignment (relevant recall group). These two questions prompted participants to: 1) indicate the date and time they would finish the

computer assignment if they completed it as far before its deadline as they typically completed assignments, and 2) describe a plausible scenario—based on their past experiences—that would result in their completing the computer assignment at their typical time. Participants' predictions were less optimistically biased in the relevant recall group than in the other two groups. Predictions in the recall group did not differ from those in the control group, even though the prediction question in the recall group asked participants to keep in mind their past experiences when making their prediction (see also Buehler and Griffin, 2003, for a similar null effect). Surprisingly, predictions were not more highly correlated with actual completion times in the relevant recall group than in the other two groups and predictions were not correlated with past completion times, indicating that participants did not incorporate their own past completion times directly into their predictions. The participants' reports of average past completion times were weakly related to actual completion times. The results suggest that people's past experience can be leveraged to reduce the degree of optimistic bias in prediction, but that people do not easily incorporate information about past completion times into their predictions.

Reference Class Forecasting

Later work directly applied the use of past distributions as a basis for prediction (as described by Kahneman and Tversky, 1979), using an approach called Reference Class Forecasting (RCF; Lovallo and Kahnemann, 2003). This approach was directly tested in the realm of major infrastructure projects (Flyvbjerg, 2008; Flyvbjerg et al., 2009), as opposed to in smaller-scale individual projects (Buehler et al., 1994). RCF circumvents the problems associated with an inside view, as well as people's reluctance to base predictions on past experience, by explicitly requiring forecasters to base their predictions on the outcomes of a distribution of comparable projects. Empirical tests of the approach support its effectiveness in reducing time and cost overruns in large-scale construction projects.

The RCF approach aims to overcome people's reluctance to base predictions on past experiences by using a structured method to impose an "outside view" on planning and prediction (Flyvbjerg et al., 2009; Lovallo & Kahneman, 2003). It requires three steps (Flyvberg, 2006):

1) Identify a relevant reference class of past, similar projects that are broad enough to be statistically meaningful but narrow enough to be truly comparable with the focal project.
2) Establish a probability distribution for the selected reference class of the parameter that is being forecasted using credible, empirical data (for a sufficient number of projects within the reference class to be able to make statistically meaningful conclusions).
3) Compare the specific case with the reference class distribution to estimate the most likely outcome for the target case. In other words, the forecaster places the current project in a statistical distribution of outcomes for a class of reference projects (e.g., group of similar past projects).

In the literature, RCF was first put into practice by Flyvbjerg and colleagues to estimate the cost of proposed transportation projects (Flvbjerg & COWI, 2004), and effective implementation followed for a variety of large-scale infrastructure projects such as hydroelectric dams, sports arenas, and oil/gas extraction projects (Flyvbjerg, 2008; Flyvbjerg et al., 2009). In 2005, the American Planning Association endorsed and recommended the adoption of RCF in addition to traditional forecasting techniques. Since then the technique has been implemented and tested for budgets of many large-scale construction projects and has gained wide acceptance in those contexts.

Application to Individual Projects

In this dissertation we test whether RCF can be applied effectively to a different context, individual predictions of task completion times. To the best of our knowledge, the approach has

not yet been tested in the domain of individual self-prediction, and instead tests have focused on budgets for large government projects. To ensure practical relevance, we targeted the kinds of projects that people carry out in everyday life – academic, work, and personal individual projects. These are also the kinds of tasks that have been targeted in much of the existing research on the planning fallacy. Increasing the accuracy of individuals' self-predictions has important time management-related benefits. For the individual, making less optimistically biased completion predictions can mean being more likely to finish work on time, being more likely to meet deadlines, being less likely to take on too many projects, being less likely to be stressed about delivering on project completion promises, as well as positive self-perceptions of one's own competencies and feelings of control over project outcomes. In the workplace, these benefits would also extend to organizations, leading to improved project management.

Although individual personal projects are smaller in scope and lower in complexity than large-scale construction projects, RCF may actually be more difficult to implement in the individual project context because people often do not keep a record of historical outcomes for their past projects. Individuals would be required to consult their knowledge/memories of relevant past experiences, which could be inaccurate and biased (e.g., Roy et al., 2005). Roy and colleagues (2005; 2008) present evidence that people tend to overestimate how long short tasks (a few minutes or less) will take and have taken in the past, and underestimate how long long tasks (longer than a few minutes) will take and have taken in the past. They theorize that memory bias is a major cause for underestimation of task completion times, but acknowledge that empirical evidence for this is sparse, especially for longer tasks. If memory bias is a major cause of underestimation, then having people recall past project completion times, and base their predictions for a future project on those past project completion times, is unlikely to yield much

more accurate predictions. On the other hand, other planning fallacy research indicates that people often have quite accurate memories of past project completion times and do, when prompted, acknowledge that previous projects were not completed as early as predicted (Buehler et al., 2010). Given that only few studies have examined memory bias in completion times for longer tasks, we have only preliminary explanations for the differing findings. One potential explanation is that the longer tasks which led Roy and colleagues to conclude that we tend to underestimate the past were shorter in duration (e.g., 5 -15 minutes, Roy & Christensen, 2008), and required less total working time (closed tasks), than the tasks examined by Buehler and colleagues (e.g., multiple-hour academic assignments, Buehler et al., 1994) which require multiple sessions over days or weeks for completion (open tasks). These differences in the task duration may mean that some participants were strictly recalling task duration (how long, in one sitting, the task took) vs. recalling time to completion (how many days until the task was finished, which includes stretches of time spent on other activities, i.e., obstacles, other demands on one's time). Because the focus of this work is task completion predictions for tasks/projects which stretch over multiple days, with working times in the magnitude of hours as opposed to minutes, and deadlines that are days to weeks away, we lean toward the expectation that memory bias is less likely to be an issue in our studies (in line with Buehler and colleagues).

Another potential challenge to the successful application of RCF in this context is that people may be resistant to the approach because they view their current project as particularly "unique", and unlike their past projects (Buehler, 2010). Because the RCF procedure requires people to assess and use the best comparison distribution available to make their prediction, the likelihood of participants discounting their past experience should be reduced. In other words, the RCF method "bypasses human bias…by cutting directly to outcomes" (Flyvberg, 2006).

Past Project Completion Times. In the context of business projects, the anticipated completion date (i.e., prediction) and project deadline will often be the same. For individual personal projects with deadlines, the predicted completion time and project deadline often differ, as demonstrated by many planning fallacy studies where the vast majority of individuals predict finishing their projects before the project deadline (e.g., Buehler & Griffin, 2003; Buehler, Griffin, & MacDonald, 1997; Buehler et al., 1994; Buehler, Messervey, & Griffin, 2005; Buehler, Peetz, & Griffin, 2010). When applying RCF in this context, it raises the question of whether individuals should be asked to recall their past project completion times in reference to the project deadline (i.e., how many days before/after the deadline they typically finished; as in Study 4 of Buehler et al., 1994) or in reference to their original prediction (i.e., how many days before/after their prediction they typically finished).

The possibility exists that thinking about past completion times in reference to predictions may be more powerful for correcting for optimistic bias because it is directly diagnostic about the accuracy of one's predictions. For example, recalling that you tend to finish projects two days after you predicted may be more likely to influence your prediction for a current project than recalling the number of days before the deadline (which wasn't set by you) that you finished. In this way, reference class forecasting may function in a similar way as metacognitive training interventions (MCT; Eichner & Berna, 2016) – helping people target fixed false beliefs (i.e., overly optimistic completion predictions) that are held with confidence, by providing unexpected corrective information (i.e., concretely reflecting on the fact that one regularly finished past projects later than predicted) that leads to a reduction in confidence and greater information seeking (Köther et al., 2017; Moritz et al., 2014).

One potential drawback of this approach is that people may be even less able to accurately recall when they originally predicted to be finished a project (compared to the deadline), especially given that they may have adjusted their prediction several times before project completion. Deadlines likely change less frequently than predictions, and memories of how close we finished to the deadline (or whether we failed to finish by the deadline) are likely stronger than our memories of when we finished relative to our (perhaps idealistic) predictions.

The Current Work

Our primary objective in this line of work is to test the effectiveness of an RCF procedure for reducing optimistic bias in personal task completion predictions. However, the work unfolds in two phases. In the first phase, before actually testing an RCF intervention, we conducted two initial studies to address questions raised earlier concerning people's knowledge and beliefs about previous task completion times. These initial studies tested whether people do hold the belief that they tend to finish tasks later than predicted. In other words, do people believe they exhibit the planning fallacy? To our knowledge, this is a novel question that has not been explored in previous research. Previous research explored how long people recalled specific tasks took to complete in the past (e.g., number of minutes, days before the deadline), whereas our studies test how long before/after their predictions (expectations for when they would finish) participants recall finishing. The initial studies also tested whether asking people to report on their knowledge and beliefs about previous completion times has an effect on their predictions for a specific upcoming task. If so, then this relatively straightforward prediction strategy may be a promising debiasing strategy and preclude the need for a more structured RCF approach. If not, the results may indicate that a more structured, formal tracking of past predictions and completion times may be required before reference class forecasting can be applied to personal completion predictions.

Study 1 participants made completion predictions for an upcoming project either before or after telling us about when they typically finish projects relative to when they expected to be finished. If participants believe that they typically finish later than predicted and are reminded of this before making their prediction, they may spontaneously use this information to make a later completion prediction. In Study 2 we manipulated people's knowledge and beliefs about the planning fallacy with the intent of strengthening the belief that they tend to finish projects later than predicted and observed whether this led to later completion predictions. These studies tested whether:

H_1: People recall that they tend to finish past projects later than predicted.

H_2: Reminding people of their past project completion times relative to prediction leads them to predict later completion times for a specific upcoming task.

H_3: The effect of past reminders will be more pronounced to the extent that people believe they tend to finish projects later than predicted; that is, the effect of reminders will be moderated by beliefs about the planning fallacy.

If our hypotheses are supported, these studies would suggest that people may not always need to be explicitly instructed to use their past completion times – that simply pointing out that they finish later than predicted is enough to lead to prediction adjustment. Based on the results of Buehler and colleagues' previous studies (1994, 2003), we might think that this is unlikely given that students in the "recall" condition (who recalled past completion times) did not make later completion predictions. However, keep in mind that these students recalled their completion times relative to deadlines, as opposed to relative to predictions.

In the second phase of our research, we introduced more structured RCF interventions and tested their effects on task completion predictions. We included several variations of the

RCF approach. In particular, we compared whether participants were prompted to recall when they finished past projects in reference to the project deadline or in reference to the predicted completion time. We also followed participants up to measure when they actually completed their project, so we could examine the accuracy of their predictions. We focused on comparing mean levels of optimistic bias (i.e., how many days after their predictions did individuals report finishing their projects) to determine which interventions led to the greatest reduction in optimistic bias, in line with the dominant way of reporting planning fallacy result in the literature (e.g., Buehler et al., 1994; Buehler & Griffin, 2003; Buehler, Griffin, & MacDonald, 1997), but also present correlational results as secondary analyses (i.e., strength of the relationship between people's predictions and actual completion times). For these studies, we had the following hypotheses:

H_4: Reference class forecasting will lead to less optimistically biased predictions than making a prediction without using this strategy (i.e., control condition).

H_5: Reference class forecasting will be more effective in reducing optimistic bias if the completion of past projects is recalled relative to predictions, as opposed to deadlines.

Finally, in Study 5, we included several additional manipulations and measures to test the process by which the RCF procedure affects task completion predictions. In particular, we asked all predictors to report memories of past completion times, and expected that:

H_6: The RCF procedure would be most effective to the extent that people recall completing past projects later than predicted.

Study 1

Study 1 tested the first two hypotheses: that people remember finishing past projects later than predicted, and that a reminder of such past project completion times (via self-report measures) leads to later completion predictions for a specific future task. In other words, if

people believe that they tend to finish later than predicted, then when we remind them of this, they will make later predictions for an upcoming project.

We also varied the type of project we asked participants to nominate – we either let participants chose their own academic, work, or personal project or instructed them to nominate a project that they *wanted to* complete, that they anticipated being fun or enjoyable to work on. Without specific instructions (based on our experiences with past studies), students tend to choose primarily academic projects that they *must* complete, and we wanted to examine whether planning fallacy beliefs and the relationship between such beliefs and predictions would differ depending on project type. In addition, we assessed beliefs both about the self and others to see if participants thought that others committed the planning fallacy more than they themselves do, and to see if self beliefs better predict completion predictions than do beliefs about others. The testing of self vs. other and want-to-do vs. have-to-do task effects is not central to our current hypotheses and thus represent subsidiary analyses that are not discussed extensively. Because we did not have specific hypotheses, comparisons between must-do and want-to-do projects are exploratory. Based on previous work showing that people tend to think that others are more biased than themselves (Buehler, Griffin, & Ross, 1994; Pronin, Lin, & Ross, 2002), we predicted that participants would give higher planning fallacy belief ratings for others than for themselves.

Method

Participants. Of the 214 undergraduate students from Wilfrid Laurier University that participated in the study (on-line for course credit) 19 failed a simple attention check (“If you are paying attention please select "Very accurate 6" for this statement.”) and an additional 4 failed to nominate a project for prediction and were thus excluded from data analysis. Of these 191 participants, an additional 5 participants failed to make a prediction for their project, making the

final sample size 186. Note that due to their nature, some want-to-do projects did not have deadlines (n = 23), so analyses which control for the deadline have a reduced sample size of 163. The final sample consisted of 154 female, 31 male, 1 unspecified participants with a mean age of 19.44 (SD = 2.33). On our ethnicity question, 64.5% identified as Caucasian, 14.0% as Asian, 4.8% as Black, 4.3% as Latino, 2.7% as Middle Eastern, and 9.7% as other.

Design. The study had a 2 Beliefs Order (beliefs first vs. predictions first) between-subjects by 2 Project Type (want-to-do vs. unspecified/must do project) between-subjects by 2 Target (self beliefs vs. other beliefs) within-subjects factorial design[1]. More specifically, participants either told us their beliefs about when they finished past projects relative to their predictions and then made a prediction for a real upcoming future project, or made a prediction for an upcoming project first and then told us their beliefs about when they finished past projects. For both the beliefs questions and prediction, participants were randomly assigned to either think about a) projects that they would consider fun or enjoyable to work on (want-to-dos) or b) we did not specify what kind of project we wanted participants to think about, and these turned out to be primarily academic projects such as assignments, essays, lab reports, group projects (i.e., must-dos). Participants rated their planning fallacy beliefs both for others and for themselves but made predictions only for themselves.

Procedure.

Beliefs. Participants' planning fallacy beliefs were elicited using the following instructions

"…think back to times when you remember having an expectation for when you would be finished a particular task or project…recall times when you had made a prediction for when (e.g., time/date) you would be done some task or project. Relative to when you expected to be finished, when do you think you actually finished these tasks or projects?" (see Appendix A1)

[1] The order of the target of the belief ratings (self vs. other) was also varied but did not influence the results and is thus not discussed further.

and rated on a scale from 1 (*finished much later than predicted*) to 7 (*finished much earlier than predicted*) with a midpoint of 4 (*finished exactly when predicted*). Participants in the want-to-do project condition received the same instructions but were instructed to think about projects they would consider fun/enjoyable to work on. For beliefs about others, participants were asked to think back to times when they remember other people having an expectation for when they would be finished projects. Participants also indicated their beliefs via a second item,

"On average, what percentage of the time would you say you underestimate, overestimate, or are accurate about how long it will take you to complete tasks or projects?"

for which they recorded three percentages that had to add to 100%. See Appendix A1 for the verbatim instructions in each condition. Based on their answer to the first beliefs question, participants were also asked an open-ended question asking them to explain why they think that they tend to finish projects later than predicted, earlier than predicted, or at the same time as predicted. Next, participants were asked to think back to times when they finished projects later than expected and presented with 10 reasons for why this could have been the case (e.g., failed to consider project obstacles, forgot to factor in procrastination). They then indicated how accurately each reason describes why they finished later than expected (1 - *not at all accurate for me*, to 7 - *extremely accurate for me*) (see Appendix A2). These items are outside the scope of the present research, have not been analyzed, and will thus not be discussed further.

Predictions. We instructed participants to think of and then briefly describe a major school, work, or personal project that they would be completing in the next three weeks that had a specific deadline. They were then asked to specify the deadline for the project (in days from today) and to make a prediction for when they would be finished (also in days from today) (see Appendix A3). The want-to-do condition included the following additional (bolded) instructions:

"It should be the kind of task or project that you look forward to completing because it will be enjoyable/pleasant to work on." In addition to our primary prediction variable, we also include a second, hypothetical, standardized completion prediction measure. All participants were asked to imagine that they were working as a research assistant and had received an approximately 8-hour long task to complete, with a deadline in 14 days. They were then asked to predict in how many days they would be finished the job (see Appendix A4). In the want-to-do project condition the project description contained the following additional sentence, "You know that you will find the work very interesting, so you look forward to and anticipate enjoying completing the job". Note that 4 participants left this question blank, and an additional 4 participants made a prediction more than three standard deviations from the mean (i.e., predicted finishing more than one week after the deadline) and were thus excluded, making the available sample size for analyses with this variable 178.

Previous knowledge. Finally, to get a sense of how many of our participants had prior knowledge of the planning fallacy, we asked "Prior to this survey, had you heard or learned about the planning fallacy and/or related biases (our tendency to predict that we will be done projects sooner than we actually are) from any other sources? (1 - *yes*, 0 - *no*), followed by an open-ended question asking where they had heard about the planning fallacy and/or related bias. 142 (76.3%) answered *no*, and 44 (23.7%) answered *yes*, indicating that most participants were unaware of the planning fallacy phenomenon, and these numbers did not differ by condition (all *p*s > .150). Predictions also didn't differ by previous knowledge of the planning fallacy, $t(184) = 0.76$, $p = .385$, $d = 0.15$.

Results

Beliefs. The first hypothesis, that people believe they tend to finish tasks later than predicted, was generally supported. Overall, collapsing across conditions, mean planning fallacy

belief ratings for the self were 3.59 (*SD* = 1.39), indicating that participants believed they tend to finish projects between slightly later than predicted (3 on the scale) and exactly when predicted (4 on the scale). This mean was significantly different from the midpoint, $t(185) = -4.06$, $p < .001$, $d = -0.30$, and the median and mode were both 3 (i.e., slightly later than predicted).

For the percent beliefs question, participants indicated they finished projects later than predicted 36% of time (*SD* = 22.93), the same time as predicted 39% of the time (*SD* = 20.22), and earlier than predicted 25% of the time (*SD* = 17.55). At least one of these percentages differed significantly from the others, $F(2, 370) = 17.00$, $p < .001$, $\eta_p^2 = .084$. Specifically, participants thought they were late significantly more often than they were early, $t(185) = 4.30$, $p < .001$, $d = 0.55$, on time more often than they were early, $t(185) = 6.49$, $p < .001$, $d = 0.75$, and late no more often than they were on time, $t(185) = 1.09$, $p = .278$, $d = 0.15$. For the analyses that follow, we also calculated a "percent bias" variable (% late - % early), capturing the extent to which participants believe their prediction errors are directional (tendency to finish later than predicted more often than earlier = a positive number, tendency to finish earlier than predicted more often than later = a negative number). Overall, mean percent bias for the self was 11.19 (*SD* = 35.48), indicating that participants believed that they tend to finish later than predicted more often than earlier than predicted.

To test whether planning fallacy beliefs differed across the conditions of the study, we conducted 2 (Beliefs Order) x 2(Project Type) x 2 (Target) ANOVAs, where Beliefs Order and Project Type are between factors and Target is a within factor (see Table 1 for means and significance tests for all beliefs variables). The ANOVA performed on the planning fallacy beliefs ratings revealed a small, but significant main effect of Target (self vs. other), $F(1, 182) = 5.77$, $p = .017$, $\eta_p^2 = .031$, such that participants rated themselves as finishing later than predicted

(M = 3.59, SD = 1.39) less often than others (M = 3.31, SD = 1.21). A significant Target by Project Type interaction also emerged, $F(1, 182) = 7.40$, $p = .007$, $\eta_p^2 = .039$, indicating that the self-other difference was significant only in the must-do project condition, $M_{Self} = 3.62$, SD = 1.40 vs. $M_{Other} = 3.03$, $SD = 1.22$, $t(94) = 3.68$, $p < .001$, $d = 0.45$, and not in the want-to-do project condition, $M_{Self} = 3.55$, $SD = 1.38$ vs. $M_{Other} = 3.58$, $SD = 1.14$, $t(90) = -0.21$, $p = .837$, $d = -0.02$. In other words, participants rated others as finishing projects later than predicted more so than themselves only for must-do projects and not for want-to-do projects. There was no main effect of Beliefs Order, $F(1, 182) = 0.07$, $p = .799$, $\eta_p^2 < .001$, indicating that beliefs did not differ depending on whether they were assessed before or after predictions. There was also no main effect of Project Type, $F(1, 182) = 2.42$, $p = .122$, $\eta_p^2 = .013$, and none of the remaining three way interactions was significant, all Fs < 1, ps > .300, η_p^2s < .006.

A similar pattern of results emerged for the percent belief question. The ANOVA performed on the percent bias variable revealed that participants rated others as finishing later (M = 17.54, SD = 32.24) more often than the self (M = 11.19, SD = 35.48), $F(1, 182) = 4.74$, p = .031, $\eta_p^2 = .025$. The relevant Target by Project type interaction was marginally significant, $F(1, 182) = 3.82$, $p = .052$, $\eta_p^2 = .021$, and subsequent t-tests revealed that self-other differences were only significant in the must-do, $M_{Self} = 8.91$, $SD = 37.53$ vs. $M_{Other} = 20.63$, $SD = 35.17$, $t(94) = -2.76$, $p = .007$, $d = -.34$, and not in the want-to-do condition, $M_{Self} = 13.57$, $SD = 33.37$ vs. $M_{Other} = 14.32$, $SD = 32.78$, $t(90) = -0.21$, $p = .838$, $d = -0.02$.

Predictions. Before testing our hypotheses about participants' predictions, we first ran a 2-way ANOVA to test whether project deadlines differed by condition. A significant effect of prediction type emerged, $F(1, 159) = 9.77$, $p = .002$, $\eta_p^2 = .058$, such that deadlines for must-do tasks (M = 14.06, SD = 7.34) were fewer days away than deadlines for want-to-do tasks (M =

18.10, *SD* = 9.07). This is not surprising as must-dos were primarily academic projects, and students tend to have a deadline for such a project coming up in the near future at most time points in the academic term, whereas want-to-dos were more likely to be personal projects which students may have fewer of and may be lower priority and thus deadlines are more likely to be further away in the future. Deadlines did not differ based on Beliefs Order, $F(1, 159) = 0.33$, $p = .567$, $\eta_p^2 = .002$, and the Beliefs Order by Project type interaction was not significant, $F(1, 159) = 2.35$, $p = .128$, $\eta_p^2 = .015$. Because of the deadline differences (and because deadlines may influence predictions), analyses examining effects on predictions control for the deadline, and thus only projects with deadlines were analyzed ($n_{must\text{-}do} = 95$, $n_{want\text{-}to\text{-}do} = 68$).

Recall that we hypothesized that reminding people of their past completion times would lead them to predict later completion times (hypothesis 2). To test the hypothesis, we submitted the predicted completion times (days after study date) to a 2 (Belief Order) x 2 (Project Type) ANCOVA with deadline days as the covariate (see Table 2 for unadjusted means[2]). There was no main effect of Beliefs Order, $F(1, 159) = 0.50$, $p = .482$, $\eta_p^2 = .003$, nor a Beliefs Order by Project type interaction, $F(1, 159) = 0.09$, $p = .767$, $\eta_p^2 = .001$, indicating that predictions didn't change depending on whether participants were reminded of their beliefs (about past completion times) before making a prediction ($M_{adj} = 12.13$, $SE = 0.43$) or not until after making their prediction ($M_{adj} = 11.70$, $SE = 0.42$). There was also no effect of Project type once deadlines were held constant, $F(1, 159) = 1.05$, $p = .308$, $\eta_p^2 = .007$. For the standardized, hypothetical prediction measure, for which participants predicted how long it would take them to complete the 8-hour job due in 14 days ($M_{Overall} = 7.42$, $SD = 4.06$), there were also no effects of Beliefs

[2] Unadjusted means are presented in the tables, adjusted means presented in text.

Order, $F(1, 174) = 0.02$, $p = .882$, $\eta_p^2 < .001$, Project type, $F(1, 174) = 0.06$, $p = .809$, $\eta_p^2 < .001$, nor a significant interaction, $F(1, 174) = 0.24$, $p = .623$, $\eta_p^2 = .001$.

Recall that we also hypothesized that the effect of past reminders would be greater to the extent that people believe they tend to finish tasks later than predicted. That is, if participants believe that they exhibit the planning fallacy, they will make later completion predictions if they are reminded of their beliefs first, as a kind of adjustment to account for their tendency to predict that they will finish sooner than they actually do. Thus, we would expect a Beliefs by Beliefs Order interaction, such that participants with greater planning fallacy beliefs in the beliefs first condition would be expected to make the latest completion predictions.

Before testing this moderation hypothesis, we examined zero-order correlations between beliefs and predictions, separately within the Self and Other conditions, to get an overall sense of the relationship between these variables (see Table 3). Self beliefs were unrelated to self-nominated and hypothetical project predictions. Other beliefs were weakly related to self-nominated project predictions, such that greater belief in planning fallacy for others was related to making later predictions (consistent with the idea that being aware that one tends to finish later than predicted results in adjusting one's prediction to be closer to the deadline). Other beliefs were unrelated to hypothetical project predictions. These results suggest that 1) there is not much of a connection between beliefs about when one finished past projects (relative to predictions) and predictions for upcoming projects, and 2) curiously, if anything, planning fallacy beliefs about others are more related to predictions (for one's own project) than planning fallacy beliefs for the self.

Next to test for the hypothesized moderation effect, we conducted multiple regression analyses with Beliefs, Beliefs Order, Project type, as well as their interactions, predicting

completion predictions. For the self-nominated tasks, the deadline date was entered first as a control variable. Based on our hypothesis, we should observe significant Beliefs by Beliefs Order interactions. This was not the case – none of the relevant interaction terms (or main effects of Beliefs or Beliefs Order) were significant, neither for self beliefs nor for other beliefs (see Table 4 – all confidence intervals contained zero).

Discussion

In Study 1 we found some support for Hypothesis 1 – students recalled finishing past projects slightly later than predicted and recalled finishing projects late more often than early. We did not find support for Hypothesis 2 or 3 – reminding students of their past project completion times did not lead them to make later predictions for an upcoming project. In addition, students' beliefs about when they finished past projects were not predictive of their project completion predictions for an upcoming future project and did not moderate the effects of the order manipulation. This null effect is surprising. If participants were motivated to use information about their past to make more accurate (realistic) future predictions, we would have expected those who were reminded of their past inaccuracy to adjust their predictions. Although there are many potential explanations for why we did not observe an effect of our order manipulation or much of a relationship between beliefs and predictions (see more in Study 2 discussion), one potential explanation, which we sought to address with a follow-up study, is that these beliefs were simply not strong enough to result in any perceived need or desire to adjust predictions. In other words, recalling that one tends to finish projects later than anticipated only a small proportion of the time, or only slightly later than predicted, may not provide a strong enough discrepancy to prompt adjustments to future predictions.

In addition, although not relevant to our main hypotheses, we found evidence that participants believe others finish projects later than predicted more often than they themselves

do, but only for must-do projects. This is in line with the idea, supported by research (e.g., Pronin, Lin, & Ross, 2002), that people believe others to be more biased than themselves. Research on the planning fallacy has also found that people make later completion predictions for others than for themselves (Buehler, Griffin, & Ross, 1994), which suggests that they may believe others to be more likely to underestimate completion times than themselves. It is interesting that this difference was observed only for must-do projects and not want-to-do projects. Although only speculation, a potential reason for this is that people are more likely to admit to themselves that they tend to finish later than predicted for want-to-do projects because these are ultimately less consequential for life success and more optional than must-do projects, and thus carry less of a threat to competence. For example, must-do tasks are more likely to have consequences for finances and career prospects whereas want-to-do task are more "for fun".

Study 2

To examine whether stronger planning fallacy beliefs would lead to a stronger relationship between beliefs and predictions, we manipulated participants' beliefs in the planning fallacy using a fictitious (Psychology Today) on-line article containing real research findings about the planning fallacy phenomenon. We assessed beliefs about the self or others (between-subjects) and, this time, asked participants to make predictions about a hypothetical academic project, either for the self or another student. We predicted that those who read the planning fallacy article would give greater planning fallacy belief ratings (self and other), make later completion predictions, and evidence a larger correlation between beliefs and predictions (due to the greater strength of and certainty about their planning fallacy beliefs).

Method

Participants. Of the 230 undergraduate students from Wilfrid Laurier University that participated in the study (on-line for course credit) 71 (30.9%) failed a simple comprehension

check ["Based on what you just read about the assignment, when is it due (i.e., what is the final deadline)?"] and were thus excluded from data analysis[3], leaving a final sample of 169. The final sample consisted of 122 female, 46 male, and 1 unspecified participants with a mean age of 19.50 (*SD* = 2.59). On our ethnicity question, 71.6% identified as Caucasian, 14.8% as Asian, 3.6% as Black, 1.8% as Latino, 2.4% as Middle Eastern, 0.6% as Aboriginal, 4.7% as other, and 0.6% did not specify.

Design. The study had a 2 (Article: planning fallacy vs. no article) by 2 (Target: self vs. other beliefs and predictions) between-subjects factorial design. More specifically, participants either read a fictitious Psychology Today article explaining that people often make optimistically biased completion predictions (exhibit the planning fallacy) or did not read any article (control condition). They then rated their planning fallacy beliefs either for the self or for the typical Laurier student. Next, they were asked to imagine having a major assignment due for a course in 14 days and to predict when they or the typical Laurier student would be finished the assignment if they were given the assignment right now, in real life.

Procedure.

Planning fallacy article. The article was entitled "Planning Fallacy", featured a picture of a calendar with a date circled as "DUE!", was formatted as a Psychology Today on-line article, and contained text explaining what the planning fallacy is and its prevalence ("There has been a significant mass of evidence collected strongly indicating that this occurs for a variety of tasks and everyday projects, such as completing large-scale industrial projects, academic assignments and personal tax forms. We tend to believe that we will finish tasks sooner than we actually do."), and an example from research (see Appendix B1).

[3] The answer to this question was bolded in the paragraph above this question. Knowing the correct answer to this question was crucial as knowledge of the deadline is important for making a completion prediction.

Article check. Participants in the article condition were asked "Based on the article you just read, on average, relative to when people expect to be finished, when do people actually finish tasks or projects?" and selected an answer one of three options "they finish later than their predictions", "they finish at the same time as their predictions", or "they finish earlier than their predictions". Of the eligible sample in that condition, 92.6% answered correctly.

Beliefs. Participants' planning fallacy beliefs were elicited using the same instructions and questions as in Study 1, with the exception that others were defined as "other students".

Predictions. We instructed participants to imagine that they, or "the typical Laurier student" (in the other condition), needed to complete a minimum 12-page research report for one of their courses, right now, in real life. The assignment had a hard deadline in 14 days and the instructor was giving an incentive for prompt completion – a 2% bonus for every day before the due date that the assignment is submitted (see Appendix B2). We chose a 12-page report because we thought that would make it a major assignment in the minds of participants, something they would imagine working on over a stretch of several days. We chose a deadline in 14 days because that seemed early enough that students would be able to imagine starting to work on it soon and far away enough that there was a reasonable amount of time available to be able to finish it on time. We chose to incorporate an incentive for finishing early to prevent a potential lack of variability in responses (the study was conducted at the end of the academic term, a very busy time for students, and we perceived a real risk that most participants could indicate that they would finish on the day of the deadline). As comprehension checks (on the same page) participants were asked to report 1) in how many days the assignment is due using a dropdown ranging from 1 days to 30 days (recall that 30.9% of the original sample answered incorrectly and these were excluded), 2) how many extra percentage points they could get for handing it in 2

days before the deadline on a 6-point scale ranging from 0% to 6% (69.6% of the original sample, 79.9% of the final sample answered correctly), and 3) the minimum number of pages the report needs to be with the option of 8, 10, or 12 pages (90.0% of the original sample, 97.6% of the final sample answered correctly).

Supplemental Measures. In addition to the primary dependent variables, we included several exploratory items assessing participants' opinions, knowledge, and beliefs about the planning fallacy. Participants rated how important they thought being accurate about completion predictions is, how motivated they are to make accurate completion predictions, and the extent to which they think underestimating when they will finish a project is problematic, on a scale from 1 (*not at all*) to 7 (*extremely*).

Participants in the planning fallacy article condition also rated the believability of the article on a scale from 1 (*I thought the information in the article was definitely true*) to 7 (*I thought the information in the article was definitely false*), with 92.6% indicating either that they thought the information was most likely true (5) or that the information in the article was definitely true (6).

Prior knowledge of the planning fallacy was assessed using the same yes-no question as in Study 1. One hundred twenty-two (72.2%) answered *no*, 46 (27.2%) answered *yes*, and 1 (0.6%) were missing, indicating that the majority of the participants did not have previous knowledge of the planning fallacy. These numbers did not differ significantly by article condition, B = -1.95, Wald = 2.82, p = .093, OR = 1.23, or Target, B = -1.68, Wald = 2.17, p = .141, OR = 0.57, but a just significant interaction emerged, B = 1.45, Wald = 383, p = .050, OR = 4.18. When graphed, the interaction pattern was that previous knowledge was greater in the article, other condition than in the other three conditions. Because these ratings were given at the

end, it is unclear whether these differences represent genuine differences in knowledge prior to exposure in this study (and thus failure of random assignment) or whether this is an effect of the manipulation, such that having learnt about the planning fallacy and having thought about how it might be true for others led participants to think that this is something they were already previously aware of. Predictions did not differ depending on whether people were previously aware of the planning fallacy or not, $t(166) = 0.47$, $p = .637$, $d = 0.07$.

Results

Beliefs. First, we examined beliefs about the planning fallacy reported by participants in the control condition. Consistent with hypothesis 1, the mean planning fallacy belief rating for the self was 3.10 (*SD* = 1.27), indicating that participants believed that they tend to finish projects slightly later than predicted (3 on the scale), similar to Study 1. This mean was significantly different from the midpoint of 4, $t(51) = -5.12$, $p < .001$, $d = -0.71$, and the median and mode were both 3. For the percent beliefs question, participants indicated they finished projects later than predicted 40% of time (*SD* = 24.89), at the same time as predicted 36% of the time (*SD* = 21.83), and earlier than predicted 24% of the time (*SD* = 18.77). At least one of these percentages differed significantly from the others, $F(2, 102) = 5.09$, $p = .008$, $\eta_p^2 = .091$. Specifically, participants thought they were later than predicted significantly more often than they were early, $t(51) = 3.08$, $p = .003$, $d = 0.73$, on time more often than they were early, $t(51) = 2.60$, $p = .012$, $d = 0.56$, and late no more often than they were on time, $t(51) = 0.80$, $p = .795$, $d = 0.20$. We again calculated a "percent bias" variable (% late - % early), capturing the extent to which participants believed their prediction errors were directional. Overall, mean percent bias for the self in the control condition was 16.37 (*SD* = 38.38), indicating that participants believed that they tend to finish later than predicted 16% more often than earlier.

Next, we submitted participants' belief ratings to a 2 (Article) x 2 (Target) ANOVA to test for effects of the manipulations (see Table 5 for means). A significant main effect of Article, $F(1, 165) = 13.24$, $p < .001$, $\eta_p^2 = .074$, indicated that participants who read the article gave greater planning fallacy belief ratings ($M = 2.44$, $SD = 1.10$) than those who did not ($M = 3.10$, $SD = 1.21$). The main effect of Target only approached significance, $F(1, 165) = 2.49$, $p = .117$, $\eta_p^2 = .015$, and a marginal Article by Target interaction emerged, $F(1, 165) = 2.76$, $p = .099$, $\eta_p^2 = .016$. Post-hoc comparisons revealed that the effect of the article manipulation was stronger for others than the self: beliefs differed significantly between those in the article and control condition for others, $M_{Article} = 2.17$, $SD = 0.99$ vs. $M_{Control} = 3.11$, $SD = 1.14$, $t(76) = 3.92$, $p < .001$, $d = 0.87$, but did not differ significantly (although trending in the same direction) for the self, $M_{Article} = 2.74$, $SD = 1.14$ vs. $M_{Control} = 3.10$, $SD = 1.27$, $t(89) = 1.37$, $p = .175$, $d = 0.29$. This result suggests that participants' planning fallacy beliefs were readily shifted by the article for others but may have been more resistant to change for the self.

A similar pattern of results emerged for the percent belief questions (see Table 5 for means and significance tests). Focusing on the percent bias score, we observed that participants believed bias was greater for others ($M = 28.81$, $SD = 29.94$) than for the self ($M = 18.04$, $SD = 35.27$), $F(1, 165) = 3.73$, $p = .055$, $\eta_p^2 = .022$. A main effect of article also emerged, $F(1, 165) = 6.11$, $p = .014$, $\eta_p^2 = .036$, with greater bias percentages in the article ($M = 29.70$, $SD = 29.69$) compared to the control ($M = 16.86$, $SD = 35.30$) condition. The Target by Project type interaction was marginally significant, $F(1, 165) = 2.86$, $p = .093$, $\eta_p^2 = .017$, such that the difference between the article and no article condition was significant only for others ($M_{Article} = 38.45$, $SD = 25.67$ vs. $M_{Control} = 17.57$, $SD = 30.96$, $t(76) = -3.26$, $p = .002$, $d = -0.72$) and not for

the self, ($M_{Article}$ = 20.28, *SD* = 31.12 vs. $M_{Control}$ = 16.37, *SD* = 38.30, *t*(89) = -0.52, *p* = .603, *d* = -0.11).

Predictions. A 2 (Article) x 2 (Target) ANOVA performed on the predicted completion times revealed a significant main effect of Article, *F*(1, 165) = 4.89, *p* = .028, η_p^2 = .029. Participants who read the article (*M* = 11.25, *SE* = 2.70) made later predictions than those who did not (*M* = 10.30, *SE* = 3.02) (see Table 6 for means). There was no main effect of Target, *F*(1, 165) = 0.97, *p* = .326, η_p^2 = .006, but a significant interaction, *F*(1, 165) = 495, *p* = .027, η_p^2 = .029. As with beliefs, this interaction indicated that the article manipulation had an effect on predictions for others, $M_{Article}$ = 11.93, *SD* = 2.57 vs. $M_{Control}$ = 9.97, *SD* = 2.57, *t*(76) = -2.91, *p* = .005, *d* = -0.66, but not for the self, $M_{Article}$ = 10.51, *SD* = 2.67 vs. $M_{Control}$ = 10.52, *SD* = 2.67, *t*(89) = 0.01, *p* = .991, *d* = -0.003. Thus, as in Study 1, there was no effect of the experimental manipulation on participants' predictions concerning their own task completion times.

Correlations. Next, we calculated zero-order correlations between beliefs and predictions, to get an overall sense of the relationship between these variables (see Table 7). Self and other beliefs were not significantly correlated with predictions, although for the Likert beliefs item the correlation trended in the expected direction (greater beliefs in finishing later than predicted = later predictions), *r* = -.17, *p* = .103 for self, *r* = -.17, *p* = .143 for other. The correlation was significant collapsed across Target, *r* = -.17, *p* = .020. As in Study 1, these results suggest there was, at best, a small connection between beliefs about when one finished past projects relative to predictions and predictions for an upcoming (hypothetical) project.

Mediation. To test whether planning fallacy beliefs mediated the effect of the article manipulation on predictions we performed a simple mediation analysis using 10,000 bootstrap samples (Model 4; Hayes, 2012), with beliefs as a mediator of the effect of the article

manipulation on predictions (see Table 8). The effect of the article on predictions was mediated through belief ratings [95% CI .0184, .5552]. This finding suggests that the article increased participants' beliefs in the planning fallacy which in turn led them to predict longer completion times. The same analysis performed on the % late and % bias items did not reveal mediation effects. Because the effect of the article on beliefs and predictions tended to be moderated by Target (self vs. other) we also ran moderated mediation analyses (Model 8; Hayes 2012) which did not provide any evidence that effects of the article on predictions through beliefs depended on whether participants were giving ratings for themselves or others (all CIs contained zero, see Table 9). This may be because the effect of the article on beliefs (the mediator) was only marginal. Simple mediation analyses for self and other separately were not significant for the self [95% CI -.0287, .5319] or for others [95% CI -.1112, .1124], perhaps suggesting a small effect or insufficient power, i.e., that the mediation was only detectable with the larger self-other combined sample size.

Supplemental measures. Participants thought that making accurate completion predictions is relatively important, $M_{overall}$ = 5.29 (*SD* = 1.28) (on a 7-point scale), reported being relatively motivated to make accurate completion predictions, $M_{overall}$ = 4.90 (*SD* = 1.27) (on a 7-point scale), and thought that underestimating when one will finish a project is more than moderately (a score of 4) problematic, $M_{overall}$ = 5.07 (*SD* = 1.39) (on a 7-point scale). These ratings did not differ significantly by Article, Target, or the Article by Target interaction, all *Fs* < 2.25, *p*s > .125. Importance of and motivation to make accurate completion predictions was unrelated to predictions, *r* = .01, *p* = .925, and *r* = -.10, *p* = .194, respectively. The extent to which participants thought underestimating when one will finish a project is problematic showed

a small negative correlation with predictions, r = -.15, p = .050, indicating that earlier predictions were related to greater problematic ratings.

Discussion

In support of Hypothesis 1 and replicating Study 1 results, participants recalled finishing past projects slightly later than predicted, indicating that they believe in the planning fallacy. We also found evidence that we strengthened participants' planning fallacy beliefs using a fictional article and lead them to make later completion predictions compared to those who did not read the article. Interestingly, these effects held true only when people were thinking about others, as opposed to themselves. This may be because 1) we can be less certain about our knowledge and the amount of information we have about others and may thus be more easily influenced by outside information, and 2) we may be more resistant to information indicating that our completion predictions are inaccurate (Pronin, Lin, & Ross, 2002) and more likely to be influenced by unrealistic optimism and desired completion times (Buehler, Griffin, & Ross, 1994). We could also be more likely to view an optimistic prediction as a motivator to finish earlier for ourselves than for others. This idea is based on the speculation that we may be more likely to identify the need for optimism as a motivation booster for ourselves than for others. On the other hand, research by Armor, Massey, and Sackett (2008) showed that both for the self and for others, in a variety of life settings (e.g., financial investment, dinner party), people believe that their predictions are optimistically biased and that they should be, i.e., that optimistically biased predictions are ideal. In fact, participants indicated that predictions should be even more optimistic for others than for the self. Regarding the link between optimistic predictions and performance, initial evidence suggests that more optimistic predictions may lead to better performance, but not as much as we expect (Tenney, Logg, & Moore, 2015). In the context of task completion predictions specifically, Buehler, Peetz, and Griffin (2010) found that optimistic

predictions may lead to finishing tasks sooner, but found evidence for this only for closed tasks (those that can be completed in a single session), and not for open tasks (those requiring many steps to be completed across different times and places). For open tasks, predictions did influence when these tasks were started (earlier predictions = earlier start times) suggesting that completion predictions have the greatest impact on the beginning phases of a project, particularly the initiation time, but that this effect diminishes over the course of a multi-stage project.

Support for a relationship between planning fallacy beliefs and completion predictions was somewhat mixed in this study – we found only a small correlation for one of the beliefs items, and only one of the three mediation analyses, testing whether reading the planning fallacy article led to later completion predictions through changes in beliefs, was significant.

Taken together, the findings of Study 1 and 2 are inconclusive in regard to hypothesis 2 – reminding people of past lateness did not consistently lead to changes in later predictions for an upcoming task. Indeed, for one's own tasks, the experimental manipulations did not have an impact on predicted completion times. This is perhaps not altogether surprising given previous findings that simply reminding people of their typical task completion times (in relation to deadlines) did not influence their self-predictions (Buehler et al., 1994; Buehler & Griffin, 2003) and that participants were not explicitly instructed to incorporate the past information into their predictions, something that will be addressed in Study 3.

One potential explanation for why we failed to observe an effect of our manipulations on self predictions, or a clear relationship between beliefs and predictions, is that participants did not believe their past project completion times are relevant for their current project. This could be because our beliefs question(s) was quite broad and did not specify project type – perhaps there was some mismatch between the types of projects participants imagined for the beliefs

question and the project they nominated or were asked to think about when making their predictions. In addition, participants may believe that the current project or their situation is different in some systematic way or that they have improved their prediction capabilities since the past projects. It is also possible that participants simply aren't motivated to try to be accurate and instead generate a prediction that is desirable, i.e., when they want to finish, or a goal to strive toward.

We also didn't assess the accuracy of people's predictions. Although previous research on the planning fallacy suggests that people who generate longer task completion predictions tend to be less biased (i.e., less prone to the planning fallacy), this was not tested in our studies. To address these potential explanations, and limitations, in the remaining studies (Studies 3 through 6) we examine the reference class forecasting (RCF) technique which provides a more structured and formalized approach to using past experience as a basis for prediction.

Study 3

Study 3 was designed to address some of the issues raised in the first two studies by using a structured prediction method, reference class forecasting, to guide participants to incorporate past completion times into future predictions. All participants self-nominated an academic project they would be completing in the next two weeks, made a prediction for when they would complete it, and were followed up to find out when they actually completed it, allowing us to assess prediction accuracy. Those who received the reference class forecasting intervention were instructed to think of and list 5-6 similar past projects, reported when they completed those projects, and then received directions for how to use this information (or not). We expected that participants who received the reference class forecasting intervention would make later, and thus less optimistically biased predictions than control participants (hypothesis 4). We reasoned that thinking about past completion times in terms of specific, similar past projects, as opposed to all

past projects in general (Studies 1 and 2) would be a more effective way of making past projects relevant for current predictions (although we do not conduct a study which does a direct comparison).

Method

Participants. Of the 305 undergraduate students from Wilfrid Laurier University that participated in the study (on-line for course credit) 45 did not nominate a project as per the instructions (29 nominated a project with a deadline that was less than 4 days away; 11 failed to nominate a project; 5 identified projects that could only be completed on a specific date in a single session), and were thus excluded from the analyses, leaving a total of 222 participants in the final sample. The final sample consisted of 161 female, 40 male, 1 other, and 1 unspecified participants with a mean age of 19.83 (SD = 2.51). On our ethnicity question, 69.0% identified as Caucasian, 21.2% as Asian, 3.4% as Caribbean, 1.5% as Latino, 1.5% as Middle Eastern, 0.5% as Black, 0.5% as Aboriginal, 2.0% as other, and 0.5% did not specify.

Design. The study had a single factor design where participants were randomly assigned to one of four conditions, a control condition, a past condition (i.e., think of similar past projects and how many days before or after the deadline you finished them and then make a prediction), and two reference class forecasting conditions. Because participants in the Past group are not instructed to incorporate their past into their predictions, we do not consider this an RCF condition. The “RCF deadline” condition is similar to the past condition but in addition presents participants with the mean completion time from their past projects and instructs them to use this mean to make their prediction, unless they have better other information available to make their prediction. In the “RCF prediction” condition, participants made an initial, intuitive completion prediction, thought of similar past projects and how many days before or after their initial prediction they finished them, and we then presented them with the mean completion time from

their past projects (in days from your original prediction) and instructed them to use this mean to adjust their original prediction, unless they had better information available to make their prediction.

We predicted that both the "RCF deadline" and "RCF prediction" condition would lead to later and less optimistically biased completion predictions than the control condition (hypothesis 4), but that of the two the "RCF prediction" condition would be most effective (hypothesis 5). We expect this because it forces participants to realize that their completion predictions tend to be wrong (i.e., that they tend to finish after they expected to), and that if they want to be accurate, they will need to change their predictions. In the "Past" and "RCF deadline" conditions the focus is not on how one has erred in the past (just on how many days before the deadline one finished in the past), so participants could more easily attribute completion times to external factors that may not be at play for the current project. We predict that the "Past" condition might fall in between the control and "RCF deadline" condition – on the one hand, making completion times of similar past projects salient could lead some participants to realize that they tend to finish relatively close to the deadline and this could be enough to shift their predictions a little later; on the other hand, because no instructions were provided for how to use this information, some participants may not feel compelled to incorporate it into their prediction. This idea is supported by previous planning fallacy research which found that participants were not influenced by recalling past completion (e.g., Buehler et al., 1994; Study 4; Buehler & Griffin, 2003), and our limited success with this approach in Study 1 and 2.

Procedure.

Project nomination (Part 1). For this study, all participants were instructed to nominate and briefly describe an academic project with a hard deadline within the next two weeks and at

least five or more days away, that they could complete any time between now and the deadline (see Appendix C1 for details). Prior to being assigned to one of our experimental conditions, they were also asked to rate: the difficulty of the project, its importance, how much control they have over it, how busy they would be with other tasks while completing it, the extent to which they would want to complete it as quickly as possible, and the importance of the quality of the final project, on a scale from 1 (*not at all*) to 7 (*extremely*). These were intended as potential covariates, and were also assessed again after the manipulation, to be able to explore whether going through the RCF procedure changed perceptions of these project characteristics.

Manipulation and predictions (Part 1). Participants in the control condition were simply asked to predict, as accurately as possible, when they would be finished the task or project they nominated in days before the deadline (see Appendix C2). Predictions in all conditions were made using a slider which ranged from "7 days or more before the deadline" to "7 days or more after the deadline" with "0 deadline day" as the midpoint. For the three experimental conditions, participants received additional instructions, including the following prelude,

> "Later in this survey we will ask you to predict, as accurately as possible, when you will be finished the project/task described above. One approach to making an accurate prediction is to think about similar past projects and when you typically completed them. We would like you to do this now. To assist you in the process, please complete the steps that follow."

In the "Past" condition, participants were instructed to list 5-6 similar past projects (see Appendix C3), to indicate how many days before the deadline they finished each of those projects (see Appendix C4), and then to make a prediction, without any information about how to incorporate the information about past projects into the prediction. The "RCF deadline" condition was the same as the "Past" condition, except that we also presented participants with their average past completion time (i.e., average days before the deadline, based on the projects they had listed) and instructed them:

"Therefore, unless you believe that you have better information available to make your prediction for this project than for your past projects, you should base your project completion prediction on this average. If your average is just before the deadline, then you should predict just before the deadline for the current project, unless you have reliable information to indicate that this project will be different." (see Appendix C5)

In the "RCF prediction" condition participants were instructed to 1) make an intuitive prediction about when they would finish their project in days before the deadline (see Appendix C6), and 2) list 5-6 similar past projects and when they finished them relative to when they had initially expected to be finished (i.e., relative to their initial predictions) (see Appendix C7). They were then 3) reminded of their initial prediction for the target project and their average past completion time, and instructed:

"Therefore, unless you believe that you have better information available to make your prediction for this project than for your past projects, you should base your project completion prediction on this average. If your average is finishing slightly later than predicted, then you should adjust your initial prediction to be somewhat later, unless you have reliable information to indicate that this project will be different." (see Appendix C8)

Operationalization of RCF. Our operationalization of the RCF procedure for the current context is worth discussing. Although we attempted to keep it as close as possible to the steps outlined in the literature (Flyvbjerg, 2006; Flyvbjerg et al., 2009; Lovallo & Kahneman, 2003), we also had to consider how feasible the steps would be for individual students without formal past project data, that were completing the exercise using on-line survey software with limited time and attentional resources, and we made adjustments accordingly:

- Step 1 of RCF involves identifying a relevant class of past, similar projects that are broad enough to be statistically meaningful but narrow enough to be truly comparable with the focal project – we believe our instructions did a good job meeting this step as we asked participants to nominate 5 to 6 past projects that

were as similar as possible to the current project in terms of scope, type, complexity, available work time, and control.

- Step 2 of RCF involves establishing a probability distribution for the selected reference class of the parameter that is being forecasted using credible, empirical data for a sufficient number of projects to be able to make statistically meaningful conclusions. Because our participants most likely did not have credible empirical data available and the number of past projects we asked them to recall was limited (five to six, a sample size that may not be statistically meaningful), our operationalization was weaker for meeting the definition of RCF for Step 2. We did give participants a visual of their past projects' distribution, using stacked sliding scales which gave them an overview of their past project completion times on one page (see Appendix C4, C7). The lack of credible empirical data available to participants is part of what makes the current research novel and interesting – posing the question of whether for individual, personal projects, RCF can still lead to less optimistically biased completion predictions, even without objective past data.
- Step 3 of RCF involves comparing the current project with the reference class distribution to estimate the most likely outcome, i.e., placing the current project in the distribution of past project outcomes. This step may arguably be quite complex, and we were not confident that we would be able to successfully instruct participants to be able to determine on their own where within their past project distribution their current project should fall because this would require them to be able to accurately and objectively judge ways in which the current project

situation is most similar to past project situations. This type of comparison is most feasible when measurable characteristics of past projects are available in records, e.g., for constructions projects the features of the project, such as the type of structure, required materials, manpower, factors that led to delays, etc., are all documented and can be referenced during the RFC process. Because such objective information is most likely not available to students, the danger exists that optimistic bias would lead them to place the current project in the earliest completed projects portion of the distribution. Therefore, we instead presented participants with an objective statistic from their past project distribution that they could easily relate to, their average past completion time, and asked them to use this mean to make their prediction (unless they had better information available to make their prediction for the project).

Exploratory measures (Part 1). Participants answered additional questions about how they arrived at their prediction, predicted their number of working hours, other project characteristics, and personality measures – these were not central to our hypotheses and thus the relevant methods and results can be found in Appendix C9.

Previous knowledge. Prior knowledge of the planning fallacy was assessed using the same yes-no question as in Study 1. One-hundred sixty-two (53.1%) answered *no*, 126 (41.3%) answered *yes*, and 17 (5.6%) were missing, indicating that a significant proportion of the participants had at least some previous knowledge of the planning fallacy phenomenon, but these numbers did not differ by condition, $X^2(3) = 3.50$, $p = .320$, $V = .110$. In addition, predictions (days before the deadline), completion times, and bias (prediction – actual completion time) did

not differ depending on whether participants were previously aware of the planning fallacy or not, all *t*s < 1.10, *p*s > .275, *d*s < 0.20.

Completion times (Part 2). Two days after the deadline for their project participants received an email inviting them to participate in Part 2 of the study. In the survey, participants were reminded of their project and deadline from Part 1 and could correct the information if applicable (i.e., if the deadline was incorrect or had changed). Next, we asked if they finished their project (*yes*/*no*) and when they finished their project (from -7 - *seven or more days before the deadline*, to 7 - *seven or more days after the deadline*). We recoded this variable for the purposes of our analyses, so that days before the deadline became positive numbers and days after the deadline became negative numbers. Participants also specified the number of hours they spent working on the project. See Appendix C10 for the verbatim instructions.

Results

Because this study contained both an initial session (Part 1: manipulation and predictions) and a brief follow-up (Part 2: completion times), there are two sets of participants for which results could be presented (all eligible participants who completed Part 1, *n* = 222, all eligible participants who completed both parts, *n* = 154). We will start by presenting the Part 1 results for all participants, and then report results for the subset of participants who completed both parts.

Part 1.

***Pre-manipulation measures*.** Participants' project deadlines did not differ by condition, $F(3,218) = 0.98$, $p = .404$, $\eta_p^2 = .013$, and were on average 10.56 days (*SD* = 4.84) from the day participants completed Part 1 of the study. Participants did not differ by condition on the following project characteristics, difficulty (*M* = 4.63, *SD* = 1.18), importance (*M* = 5.92, *SD* = 1.17), busyness with other tasks (*M* = 5.38, *SD* = 1.25), finishing as quickly as possible (*M* = 4.83, *SD* = 1.19), or quality (*M* = 6.14, *SD* = 1.01), all *F*s < 1.25, *p*s > .350, but did differ on how

much control participants thought they would have over the project, $F(3, 218) = 2.90$, $p = .036$, $\eta_p^2 = .038$, indicating that random assignment may not have been entirely successful. Participants in the control condition ($M = 5.46$, $SD = 1.24$) gave lower control ratings than participants in the RCF deadline condition ($M = 6.02$, $SD = 0.88$), $t(98.18) = -2.77$, $p = .007$, RCF prediction condition ($M = 5.89$, $SD = 1.03$), $t(109) = -1.97$, $p = .051$, and Past condition ($M = 5.78$, $SD = 1.07$), $t(103) = -1.37$, $p = .173$. None of the six project characteristics showed a significant bivariate correlation with our key dependent variable, predictions (all rs < .11, ps > .100), so only the control item was included as a covariate in the analyses that follow.[4]

Results, including means, standard deviations, F-tests, p-values, and effect sizes, for all dependent variables following the manipulation are presented in Table 10. Unadjusted means are presented in the tables and adjusted means are presented in text. For readability, statistics will be presented in text only for significant results.

Past project characteristics (1). In the three non-control conditions, participants listed past projects and when they completed them, either relative to the deadline (Past, RCF deadline) or relative to predictions (RCF prediction) [5]. For participants' average past completion times, we expected that participants would report that they tend to finish relatively close to the deadline (e.g., 1 or 2 days before, i.e., a positive number) and, based on the results of Study 1 that they tend to finish slightly later than their prediction (e.g., 1 or 2 days after, i.e., a negative number). Average past completion times in the RCF deadline and Past condition should not differ from one another (as the past projects nomination procedure is the same in the RCF deadline and Past

[4] The pattern of results for analyses with and without this control variable was the same, with only very small changes in effect sizes.

[5] Past project completion ratings from the RCF prediction condition were recoded from the way they were presented in the survey, so that positive numbers indicated finishing before the original prediction (earlier) and negative numbers indicated finishing after the original prediction (later).

conditions) and are not directly comparable to average past completion times in the RCF prediction condition because they have different reference points (deadline vs. predicted completion day). Holding constant control ratings (i.e., presenting adjusted means), participants in the RCF prediction condition reported finishing on average 0.65 days (*SE* = 0.22) after their predictions (a negative number), consistent with hypothesis 1. RCF deadline participants reported finishing on average 1.69 days (*SE* = 0.22) before the deadline (a positive number) and did not differ significantly from participants in the Past condition, who reported finishing on average 1.72 days (*SE* = 0.21) before the deadline, $t(107) = -0.03$, $p = .980$, $d = 0.02$.

***Predictions*.** Recall we predicted that participants in the control condition would have the most optimistic predictions (furthest before the deadline), followed by participants in the Past, RCF deadline, and RCF prediction condition which we expected to have the least optimistic predictions (closest to the deadline). The one-way ANCOVA revealed a significant omnibus effect of the manipulation, $F(2, 217) = 3.59$, $p = .014$, $\eta_p^2 = .047$. Planned comparisons revealed that participants in the RCF prediction condition ($M_{adj} = 1.17$, *SE* = 0.21) indeed made significantly later predictions than those in the control condition ($M_{adj} = 2.05$, *SE* = 0.21), $t(108) = -3.22$, $p = .002$, $d = -0.56$. Inconsistent with hypothesis 4, participants in the RCF deadline condition ($M_{adj} = 1.70$, *SE* = 0.20) did not make significantly later predictions than those in the control condition, $t(115) = -1.27$, $p = .206$, $d = -0.22$. Those in the Past condition ($M_{adj} = 1.96$, *SE* = 0.22) also did not make later predictions than those in the control condition, $t(102) = -0.25$, $p = .806$, $d = -0.06$. Consistent with hypothesis 2 and 4, RCF prediction participants made later completion predictions than Past participants, $t(101) = -2.85$, $p = .005$, $d = -0.51$, and RCF deadline participants, $t(114) = -2.05$, $p = .042$, $d = -0.34$.

Participants in the RCF prediction condition also made an initial, intuitive prediction before making their final prediction. Initial predictions in this condition (M_{adj} = 1.55, *SE* = 0.22) did not differ significantly, albeit marginally, from predictions in the control condition, *t*(108) = -1.70, *p* = .093, *d* = 0.31. Initial and final predictions in RCF condition did differ significantly in the expected direction such that initial predictions were more optimistic than final predictions, indicating that people adjusted their predictions as they went through the exercise, *t*(54) = 2.08, *p* = .042, *d* = 0.56.

We also compared predictions and past project completion times in the RCF deadline and Past condition (where they were in the same units, days before the deadline), to examine whether these were the same or differed. Presumably, if participants simply used their recalled past completion times, these should not differ significantly. In the Past condition, the average past completion time was 0.26 days (*SD* = 1.76) closer to the deadline than the completion prediction, but the two means did not differ significantly, *t*(48) = -1.05, *p* = .302, *d* = -0.15. In the RCF deadline condition, the average past completion time was 0.03 days (*SD* = 1.70) closer to the deadline than the completion prediction, so these means also did not differ significantly, *t*(61) = -0.15, *p* = .879, *d* = -0.02. That past completion times and predictions did not differ could mean that participants used their past completion times to make their current predictions (see also the "Correlations" section).

***Correlations*.** To better understand the relation between past project completion times and predictions, we correlated past project completion times and predictions separately within each condition. There was a medium to large positive relationship in the past condition, *r* = .41, *p* = .004 and in the RCF deadline condition, *r* = .51, *p* < .001, indicating that participants who reported finishing closer to the deadline in the past also made predictions that were closer to the

deadline this time. In the RCF prediction condition, the more days after their prediction participants reported finishing in the past, the closer to the deadline (later) their current predictions were, r = -.33, p = .012. In other words, people who tend to finish early (be it more days before the deadline, or more days before their predictions) expect to finish early again. These correlations leave open the possibility that participants made some use of their average past completion times when generating the current completion predictions. Correlations with our exploratory measures are presented in Appendix C11 (see also Table 11).

Part 2. Two days after their project's deadline, eligible participants were contacted via email with a link to our follow-up survey and 167 (75.2%) completed it. Participants were reminded of the project description and deadline they gave in Part 1 and asked to specify the deadline for the project. This was done to ensure that participants remembered the deadline for the project correctly, and in case the deadline had changed. In cases where the deadline from Part 1 and 2 mismatched by more than one day, participants were excluded from the Part 2 analyses because the completion time question that followed was expressed in days before the deadline (so if the project deadline had changed, or was incorrectly specified in Part 1, it was unclear when participants finished relative to their original prediction). Thirteen participants had mismatching deadlines, leaving 154 responses eligible for analysis. Of these, 10 specified that they did not finish their project (1 in the control, 2 in the past, 4 in the RCF deadline, and 3 in the RCF prediction condition). Also, 30 participants did not give a completion date (left that question blank) leaving only 114 participants for "actual" completion time analyses. This seemed like an unusually large number of non-responses, and upon inspection of our survey it became apparent that some participants may have left the slider at "0" (the day of the deadline) and failed to notice our instruction for that question: "Make sure you move the slider, otherwise your data

will not be recorded!" which was in black font, unlike in Part 1, where it was in bright red font. Thus, some participants may have tried to respond with the deadline day as their completion time and we failed to record it. We therefore analyzed the data both without ("sample A" in Table 12 and in text; $n = 114$) and with ("sample B" in Table 13 only; $n = 144$) the participants coded as having finished on the day of the deadline.

Completion Times and Prediction Bias. Participants reported finishing on average 1.07 ($SD = 1.65$) days before the deadline and these completion times did not differ by condition (see Table 12 and 13). Next, we calculated prediction bias by subtracting "actual" completion times from predicted completion times (i.e., difference score). See Figure 1 for a visual comparison of predicted and actual completion times by condition. Notably, within the control condition, this difference was significant ($M_{Diff} = 1.38$, $SD = 1.72$, $t(40) = 16.00$, $p < .001$, $d = 0.91$), and thus the study revealed further evidence of the planning fallacy. Collapsed across conditions, this difference was smaller, but still significant ($M_{Diff} = 0.87$, $SD = 1.79$, $t(113) = 5.19$, $p < .001$, $d = 0.49$). Planned comparisons indicated that prediction bias in the RCF prediction condition ($M_{adj} = 0.44$, $SE = 0.35$) trended toward being lower than prediction bias in the control condition ($M_{adj} = 1.21$, $SE = 0.29$), $t(54) = -1.60$, $p = .115$, $d = -0.32$. Prediction bias did not differ between the control and Past ($M_{adj} = 1.23$, $SE = 0.34$), $t(55) = 0.07$, $p = .945$, $d = 0.02$, and control and RCF deadline ($M_{adj} = 0.55$, $SE = 0.29$) conditions, $t(54) = -1.53$, $p = .131$, $d = -0.38$, although a trend toward prediction bias being somewhat smaller in the RCF deadline than in the control condition was noticeable, with an adjusted mean difference of 0.66 days ($SE = 0.42$). Prediction bias didn't differ significantly between the RCF prediction and RCF deadline condition, $t(53) = -0.26$, $p = .794$, $d = 0.07$, and the RCF prediction condition trended toward somewhat lower bias than the Past condition, $t(44) = -1.60$, $p = .117$, $d = 0.32$, with a mean difference of 0.79 days ($SE = 0.49$).

We also calculated a binary measure of prediction bias, dividing participants into those who finished by the time they predicted and those who finished later than predicted. Results for this variable were very similar to what was observed for the continuous measure. The Chi-squared test for the overall effect of condition was not significant (see Table 12, 13), but planned comparisons revealed that marginally more participants finished by the predicted time in the RCF prediction condition (60.9%) than in the control condition (38.2%), $X^2(1) = 2.82$, $p = .079$, $V = .167$. The Past (33.3%) and control condition did not differ, $X^2(1) = 0.15$, $p = .460$, $V = .030$. The RCF deadline (57.6% finished on/before) and control condition differed marginally, $X^2(1) = 2.51$, $p = .090$, $V = .140$, such that RCF deadline participants were slightly more likely to finish early/on time. The RCF deadline condition differed marginally from the Past condition, $X^2(1) = 3.28$, $p = .061$, $V = .107$, such that RCF deadline participants finished early/on time more often than Past participants. RCF prediction participants finished early/on time marginally more often than Past participants, $X^2(1) = 3.58$, $p = .054$, $V = .137$, but RCF prediction and RCF deadline participants did not differ, $X^2(1) = 0.06$, $p = .513$, $V = .029$.

As secondary analyses we also re-tested our key Part 1 measures with the Part 2 sample, compared average past completion times with completion times for the current project, tested for working time prediction bias, and analyzed reasons for finishing later than predicted (see Appendix C12; Table 12, 13).

Correlations. To examine the relationship between past project completion times, predictions, and actual completion times, we ran zero-order correlations broken down by condition (see Table 14). Past project completion times and predictions were correlated similarly in the Part 2 and Part 1 sample – a weak to moderate positive relationship, such that reporting finishing closer to the deadline was related to making later (closer to the deadline) completion

predictions. Predictions and actual completion times were weakly to moderately correlated, r = .36 to .46, in the expected direction. Past completion times and "actual" reported completion times were correlated around .3, indicating that participants who reported finishing later in the past (closer to the deadline, or later than predicted) also reported finishing later for the current project. We also correlated some of our exploratory variables of interest with our Part 2 variables, the results of which are summarized in Appendix C11 (see also Table 15).

Power analyses. Given the novelty of our manipulation, it was unclear what type of effect sizes we should expect, and we had not set sample size goals for the study and simply ran the study in the student participant pool until the end of the ongoing semester. Thus, the possibility exists that we were not able to detect effects due to insufficient power. To gain a better understanding of this potential limitation, we conducted sensitivity power analyses (see Appendix C13 for details), which suggested that the study was powered to detect medium-sized effects for the predictions ANCOVA and predictions t-tests for the Part 1 sample, medium-sized effects for the prediction bias ANCOVA for the Part 2 sample, but only had power to detect large-sized effects for the prediction bias t-tests. For studies 4, 5, and 6, we conducted a priori power analyses which aimed to allow us to detect medium-sized effects for each of our four key analyses, as far as feasibility (number of participants that could be recruited via our participant pool, financial resources available to pay MTurkers) allowed. We then conducted follow-up sensitivity analyses on the resulting final sample sizes (see also Appendix C13) which indicated power to detect medium-sized effects for most analyses. Exceptions include Study 6 with power to detect a small-medium sized effect for the Part 1 predictions ANOVA, and Study 4 and 5 with power to detect only medium-large sized effects for Part 2 prediction bias t-tests. The results of the power analyses suggest that we may fail to detect some smaller effects.

Discussion

We found some tentative evidence that reference class forecasting leads to less optimistic completion predictions than a control condition (hypothesis 3), as well as evidence that this is only the case when participants provided past completion times relative to their previous predictions (RCF prediction condition; hypothesis 5). When participants gave past project completion times relative to deadlines, although those predictions were slightly later, they did not differ significantly from those in the control condition. As previously mentioned, this difference in predictions depending on whether past completion times were recalled relative to deadlines vs. predictions may be because recalling completion times relative to one's own predictions may result in reflection or realization that one's own predictions are the problem, something that doesn't necessarily happen when one recalls completion times relative to deadlines as deadlines are often externally determined. It is possible that a focus on the past relative to predictions may make it more likely that people focus on changing their predictions, whereas a focus on deadlines may make it more likely that people try to focus on changing their behaviour to finish more quickly instead.

We also asked participants to report when they in fact finished their projects, allowing us to calculate prediction bias (how many days before/after their predictions did people finish), and found only weak evidence (marginal effects, trends) that those in the RCF prediction condition showed less prediction bias than those in the control condition (hypothesis 4). Some trends also emerged suggesting that those in the RCF deadline condition (past project complete times relative to the deadline) showed slightly less prediction bias than those in the control condition. Given the weak but suggestive evidence found in this first study and the large reduction in sample size for Part 2, we ran a replication study (Study 4) with some small changes.

Finally, we once again observed that participants recalled finishing past projects slightly later than predicted (0.66 days after their predictions, RCF prediction condition), in line with the first two studies and hypothesis 1. We found larger correlations between past project completion times and completion predictions in all of the non-control group conditions in this study compared to the first two studies, and although not necessarily causal, the results could suggest that when people are explicitly instructed to make accurate completion predictions and use specific similar past project information to do so, they are more likely to incorporate information from their past projects into their current project predictions. Comparing the relationship between past project completion times and future project completion predictions for a reference class forecasting to a control group could provide causal evidence (tested in Study 5 and 6).

Study 4

The purpose, hypotheses, and methodology of Study 4 was identical to that of Study 3 with a few exceptions: 1) we obtained a different sample: American adults (Mturkers) as opposed to students, 2) participants could nominate a work or academic project (as opposed to just academic), and 3) we included tasks with longer deadlines – we specified that the deadline could be up to 3 weeks away (as opposed to 2 weeks). These changes allowed us to observe whether a similar pattern of results would emerge with a different sample in a slightly different context.

Method

Participants. Of the 303 U.S. adults that participated in our online study via mturk.com (paid $1.50, $1 for Part 1, $0.50 for Part 2), 18 did not nominate a project as per the instructions (15 failed to describe a real project, 2 identified projects that could only be completed on a specific date in a single session, 1 nominated a project with a deadline that was more than 21 days away), and were thus excluded from the analyses. The final sample consisted of 285

participants, 173 female, 106 male, and 6 unspecified participants with a mean age of 36.45 (*SD* = 11.10). On our ethnicity question 82.1% identified as Caucasian, 5.7% as Asian, 6.0% as Black, 2.5% as Latino, 0.7% as Aboriginal, 0.4% as Caribbean, 0.7% as other, and 2.1% did not specify.

Design. Same as Study 3.

Procedure. Same as Study 3 except that participants could nominate a work or academic project and were told that the deadline could be anywhere from 5 days to 3 weeks away.

Results

We again start by presenting the Part 1 (manipulation, predictions) results for all participants, and then report results for the subset of participants who completed both Parts of the study.

Part 1.

***Pre-manipulation measures*.** Participants' project deadlines did not differ by condition, $F(3, 281) = 2.02$, $p = .111$, $\eta_p^2 = .021$, and were on average 10.02 days ($SD = 3.27$) from the day participants completed Part 1 of the study. Participants did not differ by condition on the following project characteristics, difficulty ($M = 4.65$, $SD = 1.28$), importance ($M = 6.12$, $SD = 1.19$), control ($M = 5.81$, $SD = 1.27$), busyness ($M = 5.21$, $SD = 1.38$), finishing as quickly as possible ($M = 5.65$, $SD = 1.27$), or quality ($M = 6.31$, $SD = 1.08$), all Fs < 2.25, ps > .100. Of the six project characteristics, only finishing as quickly as possible, $r = .15$, $p = .012$, and quality, $r = .11$, $p = .057$, showed significant (or marginal) zero-order correlations with completion predictions (all other rs < .06, ps > .375), so only those two items were included as covariates in the analyses that follow.[6]

[6] The pattern of results for analyses with and without this control variable was the same, with only very small changes in effect sizes.

Results, including means, standard deviations, F-tests, p-values, and effect sizes, for all dependent variables following the manipulation are presented in Table 16. Unadjusted means are presented in the tables, adjusted means are presented in text.

Past project characteristics. In the three experimental conditions, participants listed past projects and when they completed them, either relative to the deadline (Past, RCF deadline) or relative to predictions (RCF prediction).[7] For participants' average past completion times, we again expected that participants would report that they tend to finish relatively close to the deadline (e.g., 1 or 2 days before, i.e., a positive number) and that they tend to finish slightly later than their prediction (e.g., 1 or 2 days after, i.e., a negative number). Holding constant the two control variables (i.e., presenting adjusted means), participants in the RCF prediction condition reported finishing on average 0.77 days (*SE* = 0.22) before their predictions (a positive number), which is different from what we expected and what we saw in the first study. These participants indicated that they do not believe they exhibited the planning fallacy with their past projects, and this number differed significantly from 0, $t(68) = 2.99$, $p = .005$, $d = 0.47$, meaning they believe they tend to finish somewhat earlier than expected[8]. Those in the RCF deadline condition reported finishing on average 1.70 days (*SE* = 0.27) before the deadline, and this did not differ significantly from participants in the Past condition, $t(132) = 1.12$, $p = .267$, who reported finishing on average 1.29 days (*SE* = 0.26) before the deadline.

Predictions. Recall we predicted that participants in the control condition would have the most optimistic predictions (furthest before the deadline), followed by participants in the Past,

[7] Past project completion ratings from the RCF prediction condition were recoded from the way they were presented in the survey, so that positive numbers indicated finishing before the original prediction (earlier) and negative numbers indicated finishing after the original prediction (later).

[8] This result is in line with another study conducted by our lab which found that MTurk workers (American adults), on average, for unspecified projects (as in Study 1), do not believe that they exhibit the planning fallacy.

RCF deadline, and RCF prediction condition which we expected to have the least optimistic predictions (closest to the deadline). The One-way ANCOVA was marginally significant, $F(3, 278) = 2.16$, $p = .093$, $\eta_p^2 = .023$, and planned comparisons revealed that participants made significantly later predictions in the RCF prediction condition ($M_{adj} = 1.04$, $SE = 0.25$) than in the control condition ($M_{adj} = 1.82$, $SE = 0.23$), $t(144) = -2.37$, $p = .019$, $d = -0.42$, consistent with hypothesis 4.

Although the effects of the RCF intervention are in line with our hypothesis, this can be seen as surprising given that participants indicated they tended to finish past projects slightly BEFORE they had predicted, and the instructions guided participants to base their prediction on their average. Accordingly, we might have expected participants not to adjust their predictions or to adjust them to be even earlier, as they were only instructed to adjust their predictions to be somewhat later if their average is finishing slightly LATER than predicted. There are a couple of reasons why participants may have nonetheless adjusted their predictions to be later, 1) they could have (incorrectly) interpreted the negative number (representing their earlier than predicted average completion time) as indicating that they tended to finish later than predicted (although the explanation of that average explicitly stated otherwise), and/or 2) their responses may have been driven by demand characteristics because the example we gave in the instructions was about finishing later than predicted ("If your average is finishing slightly later than predicted, then you adjust your initial prediction to be somewhat later, unless you have reliable information to indicate that this project will be different"). Another possibility is that by this stage in the process participants doubted their recall of when they finished past projects (relative to initial predictions) and came to the realization that they probably do in fact tend to finish projects later than expected and therefore adjusted their predictions to be later even though that stood

somewhat in contradiction to their reports of the past projects. Related to this, yet another possibility is that participants had difficulty thinking about their completion times relative to predictions and instead simply thought of them as relative to the relevant deadlines. Thus, if they saw their past project average as indicating that they tend to finish relatively close to the deadline (as the numbers would have indicated for many participants) then they may have adjusted their prediction to be later (closer to the deadline) for this reason.

Participants in the RCF prediction condition also made an initial, intuitive prediction before making their final prediction. Initial predictions in this condition (M_{adj} = 1.52, SE = 0.25) did not differ significantly from predictions in the control condition (M_{adj} = 1.83, SE = 0.23), $t(144) = -0.90$, $p = .371$, $d = -0.16$. Initial (M = 1.49, SD = 2.15) and final (M = 1.03, SD = 1.96) predictions in the RCF prediction condition did differ significantly in the expected direction, $t(68) = 2.52$, $p = .014$, $d = 0.22$; initial predictions were more optimistic than final predictions, suggesting that people adjusted their predictions as they went through the exercise.

Participants in the RCF deadline condition (M_{adj} = 1.48, SE = 0.26) did not make significantly later predictions than those in the control condition, $t(140) = -0.98$, $p = .327$, $d = -0.16$, but predictions in the RCF past (M_{adj} = 1.14, SE = 0.24) condition did differ significantly from those in the control condition (in the expected direction), $t(148) = -2.05$, $p = .042$, $d = -0.33$. RCF prediction participants did not make significantly later completion predictions than RCF deadline participants, $t(128) = -1.38$, $p = .170$, $d = 0.22$, and made nearly identical predictions as Past participants, $t(136) = -0.13$, $p = .901$, $d = 0.05$.

We also compared predictions and past project completion times in the RCF deadline and Past condition (where they were in the same units, days before the deadline), to examine whether these were the same or differed. Presumably, if participants simply used their recalled past

completion times, these should not differ significantly. In the Past condition, the average past completion time (M = 1.28, SD = 2.06) was 0.16 days (SD = 2.27) closer to the deadline than the completion prediction (M = 1.11, SD = 2.34), which was not a significant difference, $t(71) = 0.62$, $p = .540$, $d = 0.08$. In the RCF deadline condition, the average past completion time (M = 1.71, SD = 2.32) was 0.22 days (SD = 1.68) closer to the deadline than the completion prediction (M = 1.48, SD = 1.98), which was also not a significant difference, $t(63) = 1.05$, $p = .297$, $d = 0.11$. That past completion times and predictions did not differ suggests that participants may have used their past completion times to make their current predictions (see "Correlations" for the relatedness of the variables).

***Exploratory measures*.** All the variables that followed and their analyses were exploratory in nature. Results are summarized in Appendix D1 (see also Table 16).

***Correlations*.** To better understand the extent to which past project completion times may have influenced predictions, we correlated past project completion times and predictions separately within each condition. In the past condition, $r = .47$, $p < .001$, and in the RCF deadline condition, $r = .71$, $p < .001$, participants who reported finishing closer to the deadline in the past (later) also made predictions that were closer to the deadline (later). In the RCF prediction condition, this correlation represents the relationship between the number of days before/after predictions participants recalled finishing in the past, and the number of days before the deadline they anticipate finishing for this project. This correlation was medium to large and positive, $r = .47$, $p < .001$, indicating that participants who finished later (more days after their prediction or fewer days before their prediction) in the past, predicted finishing later (closer to the deadline) for this project, consistent with Study 3.

We also ran zero-order correlations, collapsed across conditions, between completion predictions and our additional measures (see Appendix D2; Table 17).

Part 2. Two days after the project's deadline, eligible participants were contacted via email with a link to our follow-up survey and 198 (69.5%) participated. Participants were reminded of the project description and deadline they gave in Part 1 and asked to specify the deadline for the project. This was done to ensure that participants remembered the deadline for the project correctly, and in case the deadline had changed. In cases where the deadline from Part 1 and 2 mismatched by more than one day, participants were excluded from the Part 2 analyses because the completion time question that followed was expressed in days before the deadline. Twenty-one participants had mismatching deadlines, so 177 Part 2 responses were eligible for analysis. Of these, 25 specified that they did not finish their project (7 in the control, 7 in the past, 6 in the RCF deadline, and 5 in the RCF prediction condition), leaving 152 "actual" completion dates (in days before the deadline) for analysis.

Completion Times and Prediction Bias. Unexpectedly, self-reported completion times differed significantly by condition, $F(3, 146) = 5.55$, $p = .001$, $\eta_p^2 = .103$ (see Table 18). RCF prediction participants ($M_{adj} = 1.87$, $SE = 0.25$) reported finishing their project sooner (more days before the deadline) than participants in the control ($M_{adj} = 0.86$, $SE = 0.22$), $t(79) = -3.09$, $p = .003$, $d = 0.50$, Past ($M_{adj} = 0.77$, $SE = 0.23$), $t(77) = -3.76$, $p < .001$, $d = -0.54$, and RCF deadline ($M_{adj} = 0.69$, $SE = 0.25$) conditions, $t(71) = -3.26$, $p = .002$, $d = -0.57$. It is unclear why our manipulation, or the RCF prediction condition specifically, would influence actual completion times, which we would expect to be driven primarily by external factors such as the size, complexity, and difficulty of the project, how much time participants have available to complete the project, and chronic motivational tendencies, such as eagerness or procrastination. One

potential explanation could be that reflecting in depth about past completion times, especially in reference to one's own expectations (as opposed to external deadlines), motivated participants to finish as soon as possible in order to meet and exceed their expectations. Such a possibility is highly speculative, and we would want to see this result replicated to believe that it is robust. The control, Past, and RCF deadline conditions did not differ significantly from one another (all *p*s > .600).

Next, we calculated prediction bias (difference score) by subtracting "actual" completion times from predicted completion times, representing how many days before or after their prediction participants finished their project. See Figure 2 for a visual comparison of predicted and actual completion times by condition. Notably, within the control condition, participants finished 1.29 days later than predicted, $t(40) = 3.74$, $p = .001$, $d = 0.75$; thus once again the study yielded evidence of the planning fallacy. Collapsed across conditions, participants finished their projects on average, 0.36 days ($SD = 2.17$) later than predicted, $t(151) = 1.84$, $p = .068$, $d = 0.17$.

The ANCOVA performed on this difference score revealed a significant main effect of condition, $F(3, 146) = 4.84$, $p = .003$, $\eta_p^2 = .091$. Bias was lower in each of the intervention conditions than in the control condition ($M_{adj} = 1.28$, $SE = 0.33$) and this difference was significant for the RCF prediction condition ($M_{adj} = -0.47$, $SE = 0.34$), $t(79) = -3.85$, $p < .001$, $d = -0.82$, Past condition ($M_{adj} = 0.12$, $SE = 0.34$), $t(79) = -2.40$, $p = .019$, $d = 0.54$, but not for the RCF deadline condition ($M_{adj} = 0.48$, $SE = 0.37$), $t(73) = -1.53$, $p = .132$, $d = 0.37$. Prediction bias was not significantly lower for RCF prediction than Past participants, $t(77) = -1.23$, $p = .223$, $d = -0.28$, and marginally lower for RCF prediction than RCF deadline participants, $t(71) = -1.94$, $p = .057$, $d = -0.44$.

We again calculated a binary measure of prediction bias, dividing participants into those who finished by the time they predicted and those who finished later than predicted. Results for this variable were similar to what was observed for the continuous measure. The Chi-squared test for the overall effect of condition was significant, $X^2(3) = 9.23$, $p = .026$, $V = .229$. Planned comparisons indicated that participants were more likely to finish by the predicted time in the RCF prediction (74.4%) than in the control condition (43.8%), $X^2(1) = 8.77$, $p = .003$, $V = .310$. The control condition did not differ from the Past condition (54.3%), $X^2(1) = 1.06$, $p = .304$, $V = .106$, or the RCF deadline condition (51.3%), $X^2(1) = 0.49$, $p = .484$, $V = .075$. RCF prediction participants finished on time more often than Past participants, $X^2(1) = 3.89$, $p = .049$, $V = .209$, and RCF deadline participants, $X^2(1) = 4.72$, $p = .030$, $V = .240$.

As secondary analyses we also re-tested our key Part 1 measures with the Part 2 sample, compared average past completion times with completion times for the current project, tested for working time prediction bias, and reasons for finishing later than predicted (see Appendix D3; Table 18).

Correlations. To examine the relationships between past project completion times, predictions, and actual completion times we ran zero-order correlations broken down by condition (see Table 19). Past project completion times and predictions were positively correlated, meaning that participants who indicated that they tended to finish closer to the deadline also made a prediction that was closer to the deadline. Interestingly, this correlation was stronger in the Past condition ($r = .80$) than in the RCF deadline condition ($r = .48$), perhaps suggesting that participants based predictions more on their past in the Past condition than in the RCF deadline condition. In the RCF prediction condition, past project completion times were only marginally correlated with predictions ($r = .23$) which is not surprising given that past

completion times are relative to predictions and predictions are relative to the deadline. Predictions and actual completion times were significantly positively correlated in the Past condition (r = .37), marginally positively correlated in the RCF prediction condition (r = .36), and not significantly correlated in the RCF deadline (r = .25) and control (r = .13) conditions. This pattern of correlations is roughly in line with the results for mean bias across conditions: RCF prediction and Past participants showed the lowest levels of optimistic bias and the highest degree of correlational accuracy. Past completion times and "actual" reported completion times were not correlated.

Correlations between our additional variables of interest and Part 2 variables are summarized in Appendix D2 (see also Table 20).

Discussion

Using a sample with greater age variability (MTurkers), project variability (work and academic projects), and deadline variability (deadlines up to three weeks away), we again found that reference class forecasting led to less optimistic and less biased completion predictions than a control condition (hypothesis 4), and that this was particularly true when the procedure was based on recalling past experiences in relation to predictions (hypothesis 5).

However, results were not entirely consistent with those of Study 3. In Study 4, all the intervention groups, regardless of whether past completion times were in reference to deadlines (RCF deadline) or predictions (RCF prediction), made less optimistic completion predictions than the control group (whereas in Study 3 it was only the RCF prediction condition that differed from the control group). In addition, a surprising result emerged: participants in the RCF prediction condition reported that they tended to finish past projects before their predictions, indicating that they did not believe they exhibit the planning fallacy (contrary to hypothesis 1), and yet these participants made later completion predictions (compared to the control group and

compared to their initial, intuitive predictions) as a result of the RCF exercise. Also unexpected (and unlike in Study 3), participants who completed the RCF prediction exercise reported finishing their current projects sooner (more days before the deadline) than all other participants. Finally, participants in two of the three intervention groups (RCF prediction and Past) were less optimistically biased than control participants, consistent with hypothesis 4. We did not find support for hypothesis 5 – RCF was not more effective in reducing bias when past project completion times were recalled relative to predictions vs. deadlines, unlike in Study 3.

Study 5

Summarizing our results so far, we found evidence in three of our four studies that people believe they exhibit the planning fallacy at least slightly (hypothesis 1). In Study 1 and 2 we found little evidence that simply reminding people of their overall past completion times leads them to make less optimistic completion predictions (hypothesis 2), including people who recall finishing later than predicted (hypothesis 3). When we instructed participants to incorporate past completion times into their predictions more explicitly using the more structured RCF method in Study 3 and 4, we found that they indeed made somewhat later completion predictions (hypothesis 4). Study 3 also suggested that recalling past completion times in reference to predictions as opposed to deadlines results in less optimistic completion predictions (hypothesis 5), but this result was not replicated in Study 4.

Given the somewhat mixed evidence that people believe in the planning fallacy and tentative evidence that the RCF procedure leads to later, less biased completion predictions, we decided to run a fifth study examining the RCF intervention. The main purpose of the study was to test the effect of RCF on completion predictions and optimistic bias. We also sought to explore further the processes underlying the effects of the intervention. In particular, the study provided another test of students' beliefs about the planning fallacy (in studies so far students

have reported stronger planning fallacy beliefs than Mturk samples), and allowed us to test whether RCF is more effective to the extent that people recall completing past projects later than predicted.

The study was also designed to explore which specific components of our RCF procedure are particularly important (i.e., the active ingredients) for leading to later completion predictions. To do so, we focused on the most promising RCF group thus far (i.e., the RCF prediction condition from Study 3 and 4, where participants recalled past completion times in reference to predictions). One feature of that condition is that forecasters are presented with a calculated mean that summarizes their reports of past completion times along with explicit instructions for how to use it (i.e., use the mean unless you have better information available). Is it important to provide the calculated mean and instruction to use it, or would the intervention be equally effective if participants were simply prompted to list when similar past projects were finished relative to prediction? To answer this question, we varied whether participants were provided with the mean and instructions to use it. Another noteworthy feature of the intervention is that participants are asked to make an initial and then a final prediction. Conceivably, this in itself is enough to prompt later completion predictions. To test this possibility, we included an additional control condition where participants made an initial and then a final prediction without any other RCF procedures occurring in between. In sum, we created prediction conditions designed to isolate the key active ingredient in the RCF prediction procedure.

Method

Participants. Of the 437 undergraduate students from Wilfrid Laurier University who participated in the study on-line for course credit 423 participants completed the survey up to the key variable of interest (i.e., completion prediction). Of these, 45 did not nominate a project as per the instructions (15 nominated a project with a deadline that was less than 4 days away; 6

failed to nominate a project; 1 identified a project that could only be completed on a specific date in a single session), and were thus excluded from the analyses, leaving a total of 401 participants. The final sample consisted of 298 female, 98 male, and 5 unspecified participants with a mean age of 19.67 (*SD* = 2.57). On our ethnicity question, 64.3% identified as Caucasian, 16.7% as Asian, 2.2% as Caribbean, 1.7% as Latino, 3.2% as Middle Eastern, 3.0% as Black, 1.2% as Aboriginal, 6.0% as other, and 1.5% did not specify.

Design. The study had primarily a single-factor design where participants were randomly assigned to one of six prediction conditions: two control conditions and four RCF conditions (see "Manipulation" section).

Procedure.

Project nomination (Part 1). All participants were instructed to nominate and briefly describe an academic project with a hard deadline in the next three weeks, that they could complete any time between now and the deadline (see Appendix E1 for details).

Manipulation and predictions (Part 1). Next, participants encountered one of two control conditions,

Condition 1 (control group 1) = simply make a prediction (see Appendix E2),

Condition 2 (control group 2) = make an initial prediction and then make a final prediction (see Appendix E3),

or one of the four reference class forecasting conditions, which had identical past project nomination and completion prediction instructions (see Appendix E4), but otherwise differ as follows:

Condition 3 (RCF group 1) = make an initial prediction, think of similar past projects and when you finished them relative to when you initially expected, receive the mean for

your past projects and instructions for how to use this mean to make your final prediction (all RCF prediction components; see Appendix E5),

Condition 4 (RCF group 2) = make an initial prediction, think of similar past projects and when you finished them relative to when you initially expected, make a final prediction using the information from the past projects (no mean, no instructions) (Appendix E6),

Condition 5 (RCF group 3) = think of similar past projects and when you finished them relative to when you initially expected, receive the mean for your past projects and instructions for how to use this mean to make your final prediction (no initial prediction) (Appendix E7),

Condition 6 (RCF group 4) = think of similar past projects and when you finished them relative to when you initially expected, make a final prediction using the information from the past projects (no initial prediction; no mean, no instructions) (Appendix E8).

Table 21 summarizes the components of the prediction intervention condition included in each condition.

It is also worth noting that, the four reference class forecasting conditions additionally fell into a 2 Initial Prediction (yes vs. no) x 2 Past Means and Instructions (yes vs. no) factorial design, which would allow us to test the effects of each factor separately and also possible interaction effects. Participants in the two control conditions, who did not recall past projects and their completion times prior to prediction, were asked to nominate similar past projects and completion times using the same materials as the RCF groups after making their (final) predictions, allowing us to compare recalled past project completion times between the RCF and control groups (see Appendix E9). All completion predictions were made in days before the deadline.

Other changes between Studies 3, 4 and Study 5. Although our RCF procedure for Study 5 mimicked those of Study 3 and 4 (RCF prediction conditions) quite closely, a couple of additional small changes were made across the relevant groups that are worth mentioning. One, we switched from sliders to dropdown menus for the recollection of past completion times. This was done to highlight the distinction and avoid any potential confusion between the past completion times which are relative to predictions (where 0 = finishing same day as predicted) and completion predictions which are relative to deadlines (where 0 = day of the deadline) (Appendix C5 vs. Appendix E4). Two, we changed how participants received information about their mean past completion times: in the earlier studies we simply piped in the mean completion time which was either positive or negative, and then indicated that positive numbers meant days AFTER their prediction and negative numbers meant days BEFORE their prediction (see Appendix C8). Although this may at first seem counterintuitive, participants receive only one mean and we assumed that most people would indicate having finished slightly after their predictions, so coding past completion times that were after predictions as positive would mean that most of the numbers presented would be positive and it would be most natural to read number of days after predicted as a positive number (e.g., "On average you finish 2 days after your prediction…"). The presentation of negative numbers could be confusing if participant didn't fully read our instructions [e.g., "On average you finish -2 days after your prediction (if the number is positive)/before your prediction (if the number is negative)]. Thus in Study 5 we used Qualtrics coding and branching to present participants with different instructions depending on whether their mean past completion times were later, earlier, or the same as predicted, so that participants wouldn't have to interpret the (sign of) the mean on their own (see Appendix E5 and

E7 for examples). We hoped that these changes would decrease the likelihood of participants misunderstanding our instructions and increase response accuracy.

Additional measures (Part 1). As in Study 3 and 4 participants were asked to write a few sentences explaining how they arrived at their prediction, predicted how many hours they would spend working on the project, rated the extent to which they based their predictions on five factors, and completed propensity to plan and conscientiousness scales.

Part 2. Two days after their project's deadline, eligible participants were contacted via email with a link to our follow-up survey (see Appendix E10). Participants were reminded of the project description and deadline they gave in Part 1 and asked to specify if the deadline for the project had changed. Those whose deadline had changed were excluded from the Part 2 analyses (as their predictions were for a different deadline than their actual completion times). Next, we asked if they finished their project, when they finished it (days before the deadline), and how many hours they spent working on it (same as Study 3, 4).

Finally, participants indicated their prior knowledge of the planning fallacy using the same yes-no question as in Study 1. 131 (54.6%) answered *no*, 109 (45.4%) answered *yes*, indicating that a significant proportion of the participants had at least some previous knowledge of the planning fallacy phenomenon, but these numbers did not differ by condition, $X^2(5) = 4.98$, $p = .418$, $V = .144$. In addition, predictions, completion times, and bias (prediction – actual completion time) did not differ depending on whether people were previously aware of the planning fallacy or not, all ts < 1.60, ps > .120, ds < 0.26.

Results

We again start by presenting the Part 1 results for all participants, and then report results for the subset of participants who completed both Part 1 and 2.

Part 1. Due to an error in our Qualtrics programming, predictions were not recorded in two of our experimental groups (RCF 1, 3) including the condition that contained the full RCF intervention used in previous studies (RCF 1), so only 268 completion predictions were available for analysis. Thus, for the main dependent variables (predictions, prediction bias) analyses are conducted to compare the four remaining conditions, including RCF condition 2 (initial prediction, no mean and instructions) and 4 (no initial prediction, no mean and instructions). Means, standard deviations, F-tests, p-values, and effect sizes, for all dependent variables can be found in Table 22.

***Pre-manipulation measures*.** Participants' project deadlines did not differ by condition, $F(3, 395) = 0.71$, $p = .616$, $\eta_p^2 = .009$, and were on average 12.59 days ($SD = 5.81$) from the day participants completed Part 1 of the study. Deadlines and predictions were not significantly correlated, $r = .10$, $p = .082$.

***Past project characteristics*.** Participants listed past projects and when they completed them (relative to predictions) either before (RCF conditions) or after (control conditions) making predictions. For participants' average past completion times, we again expected that they would report that they tend to finish slightly later than predicted (e.g., 1 day after predicted). Contrary to our expectations, and in line with the results of Study 4, participants reported finishing on average 0.27 days (SD = 1.47) before their predictions (a negative number), indicating that on average, they did not believe they exhibited the planning fallacy. This number differed significantly from 0, $t(400) = 3.63$, $p < .001$, $d = -0.18$. Interestingly, the one-way ANOVA was significant, $F(5, 395) = 11.74$, $p < .001$, $\eta_p^2 = .129$, indicating that the past completion time means differed by condition. Examination of the means showed that participants in the two control groups indicated finishing earliest, $M_{Control1} = -0.89$ ($SD = 1.43$), $M_{Control2} = -1.05$ ($SD =$

1.36), and those in RCF group 3 and 4 indicated finishing latest (albeit still less than half a day after their predictions), M_{RCF4} = 0.40 (*SD* = 1.75), M_{RCF5} = 0.23 (*SD* = 1.21), and RCF groups 1 and 2 fell in between, M_{RCF1} = -0.13 (*SD* = 1.30), M_{RCF2} = -0.15 (*SD* = 1.23). As we had no a priori hypotheses, we adjusted for multiple comparisons (α = .05/15 comparisons = .003) for our post-hoc comparisons, and found that the control groups differed significantly from all of the RCF groups (all *p*s < .003), whereas the RCF groups did not differ significantly using the *p* < .003 cut-off (see Table 23). A potential explanation for these results is that people's perceptions of past completion times are influenced by making predictions. An alternative explanation is demand characteristics – that people in the RCF groups felt that they should report later past project completion times because of the instructions which stated that the purpose of recalling the past was to help them make an accurate completion prediction.

Predictions. The one-way ANOVA based on the four conditions did not reveal a significant omnibus effect of the manipulation on completion predictions, $F(3, 264) = 1.53$, $p = .207$, $\eta_p^2 = .017$. Because we had a priori hypotheses we nonetheless conducted the relevant planned comparisons, which revealed that RCF 2 participants, *M* = 1.27 (SD = 1.38), made significantly later predictions than control 1 participants, *M* = 1.83 (*SD* = 1.67), $t(134) = -2.15$, $p = .034$, $d = -0.36$. Those in the RCF 4 condition, *M* = 1.38 (*SD* = 1.66), did not quite differ significantly from those in control condition 1, $t(128) = -1.53$, $p = .129$, $d = -0.27$. The mean for control condition 2 (with initial and final prediction), *M* = 1.51 (*SD* = 1.70) fell in between that of control group 1 and the RCF groups, and did not differ from either, $t(128)_{Control1} = -1.53$, $p = .129$, $d = -0.15$, $t(128)_{RCF2} = -1.53$, $p = .129$, $d = 0.16$. The two RCF groups also did not differ significantly, $t(134) = -0.45$, $p = .645$, $d = -0.14$.

Because condition 1 and condition 3 predictions were missing, we could not run the 2 by 2 ANOVA testing for effects of initial predictions, means and instructions, and their interaction. We did conduct a more general control vs. RCF groups test, pooling RCF predictions into two groups. An independent samples t-test revealed that predictions were marginally later in the RCF groups (M = 1.32, SD = 1.51) than in the control groups (M = 1.68, SD = 1.69), $t(266)$ = 1.75, p = .081, d = 0.22, suggesting that RCF condition participants made slightly later completion predictions than those in the control groups.

Participants in the Control 2 (M = 1.82, SD = 1.87) and the RCF 3 (M = 1.86, SD = 1.84) condition made an initial intuitive prediction before making their final prediction, and these initial predictions did not differ between the two groups, $t(136)$ = -0.12, p = .904, d = -0.02, or from predictions in the control 1 condition (M = 1.83, SD = 1.67), $t(130)_{\text{Control1}}$ = 0.03, p = .975, d = -0.01, $t(134)_{\text{RCF3}}$ = -0.09, p = .925, d = -0.02, respectively. Regardless of whether participants simply made two predictions or went through the RCF procedure, final predictions were later than initial predictions, Control 2 M_{Diff} = 0.31, SD = 0.86, $t(66)$ = 3.00, p = .004, d = 0.17, RCF 3 M_{Diff} = 0.59, SD = 1.49, $t(70)$ = 3.35, p = .001, d = 0.32. Although the difference looks somewhat bigger in the RCF 3 condition than in the Control 2 condition, this difference between the two was not significant, M_{Diff} = -0.28, SE = 0.21, $t(112.97)$ = -1.35, p = .178, d = -0.23.

Moderation and mediation tests. Next, we tested whether participants were especially likely to make later completion predictions if they reported later past completion times (hypothesis 6). We did not find support for this moderation hypothesis – all confidence intervals for the relevant condition by past completion time interaction terms contained 0 (see Table 24).

Instead, for significant (control 1 vs. RCF 2 condition) or marginal (control vs. RCF conditions combined) main effects of condition, we observe a substantial reduction in those effects when past completion times were included in the regression, which could suggest that the effect of the RCF procedure on predictions is mediated by recalling different past completion times. In other words, this interpretation would mean that RCF procedure led participants to recall later past completion times (than not going through an RCF procedure) which in turn led them to make later predictions. Alternatively, it could instead be the case that having in mind a later prediction leads to the recall of later past completion times. Simple mediation analyses (10,000 bootstrap samples, Model 4; Hayes, 2012) supported both possibilities, that RCF 2 led to later completion predictions than the control through past completion times [95% CI -.531, .067], and that RCF 2 led to later past completion times through predictions [95% CI .010, .315] (see Figure 3). A mediation model testing the combined control condition against the combined RCF conditions also yielded evidence that the RCF groups led to later predictions through past completion times [95% CI -.600, -.171], but the reverse mediation model was not significant [95% CI -.009, .205] (see Figure 4). Keep in mind, that because participants were asked to report past completion times later in the control group than in the RCF groups (after vs. before predictions), the possibility exists that differences in past completion times could be driven by an order effect as opposed to the RCF procedure.

***Exploratory measures*.** The results for all the exploratory variables that followed are summarized in Appendix E11 (see also Table 22).

Correlations. As in Studies 3 and 4, past completion times and predictions were generally negatively correlated, meaning that for the most part, the earlier participants recalled finishing past projects, the earlier (more optimistic) their predictions: Control 1 r = -.31, p =.013;

Control 2 r = -.44, p < .001; RCF 2 r = -.37, p =.002; and RCF 4 r = -.12, p =.330. The correlation wasn't higher in the RCF groups than the control groups.

We also ran zero-order correlations, collapsed across conditions, between completion predictions and our additional measures (see Appendix E12; Table 25).

Part 2. Two days after their project's deadline, eligible participants were contacted via email with a link to our follow-up survey and 279 (69.6%) completed the survey. Participants were reminded of the project description and deadline they gave in Part 1 and asked to specify if the deadline for the project had changed. Those whose deadline had changed, n = 39, were excluded from the Part 2 analyses (as their predictions were for a different deadline than their actual completion times). This left 240 participants and 13 of these indicated that they did not finish their project (2 in Condition 1, 0 in Condition 2, 4 in Condition 3, 1 in Condition 4, 2 in Condition 5, and 4 in Condition 6), and were thus also not included in the main analyses (as their completion date was unknown).

Completion Times and Prediction Bias. As expected, self-reported completion times (M_{Overall} = -0.37, SD = 1.32) did not differ by condition, $F(5, 218) = 0.25$, p = .942, η_p^2 = .006. Next, we calculated prediction bias (difference score) by subtracting "actual" completion times from predicted completion times (see Table 26). Participants in control group 1 finished on average 0.73 days (SD = 1.66) later than predicted, $t(39) = 2.76$, p = .009, d = -0.44; thus once again the study yielded evidence of the planning fallacy. Collapsed across conditions, participants finished their projects on average, 0.44 days (SD = 1.53) later than predicted, $t(147)$ = -3.50, p = .001, d = -0.29.

The ANOVA performed on this difference score (prediction bias) did not yield a significant omnibus effect, $F(5, 218) = 1.58$, p = .197, η_p^2 = .032. Because we had a priori

hypotheses, we nonetheless conducted planned comparisons: prediction bias was marginally lower in the RCF 4 (M = -0.07, SD = 1.75) than in the Control 1 (M = 0.73, SD = 1.66) condition, $t(67) = 1.91$, $p < .060$, d =0.47, but was not lower in the Control 1 than the RCF 2 (M = 0.46, SD = 1.45) condition, $t(79) = 0.76$, $p = .453$, d =0.17. The Control 2 (M = 0.46, SD = 1.45) condition did not differ from the other three groups on prediction bias (all ps > .100), and the two RCF groups did not differ ($p = .170$). See also Figure 5 for a visual comparison of predicted and actual completion times by condition.

For the binary measure of prediction bias, participants were divided into those who 1) finished by the time they predicted and, 2) finished later than predicted. The Chi-squared test for the overall effect of condition was not significant, $X^2(3) = 0.88$, $p = .830$, $V = .077$. Planned comparisons indicated that participants were no more likely to finish by the predicted time in the RCF 2 (56.1%) than in the control 1 condition (55.0%), $X^2(1) = 0.01$, $p = .921$, $V = .011$, or in the RCF 4 (65.5%) than in the control 1 condition, $X^2(1) = 0.77$, $p = .380$, $V = .106$. Those in the RCF conditions were also no more likely than those in control condition 2 (57.9%) to finish by the predicted time (ps > .500).

Thus, given that only one of the RCF groups differed marginally from one of the control groups, and no other significant differences emerged, we did not find substantial evidence for RCF reducing optimistic bias, at least not for the RCF conditions that we could examine in this study.

As secondary analyses we also re-tested our key Part 1 measures with the Part 2 sample, compared average past completion times with completion times for the current project, and tested for working time prediction bias. (see Appendix E13; Table 26).

Correlations. To examine the relationship between past project completion times, predictions, and actual completion times we ran zero-order correlations broken down by condition (see Table 27). Past project completion times and predictions were significantly correlated in control condition 1, r = -.46, and control condition 2, r = -.54, meaning that participants who made a prediction that was closer to the deadline also indicated that they tended to finish closer to their predictions in the past. Unlike in Part 1, past completion times and predictions were not significantly related in the RCF 2, r = -.01, or RCF 4, r = -.29, condition in Part 2. Predictions and actual completion times were significantly positively related in all groups (r = .38 to .66), indicating that people who predicted finishing earlier (more days before predicted) also reported finishing earlier (more days before the deadline).

Discussion

Although Study 5 was limited by an error with our survey software programming, it nonetheless allowed for additional tests of our hypotheses. As in Study 4, and in contrast to Study 1, 2, and 3, on average, the students in this study indicated that they did not believe they exhibited the planning fallacy in their past projects (hypothesis 1).

We were also able to test whether some variations of an RCF intervention – albeit those that did not include all of the components used previously – had an impact on prediction and prediction bias. We found only weak, inconsistent evidence that these particular variants of the RCF intervention led to later, less optimistically biased completion predictions (hypothesis 4), as only one of the RCF groups (RCF 2: initial, final prediction; no mean or instructions) made predictions that differed significantly in the expected direction from those in the main control group, and only for the Part 1 sample. The other RCF group (4: no initial, final prediction, no mean or instructions) did not evidence later completion predictions than the main control group in Part 1 but had marginally lower optimistic prediction bias in Part 2. The RCF groups did not

differ significantly from the control group where participants simply made an initial and final prediction without instructions, and this control group also didn't differ from the main control group. We also did not find evidence for our moderation hypothesis (6) – the RCF procedure was not more effective to the extent that people recalled completing past projects later than predicted.

In this study we also asked control condition participants to give ratings of their past completion times (using the same procedure as in the RCF groups, but after their predictions) and these were earlier than those in RCF groups 3 and 4 groups, which did not make initial completion predictions. RCF groups 1 and 2, which made initial predictions before giving ratings of similar past project completion times, fell in between. These results could suggest that a) making a completion prediction for a current project changes recall of past project completion times, and/or b) that recalling past project completion times with the purpose of making a more accurate completion predictions leads to the recall of less optimistic past project completion times (which in turn leads to less optimistic predictions).

The fact that correlations between past and predictions, and predictions and actual completion times were not greater in the RCF groups than in the control groups seems to suggest that the RCF intervention may not be leading people to make more accurate predictions at a correlational level. Instead, when RCF leads to later completion predictions, and to reduced bias at the mean level, it may be that the procedure prompts people to make crude adjustments in an attempt to account for bias, adjustments that are not sensitive to individual variation in past completion times.

Study 6

Study 6 used the same design as Study 5, with two control and four RCF experimental groups, but a different sample, adult Americans (MTurkers), as opposed to students. We were

unable to test all of our Study 5 hypotheses due to losing predictions in two of our RCF conditions, so we ran the study again.

Method

Participants. Of the 511 American adults that participated in the study via mturk.com (compensated $1.00 USD for Part 1 and $1.00 USD for Part 2 participation) 19 did not nominate a project as per the instructions (1 nominated a project with a deadline that was less than 4 days away; 11 failed to nominate a project; 4 identified a project that could only be completed on a specific date in a single session) and 11 did not generate any past projects, and were thus excluded from the analyses, leaving a total of 484 participants in the final Part 1 sample. The final sample consisted of 263 female, 215 male, 1 other, and 5 unspecified participants with a mean age of 37.47 (*SD* = 12.05). On our ethnicity question, 76.7% identified as Caucasian, 8.7% as Asian, 6.4% as Black, 3.1% as Latino, 0.6% as Caribbean, 0.6% as Middle Eastern, 0.2% as Aboriginal, 2.7% as other, and 1.0% did not specify.

Design. The study design and procedure were identical to that of Study 5 with a few minor exceptions (for verbatim instructions see Appendix F1-9). Participants could nominate a work, personal, or academic project (whereas an academic project was the only option in Study 5). We also altered the instructions for recalling past completion times to address a potential problem. In Study 3 and 4, all conditions were presented with two examples of predictions as part of the instructions for using the past completion time scale (experimental groups) or the prediction scale (control group), and the examples were finishing 6 days before and 3 days after the deadline (or prediction in RCF prediction condition) (see Appendix C2, C4). These examples were originally chosen because they seemed like relatively unlikely, but not impossible, completion times for student projects, and would thus be unlikely to provide a cue about average

project completion times for participants to gravitate toward. On the other hand, the example of 6 days early is quite early and might act as an anchor that could lead participants to report early past completion times and predictions. Therefore, in Study 5, when we switched to dropdown menus for recalling past completion times, we changed the example to a symmetric 3 days before and 3 days after the deadline. Unfortunately, we neglected to make this change in the control conditions. We fixed this in Study 6 so that the examples were the same (3 days before and 3 days after) in all conditions. The measures of factors perceived to influence completion predictions were not included in this study.

At the end of the Part 2 survey (see Appendix F10), participants indicated their prior knowledge of the planning fallacy using the same yes-no question as in Study 1. 266 (55.0%) answered *no*, 104 (21.5%) answered *yes*, and 114 failed to reply (23.6%), suggesting that the majority of participants did not have previous knowledge of the planning fallacy phenomenon. These numbers did differ by condition, $X^2(10) = 23.78$, $p = .008$, $V = .157$. The pattern of counts indicated that this difference was driven by a greater proportion of participants in the control groups (Control 1 = 23.1%, Control 2 = 23.1%) indicating that they had previous knowledge of the planning fallacy than in the RCF groups (RCF 1 = 21.1%, RCF 2 = 15.0%, RCF 3 = 19.9%, RCF 4 = 17.3%). Predictions, completion times, and prediction bias did not differ depending on whether people were previously aware of the planning fallacy or not, all *t*s < 1.40, *p*s > .180.

Results

Part 1. Results, including means, standard deviations, F-tests, p-values, and effect sizes for all dependent variables can be found in Table 29.

***Pre-manipulation measures*.** Participants' project deadlines did not differ by condition, $F(5, 478) = 1.51$, $p = .185$, $\eta_p^2 = .016$, and were on average 13.47 days ($SD = 4.37$) from the day

participants completed Part 1 of the study. Deadlines and predictions were not significantly correlated, $r = 0.03$, $p = .589$.

Past project characteristics. Participants listed past projects and when they completed them (relative to predictions) either before (RCF conditions) or after (control conditions) making predictions. We again expected that participants would report they tend to finish slightly later than predicted. This time, unlike the previous study, participants did report finishing on average 0.25 days (SD = 2.08) after their predictions, and this average differed significantly from 0, $t(482) = 2.61$, $p = .001$, $d = 0.12$, indicating that participants believed they exhibited a small amount of planning fallacy. The ANOVA performed on the mean past completion time was significant, $F(5, 477) = 2.70$, $p = .020$, $\eta_p^2 = .028$, indicating that reports of past completion time differed by condition. As in Study 5, participants in the two control groups ($M = -0.23$, $SD = 1.73$) reported finishing earlier than those in RCF groups ($M = 0.49$, $SD = 2.20$), see Table 30 for t-tests.

Predictions. The one-way ANOVA performed on completion time predictions revealed a significant omnibus effect, $F(5, 477) = 4.85$, $p < .001$, $\eta_p^2 = .048$. Consistent with hypothesis 4, planned comparisons showed that participants in all RCF groups (combined $M = 0.59$, $SD = 2.28$; see Table 29 for means) made significantly later predictions than those in control condition 1 ($M = 1.76$, $SD = 2.03$), all ts > 2.20, all ps < .030 (see Table 31 for means and t-test statistics). RCF 1, 3, and 4 also differed significantly from control group 2 ($M = 1.31$, $SD = 1.69$), all ts > 2.00, all ps < .050, but RCF 2 ($M = 0.83$, $SD = 2.33$) only differed marginally from control group 2, $t(157) = -1.69$, $p = .095$. The four RCF groups did not differ significantly from one another (all ps > .700).

The 2 (Initial Prediction) by 2 (Past Means and Instructions) ANOVA yielded no main effect of initial predictions, $F(1, 318) = 0.34$, $p = .563$, $\eta_p^2 = .001$, no main effect of past mean and instructions, $F(1, 318) = 0.04$, $p = .851$, $\eta_p^2 < .001$, nor a significant interaction, $F(1, 318) = 1.25$, $p = .264$, $\eta_p^2 = .004$. These results indicate that all of the RCF conditions were similarly effective for facilitating later completion predictions.

Participants in the Control 2 ($M = 1.34$, $SD = 1.83$), RCF 1 ($M = 1.67$, $SD = 2.08$), and RCF 2 ($M = 1.38$, $SD = 2.53$) condition made an initial intuitive prediction before making their final prediction, and these did not differ from each other, $F(2, 239) = 0.56$, $p = .573$, $\eta_p^2 = .005$, or from predictions in the control group (all ps > .200). Unlike in Study 5, initial ($M = 1.34$, $SD = 1.83$) and final ($M = 1.31$ $SD = 1.69$) predictions did not differ in control group 2, $t(79) = 0.21$, $p = .836$, $d = 0.01$. Participants in both RCF groups with initial predictions made later final predictions than initial predictions, RCF 1 $M_{Diff} = 1.17$, $SD = 2.16$, $t(83) = 4.95$, $p < .001$, $d = 0.52$, RCF 2 $M_{Diff} = 0.55$, $SD = 1.75$, $t(77) = 3.35$, $p = .007$, $d = 0.22$, indicating that these participants adjusted their predictions to be later following the RCF procedure.

Moderation and mediation tests. Next, we tested whether participants were especially likely to make later completion predictions if they reported later past completion times (hypothesis 6). In contrast to Study 5, we found some support for this hypothesis. The condition by past completion time interaction terms were significant for the RCF 1 vs. control group 1 (and 2) comparison, and marginally significant for the RCF 2 vs. control group 1 (and 2) comparison, but not for any of the other groups (see Table 32). The interaction pattern was the same in each case – in the RCF groups later recalled past completion times were more predictive of later completion predictions (steeper slope) than in the control groups (see Figure 6 and 7).

As in Study 5, we also observed a substantial reduction in the main effect of condition on predictions when past completion times were included in the regressions (see Table 32), which could suggest that the effect of the RCF procedure on predictions is mediated by past completion times, or that the effect of RCF past completion times is mediated by predictions. Simple mediation analyses (10,000 bootstrap samples, Model 4; Hayes, 2012) supported both possibilities (see Table 33). The RCF conditions, with the exception of RCF 4, led to later completion predictions than the control through past completion times [RCF vs. Control groups combined 95% CI -.3953, -.1066], and the control groups led to earlier past completion times through predictions [RCF vs. Control groups combined 95% CI .1412, .4649].

Correlations. As in the previous three studies, past completion times and predictions were generally negatively correlated, meaning that for the most part, the earlier participants recalled finishing past projects, the earlier (more optimistic) their predictions: Control 1 r = -.16, p =.144; Control 2 r = -.27, p = .017; RCF 1 r = -.57, p < .001; RCF 2 r = -.49, p <.001, RCF 3 r = -.19, p = .087, and RCF 4 r = -.08, p < .458. The correlation was strongest in the RCF 1 and 2 groups (RCF with initial predictions).

Part 2. Two days after their project's deadline, eligible participants were contacted via email with a link to our follow-up survey and 372 (72.8%) completed the survey. Participants were reminded of the project description and deadline they gave in Part 1 and asked to specify if the deadline for the project had changed. Those whose deadline had changed, n = 84, were excluded from the Part 2 analyses (as their predictions were for a different deadline than their actual completion times). This left 288 participants and 15 of these indicated that they did not finish their project (2 in Condition 1, 2 in Condition 2, 6 in Condition 3, 1 in Condition 4, 2 in

Condition 5, and 2 in Condition 6), and were thus also not included in the main analyses (as their completion date was unknown), leaving a final sample size of 273.

Completion Times and Prediction Bias. Self-reported completion times ($M_{Overall}$ = 1.59, *SD* = 1.93) did not differ by condition, $F(5, 267) = 1.26$, $p = .282$, $\eta_p^2 = .023$ (see also Table 34).

Participants in Control group 1 finished on average 0.67 days (*SD* = 2.40) later than predicted, $t(45) = 1.90$, $p = .064$, $d = 0.29$; thus yielding marginal evidence of prediction bias. Collapsed across conditions, participants finished their projects on average, 0.14 days (*SD* = 2.78) earlier than predicted, $t(272) = -1.04$, $p = .301$, $d = -0.07$.

The ANOVA for the prediction-actual difference score (i.e., prediction bias) was marginally significant, $F(5, 267) = 2.09$, $p = .068$, $\eta_p^2 = .038$. Post-hoc comparisons revealed that prediction bias was significantly lower in each of the RCF groups than in control group 1 (all *p*s < .050). Control group 2 (*M* = 0.19, *SD* = 2.13) did not differ from the RCF groups on prediction bias (all *p*s > .100), and the RCF groups did not differ from each other (all *p*s > .100). The combined RCF (*M* = -0.42, *SD* = 2.23) and control (*M* = 0.44, *SD* = 2.28) groups differed significantly, $t(272) = 2.97$, $p = .003$, $d = -0.38$, indicating that the RCF procedure led to less optimistically biased completion predictions. See Figure 8 for a visual comparison of predicted and actual completion times by condition.

For the binary measure of prediction bias, the Chi-squared test for the overall effect of condition was not significant, $X^2(5) = 5.25$, $p = .386$, $V = .139$. Given our a priori hypotheses, we nonetheless conducted planned comparisons which indicated that participants were significantly more likely to finish early or on time in the RCF 3 (78.4%) condition than in the control condition 1 (60.0%), $X^2(1) = 3.85$, $p = .050$, $V = .200$, but none of the other RCF conditions differed from control group 1 (all *p*s > .150), and none of the RCF groups differed from control

group 2 (all *p*s > .090). Overall, the difference between the combined RCF (72.0%) and control groups (61.4%) was marginally significant, $X^2(1) = 3.10$, $p = .078$, $V = .107$.

We conducted secondary analyses that re-tested our key Part 1 measures with the Part 2 sample, compared average past completion times with completion times for the current project, and tested for working time prediction bias. (see Appendix F11; Table 29, 34).

Correlations. To examine the relationship between past project completion times, predictions, and actual completion times we ran zero-order correlations broken down by condition (see Table 35). Past project completion times and predictions were significantly correlated only in control condition 2, $r = -.32$, RCF 3 $r = -.54$, and RCF 2 $r = -.64$, meaning that for participants in these groups (all of which had initial predictions) later predictions were associated with later past completion times. Predictions and actual completion times were significantly positively related in all groups ($r = .33$ to $.50$), except for in the RCF 1 group ($r = .09$), indicating that people who predicted finishing earlier (more days before predicted) actually finished earlier (more days before the deadline). Correlations between our additional variables of interest and Part 2 variables can be seen in Table 36.

Discussion

Although Study 6 participants believed that they exhibited only a very small amount of planning fallacy in past projects (hypothesis 1), those who went through the reference forecasting interventions gave later, less optimistically biased completion predictions than those in the control groups (hypothesis 4). The control group that included an initial and final prediction, but no RCF intervention, did not lead to later, less optimistically biased completion predictions. However, it did evidence descriptively later completion predictions than the prediction only control group and did not consistently differ from the individual RCF groups. One possible reason for this is that the condition creates demand characteristics that prompt some participants

to make a slightly later final prediction. Unlike in Study 5, we found some support for the moderation hypothesis (hypothesis 6): in the RCF groups with initial and final predictions, RCF was most effective to the extent that people recalled completing past projects later than predicted.

We also manipulated two factors to help us better understand the process by which RCF procedure leads to less optimistic predictions – exploring whether making an initial and final prediction is important and whether presenting participants with a mean and instructions to use that mean to make their prediction is important. We did not find evidence that either of these factors altered the effectiveness of the RCF intervention. Although this result has yet to be replicated, it suggests that a fairly minimal intervention that instructs participants on how to identify a distribution of relevant prior project outcomes (without also giving explicit instruction for how to use that information or a reference point for what their intuitive, unadjusted prediction is), may be sufficient to help participants make less optimistically biased completion predictions.

On the other hand, we did observe stronger correlations between past completion times and completion predictions in those RCF groups that included initial and final predictions, suggesting that the link between past completion times and predictions may be somehow enhanced by this extra step. Although past project completion times and actual completion times were not correlated in this study, in circumstances where past completion times are more strongly related to actual completion times (e.g., when objective past completion times are available), this factor may be more important.

General Discussion

The purpose of the current work was to examine people's beliefs and recollections of their personal past project completion times and to design and test an intervention that would help them to use this information to make less optimistically biased predictions. In all six studies

(N = 1,747) we measured people's beliefs or recollections of their past project completion times, to determine whether people recall finishing past projects later than predicted (i.e., if they believe they exhibit the planning fallacy). The results were generally in line with our hypothesis that people believe they tend to underestimate completion times (Hypothesis 1), with studies 1, 2, 3, and 6 indicating that, on average, people recall finishing projects slightly later than predicted, and studies 4 and 5 indicating the contrary, that people believe they finish projects slightly earlier than predicted. A summary, including effect sizes and a mini meta-analysis (Goh, Hall & Rosenthal, 2016), of the planning fallacy beliefs ratings across the six studies is presented in Table 37. It's unclear why these differences between the studies emerged. In Study 5, this may reflect the fact that the participants who reported finishing earlier than predicted in the past were influenced by having just made their prediction for an upcoming task. The findings extend previous work on the planning fallacy which has often documented people's tendency to underestimate completion times for a specific target task, but rarely examined people's overall beliefs or theories about the accuracy of their task completion predictions.

Study 1 and 2 tested whether simply reminding people of their past project completion times, which on average were slightly later than original predictions, would lead participants to make later predictions for an upcoming project (Hypothesis 2). The results did not support this hypothesis: recalling past project completion times did not lead to later completion predictions. Although this finding is in line with the results of Buehler and colleagues (1994), and Buehler and Griffin (2003), we had expected it might differ because, unlike the previous studies, participants were prompted to report their past completion times in reference to predictions, rather than to task deadlines. Furthermore, it was also not the case that the effect of past reminders was moderated by beliefs about the planning fallacy, that is, that reminders of past

project completion times led to later predictions to the extent that people believed they tended to finish projects later than predicted (Hypothesis 3). These findings lend further support to the idea that people are not easily led to incorporate knowledge of past completion times into their predictions for a specific upcoming task (Buehler et al., 1994; Flyvbjerg et al., 2005; Kahneman & Lovallo, 1993; Kahneman & Tversky, 1979).

In Studies 3 through 6 we tested whether introducing a more explicit, structured approach to incorporating past project completion times into predictions, reference class forecasting (RCF; Lovallo & Kahnemann, 2003; Flyvbjerg 2006) would lead to less optimistically biased predictions for individual projects (Hypothesis 4). To the best of our knowledge, our studies are the first to systematically test the effectiveness of reference class forecasting for reducing optimistic prediction bias in the domain of individual, personal projects. Although the results were not always consistent across all analyses, we did find evidence that RCF interventions lead to later, less optimistically biased completion predictions. In three of the four studies, RCF interventions led to less optimistic predictions, and eliminated the optimistic prediction bias completely in two studies (Study 4, 6).

In Study 3 and 4 we tested hypothesis 5, that recalling past project completion times in reference to predictions would be more effective for leading to later completion predictions than recalling them in reference to deadlines. We found evidence in support of this hypothesis in Study 3, but not in Study 4, indicating that further research is needed to test this hypothesis.

Finally, in Study 5 and 6, we tested hypothesis 6, that the RCF procedure would be most effective to the extent that people recall completing past projects later than predicted. We did not find that the effect of RCF was moderated by past project completion times in Study 5 but found support for this in 2 out of 4 experimental groups in Study 6.

In addition, we further examined the most effective RCF group from studies 3 and 4 and tested whether certain components of the intervention were particularly important for leading to later completion predictions. In particular, we examined whether making an initial and then a final prediction was important, and whether receiving the mean of one's past project and instructions to use that mean was important. The results suggested that regardless of whether these components were included, the RCF procedure was similarly effective. The pattern of findings suggests that the most important aspect of the RCF procedure may be the recall of multiple similar past projects, their completion times in relation to predictions, and the more general instruction that this information can be used to make a more accurate completion prediction. A high-level overview of the main results for each study, showing whether each of our hypotheses were supported or not can be found in Table 38.

When we compared participants' recall of past completion times in reference to predictions with the amount of bias in their prediction for the target task in our RCF groups, we saw that, for the most part, participants' reports of previous prediction bias did not differ significantly from the bias observed for the target task. That is, there was no evidence of substantial memory bias for the one target task assessed in each study. Whereas Roy and colleagues (2005; 2008) found evidence of memory bias, wherein people tended to underestimate how long long tasks took in the past, our recall measures indicated that people did not systematically underestimate how long it took to complete past projects, at least compared to the single target task assessed by us. Of course, it is possible that task was not comparable to the set of projects participants reported on, so more detailed records of past project completion times would ultimately be required to be able to draw strong conclusions. Nonetheless, one reason why

recall in our studies may be less likely to be biased is that our participants recalled past times primarily in relation to predictions.

Although the results of the RCF interventions might seem to suggest that recalling similar past project completion times relative to predictions is an effective way to help people link their past completion times to their predictions, the relatively small or moderate, inconsistent correlations between past project completion times in our studies casts doubts on this. In fact, a complete lack of correlation between mean past completion times and predictions in many of our most effective RCF groups may suggest that the RCF procedure led to later completion prediction via a kind of blunt, uncalibrated adjustment process. This potential explanation is in line with Buehler and colleagues' (1994) findings that people's past experience can be leveraged to reduce the degree of optimistic bias in prediction, but that people do not easily incorporate information about past completion times into their predictions. Alternatively, the possibility exists that participants were incorporating other more sophisticated information into their predictions, potentially even about their past projects or the past project distribution (e.g., perhaps they spontaneously place the current project somewhere within their distribution of past projects).

Interestingly, we also observed that predictions in some RCF prediction groups were slightly later that their actual completion times (i.e., slightly pessimistic; the RCF prediction condition in Study 4, the RCF prediction conditions in Study 6), suggesting that the RCF procedure may also have the potential to lead to over-adjustment or overcorrection, something that has been observed in other contexts, e.g., overestimation bias in future personal spending (Peetz, Buehler, Koehler, & Moher, 2015), overcorrection for misinformation influence on

eyewitness testimony (Echterhoff, Groll, & Hirst, 2007), overcorrection of bias in personnel decisions (e.g., Tetlock, Mitchell, & Murray, 2008), and may warrant further investigation.

Limitations and Future Directions

One limitation of the current work is that we cannot know whether the RCF interventions led to later completion predictions due to demand characteristics. Conceivably, the more elaborate the procedure, the more demand participants feel to alter their responses to be in line with what they believe the researchers want. In part to address this concern, we introduced control conditions where participants engaged in procedures that might be expected to create demand characteristics. In particular, participants were asked to make an initial prediction and then later asked to make their final prediction. These conditions did not lead participants to adjust their predictions to the same degree as the RCF conditions, which may help to allay concerns about demand characteristics. On the other hand, these control groups did not always differ significantly from the RCF groups. Moreover, the procedures were not as elaborate as in the RCF conditions and thus may not have produced equally strong experimental demand. Future research could attempt to hold constant how elaborate the procedure in each group is and try to measure whether and how participants are actually using the past project information while forming their prediction. For example, this could be done using think-aloud procedures or asking participants to indicate where in their distribution of past projects they believe the current project falls and why.

Relatedly, because our RCF studies did not include a true past recall control condition, where we measured past project completion times before predictions but outside the context of an RCF procedure, we do not know whether the RCF procedure led to later predictions through unbiased past completion time recalls or whether knowing that the purpose of the exercise was to make an accurate completion prediction led participants to change their past recall.

Given the more applied nature of this work it seems worthwhile, beyond the statistical significance, to discuss the practical significance of the RCF intervention results. Table 38 shows a summary of the effects of our RCF (prediction groups) interventions for studies 4-6, including mean differences effect sizes for predictions and prediction bias, as well as the meta-analytic effect size (Goh, Hall, & Rosenthal, 2016) and common language effect size (Grissom and Kim, 2005; McGraw and Wong, 1992). In the aggregate, predictions were about three-quarters of a day later in the RCF prediction groups than main control groups, and the RCF prediction groups evidenced about a half a day less prediction bias than the main control groups. These are not usually large effect sizes, but do keep in mind that these results were obtained with participants who experienced only a small degree of planning fallacy, that this is only a first round of intervention design for RCF in this context, and that even small optimistic errors in prediction can lead to negative consequences especially when someone has multiple projects and other life events on the go. In a real-world context, we would expect to target people that can be identified as having significant, chronic planning fallacy problems with RCF interventions, and thus they may show larger benefits from this approach. Of course, RCF's efficacy with such groups of people has yet to be tested.

Another limitation of this work involves our operationalization of RCF in the context of individual personal projects. We had to make a number of difficult decisions in the process of trying to make the technique feasible for participants in a relatively short one-session online survey study while still remaining true to the fundamental definition and components of RCF (Flyvbjerg, 2006; Lovallo & Kahneman, 2003). For example, to meet step 1 of RCF, similar past projects should be broad enough to be statistically meaningful but narrow enough to be truly comparable. Is five to six similar past projects enough to be statistically meaningful in our

context? We don't know. For step 2, credible empirical data should be used to establish the probability distribution, but our participants likely wouldn't have had empirical data, so we had to do the next best thing – ask them to try to recall as accurately as possible. For step 3, the forecaster should place the current project in the distribution of outcomes for the group of similar past projects, but asking our participants to be able to assess their projects' characteristics relatively to his/her past projects in a matter of a few minutes, and without any concrete instructions as to what constitutes relevant characteristics or data on those characteristics seemed unreasonable – thus we either asked participants to use the mean (mean and instructions groups) or left it up to them to decide what to do with the past project completion item information (no mean, instructions groups).

Given these adjustments to the procedure, it wouldn't be unreasonable for someone to question whether our procedure still meets the definition of RCF. We believe it does because it preserved the three essential elements of the RCF procedure, 1) identifying a reference class of similar past projects (via our past projects nomination instructions), 2) establishing a probability distribution for the selected reference class (via our past projects completion time recall instructions), and 3) having the forecaster place the current project in the distribution of past project outcomes (via our instructions for making a completion prediction using past project outcomes), as outlined by Flyvbjerg (2006), and Lovallo and Kahneman (2003). Further research will be needed to learn how to best instruct participants and what degree of data, analysis and detail is required for the procedure to achieve the best possible balance of feasibility and accuracy. This is especially true for eliciting the list of similar past project completion times. Future studies could ask people to keep a record of their predictions and completions times for a

time (e.g., using daily diary procedures, or time-use applications) and then use this record to construct a relevant distribution.

Other interesting potential avenues for future research include comparing the effectiveness of RCF for generating less optimistically biased predictions to other debiasing strategies, such as unpacking (Peetz et al., 2015), considering alternative scenarios for how a task may unfold (Newby-Clark et al., 2000), using third-person imagery (Buehler et al., 2012), and backward planning (Wiese, Buehler, & Griffin, 2016). In addition, we don't know much about whether and how completion predictions change over time, as a project progresses. For example, people may change their predictions to fit what is desirable and feasible in their minds following the generation of a prediction using RCF. Future research could explore how project completion predictions change over the course of a project and whether interventions such as RCF have permanent or only a temporary influence on predictions and related project plans.

Although our work provides only initial evidence of the promise of RCF for helping individuals generate less optimistically biased completion predictions, if future works confirms its efficacy, there are a number of potential important applied benefits, included reduced likelihood of late project completion, lower likelihood of taking on too many projects or being too ambitious about project outcomes, reduced overtime work and fatigue, and reduced negative affect and stress about timely project completion. In addition, with time management technologies, such as apps and online calendars, the tracking of both predictions and project completion times can be easy and convenient, and allow for flexible adjustments to predictions based on project characteristics and other demands on one's time that are expected to influence project completion times.

Conclusions

Being able to predict accurately when we will be finished work, personal, or academic projects is an important factor for success in each of these domains. Underestimating the time needed to finish tasks can have real negative consequences, from as small as having to work a little bit of overtime, to as large as being fired. In this research, we have attempted to adapt reference class forecasting, a technique used in large scale organizational projects, to the domain of personal predictions, and found initial evidence that this approach can help individuals make less optimistically biased completion predictions. However, our results were not always consistent and point to the need for future work to replicate these effects, to identify boundary conditions, and to better understand the mechanisms by which RCF can lead to less optimistically biased and more accurate completion predictions.

References

American Planning Association (APA) (2005). JAPA article calls on planners to help end inaccuracies in public project revenue forecasting. Available at http://www.planning.org/newsreleases/2005/ftp040705.htm (accessed 15 November 2007).

Armor, D. A., Massey, C., & Sackett, A. M. (2008). Prescribed optimism: Is it right to be wrong about the future? *Psychological Science*, *19*, 329-331.

Buehler, R., & Griffin, D. (2003). Planning, personality, and prediction: The role of future focus in optimistic time predictions. *Organizational Behavior and Human Decision Processes*, *92*, 80-90.

Buehler, R., & Griffin, D. (2015). The planning fallacy: When plans lead to optimistic forecasts. In M. D. Mumford & M. Frese (Eds.), *The psychology of planning in organizations: Research and applications (pp. 31-57).* New York: Routledge.

Buehler, R., Griffin, D., & Peetz, J. (2010). The planning fallacy: Cognitive, motivational, and social origins. In M. P. Zanna & J. M. Olson (Eds.), *Advances in experimental social psychology* (Volume 43, pp. 1-62). Burlington: Academic Press.

Buehler, R., Griffin, D., Lam, K. C. H., & Deslauriers, J. (2012). Perspectives on prediction: Does third-person imagery improve task completion estimates? *Organizational Behavior and Human Decision Processes, 17*, 138-149.

Buehler, R., Peetz, J., & Griffin, D. (2010). Finishing on time: When do predictions influence completion times? *Organizational Behavior and Human Decision Processes, 111*, 23-32.

Buehler, R., Griffin, D., & MacDonald, H. (1997). The role of motivated reasoning in optimistic time predictions. *Personality and Social Psychology Bulletin, 23*, 238-247.

Buehler, R., Griffin, D., & Ross, M. (1994). Exploring the "planning fallacy": Why people

underestimate their task completion times. *Journal of Personality and Social Psychology, 6,* 366-381.

Buehler, R., Messervey, D., & Griffin, D. (2005). Collaborative planning and prediction: Does group discussion affect optimistic biases in time estimation? *Organizational Behavior and Human Decision Processes, 97,* 47-63.

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.

Costa, P. T., & Mac Crae, R. R. (1992). *Neo personality inventory-revised (NEO PI-R)*. Odessa, FL: Psychological Assessment Resources.

Dunning, D. (2007). Prediction: The inside view. In A. W. Kruglanski and E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (Second edition). New York: Guilford Press.

Echterhoff, G., Groll, S., & Hirst, W. (2007). Tainted truth: Overcorrection for misinformation influence on eyewitness memory. *Social Cognition*, *25,* 367-409.

Eichner, C., & Berna, F. (2016). Acceptance and efficacy of metacognitive training (MCT) on positive symptoms and delusions in patients with schizophrenia: A meta-analysis taking into account important moderators. *Schizophrenia Bulletin, 42,* 952-962.

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

Flyvbjerg, B. (2006). From Nobel Prize to project management: Getting risks right. *Project Management Journal*, *37*(3), 5-15.

Flyvbjerg, B. (2008). Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice. *European Planning Studies, 16*, 3-21.

Flyvbjerg, B., & COWI (2004). *Procedures for dealing with optimism bias in transport planning: Guidance document*. London: UK, Department for Transport.

Flyvbjerg, B., Garbuio, M., & Lovallo, D. (2009). Delusion and deception in large infrastructure projects: Two models for explaining and preventing executive disaster. *California Management Review, 51*, 170-194.

Flyvbjerg, B., Holm, M. K. S., & Buhl, S. L. (2005). How (in)accurate are demand forecasts in public works projects? *Journal of the American Planning Association, 71*, 131-146.

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10*, 535-549.

Griffin, D., & Buehler, R. (1999). Frequency, probability, and prediction: Easy solutions to cognitive illusions? *Cognitive Psychology, 38*, 48-78.

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Lawrence Erlbaum Associates Publishers.

Halkjelsvik, T., & Jørgensen, M. (2012). From origami to software development: A review of studies on judgment-based predictions of performance time. *Psychological Bulletin, 138*, 238-271.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from http://www.afhayes.com/ public/process2012.pdf

Jørgensen, M. (2004). A review of studies on expert estimation of software development effort. *Journal of Systems and Software, 70*, 37-60.

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

Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *TIMS Studies in Management Science, 12*, 313-327.

Köther, U., Vettorazzi, E., Veckenstedt, R., Hottenrott, B., Bohn, F., Scheu, F., ... & Moritz, S. (2017). Bayesian analyses of the effect of metacognitive training on social cognition deficits and overconfidence in errors. *Journal of Experimental Psychopathology, 8*, 158-174.

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.

Liberman, N., Trope, Y., & Stephan, E. (2007). Psychological distance. In A. W. Kruglanski & E. T. Higgins & (Eds.). *Social psychology: Handbook of basic principles*. New York: Guilford Press.

Lovallo, D., & Kahneman, D. (2003). Delusions of success: How optimism undermines executives' decisions. *Harvard Business Review, 81*, 56-63.

Lynch Jr, J. G., Netemeyer, R. G., Spiller, S. A., & Zammit, A. (2009). A generalizable scale of propensity to plan: The long and the short of planning for time and for money. *Journal of Consumer Research, 37*, 108-128.

McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin, 111*(2), 361-365.

Min, K. S., & Arkes, H. R. (2012). When is difficult planning good planning? The effect of scenario-based planning on optimistic prediction bias. *Journal of Applied Social Psychology, 42,* 2701-2729.

Moritz, S., Andreou, C., Schneider, B. C., Wittekind, C. E., Menon, M., Balzan, R. P., & Woodward, T. S. (2014). Sowing the seeds of doubt: a narrative review on metacognitive training in schizophrenia. *Clinical Psychology Review, 34,* 358-366.

Newby-Clark, I. R., Ross, M., Buehler, R., Koehler, D. J., & Griffin, D. (2000). People focus on optimistic and ignore pessimistic scenarios while predicting their task completion times. *Journal of Experimental Psychology: Applied, 6,* 171-182.

Norem, J. K., & Cantor, N. (1986). Defensive pessimism: Harnessing anxiety as motivation. *Journal of Personality and Social Psychology, 51,* 1208-1217.

Peetz, J., Buehler, R., Koehler, D., & Moher, E. (2015). Bigger not better: Unpacking future expenses inflates spending predictions. *Basic and Applied Social Psychology, 37(1),* 19-30.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28,* 369-381.

Roy, M. M., Christenfeld, N. J. S., & McKenzie, C. R. M. (2005). Underestimating the duration of future events: Memory incorrectly utilized or memory bias? *Psychological Bulletin, 131,* 738-756.

Roy, M. M., Mitten, S. T., & Christenfeld, N. J. (2008). Correcting memory improves accuracy of predicted task duration. *Journal of Experimental Psychology: Applied, 14*(3), 266-275.

Tenney, E. R., Logg, J. M., & Moore, D. A. (2015). (Too) optimistic about optimism: The belief that optimism improves performance. *Journal of Personality and Social Psychology, 108,* 377-399.

Tetlock, P. E., Mitchell, G., & Murray, T. L. (2008). The challenge of debiasing personnel decisions: Avoiding both under-and overcorrection. *Industrial and Organizational Psychology*, *1*(4), 439-443.

Wiese, J., Buehler, R., & Griffin, D. (2016). Backward planning: Effects of planning direction on predictions of task completion time. *Judgment and Decision Making*, *11*(2), 147-167.

Appendix A1

Planning Fallacy Beliefs Ratings - Condition: Must-do, Self

In this section, we would like you to **think back to times when you remember having an expectation for when you would be finished a particular task or project** (e.g., an academic, work, or personal project).

In other words, try to recall times when you had made a prediction for when (e.g., time/date) you would be done some task or project.

Relative to when you expected to be finished, when did you actually finish these tasks or projects (on average)?

Finished much later than predicted	Finished later than predicted	Finished slightly later than predicted	Finished exactly when I predicted	Finished slightly earlier than predicted	Finished earlier than predicted	Finished much earlier than predicted
○	○	○	○	○	○	○

Sometimes people incorrectly predict when they will complete tasks or projects, finishing later than predicted or earlier than predicted. Other times people finish when they expected to be finished.

On average, what percentage of the time would you say YOU underestimate, overestimate, or are accurate about how long it will take you to complete tasks or projects?
(these must add to 100%)?

0 % of the time - I finish before my original prediction for when I would complete the task/project (I overestimate how long it will take)

0 % of the time - I finish at the same time as my original prediction (I accurately predict how long it will take to complete the task/project)

0 % of the time - I finish after my original prediction (I underestimate how long it will take)

Appendix A1

Planning Fallacy Beliefs Ratings - Condition: Must-do, Other

In this section, we would like you to **think back to times when you remember other people having an expectation for when they would be finished a particular task or project** (e.g., an academic, work, or personal project).

In other words, try to recall times when other people had made a prediction for when (e.g., time/date) they would be done some task or project.

Relative to when people expected to be finished, when do you think they actually finished these tasks or projects (on average)?

Finished much later than predicted	Finished later than predicted	Finished slightly later than predicted	Finished exactly when they predicted	Finished slightly earlier than predicted	Finished earlier than predicted	Finished much earlier than predicted
○	○	○	○	○	○	○

Sometimes people incorrectly predict when they will complete tasks or projects, finishing later than predicted or earlier than predicted. Other times people finish when they expected to be finished.

On average, what percentage of the time would you say people underestimate, overestimate, or are accurate about how long it will take to complete tasks or projects?
(these must add to 100%)?

0 % of the time - People finish before their original prediction for when they would complete the task/project (they overestimate how long it will take)

0 % of the time - People finish at the same time as their original prediction (they accurately predict how long it will take to complete the task/project)

0 % of the time - People finish after their original prediction (they underestimate how long it will take)

Appendix A1

Planning Fallacy Beliefs Ratings - Condition: Want-to-do, Self

In this section, we would like you to **think back to times when you remember having an expectation for when you would be finished a particular task or project** (e.g., an academic, work, or personal project).

Specifically, we would like you to think about tasks or projects that you you would consider **enjoyable/fun** to work on. In other words, try to recall times when you had made a prediction for when (e.g., time/date) you would be done some task or project you considered pleasant to work on.

Relative to when you expected to be finished, when did you actually finish these enjoyable tasks or projects (on average)?

Finished much later than predicted	Finished later than predicted	Finished slightly later than predicted	Finished exactly when I predicted	Finished slightly earlier than predicted	Finished earlier than predicted	Finished much earlier than predicted
○	○	○	○	○	○	○

Sometimes people incorrectly predict when they will complete tasks or projects, finishing later than predicted or earlier than predicted. Other times people finish when they expected to be finished.

On average, what percentage of the time would you say YOU underestimate, overestimate, or are accurate about how long it will take you to complete enjoyable tasks or projects?
(these must add to 100%)?

0 % of the time - I finish before my original prediction for when I would complete the task/project (I overestimate how long it will take)

0 % of the time - I finish at the same time as my original prediction (I accurately predict how long it will take to complete the task/project)

0 % of the time - I finish after my original prediction (I underestimate how long it will take)

Appendix A1

Planning Fallacy Beliefs Ratings - Condition: Want-to-do, Other

In this section, we would like you to **think back to times when you remember other people having an expectation for when they would be finished a particular task or project** (e.g., an academic, work, or personal project).

Specifically, we would like you to think about tasks or projects that you they would consider **enjoyable/fun** to work on. In other words, try to recall times when other people had made a prediction for when (e.g., time/date) they would be done some task or project they considered pleasant to work on.

Relative to when people expected to be finished, when do you think they actually finished these enjoyable tasks or projects (on average)?

Finished much later than predicted	Finished later than predicted	Finished slightly later than predicted	Finished exactly when they predicted	Finished slightly earlier than predicted	Finished earlier than predicted	Finished much earlier than predicted
○	○	○	○	○	○	○

Sometimes people incorrectly predict when they will complete tasks or projects, finishing later than predicted or earlier than predicted. Other times people finish when they expected to be finished.

On average, what percentage of the time would you say people underestimate, overestimate, or are accurate about how long it will take to complete enjoyable tasks or projects? (these must add to 100%)?

0 % of the time - People finish before their original prediction for when they would complete the task/project (they overestimate how long it will take)

0 % of the time - People finish at the same time as their original prediction (they accurately predict how long it will take to complete the task/project)

0 % of the time - People finish after their original prediction (they underestimate how long it will take)

Appendix A2

Reasons for Planning Fallacy Items

For the questions below we would like you to think about **times when you finished tasks/projects later than anticipated** (i.e., when you predicted that you would be done earlier than you actually were).

Listed below are a number of reasons why people might finish a task later than expected. We want to know which of these you think best describes when you finished task that you **had to do** later than expected. That is, please rate how accurately each reason describes why YOU finished these tasks later than expected.

	Not at all accurate for me 1	Slightly accurate 2	Somewhat accurate 3	Moderately accurate 4	Accurate 5	Very accurate 6	Extremely accurate for me 7
1. I focused too much on when I WANTED to be finished, instead of focusing on being realistic about when I would be finished.	○	○	○	○	○	○	○
2. I neglected to consider my PAST EXPERIENCES of how long it typically takes me to complete similar tasks (which would have led me to a more accurate prediction).	○	○	○	○	○	○	○
3. I focused too much on the steps required to complete the project that were salient at the time (and how long each of them would take) and neglected to consider other important factors.	○	○	○	○	○	○	○
4. I failed to anticipate a number of important steps or sub-steps of the task/project itself and how long they would take.	○	○	○	○	○	○	○

	Not at all accurate for me 1	Slightly accurate 2	Somewhat accurate 3	Moderately accurate 4	Accurate 5	Very accurate 6	Extremely accurate for me 7
If you are paying attention please select "Very accurate 6" for this statement.	○	○	○	○	○	○	○
5. I failed to anticipate project obstacles that would slow me down.	○	○	○	○	○	○	○
6. I failed to consider outside factors and competing priorities (things not relevant to the task/project itself) that would be taking up my time.	○	○	○	○	○	○	○
7. I based my expectations on past projects that went better than average and not representative of how my projects usually unfold.	○	○	○	○	○	○	○
8. I told myself that the project is unique or would be different from other projects and that therefore I would be done sooner than in the past.	○	○	○	○	○	○	○
9. I based my prediction on how long it would take if I were highly motivated to complete the project/task at all times and forgot to factor in that I would be slow or procrastinate at times due to a lack of motivation.	○	○	○	○	○	○	○
10. I purposely made an optimistic prediction about when I would be done, in order to motivate myself to get working and finish the project faster.	○	○	○	○	○	○	○
11. Other reason (please specify)	○	○	○	○	○	○	○

Appendix A3

Self-nominated Project Completion Prediction Instructions

For this question, please **think of a task or project that you will be completing in the next three weeks**. It could be a school, work, or personal project.

It should also have the following other characteristics:

1) It is a fairly major project that will require considerable time and effort to complete (i.e., not something that is simply part of your everyday routine).

2) It is a project that is specific and concrete, so that it would be easy to know exactly when it was completed (e.g., read "Chapters 5 and 6 of Biology textbook" not "do some reading").

3) It is a project that you are free to complete whenever you choose within the next 3 weeks (e.g., writing an essay), rather than a project that can only be done at a specified time (e.g., giving a speech).

4) It is a project that you have not yet started but need to finish within the next 3 weeks.

5) The project has a specific final deadline by which it must be finished.

Once you have thought of a task or project with these characteristics, please briefly describe it in a few words in the box below:

Please tell us how many days from today the final deadline for the task or project is. That is, **in how many days is the task or project due**? (please select an option from the drop down list)

Feel free to use the calendar below to count how many days from today the deadline for the project is. If you are having trouble remembering the deadline, please feel free to take a minute to look it up via your calendar or planner, or use any other on-line resource (e.g., course webpage) to help you find it.

Next, please think about when you will be completing the task or project. We would like you to **make a prediction for when you will be finished the project**. That is, how many days from today do you anticipate having completed the project? (please select an option from the drop down list below)

Appendix A4

Hypothetical, Standardized Project Completion Predictions Instructions

Imagine that you have been hired by a research group at the university to do some temporary work for their research project. The job they would like you to do will take approximately 8 hours to complete and must be completed within the next two weeks (the deadline to submit the work is in 14 days). You know the experience will be very valuable, so you have every intention to complete the job. Keeping in mind any plans that you already have for your time, please try to predict as accurately as possible when you would be finished the work. I predict I would be finished the job (please select an option from the drop-down list):

Appendix B1
Planning Fallacy Article

Planning Fallacy

Posted Nov 15, 2017

A large body of recent research has looked at people's predictions of task completion times (i.e., when they expect to finish various tasks or projects). These predictions are estimates of how much time we need to complete tasks. Studies and experiments demonstrate that people's predictions of the duration needed to complete tasks are often overly optimistic. Researchers, such as Dr. Daniel Kahneman, have found that when we think about some project we need to complete, we tend to underestimate how long it will take to finish. This phenomenon is referred to as the planning fallacy.

There has been a significant mass of evidence collected strongly indicating that this occurs for a variety of tasks and everyday projects, such as completing large-scale industrial projects, academic assignments and personal tax forms. We tend to believe that we will finish tasks sooner than we actually do. Oftentimes, we expect to finish tasks many days before the deadline. In reality, people tend to finish tasks pretty close to the deadline. In one study, a class of graduate students were asked to estimate the date at which they will finish their thesis. They predicted that they will complete their thesis is 34 days. However, they actually completed their thesis in 56 days, demonstrating that people tend to be overly optimistic about when they expect to finish tasks.

Appendix B2

Hypothetical Assignment

We would now like you to imagine yourself in a hypothetical situation. Please do your best to imagine what you would do if you encountered this situation in real life.

Imagine that you need to complete a major assignment for one of your courses that is due in 14 days (i.e., it has a hard deadline that is two weeks away).

For this assignment, you are required to write a minimum 12-page research report that includes at least 8 references (four from books available only in the library).

This assignment falls at a time of year that is very busy for students, and, as an incentive to have it done promptly, the instructor is awarding an extra 2% for every day before the due date that the assignment is submitted.

Appendix C1

Project Nomination Instructions

For this study, we would like you to think of an **academic task or project** that you will be completing in the future.

This project needs to fit the following 5 criteria:

1) The project has to be completed some time in the next two weeks. That is, it has a firm **final deadline in the next two weeks**. The deadline should also be 5 or more days away.

2) It is a project that you are **free to complete whenever you choose** within the next 2 weeks (e.g., writing an essay, study for a test), rather than a project that can only be done at a specified time (e.g., attending a workshop to complete a certificate, take a test).

3) You are **hoping to finish as soon as possible** (ideally, you would like to finish well before the deadine).

4) It is a fairly **major project** that will require considerable time and effort to complete (i.e., not something that is simply part of your everyday routine).

5) It is a project that is specific and concrete, so that it would be easy to know exactly when it was completed (e.g., "complete 5-page psychology essay" not "make progress on biology research report").

Once you have thought of a task or project with the 5 characteristics listed above, please briefly describe it in the box below:

Next, please tell us the **final deadline** for the project using the date picker below:

If you are having trouble remembering the deadline, please feel free to take a minute to look it up via your calendar or planner, or use any other on-line resource (e.g., course webpage).

← January 2018 →

Su	Mo	Tu	We	Th	Fr	Sa
31	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31	1	2	3
4	5	6	7	8	9	10

Appendix C2

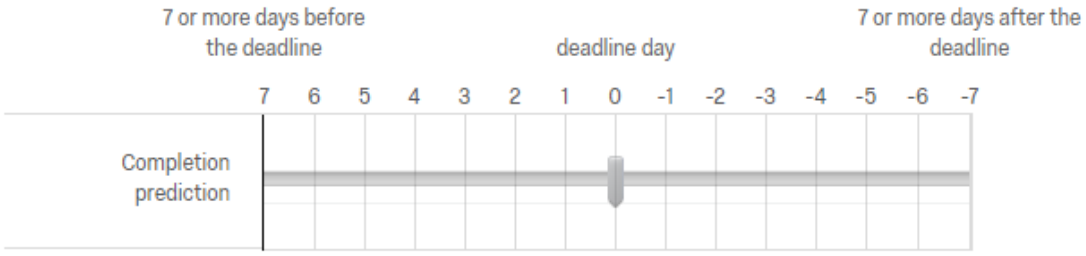
Control, "Past" Condition – (Final) Prediction Instructions

Step 3 - Completion prediction for current project

We would now like you to predict, as accurately as possible, when you will be finished the project/task you just nominated.

For example, if you think you will finish the project **on the day of the deadline**, put the slider at **"0",** if you think you will finish the project **6 days early**, put the slider at **"6"**, if you think you will finish the project **3 days late** then put the slider at **"-3"**.

Using the slider below, give your prediction of when you expect to be finished the current project. **Make sure you click on the slider (even to answer with 0), otherwise your data will not be recorded!**



Appendix C3

"Past", "RCF deadline", "RCF prediction" Condition - Past Projects Nomination

Later in this survey we will ask you to predict, as accurately as possible, when you will be finished the project/task described above.

One approach to making an accurate prediction is to think about similar past projects and when you typically completed them. We would like you to do this now. To assist you in the process, please complete the three steps that follow.

Step 2 - Identify Similar Past Projects

Next, we would like you to think of 5 to 6 projects that are as similar as possible to the one you just nominated, that you remember completing.

For example, these other projects may be similar in terms of project scope, type, complexity, amount of time you have to work on them, how much control you have over their outcome, etc.

In the boxes below, briefly (in 1-2 lines) describe each past project you thought of.

We want everyone to take at least a couple of minutes to think of these past projects, so the "continue" button to move to the next page will not appear until 2 minutes have elapsed.

Past Project 1

Past Project 2

Past Project 3

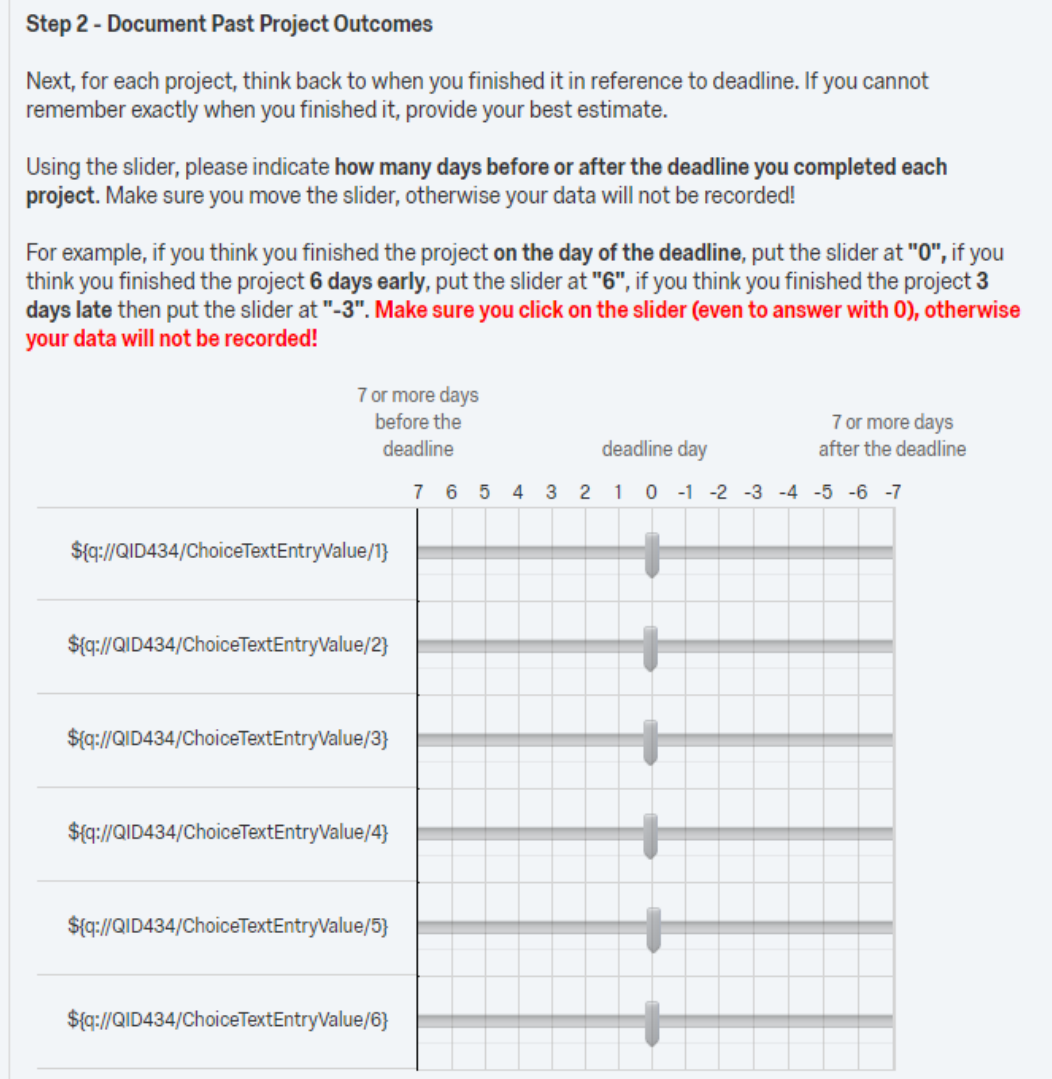
Past Project 4

Past Project 5

Past Project 6

Appendix C4

"Past", "RCF deadline" Condition - Past Project Completion Ratings



Appendix C5

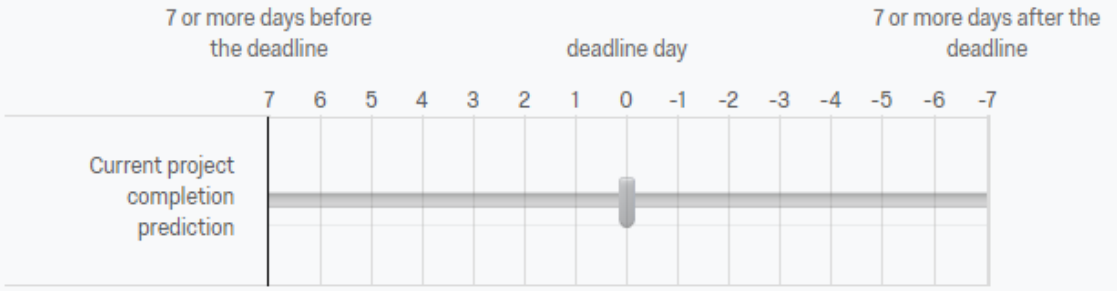
"RCF deadline" Condition – (Final) Prediction Instructions

Step 3 - Make a completion prediction using past project outcomes

Based on your past projects, on average **you finish ${e://Field/Mean2} days before the deadline** (or after if the number is negative).

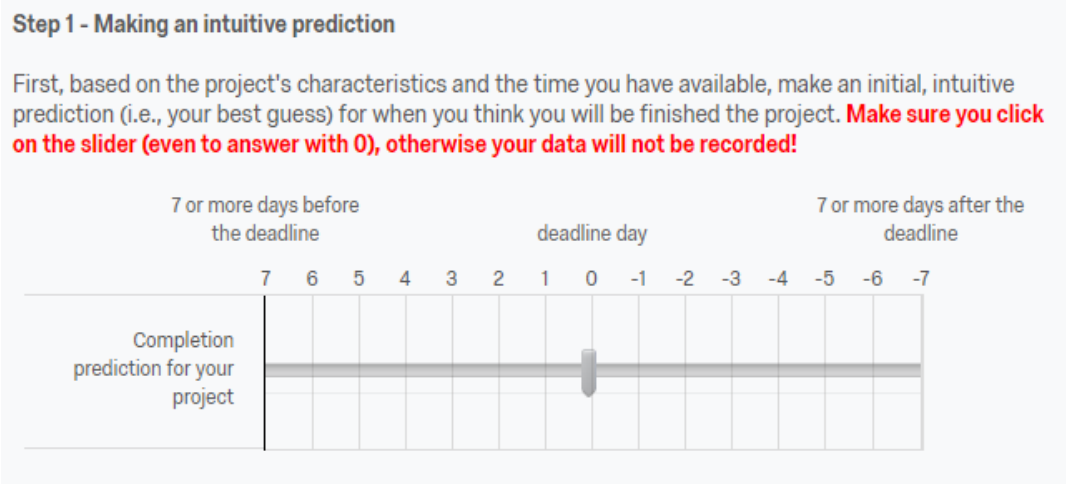
Therefore, unless you believe that you have better information available to make your prediction for this project than for your past projects, you should base your project completion prediction on this average. If your average is just before the deadline, then you should predict just before the deadline for the current project, unless you have reliable information to indicate that this project will be different.

Using the slider below, give your prediction for when you anticipate having completed the project. **Make sure you click on the slider (even to answer with 0), otherwise your data will not be recorded!**



Appendix C6

"RCF prediction" Condition - Initial, Intuitive Prediction



Appendix C7

"RCF prediction" Condition - Past Project Completion Ratings

Step 3 - Document Predictions & Outcomes for Past Projects

Next, for each project, think back to when you first formed an expectation (prediction) for when you would be finished and when you actually finished the project. If you cannot remember exactly, provide your best estimate.

Using the sliders, please indicate **how many days before or after your predicted completion date you actually finished.** Make sure you move the slider, otherwise your data will not be recorded!

For example, if you finished on the day you initially expected to finish put the slider at **"0"**, if you think you finished the project **6 days before you anticipated finishing**, put the slider at **"-6"**, if you think you finished the project **3 days after** your initial prediction then put the slider at **"3"**. **Make sure you click on the slider (even to answer with 0), otherwise your data will not be recorded!**

	7 days or more before the prediction						original completion prediction						7 days or more after the prediction		
	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
${q://QID407/ChoiceTextEntryValue/1}															
${q://QID407/ChoiceTextEntryValue/2}															
${q://QID407/ChoiceTextEntryValue/3}															
${q://QID407/ChoiceTextEntryValue/4}															
${q://QID407/ChoiceTextEntryValue/5}															
${q://QID407/ChoiceTextEntryValue/6}															

Appendix C8

"RCF prediction" Condition – Final Prediction Instructions

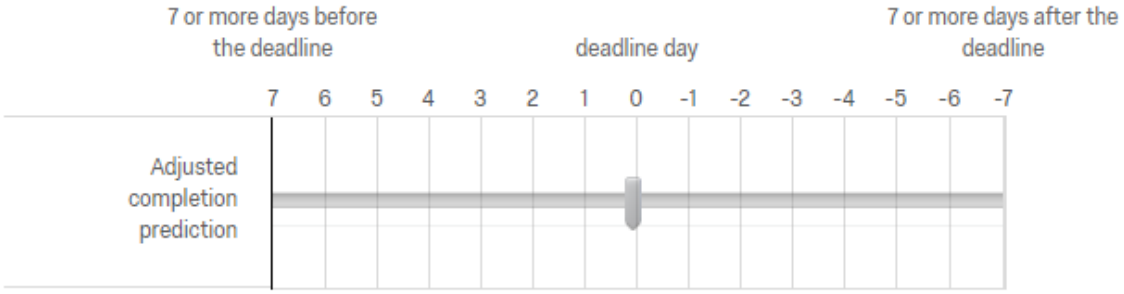
Step 4 - Correcting your initial prediction

Adjust your initial prediction by using the data from the past projects you listed above.

Your **initial prediction** for finishing the project was **${q://QID414/ChoiceNumericEntryValue/1} days before the deadline**. Based on your past projects, **on average you finish ${e://Field/Mean} days after your prediction** (if the number is positive)**/before your prediction** (if the number is negative).

Therefore, unless you believe that you have better information available to make your prediction for this project than for your past projects, you should base your project completion prediction on this average. If your average is finishing slightly later than predicted, then you adjust your initial prediction to be somewhat later, unless you have reliable information to indicate that this project will be different.

Using the slider below, indicate your final, adjusted prediction for when you anticipate having completed the project. **Make sure you click on the slider (even to answer with 0), otherwise your data will not be recorded!**



Appendix C9

Exploratory Measures Part 1 (Study 3)

Method

Part 1. Participants were asked to write a few sentences explaining how they arrived at their prediction, the factors they considered, what they based their prediction on. Participants not in in the control condition were also asked to write a few sentences about how they found the exercise for making an accurate completion prediction that we asked them to engage in (e.g., whether they found it useful, whether they liked or disliked it).

Predicted working time. Next, participants predicted how many hours of actual working time they would spend on the project.

Factors perceived to influence completion predictions. We asked participants to rate the extent to which they based their prediction on five factors: thoughts about when they would ideally like to be finished, thoughts about the steps needed to complete the project and how long each step would take, thoughts about potential obstacles that might delay progress on the task, thoughts about other demands on their time (i.e., competing events or activities), and thoughts about their own past experiences with similar projects (i.e., when you typically finish projects), on a scale from 1 (*not at all*) to 7 (*a great deal*). Given the explicit focus on multiple past projects in the experimental conditions, we thought that participants in these conditions might give higher ratings on the past experiences question than those in the control condition.

Project characteristics. Participants again rated the project on: difficulty, importance, control, how busy they would be with other tasks while completing it, desire to complete as quickly as possible, and importance of the quality of the final project (1 - *not at all*, to 7 - *extremely*).

Confidence. Participants rated how confident they were that their completion prediction is accurate (that they will finish on the day they predicted) on a scale from 1 (*not at all confident*) to 7 (*extremely confident).*

Similarity of past projects to current project. Participants in the three intervention conditions also rated how the past projects they listed compared to the current project in terms of similarity (1 - *not at all similar*, to 7 - *extremely similar*), complexity (1 – *past projects much less complex*, to 7 – *past projects much more complex*), scope (1 – *past projects had a much smaller scope*, to 7 – *past projects had a much larger scope*), and amount of time available to work on the project (1 - *much more time available for past projects*, to 7 – *much less time available for past projects*). Participants also rated how relevant they thought the past projects were for helping them to make a prediction for when they would be finished the current project (1 - *not at all relevant*, to 7 - *extremely relevant).*

Past projects exercise ratings. Participants rated the exercise they engaged in (using past projects to make a more accurate completion prediction) on the following characteristics: straightforward, difficult, useful, annoying, on a scale from 1 (*not at all)* to 7 (*extremely).*

Personality characteristics. We measured participants' general tendency to engage in planning for their time using Lynch, Netemeyer, Spiller, and Zammit's (2010) propensity to plan scale (e.g., "I consult my planner to see how much time I have for the next few weeks", 1 - *strongly disagree*, to 7 - *strongly agree*) ($\alpha = .90$) and conscientiousness (NEO PI-R; Costa & Mac Crae, 1992) (e.g., "I am always prepared", 1 - *strongly disagree*, to 7 – *strongly agree*) ($\alpha = .86$).

Opinions. Participants rated how important they thought being accurate about completion predictions is, how motivated they are to make accurate completion predictions, and the extent to

which they think underestimating when they will finish a project is problematic, on a scale from 1 (*not at all*) to 7 (*extremely*).

Results

Predicted Working Time. No differences between conditions emerged on the predicted number of working hours.

Factors Perceived to Influence Completion Predictions. Endorsement of the five factors that participants could have based their predictions on did not differ by condition (see Table 10). Notably, this was the case even for the past experience item – participants who were instructed to think about past projects did not indicate that they based their prediction on past experience significantly more than those in the control condition. There was also no significant difference in item endorsement between items, $F(4, 880) = 0.80$, $p = .525$, $\eta_p^2 = .004$. Overall endorsement was as follows: based prediction on when I want to be finished, $M = 5.05$ ($SD = 1.57$), based prediction on the steps needed to complete the project and how long they will take, $M = 5.08$ ($SD = 1.48$), based prediction on possible obstacles that may delay progress, $M = 5.15$ ($SD = 1.52$), based prediction on other demands on my time, $M = 5.54$ ($SD = 1.30$), and based prediction on past experiences with similar projects, $M = 5.64$ ($SD = 1.30$).

Past Project Characteristics. The seven project characteristics were assessed both before and after the manipulation and thus repeated measures ANOVAs were conducted. No pre-post differences, main effects of condition, or (time x condition) interaction emerged for six out of the seven variables. Pre-manipulation ratings of level of control over the project ($M = 5.79$, $SD = 1.07$) were marginally higher than post-manipulation ratings ($M = 5.67$, $SD = 1.14$), $F(1, 218) = 3.21$, $p = .075$, $\eta_p^2 = .015$. The overall main effect of condition was not significant, $F(3, 218) = 1.90$, $p = .131$, $\eta_p^2 = .025$, but a trend of higher control ratings in the RCF deadline ($M = 5.87$, *SE*

= 0.13) and RCF prediction condition (M = 5.84, SE = 0.13) compared to the control condition (M = 5.47, SE = 0.13) was visible. The interaction was not significant.

Confidence. Participants' confidence in the accuracy of their completion predictions did not differ by condition and the overall mean was 5.50 (SD = 1.07).

Similarity of Past Projects to Current Project. Past projects did not differ by condition on rated similarity, with an overall mean of 5.41 (SD = 1.05) on the 7-point rating scale. Past projects were rated to be quite similar (a score of 4 = the same) to the current project in terms of complexity ($M_{overall}$ = 4.35, SD = 1.01), scope ($M_{overall}$ = 4.23, SD = 1.08), and time available to work on them ($M_{overall}$ = 4.41, SD = 1.12). Scope and time available did not differ by condition, but for complexity a marginal effect emerged, $F(2, 162) = 2.79$, $p = .065$, $\eta_p^2 = .035$, such that participants in the Past condition (M_{adj} = 4.56, SE = 0.15) rated the past projects as slightly more complex relative to the current project than participants in the RCF deadline condition (M_{adj} = 4.11, SE = 0.13), $t(107) = -2.06$, $p = .042$, d = 0.43. The RCF prediction condition (M_{adj} = 4.42, SE = 0.14) fell in between and didn't significantly from the other groups. Past project relevancy for helping participants to make a prediction for the current project did not differ by condition, and was relatively high, with an overall mean of 5.27 (SD = 1.20).

Past Projects Exercise Ratings. Participants found the past projects exercise relatively straightforward ($M_{overall}$ = 5.20, SD = 1.09), useful ($M_{overall}$ = 4.93, SD = 1.19), and not very difficult ($M_{overall}$ = 3.66, SD = 1.61), and these ratings did not differ by condition. These results are encouraging as they suggest that participants don't find reference class forecasting to be too difficult and that they see some value in it. A marginal effect emerged for the annoying item, $F(2, 162) = 2.80$, $p = .064$, $\eta_p^2 = .034$, such that participants in the RCF deadline condition (M_{adj} = 4.00, SE = 0.23) rated the exercise as slightly more annoying than those in the Past condition

(M_{adj} = 3.22, *SE* = 0.25), *t*(107) = 2.26, *p* = .026, *d* = 0.43. The RCF prediction condition fell in between (M_{adj} = 3.60, *SE* = 0.23) and did not differ from the other groups.

Personality Characteristics. The overall mean score on the 7-point propensity to plan scale was 5.08 (*SD* = 1.27) and did not differ by condition, *F*(3, 218) = 0.84, *p* = .472, η_p^2 = .012. Conscientiousness scale scores also did not differ by condition, *F*(3, 218) = 0.76, *p* = .516, η_p^2 = .011, with an overall mean of 4.67 (*SD* = 0.93).

Exploratory multiple regression analyses revealed that conscientiousness and propensity to plan did not interact with condition (dummy or contrast coding) to predict completion predictions, suggesting that they did not act as moderators in this study.

Opinions. Participants thought that making accurate completion predictions is relatively important, $M_{overall}$ = 5.37 (*SD* = 1.15) (on a 7-point scale), reported being relatively motivated to make accurate completion predictions, $M_{overall}$ = 5.14 (*SD* = 1.27) (on a 7-point scale), and thought that underestimating when one will finish a project is a bit more than moderately (a score of 4) problematic, $M_{overall}$ = 4.90 (*SD* = 1.59) (on a 7-point scale). These ratings did not differ by condition for the "motivated to be accurate" and "underestimation problematic" items, and differed marginally for the "importance of making accurate predictions" item, *F*(3, 218) = 2.14, *p* = .097, η_p^2 = .029, such that participants in the Past condition (M_{adj} = 5.63, *SE* = 0.16) gave somewhat higher ratings than participants in the RCF deadline (M_{adj} = 5.15, *SE* = 0.15), *t*(108) = -2.36, *p* = .020, *d* = -.41, and RCF prediction condition (M_{adj} = 5.23, *SE* = 0.16), *t*(100) = -1.76, *p* = .081, *d* = -.34, and the control condition fell in between (M_{adj} = 5.51, *SE* = 0.16), *t*(103) = -0.66, *p* = .509, *d* = -.10.

Appendix C10
Completion Time Questions

In the first part of the study you nominated a task or project that you would be completing in the next three weeks. Please briefly describe what the task or project was in the box below.

The information we have from the P1 survey is: ${e://Field/Project}

What was the final deadline for this project?

The information we have from the P1 survey is: ${e://Field/Deadline}

Did you finish this project?

- Yes
- No

When did you finish the project?

Make sure you move the slider, otherwise your data will not be recorded!

7 or more days before the deadline | on the deadline day | 7 or more days after the deadline

-7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7

How long you did you spend working on the project itself? In the box below, please indicate **how many hours you spent working on the project.** Please enter a numeric value (for half hours enter .5).

Appendix C11

Correlations between Key Variables and Exploratory Measures (Study 3)

Part 1. We also ran zero-order correlations, collapsed across conditions, between completion predictions and our additional measures (see Table 11). Most notably, participants who endorsed basing their prediction on when they ideally want to be finished more made earlier (further from the deadline) predictions, participants who endorsed basing their predictions on their past experiences more made later (closer to the deadline) predictions, participants who found the planning exercise more useful made earlier predictions, and participants who found making accurate completion predictions more important made earlier predictions. Interestingly, ratings of how focused participants were on finishing as quickly as possible were not related to predictions.

Both propensity to plan and conscientiousness showed small positive correlations with predictions, $r = .15$, $p = .023$ and $r = .19$, $p = .006$, respectively, indicating that a greater tendency to plan for one's time and higher levels of conscientiousness were associated with more optimistic (further before the deadline) completion predictions. Similar correlations emerged with mean past project completion times. In the Past condition, propensity to plan and conscientiousness were positively related to past project completion times (finishing earlier), $r = .35$, $p = .014$ and $r = .33$, $p = .024$, and in the RCF deadline condition propensity to plan and conscientiousness were also positively related to past completion times (finishing earlier), $r = .21$, $p = .08$ and $r = .32$, $p = .010$. These results could mean that greater conscientiousness and propensity to plan are related to actual earlier completion times or to particularly optimistic predictions and potentially biased recollection of past completion times (remembering finishing earlier than one actually did). Analyses with actual completion time should shed some light on these possibilities. In the RCF prediction condition propensity to plan was unrelated to past

project completion times, r = -0.08, p = .548, and marginally negatively related to conscientiousness, r = -.25, p = .066, indicating that participants who scored higher on conscientiousness tended to report finishing slightly later (relative to their predictions).

Part 2. Reported working time was not related to completion times suggesting that these predictions domains may track different types of information. The extent to which participants reported basing their predictions on when they ideally want to be finished was positively related to both predictions and "actual" completion times, indicating that people who based their prediction on when they wanted to be finished were also finishing sooner. Being focused on finishing as quickly as possible was related to more optimistic predictions, and marginally faster "actual" completion times. Greater confidence in having made an accurate completion prediction was related to predicting that one will finish closer to the deadline, and unrelated to actual completion times. Having more control over the project was related to finishing sooner.

Using the Part 2 sample (A), conscientiousness (r = .17, p = .079) and propensity to plan (r = .23, p = .014) were still positively related to completion predictions. Both were positively related to actual completion times, such that greater conscientiousness (r = .20, p = .032) and propensity to plan (r = .31, p = .001) were related to finishing earlier (more days before the deadline). Both were unrelated to bias (conscientiousness r = -.04, p = .639; propensity to plan r = -.09, p = .359). These results suggest that participants higher in conscientiousness and propensity to plan both made earlier predictions and finished their projects earlier than those lower on these personality traits. See Table 15 for details.

Appendix C12

Secondary Analyses for the Part 2 Sample (Study 3)

Method

Reasons for finishing later than expected. Following completion time ratings participants who finished the project later than expected were asked to rate six reasons on the extent to which they described why they finished later than expected (on a scale from 1 - *not at all*, to 7 - *a great deal*): underestimated how long it would take to complete each step needed to finish the project, failed to anticipate all the steps that would be required to complete the project, failed to anticipate project obstacles, didn't anticipate other demands on my time enough, priorities changed (this project became lower priority), my motivation declined more than I anticipated (e.g., procrastination).

Results

Past Completion Times, Predictions, Predicted Work Time. Past completion times showed the same pattern of results as the Part 1 sample (see Table 12). However, the ANCOVA revealed that the overall effect for predictions was no longer statistically significant, and planned comparisons revealed only a marginal difference between the RCF prediction (M_{adj} = 1.38, *SE* = 0.32) and control (M_{adj} = 2.10, *SE* = 0.27) condition, *t*(54) = -1.74, *p* = .088, *d* = -.032. As before, the RCF deadline (M_{adj} = 1.87, *SE* = 0.27) and past condition (M_{adj} = 2.34, *SE* = 0.31) did not differ from the control condition. The RCF prediction condition differed significantly from the Past condition, *t*(44) = -2.16, *p* = .037, *d* = 0.42, and did not differ from the RCF deadline condition, *t*(53) = -1.25, *p* = .216, d = -0.22.

As in Part 1, the number of predicted work hours did not differ by condition.

We also compared reports of average past completion times with completion times for the current project. In the Past condition, the average past completion time (*M* = 2.29, *SD* = 1.64)

was earlier (further from the deadline) than the "actual" completion time (M = 1.08, SD = 1.74), $t(23) = 3.02$, $p = .006$, $d = 0.69$. In the RCF deadline condition, the average past completion time (M = 1.85, SD = 1.76) was 0.43 days earlier than the "actual" completion time (M = 1.42, SD = 1.70), but this difference was not significant, $t(32) = 1.20$, $p = .237$, $d = 0.24$. If participants' completion times for the current project are representative of their typical completion times, then the results for the Past and RCF deadline condition suggest that participants may be underestimating a little bit just how close to the deadline they typically finish. In the RCF prediction condition, because we asked for past completion times in reference to predictions, we compared to prediction bias – this compares whether how much people say they tend to finish after they predicted is the same or different from how much later they finish from their current predictions. Participants had indicated that they tend to finish 0.80 days after their predictions and finished 0.30 days after their current prediction, a non-significant difference, $t(22) = 1.20$, $p = .245$, $d = 0.27$. Thus, although the means suggest a small reduction in the prediction bias compared to past projects, this difference was not statistically different. On the other hand, the bias for the current project was not significantly different from 0, $t(22) = 0.85$, $p = .404$, $d = 0.17$, but it was for previous projects.

Working Time, Work Time Bias. Participants did not differ by condition on the number of hours they reported having worked on the project (see Table 12, 13). Overall, participants overestimated the time they would spend working on the project by 3.70 hours (SD = 11.41), and this was significantly different from 0, $t(113) = 3.45$, $p = .001$, $d = 2.15$. This (overestimation) bias did not differ by condition.

Reasons for Finishing Later Than Predicted. Participants who finished their project later than expected did not differ by condition on the six reasons for finishing later than

predicted, all *F*s < 0.75, *p*s > .700. There was also no significant differences in item endorsement between items, $F(5, 555) = 0.84$, $p = .520$, $\eta_p^2 = .008$. Overall endorsement was as follows: underestimated how long it would take to complete each step needed to finish the project, $M = 3.99$ ($SD = 1.98$), failed to anticipate all the steps that would be required to complete the project, $M = 3.38$ ($SD = 1.79$), failed to anticipate project obstacles, $M = 3.72$ ($SD = 1.90$), didn't anticipate other demands on my time enough, $M = 3.87$ ($SD = 2.04$), priorities changed (this project became lower priority), $M = 3.68$ ($SD = 1.99$), my motivation declined more than I anticipated (e.g., procrastination), $M = 4.17$ ($SD = 2.02$).

Appendix C13

Power Analyses for Studies 3-6

Two types of power analyses are presented: 1) sensitivity analyses, i.e., given the achieved sample size and design, what effect size can we expect to detect, and 2) a priori power analyses, based on the size of effect we want to be able to detect and design, what sample size is required. In our case, both size of the effect and feasibility (students available for studies, cost) determined sample sizes.

Study #	**Power Analysis**		**Predictions** ANOVA/ANCOVA	**Predictions** Independent t-tests	**Prediction Bias** ANOVA/ANCOVA	**Prediction Bias** Independent t-tests
Study 3 (students)	Sensitivity	*N*	222	104 (comparing the two smallest *n*s)	154	47 (comparing the two smallest ns)
		Effect Size (power to detect)	f = 0.23 (medium)	d = 0.56 (medium)	f = 0.28 (medium)	d = 0.84 (large)
Study 4 (mturkers)	A Priori	*N*	280 (pay for 300, expect to exclude ~20)	140 (two groups)	196 (estimated 70% response rate)	98 (two groups)
		Effect Size (power to detect)	f = 0.20 (medium)	d = 0.48 (medium)	f = 0.25 (medium)	d = 0.58 (medium)
	Sensitivity	*N*	285	133 (comparing the two smallest *n*s)	152	72 (comparing the two smallest *n*s)
		Effect Size (power to detect)	f = 0.21 (medium)	d = .49	f = 0.28 (medium)	d = .67 (medium to large)
Study 5 (students, 2 groups missing)	A Priori	*N*	350 (estimate; recruit as many as possible in a term)	117 (two groups)	210 (estimated 60% response rate)	67 (two groups)

		Effect Size (power to detect)	f = 0.19 (small to medium)	d = 0.52 (medium)	f = 0.25 (medium)	d = 0.69 (medium to large)
	Sensitivity	*N*	268 (four groups)	130 (smallest two groups)	148 (four groups)	67 (smallest two groups)
		Effect Size (power to detect)	f = 0.20 (medium)	d = 0.50 (medium)	f = 0.28 (medium)	d = 0.70 (medium to large)
Study 6 (mturkers)	A Priori	*N*	500 (pay for 525, expect to exclude ~25)	167 (two groups)	350 (estimated 70% response rate)	116 (two groups)
		Effect Size (power to detect)	f = 0.16 (small to medium)	d = 0.44 (medium)	f = 0.19 (small to medium)	d = 0.52 (medium)
	Sensitivity	*N*	484	158 (smallest two groups)	273	84 (smallest two groups)
		Effect Size (power to detect)	f = 0.16 (small to medium)	d = 0.45 (medium)	f = 0.22 (medium)	d = 0.62 (medium)

Notes. Analyses were conducted at α = .05, power = 0.80, using G*Power (Erdfelder, Faul, & Buchner, 1996). Effect sizes were categorized via Cohen's (1998) suggested values.

Appendix D1

Effects of Condition and Descriptive Statistics for Exploratory Measures (Study 4)

Predicted Working Time

No differences between conditions emerged on the predicted number of working hours (see Table 16).

Factors Perceived to Influence Completion Predictions

Endorsement of the five factors that participants could have used to make or factor into their completion predictions did not differ by condition, with the exception of the past experience item for which the effect of condition was marginal, $F(3, 278) = 2.56$, $p = .055$, $\eta_p^2 = .027$. Here, we might expect participants in the three intervention conditions to give higher ratings than control participants, as they were instructed to consider similar past projects for making an accurate completion prediction. This expectation was not supported. Instead, participants in the RCF prediction condition ($M_{adj} = 5.63$, $SE = 0.15$) gave the lowest ratings, significantly lower than those in the Past condition ($M_{adj} = 6.18$, $SE = 0.14$), $p = .008$, $d = -0.45$, and marginally lower than those in the Control condition ($M_{adj} = 5.97$, $SE = 0.15$), $p = .088$, $d = -0.27$. In addition, ratings were marginally lower in the RCF deadline condition ($M_{adj} = 5.82$, $SE = 0.15$) than in the Past condition (which evidenced the highest mean rating), $p = .086$, $d = -0.30$. No other differences were significant.

There were no significant differences in item endorsement between items, $F(4, 1104) = 1.35$, $p = .251$, $\eta_p^2 = .005$. Overall endorsement was as follows: based prediction on when I want to be finished, $M = 5.53$ ($SD = 1.51$), based prediction on the steps needed to complete the project and how long they will take, $M = 5.81$ ($SD = 1.33$), based prediction on possible obstacles that may delay progress, $M = 5.38$ ($SD = 1.51$), based prediction on other demands on

my time, M = 5.58 (SD = 1.36), and based prediction on past experiences with similar projects, M = 5.90 (SD = 1.28).

Past Project Characteristics

The seven project characteristics were assessed both before and after the manipulation and thus repeated measures ANOVAs were conducted. No pre-post differences, main effects of condition, or (time x condition) interaction emerged for the difficulty or control item. On the importance item, both a within-subjects effect of time and time by condition interaction emerged (see Table 16). When examined, the difference in project importance between pre- (M_{adj} = 6.12, SE = 0.06) and post-manipulation (M_{adj} = 6.06, SE = 0.05) was very small (M_{diff} = 0.06, SE = 0.05) and not significant, p = .231. When broken down, the time by condition interaction was that participants in the Past and RCF deadline condition reported slightly lower levels of project importance post-manipulation (Past M_{adj} = 5.87, SE = 0.11; RCF deadline M_{adj} = 6.18, SE = 0.11), compared to pre-manipulation (Past M_{adj} = 6.08, SE = 0.12; RCF deadline M_{adj} = 6.42, SE = 0.12), p = .037 and .025, respectively. For busyness, both a main effect of time and marginal time by condition interaction emerged (see Table 16). The time effect was that participants gave slightly higher busyness ratings post-, M_{adj} = 5.36, SE = 0.08, compared to pre-manipulation, M_{adj} = 5.21, SE = 0.08, p = .030. Furthermore, this difference was largest, and reached significance only in the RCF prediction condition, pre-manipulation, M_{adj} = 4.92, SE = 0.16 vs. post-manipulation, M_{adj} = 5.30, SE = 0.16, p = .006, stats. For the "complete the project as quickly as possible" item, only the within-subjects effect of time was significant (see Table 16), but post-hoc tests revealed no differences between the means collapsed across condition, pre-manipulation M_{adj} = 5.66, SE = 0.06 vs. post-manipulation M_{adj} = 5.61, SE = 0.07, p = .457. For importance of the quality of the final product, only a marginal effect of condition emerged (see

Table 16), such that participants in the Past condition (M_{adj} = 6.09, *SE* = 0.10) gave slightly lower ratings than participants in the Control (M_{adj} = 6.40, *SE* = 0.10, *p* = .023), RCF deadline (M_{adj} = 6.36, *SE* = 0.10, *p* = .057), and RCF prediction (M_{adj} = 6.40, *SE* = 0.10, *p* = .030) conditions, and the other conditions did not differ from one another.

Confidence

In this study, participants' confidence in the accuracy of their completion predictions differed marginally by condition, with participants in the RCF prediction condition giving the highest confidence ratings, M_{adj} = 5.95, *SE* = 0.13, significantly higher than those in the RCF deadline condition, M_{adj} = 5.51, *SE* = 0.14, *p* = .023, *d* = 0.40, and control condition, M_{adj} = 5.55, *SE* = 0.12, *p* = .029, *d* = 0.38, but not significantly higher than those in the Past condition, M_{adj} = 5.68, *SE* = 0.12, *p* = .153, *d* = 0.26. No other differences were significant. Interestingly then, those participants who gave the least optimistic completion predictions also gave the highest confidence ratings, suggesting that participants may have been aware that they adjusted their predictions the most in this condition and that this adjustment would make their predictions more accurate.

Similarity of Past Projects to Current Project

These ratings were available only in the intervention conditions and were not expected to differ by condition given that past project nomination instructions were the same in the three conditions. Past projects did not differ by condition on rated similarity, with an overall mean of 5.54 (*SD* = 1.19) on the 7-point rating scale. Past projects were rated to be quite similar (a score of 4 = the same) to the current project in terms of complexity ($M_{overall}$ = 4.18, *SD* = 1.08), scope ($M_{overall}$ = 4.11, *SD* = 1.15), and time available to work on them ($M_{overall}$ = 4.06, *SD* = 1.32), and these ratings did not differ by condition. Past project relevancy for helping participants to make a

prediction for the current project did not differ by condition and was relatively high with an overall mean of 5.43 (*SD* = 1.33).

Past Projects Exercise Ratings

Participants found the exercise relatively straightforward ($M_{overall}$ = 5.53, *SD* = 1.30), useful ($M_{overall}$ = 5.29, *SD* = 1.43), and not particularly difficult ($M_{overall}$ = 3.67, *SD* = 1.76) or annoying ($M_{overall}$ = 2.83, *SD* = 1.76), and these ratings did not differ by condition (see Table 16). These results suggest that participants generally found reference class forecasting as reasonably easy to do and purposeful.

Personality Characteristics

The overall mean score on the 7-point propensity to plan scale was 5.22 (*SD* = 1.28) and did not differ by condition, $F(3, 278) = 0.13$, $p = .944$, $\eta_p^2 = .001$. Conscientiousness scale scores also did not differ by condition, $F(3, 278) = 0.33$, $p = .806$, $\eta_p^2 = .004$, with an overall mean of 5.56 (*SD* = 1.12).

Exploratory multiple regression analyses revealed that conscientiousness and propensity to plan did not interact with condition (dummy or contrast coding) to predict completion predictions, suggesting that they did not act as moderators in this study.

Opinions

Participants thought that making accurate completion predictions is relatively important, $M_{overall}$ = 5.83 (*SD* = 1.20) (on a 7-point scale), reported being relatively motivated to make accurate completion predictions, $M_{overall}$ = 5.89 (*SD* = 1.12) (on a 7-point scale), and thought that underestimating when one will finish a project is a bit more than moderately (a score of 4) problematic, $M_{overall}$ = 4.23 (*SD* = 2.03) (on a 7-point scale). These ratings did not differ by condition.

Appendix D2

Correlations between Key and Exploratory Variables (Study 4)

Part 1. Unlike in Study 3, the number of predicted working hours was not related to completion predictions, meaning it was not the case that the more hours participants expected to work on the project, the closer to the deadline it would be finished (see also Table 17). Participants who gave higher ratings on basing their prediction on when they want to be finished gave earlier predictions, as in Study 3. This time, there was no correlation between ratings of the extent to which participants based their prediction on past experiences and predictions. Unlike Study 3, ratings of how focused participants were on finishing as quickly as possible were now related to predictions in the expected direction (the more quickly participants said they wanted to finish, the earlier their predictions). Importance of, or motivation to make accurate completion predictions was unrelated to predictions.

Unlike the previous study, propensity to plan was not related to completion predictions in this study, $r = .06$, $p = .340$, but conscientiousness showed a small positive correlation with predictions as in Study 3, $r = .16$, $p = .007$, indicating that higher levels of conscientiousness were associated with slightly more optimistic (further before the deadline) completion predictions. Similar correlations emerged with mean past project completion times. In the Past condition, propensity to plan was not significantly related to past project completion times, $r = .16$, $p = .196$, and conscientiousness was significantly positively related to past project completion times (finishing earlier), $r = .26$, $p = .029$. In the RCF deadline condition propensity to plan and conscientiousness were not significantly correlated with past completion times, $r = .10$, $p = .414$, and $r = .13$, $p = .304$. Similarly, in the RCF prediction condition propensity to plan

was unrelated to past project completion times, $r = .00$, $p = 1.00$, and so was conscientiousness, $r = .14$, $p = .243$.

Part 2. Reported working time was not related to completion times again, perhaps suggesting that these predictions domains are tracking different types of information (see Table 20). The extent to which participants reported basing their predictions on when they ideally want to be finished was positively related to predictions and unrelated to completion times, indicating that people who based their prediction more on when they wanted to be finished thought that they would finish sooner. Greater confidence in having made an accurate completion prediction was associated with later (less optimistic) predictions, and unrelated to actual completion times.

Using the Part 2 sample, conscientiousness and propensity to plan were not related to predictions or completion times. (all rs < .11, ps > .200).

Appendix D3

Secondary Analyses for the Part 2 Sample (Study 4)

Past Completion Times, Predictions, Predicted Work Time. The main dependent variables for the Part 2 sample can be seen in Table 18. Past completion times and did not differ between the Past and RCF deadline conditions, $t(70) = -0.27$, $p = .790$, $d = -0.07$. The ANCOVA for predictions was statistically significant, $F(3, 146) = 3.24$, $p = .024$, $\eta_p^2 = .063$, and planned comparisons revealed differences between the control condition and each of the three intervention conditions, such that participants in the control condition made more optimistic (earlier) predictions than those in the intervention conditions. More specifically, the control condition ($M_{adj} = 2.14$, $SE = 0.29$) differed significantly from the Past condition ($M_{adj} = 0.89$, $SE = 0.30$), $t(77) = -2.77$, $p = .007$, $d = 0.48$, RCF deadline condition ($M_{adj} = 1.17$, $SE = 0.32$), $t(73) = -2.25$, $p = .028$, $d = 0.37$, and RCF prediction condition ($M_{adj} = 1.40$, $SE = 0.31$), $t(79) = -2.02$, $p = .047$, $d = 0.28$. The three intervention conditions did not differ significantly from one another (all ps > .200). As in the Part 1 sample, the number of predicted work hours did not differ by condition (see Table 18).

Comparing mean past completion times to actual completion times, in the Past condition, the average reported past completion time was 0.48 days ($SD = 194$) earlier (further from the deadline) than the "actual" completion time, but this difference was not significant, $t(38) = 1.56$, $p = .128$, $d = 0.30$. In the RCF deadline condition, the average reported past completion time was 0.66 ($SD = 2.68$) days earlier than the "actual" completion time, but this difference was also not significant, $t(32) = 1.42$, $p = .167$, $d = 0.32$. If participants' completion times for the current project are representative of their typical completion times, then the results for the Past and RCF deadline condition suggest that this sample of participants is generally consistent in when they

finish and accurate in their recall of how close to the deadline they typically finish. In the RCF prediction condition, because we asked for past completion times in reference to predictions, we compared to prediction bias – this compares whether how much people say they tend to finish before or after they predicted is the same or different from how much earlier or later they finish from their current predictions. Participants had reported that they tend to finish 1.03 days before their predictions, and for this project finished 0.47 days before their current prediction (M_{Diff} = 0.56, SD = 3.09), a non-significant difference, $t(38) = 0.76$, $p = .454$, $d = 0.18$, indicating that participants finished this project similarly early as past projects.

Working Time, Work Time Bias. Overall, participants reported spending on average 18.92 (SD = 19.61) hours working on their projects. The ANCOVA performed on this measure revealed a marginal effect of condition, $F(3, 146) = 2.15$, $p = .096$, $\eta_p^2 = .043$. Participants in the Past (M_{adj} = 21.38, SE = 3.10) and RCF deadline (M_{adj} = 23.93, SE = 3.38) conditions reported a greater number of working hours than control participants (M_{adj} = 13.18, SE = 3.03), $t(77) = 1.85$, $p = .068$, $d = 0.44$, and $t(70) = 2.45$, $p = .017$, $d = 0.55$, respectively, and the RCF prediction condition fell in between (M_{adj} = 18.24, SE = 3.14), not differing from any of the other conditions (all ps > .200). Overall, participants overestimated the time they would spend working on the project by 1.11 hours (SD = 18.30), but the bias score for this measure was not significantly different from 0, $t(151) = 0.75$, $p = .455$, $d = 0.06$, and did not differ significantly across conditions.

Reasons for Finishing Later Than Predicted. For participants who finished later than predicted, their ratings of the reasons why they finished late were each submitted to a one-way MANOVA. Ratings did not differ by condition for five out of the six reasons, all Fs < 2.00, ps > .120. A significant effect did emerge on the "underestimating how long it would take to complete

each project step" item, $F(3, 87) = 3.64$, $p = .016$, $\eta_p^2 = .114$, with the pattern being that participants in the control ($M_{adj} = 4.25$, $SE = 0.42$) and Past ($M_{adj} = 4.19$, $SE = 0.46$) condition gave higher ratings than participants in the RCF deadline ($M_{adj} = 2.59$, $SE = 0.44$) and RCF prediction ($M_{adj} = 2.99$, $SE = 0.47$). There was no significant differences in item endorsement between items, $F(5, 440) = 0.91$, $p = .476$, $\eta_p^2 = .010$. Overall endorsement was as follows: underestimated how long it would take to complete each step needed to finish the project, $M = 3.54$ ($SD = 2.20$), failed to anticipate all the steps that would be required to complete the project, $M = 3.00$ ($SD = 1.94$), failed to anticipate project obstacles, $M = 3.03$ ($SD = 1.92$), didn't anticipate other demands on my time enough, $M = 3.23$ ($SD = 1.96$), priorities changed (this project became lower priority), $M = 3.11$ ($SD = 2.10$), my motivation declined more than I anticipated (e.g., procrastination), $M = 2.76$ ($SD = 1.88$).

Appendix E1

Project Nomination Instructions

Welcome to the study!

For this study, we would like you to think of an **academic task or project** that you will be completing in the future.

This project needs to fit the following 5 criteria:

---

1) The project has to be completed some time in the next three weeks. That is, it has a firm **final deadline in the next three weeks**. **The deadline must** also **be 5 or more days away** (the project cannot be due in the next 5 days).

2) It is a project that you are **free to complete whenever you choose** within the next 3 weeks (e.g., writing an essay, studying for a test), rather than a project that can only be done at a specified time (e.g., attending a workshop to complete a certificate, taking a test).

3) You are **hoping to finish as soon as possible** (ideally, you would like to finish before the deadline).

4) It is a fairly **major project** that will require considerable time and effort to complete (i.e., not something that is simply part of your everyday routine).

5) It is a project that is specific and concrete, so that it would be easy to know exactly when it was completed (e.g., "complete 5-page psychology essay" not "make progress on biology research report").

Once you have thought of a task or project with the 5 characteristics listed above, please briefly describe it in the box below:

Next, please tell us the **final deadline** for the project using the date picker below:

If you are having trouble remembering the deadline, please feel free to take a minute to look it up via your calendar or planner, or use any other on-line resource.

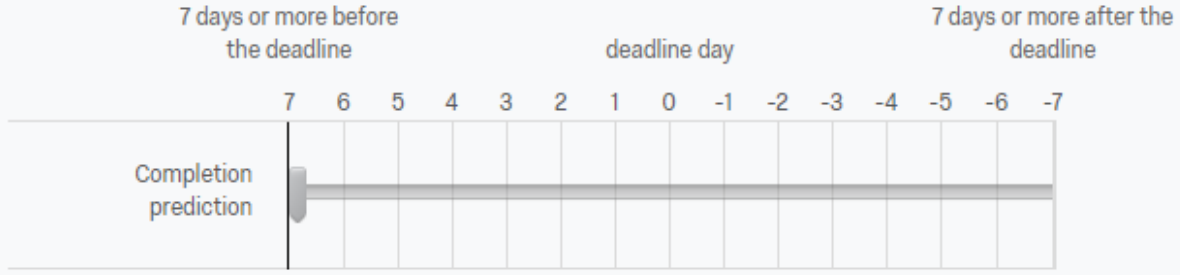
Appendix E2

Condition 1 (Control 1) – (Final) Prediction Instructions

We would now like you to predict, as accurately as possible, when you will be finished the project you just nominated.

For example, if you think you will finish the project **on the day of the deadline**, put the slider at **"0",** if you think you will finish the project **6 days early**, put the slider at **"6"**, if you think you will finish the project **3 days late** then put the slider at **"-3"**.

Using the slider below, give your prediction of when you expect to be finished the current project. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix E3

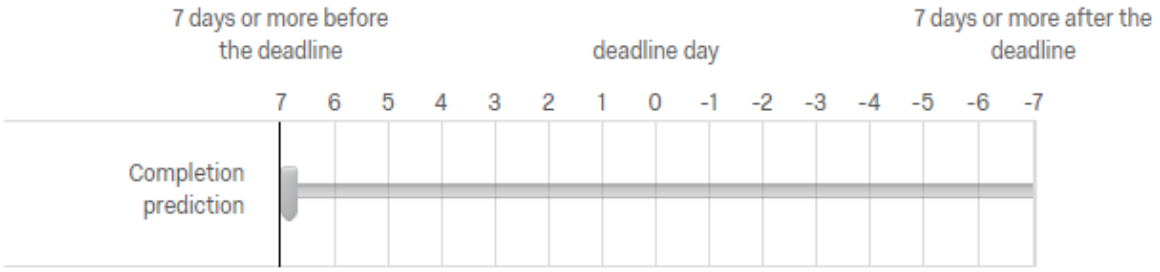
Initial Prediction for Condition 2, 3, 4 (Control 1, RCF 1, RCF 2); Final Prediction for Condition 1 - Instructions

We would now like you to predict, as accurately as possible, when you will be finished the project you just nominated.

Step 1 - Making an intuitive prediction

First, based on the project's characteristics and the time you have available, make an initial, intuitive prediction (i.e., your best guess) for when you think you will be finished the project.

For example, if you think you will finish the project **on the day of the deadline**, put the slider at **"0",** if you think you will finish the project **6 days early**, put the slider at **"6"**, if you think you will finish the project **3 days late** then put the slider at **"-3"**. Make sure you click on the slider, otherwise your data will not be recorded.

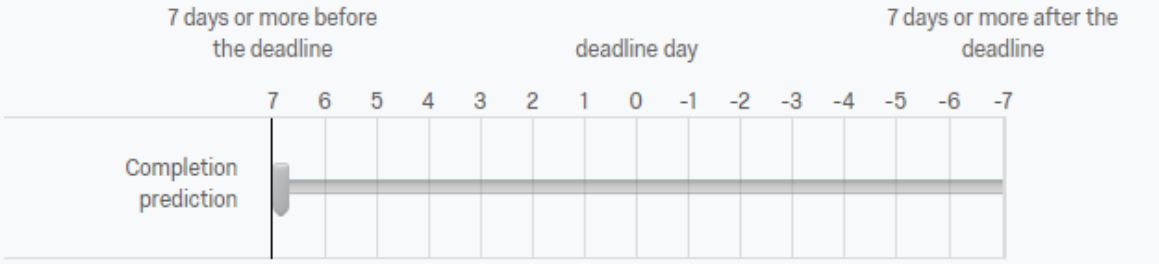


Step 2 - Final prediction

Now, we would like you to make your final prediction for when you will be finished the project.

Your final prediction can be the same as, or different from, your initial prediction. You may choose to change or keep your prediction the same for any reason, keeping in mind that you want it to be as accurate as possible.

Using the slider below, give your prediction of when you expect to be finished the current project. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix E4

RCF groups (RCF 1, 2, 3, 4) Past Projects Nomination, Completion Time Recall Instructions

Step 2 - Identify Similar Past Projects

One approach to making an accurate prediction is to think about similar past projects and compare when you thought you would be finished and when you were actually finished. We would like you to do this now. To assist you in the process, please complete the four steps that follow.

Next, we would like you to think of 5 to 6 projects that are as similar as possible to the one you just nominated, that you remember completing.

For example, these other projects may be similar in terms of project scope, type, complexity, amount of time you have to work on them, how much control you have over their outcome, etc.

In the boxes below, briefly (in 1-2 lines) describe each past project you thought of.

We want everyone to take at least a couple of minutes to think of these past projects, so the "continue" button to move to the next page will not appear until 2 minutes have elapsed.

Past Project 1

Past Project 2

Past Project 3

Past Project 4

Past Project 5

Past Project 6

Step 3 - Document Predictions & Outcomes for Past Projects

Next, for each project, think back and try to remember when you initially expected you would be finished. Then, try to remember when you actually finished the project and whether this was earlier or later than you had initially predicted.

Using the dropdowns, please indicate how many days before or after your initial prediction you actually finished the project. If you cannot remember exactly provide your best estimate.

For example, if you finished on the day you initially expected to finish then select **"0",** if you finished the project **3 days after you anticipated finishing** then select **"3 days later than predicted"**, if you finished the project **3 days before you anticipated finishing** then select **"3 days earlier than predicted"**.

↳ ${q://QID407/ChoiceTextEntryValue/1}

↳ ${q://QID407/ChoiceTextEntryValue/2}

↳ ${q://QID407/ChoiceTextEntryValue/3}

↳ ${q://QID407/ChoiceTextEntryValue/4}

↳ ${q://QID407/ChoiceTextEntryValue/5}

↳ ${q://QID407/ChoiceTextEntryValue/6}

Appendix E5

RCF Condition 1 with Mean and Instructions and Initial Predictions – Final Prediction Instructions

Step 4 - Correcting your initial prediction

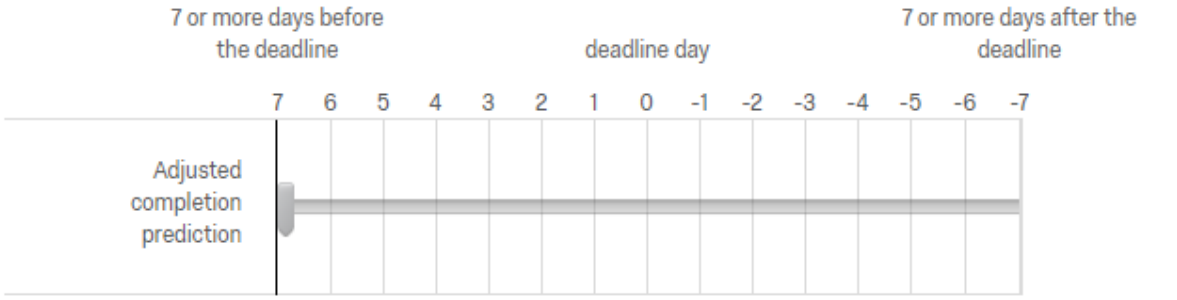
Adjust your initial prediction by using the information from the past projects you listed above.

Your **initial prediction** for finishing the project was **${q://QID414/ChoiceNumericEntryValue/1} days before the deadline**.

Based on your past projects, **on average you finish ${e://Field/Mean33} days after your prediction. Therefore, unless you believe that you have better information available to make your prediction for this project, you should base your project completion prediction on this average.**

If your average is finishing slightly later than predicted, then you should adjust your initial prediction to be somewhat later, if your average is finishing slightly earlier than predicted, then you should adjust your initial prediction to be somewhat earlier - unless you have reliable information to indicate that this project will be different.

Using the slider below, indicate your final, adjusted prediction for when you anticipate having completed the project. Make sure you click on the slider, otherwise your data will not be recorded.



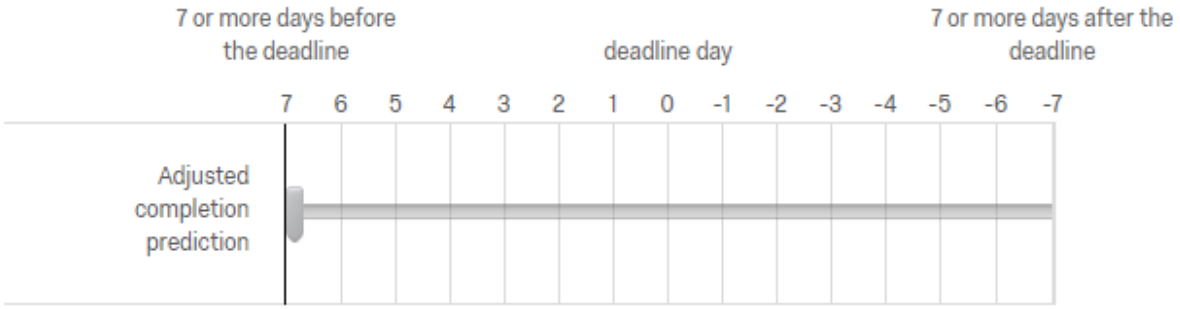
Appendix E6

RCF Condition 2 with Initial Prediction (No mean or instructions) – Final Prediction Instructions

Step 4 - Correcting your initial prediction

Adjust your initial prediction by using the information from the past projects you listed above to make a final prediction, keeping in mind that you want it to be as accurate as possible.

Using the slider below, indicate your final, adjusted prediction for when you anticipate having completed the project. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix E7

RCF Condition 3 with Mean and Instructions (no initial prediction) – Final Prediction Instructions

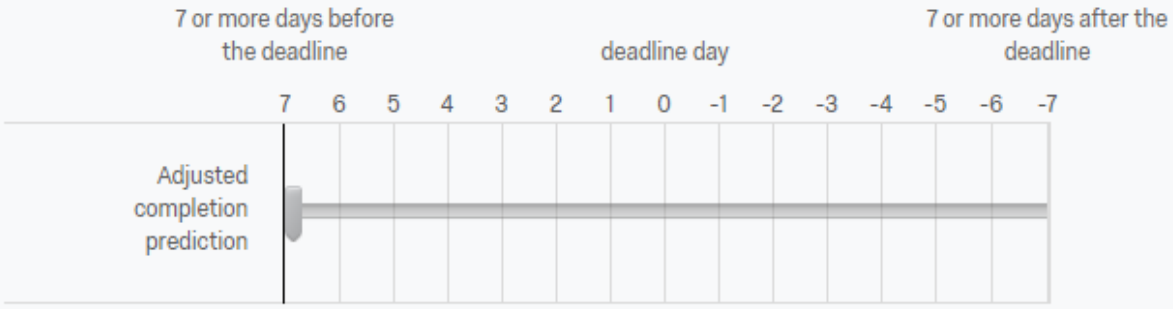
Step 3 - Final Prediction

Now, we would like you to make your prediction for when you will be finished the current project, keeping in mind that you want it to be as accurate as possible. Adjust your prediction using the information from the past projects you listed above.

Based on your past projects, **on average you finish $e{ abs(${e://Field/Mean44}) } days before your prediction.**

Therefore, unless you believe that you have better information available to make your prediction for this project than for your past projects, you should base your project completion prediction on this average. If your average is finishing slightly later than predicted, then you should adjust your prediction to be somewhat later, unless you have reliable information to indicate that this project will be different.

Using the slider below, indicate your final prediction for when you anticipate having completed the project. Make sure you click on the slider, otherwise your data will not be recorded.



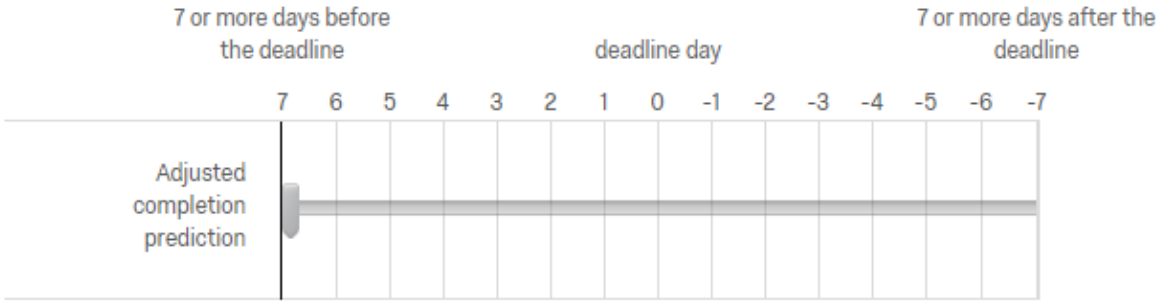
Appendix E8

RCF Condition 4 (no mean and instructions, no initial prediction) – Final Prediction Instructions

Step 3 - Final Prediction

Now, we would like you to make your prediction for when you will be finished the current project, keeping in mind that you want it to be as accurate as possible. Adjust your prediction using the information from the past projects you listed above.

Using the slider below, indicate your final prediction for when you anticipate having completed the project. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix E9

Control Condition (1, 2) Past Project Nomination and Completion Time Recall – Instructions

Recall Similar Past Projects

In this section we would like you to think of 5 to 6 projects that are as similar as possible to the one you just nominated, that you remember completing.

For example, these other projects may be similar in terms of project scope, type, complexity, amount of time you have to work on them, how much control you have over their outcome, etc.

In the boxes below, briefly (in 1-2 lines) describe each past project you thought of.

We want everyone to take at least a couple of minutes to think of these past projects, so the "continue" button to move to the next page will not appear until 2 minutes have elapsed.

Past Project 1

Past Project 2

Past Project 3

Past Project 4

Past Project 5

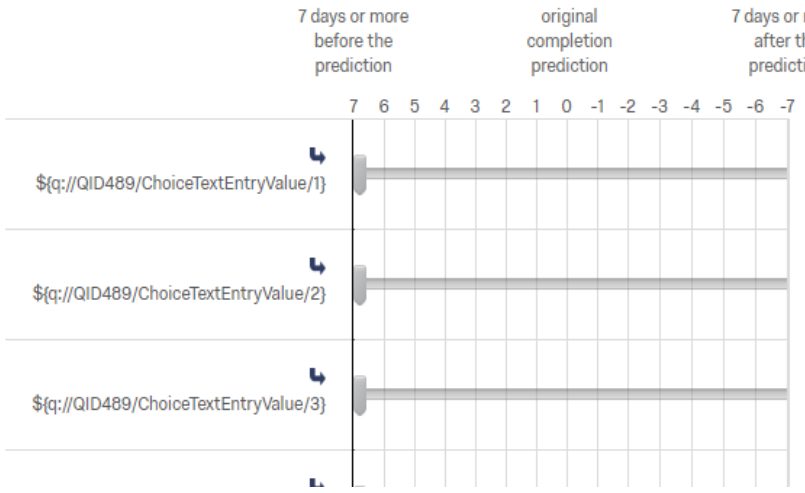
Past Project 6

Document Predictions & Outcomes for Past Projects

Next, for each project, think back to when you first formed an expectation (prediction) for when you would be finished and when you actually finished the project. If you cannot remember exactly, provide your best estimate.

Using the sliders, please indicate **how many days before or after your predicted completion date you actually finished.**

For example, if you finished on the day you initially expected to finish put the slider at **"0",** if you finished the project **6 days before you anticipated finishing**, put the slider at **"6"**, if you finished the project **3 days after** your initial prediction then put the slider at **"-3"**. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix E10

Part 2 Survey ("Actual" Completion Times)

In the first part of the study you nominated a task or project that you would be completing in the next three weeks. Please briefly describe what the task or project was in the box below.

The information we have from the P1 survey is: ${e://Field/Project3}

Based on the information we have from the Part 1 survey, the final deadline for this project was:

${e://Field/Deadline3}

Did this deadline change since the Part 1 survey?

- Yes
- No

Display This Question:
If Did this deadline change since the Part 1 survey? Yes Is Selected

What was the final deadline for this project?

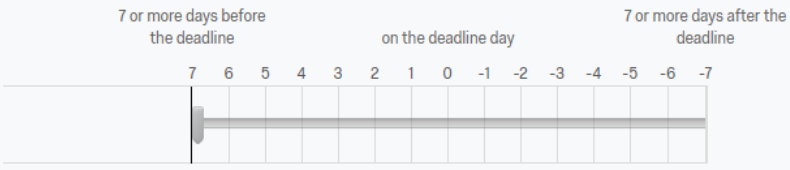
Did you finish this project?

- Yes
- No

Display This Question:
If Did you finish this project? Yes Is Selected

When did you finish the project?

Make sure you move the slider, otherwise your data will not be recorded!



Page Break

How long you did you spend working on the project itself? In the box below, please indicate **how many hours you spent working on the project.** Please enter a numeric value (for half hours enter .5).

Appendix E11

Effects of Condition and Descriptive Statistics for Exploratory Measures (Study 5)

Predicted Working Time

No differences between conditions emerged on the predicted number of working hours (see Table 22).

Factors Perceived to Influence Completion Predictions

Endorsement of the five factors that participants could have used to make or factor into their completion predictions did not differ by condition. Again, perhaps surprisingly, those in the RCF conditions who were explicitly instructed to recall and make use of their past, did not rate basing their predictions more on past experiences with similar projects than those in the control groups, and the ANOVA was not significant, $F(5, 395) = 0.08$, $p = .996$, $\eta_p^2 = .001$. There were significant differences in item endorsement between items, $F(4, 1008.972) = 13.19$, $p < .001$, $\eta_p^2 = .047$. Overall endorsement was as follows: based prediction on when I would ideally like to be finished, $M = 5.38$ ($SD = 1.53$), based prediction on the steps needed to complete the project and how long they will take, $M = 5.13$ ($SD = 1.48$), based prediction on possible obstacles that may delay progress, $M = 5.00$ ($SD = 1.66$), based prediction on other demands on my time, $M = 5.56$ ($SD = 1.48$), and based prediction on past experiences with similar projects, $M = 5.70$ ($SD = 1.36$). Adjusting for multiple comparisons ($\alpha = 0.05/20$ comparisons $= .003$), participants rated that they based their predictions more on their own past experiences with similar projects than obstacles ($p < .001$) and steps needed to complete the project ($p < .001$), and didn't significantly differ in their ratings from ideal finish time ($p = .007$), and other demands on their time ($p = .192$). Participants indicated that they based their predictions on other demands on their time more than obstacles ($p < .001$) and steps need to complete the project ($p < .001$), but not when they would ideally like to be finished ($p = .128$). They also rated that they based their predictions

more on when they would ideally like to be finished than obstacles (p = .002). No other comparison were significant (all ps > .020).

Personality Characteristics

The overall mean score on the 7-point propensity to plan scale was 5.00 (SD = 1.29) and did not differ by condition, $F(5, 395) = 0.59$, p = .622, η_p^2 = .007. Conscientiousness scale scores also did not differ by condition, $F(5, 395) = 0.27$, p = .845, η_p^2 = .003, with an overall mean of 4.59 (SD = 1.01).

Appendix E12

Correlations between Main and Exploratory Variables (Study 5)

Part 1. As in Study 4 (and unlike in Study 3), the number of predicted working hours was not related to completion predictions. Participants who indicated basing their predictions more on when they would ideally like to be finished and the steps needed to complete the project gave earlier predictions, as in Study 3 and 4. As in Study 4, there was no correlation between predictions and ratings of the extent to which participants based their predictions on past experiences, obstacles, or competing demands (see Table 25).

As in Study 3, propensity to plan was related to completion predictions in this study, r = .23, $p < .001$, and conscientiousness also showed a small positive correlation with predictions, r = .28, $p < .001$, indicating that higher levels of propensity to plan and conscientiousness were associated with slightly more optimistic (further before the deadline) completion predictions. Similarly, the correlation with past project completion times was negative, r = -.20, $p < .001$ for propensity to plan, r = -.24, $p < .001$, for conscientiousness, indicating that more conscientious, higher in propensity to plan participants recalled finishing projects earlier.

Part 2. Correlations between our additional variables of interest and Part 2 variables can be seen in Table 28.

Appendix E13

Secondary Analyses Part 2 (Study 5)

Past Completion Times, Predictions. The dependent variables for the Part 2 sample can be seen in Table 26. As in the Part 1 sample, past completion times differed by condition, $F(5, 218) = 7.27$, $p < .001$, $\eta_p^2 = .134$, with the same pattern of control participants reporting finishing earliest, RCF 3 and 4 reporting finishing latest, and RCF 1 and 2 falling in between. The ANOVA for predictions was not significant again, $F(3, 144) = 0.89$, $p = .443$, $\eta_p^2 = .018$. This time, although the pattern of means was the same as for Part 1, planned comparisons did not reveal any significant differences between the control condition 1 ($M = 1.83$, $SD = 1.52$) and the two intervention conditions, for condition 2, $M = 1.39$, $SD = 1.05$, $t(79) = 1.51$, $p = .136$, $d = 0.33$; for condition 4, $M = 1.30$, $SD = 1.47$, $t(67) = 1.37$, $p = .174$, $d = 0.35$. The two intervention conditions also did not differ significantly from one another or from control condition 2 (all ps > .500). Comparing the pooled RCF ($M = 1.35$, $SD = 1.33$) and control ($M = 1.67$, $SD = 1.58$) groups, the difference was no longer marginally significant, $t(148) = -1.33$, $p = .187$, $d = -0.22$.

Comparing recalled mean past completion times to actual completion times, because we asked for past completion times in reference to predictions, we compared to prediction bias – this compares whether how much people say they tend to finish after they predicted is the same or different from how much later they finish from their current predictions. Participants had indicated that they tend to finish 0.57 days ($SD = 1.33$) before their predictions and reported, for this project, finishing 0.44 days ($SD = 1.53$) after their predictions, a significant difference, $t(147) = -6.14$, $p < .001$, $d = -0.70$, indicating that participants finished this project slightly later than they recalled completing past projects. This difference differed by condition, $F(3, 144) = 6.30$, $p < .001$, $\eta_p^2 = .116$, with participants in the control groups, $M_{Control1} = 1.70$ ($SD = 2.11$),

$M_{\text{Control 2}}$ = 1.42 (*SD* = 1.87), reporting completing their current project later than they recalled finishing in the past compared to the RCF 4 group, M_{RCF4} = -0.14 (*SD* = 2.26), $t(67)_{\text{RCF4vs.Control1}}$ = 3.47, *p* = .001, *d* = 0.84, and $t(65)_{\text{RCF4vs.Control2}}$ = 3.10, *p* = .003, *d* = 0.75, respectively. The RCF group 2, M_{RCF2} = 0.73 (*SD* = 1.27), fell in between, reporting completing the project marginally earlier than those in the control group 1, *t*(67.78) = 2.50, *p* = .015, *d* = 0.55, marginally earlier than those in the control group 2, *t*(64.62) = 1.90, *p* = .062, *d* = 0.43, and marginally later than those in RCF group 4, *t*(4.51) = 1.88, *p* = .067, *d* = 0.49.

Working Time, Work Time Bias. Overall, participants reported spending on average 11.85 (*SD* = 8.92) hours working on their projects, and these did not differ by condition, *F*(5, 218) = 1.14, *p* = .342, η_p^2 = .025. They overestimated the time they would spend working on the project by 3.11 hours (*SD* = 7.35), and the bias score was significantly different from 0, *t*(223) = 6.37, *p* < .001, *d* = 0.37, but did not differ significantly across conditions, *F*(5, 218) = 0.70, *p* = .625, η_p^2 = .016.

Appendix F1
Project Nomination Instructions

Nominate Your Project

Welcome to the study! In this section we would like you to think of a specific work, personal, or academic task or project that you will be completing in the future. This project needs to fit the following 5 criteria:

1) The project has to be completed some time in the next three weeks. That is, it has a firm **final deadline in the next 3 weeks**.

2) The **deadline must be 5 or more days away** (the project cannot be due within the next 5 days).

2) It is a project that you are **free to complete whenever you choose** within the next 3 weeks (e.g., writing a report, building a shed, studying for a test), rather than a project that can only be done at a specified time (e.g., attending a workshop to complete a certificate, moving into a new apartment on a specific day, taking a test).

3) You are **hoping to finish as soon as possible.**

4) It is a fairly **major project** that will require considerable time and effort to complete (i.e., not something that is simply part of your everyday routine) and has multiple steps or parts.

5) It is a project that is specific and concrete, so that it would be easy to know exactly when it was completed (e.g., "complete 10-page inventory report", NOT "make progress on my home renovation project").

Once you have thought of a task or project that meets these 5 criteria, please briefly describe it in the box below:

Next, please tell us the **final deadline** for the project using the date picker below:

If you are having trouble remembering the deadline, please feel free to take a minute to look it up via your calendar or planner, or use any other on-line resource.

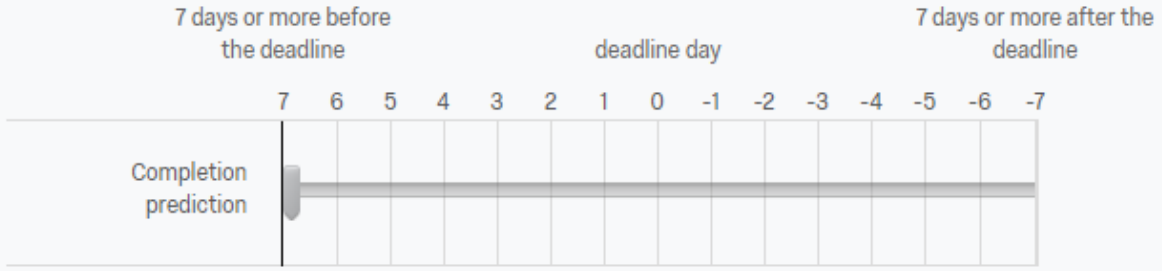
Appendix F2

Condition 1 (Control 1) – (Final) Prediction Instructions

Make a Completion Prediction

We would now like you to predict, as accurately as possible, when you will be finished the project you just nominated.

For example, if you think you will finish the project **on the day of the deadline**, put the slider at **"0",** if you think you will finish the project **6 days early**, put the slider at **"6"**, if you think you will finish the project **3 days late** then put the slider at **"-3"**. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix F3

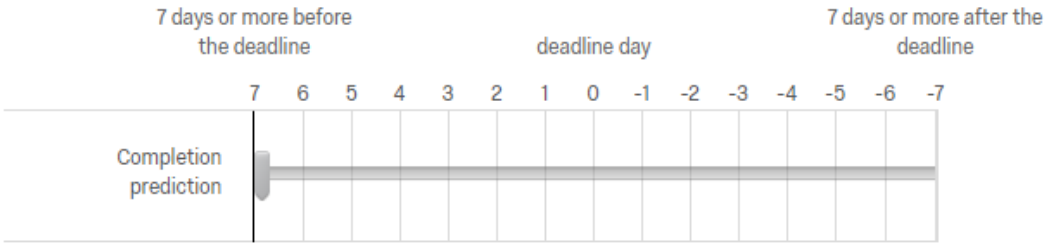
Initial Prediction for Condition 2, 3, 4 (Control 2, RCF 1, 2) and Final Prediction for Condition 2 (Control 2) - Instructions

We would like you to predict, as accurately as possible, when you will be finished the project you just nominated.

Step 1 - Making an intuitive prediction

First, based on the project's characteristics and the time you have available, make an initial, intuitive prediction (i.e., your best guess) for when you think you will be finished the project.

For example, if you think you will finish the project **on the day of the deadline**, put the slider at **"0",** if you think you will finish the project **6 days early**, put the slider at **"6"**, if you think you will finish the project **3 days late** then put the slider at **"-3"**. Make sure you click on the slider, otherwise your data will not be recorded.



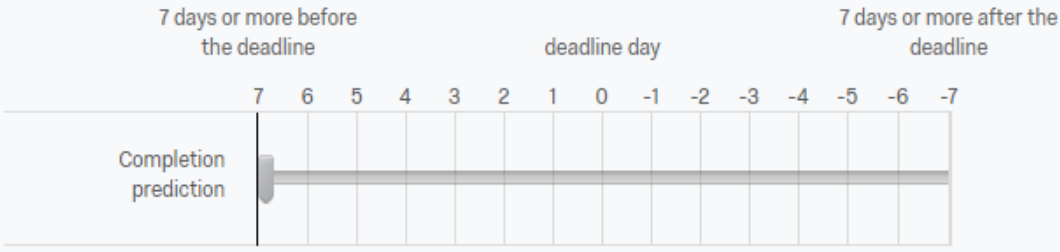
Page Break

Step 2 - Final prediction

Now, we would like you to make your final prediction for when you will be finished the project.

Your final prediction can be the same as, or different from, your initial prediction. You may choose to change or keep your prediction the same for any reason, keeping in mind that you want it to be as accurate as possible.

Using the slider below, give your prediction of when you expect to be finished the current project. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix F4

RCF Conditions (RCF 1, 2, 3, 4) – Past Projects Nomination, Completion Time Recall Instructions

Step 2 - Recall Similar Past Projects

One approach to making an accurate prediction is to think about similar past projects and compare when you thought you would be finished and when you were actually finished.

Think of 5 to 6 projects that are as similar as possible to the one you just nominated, that you remember completing in the past.

For example, these other projects may be similar in terms of project scope, type, complexity, amount of time you have to work on them, how much control you have over their outcome.

In the boxes below, briefly (in 1-2 lines) describe each past project you thought of. This part is very important - the more projects you can remember the better - input AT LEAST 4 past projects!

Past Project 1

Past Project 2

Past Project 3

Past Project 4

Past Project 5

Past Project 6

Step 3 - Recalling Completion Times for Past Projects (in reference to predictions)

Next, for each project, think back and try to remember when you initially expected you would be finished. Then, try to remember when you actually finished the project and whether this was earlier or later than you had initially predicted.

Using the dropdowns, please indicate how many days before or after your initial prediction you actually finished the project. If you cannot remember exactly provide your best estimate.

For example, if you finished on the day you initially expected to finish then select **"0"**, if you finished the project **3 days after you anticipated finishing** then select **"3 days later than predicted"**, if you finished the project **3 days before you anticipated finishing** then select **"3 days earlier than predicted"**.

↳ ${q://QID407/ChoiceTextEntryValue/1}	7 or more days later than predicted ▼
↳ ${q://QID407/ChoiceTextEntryValue/2}	3 days later than predicted ▼
↳ ${q://QID407/ChoiceTextEntryValue/3}	0 Same day as predicted ▼
↳ ${q://QID407/ChoiceTextEntryValue/4}	1 day earlier than predicted ▼
↳ ${q://QID407/ChoiceTextEntryValue/5}	4 days earlier than predicted ▼
↳ ${q://QID407/ChoiceTextEntryValue/6}	7 or more days earlier than predicted ▼

Appendix F5

Condition 3 (RCF 1) with Mean and Instructions and Initial Predictions – Final Prediction Instructions

Step 4 - Correcting your initial prediction

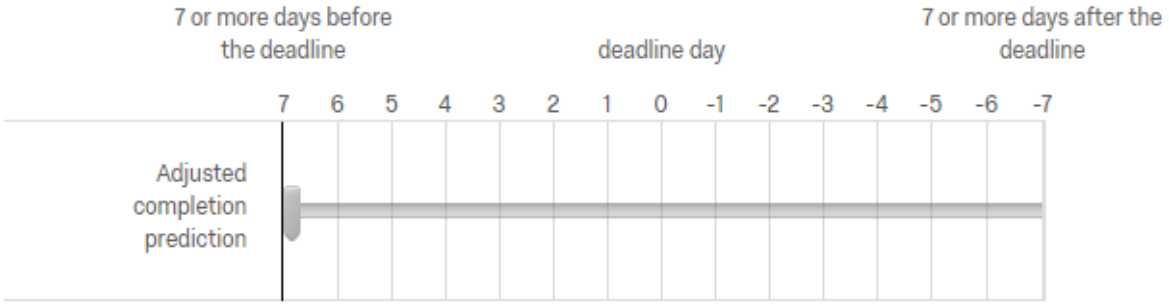
Adjust your initial prediction by using the information from the past projects you listed above.

Your **initial prediction** for finishing the project was **${q://QID414/ChoiceNumericEntryValue/1} days before the deadline**.

Based on your past projects, **on average you finish ${e://Field/Mean33} days after your prediction. Therefore, unless you believe that you have better information available to make your prediction for this project, you should base your project completion prediction on this average.**

If your average is finishing slightly later than predicted, then you should adjust your initial prediction to be somewhat later, if your average is finishing slightly earlier than predicted, then you should adjust your initial prediction to be somewhat earlier - unless you have reliable information to indicate that this project will be different.

Using the slider below, indicate your final, adjusted prediction for when you anticipate having completed the project. Make sure you click on the slider, otherwise your data will not be recorded.



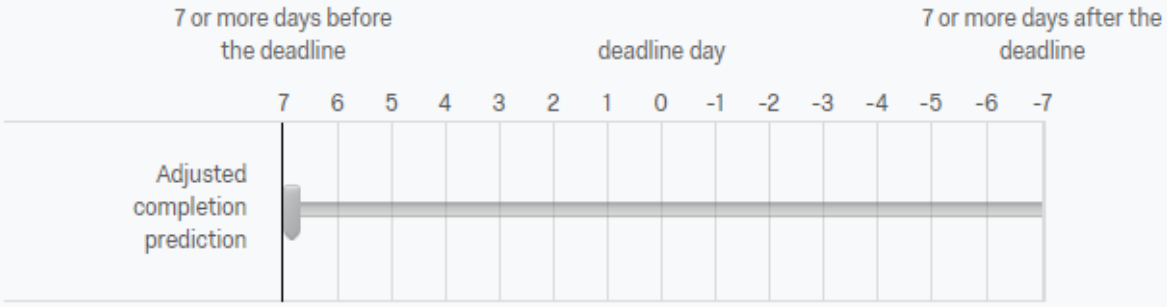
Appendix F6

Condition 4 (RCF 2) with Initial Prediction (No mean or instructions) – Final Prediction Instructions

Step 4 - Correcting your initial prediction

Adjust your initial prediction by using the information from the past projects you listed above to make a final prediction, keeping in mind that you want it to be as accurate as possible.

Using the slider below, indicate your final, adjusted prediction for when you anticipate having completed the project. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix F7

Condition 5 (RCF 3) with Mean and Instructions (no initial prediction) – Final Prediction Instructions

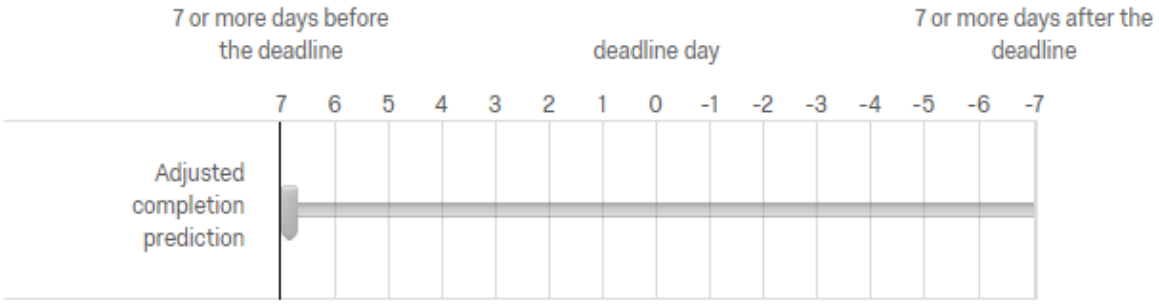
Step 3 - Final Prediction

Now, we would like you to make your prediction for when you will be finished the current project, keeping in mind that you want it to be as accurate as possible. Adjust your prediction using the information from the past projects you listed above.

Based on your past projects, **on average you finish $e{ abs(${e://Field/Mean44}) } days before your prediction. Therefore, unless you believe that you have better information available to make your prediction for this project than for your past projects, you should base your project completion prediction on this average.**

If your average is finishing slightly earlier than predicted, then you should adjust your initial prediction to be somewhat earlier, if your average is finishing slightly later than predicted, then you should adjust your initial prediction to be somewhat later - unless you have reliable information to indicate that this project will be different.

Using the slider below, indicate your final prediction for when you anticipate having completed the project. Make sure you click on the slider, otherwise your data will not be recorded.



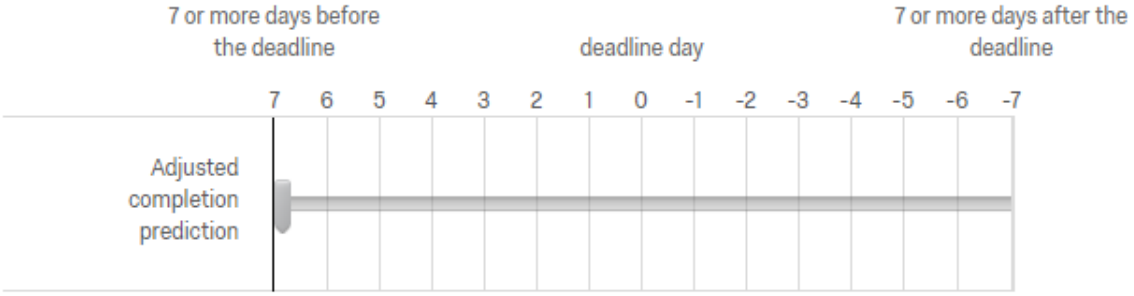
Appendix F8

Condition 6 (RCF 4) (no mean and instructions, no initial prediction) – Final Prediction Instructions

Step 3 - Final Prediction

Now, we would like you to make your prediction for when you will be finished the current project. Adjust your prediction using the information from the past projects you listed above.

For example, if you think you will finish the project **on the day of the deadline**, put the slider at **"0",** if you think you will finish the project **6 days early**, put the slider at **"6"**, if you think you will finish the project **3 days late** then put the slider at **"-3"**. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix F9

Conditions 1, 2 (Control 1, 2) Past Project Nomination and Completion Time Recall – Instructions

Recall Similar Past Projects

Think of 5 to 6 projects that are as similar as possible to the one you just nominated, that you remember completing in the past.

For example, these other projects may be similar in terms of project scope, type, complexity, amount of time you have to work on them, how much control you have over their outcome.

In the boxes below, briefly (in 1-2 lines) describe each past project you thought of. This part is very important - the more projects you can remember the better - input AT LEAST 4 past projects!

Past Project 1

Past Project 2

Past Project 3

Past Project 4

Past Project 5

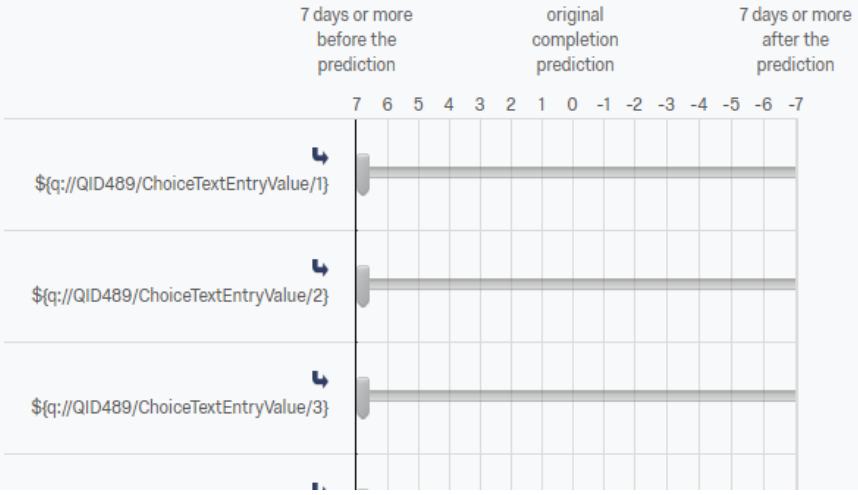
Past Project 6

Recalling Completion Times for Past Projects (in reference to predictions)

Next, for each project, think back to when you first formed an expectation (prediction) for when you would be finished and when you actually finished the project. If you cannot remember exactly, provide your best estimate.

Using the sliders, please indicate **how many days before or after your predicted completion date you actually finished.**

For example, if you finished on the day you initially expected to finish put the slider at **"0",** if you finished the project **6 days before you anticipated finishing**, put the slider at **"6"**, if you finished the project **3 days after** your initial prediction then put the slider at **"-3"**. Make sure you click on the slider, otherwise your data will not be recorded.



Appendix F10

Part 2 Survey ("Actual" Completion Times)

1. In the first part of the study you nominated a task or project that you would be completing in the next three weeks. Please briefly describe what the task or project was in the box below.

The information we have from the P1 survey is: ${e://Field/Project3}

2. Based on the information we have from the Part 1 survey, the final deadline for this project was:

${e://Field/Deadline3}

Did this deadline change since the Part 1 survey?

- Yes
- No

If the deadline has changed, or if you did not see a deadline displayed above, please type the final deadline for the project in the box below.

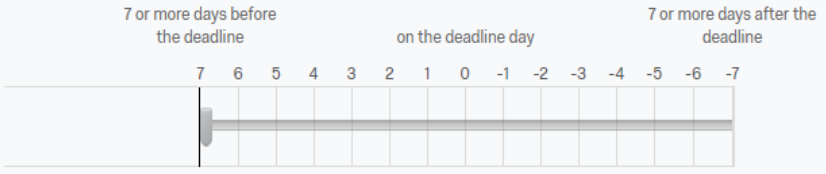
3. Did you finish this project?

- Yes
- No

Display This Question:
If 3. Did you finish this project? Yes Is Selected

When did you finish the project?

Make sure you move the slider, otherwise your data will not be recorded!



Page Break

How long you did you spend working on the project itself? In the box below, please indicate **how many hours you spent working on the project.** Please enter a numeric value.

Appendix F11

Secondary Analyses (Study 6)

Part 1.

***Exploratory measures*.** The predicted number of working hours differed depending on experimental group, $F(5, 477) = 3.13$, $p = .009$, $\eta_p^2 = .032$ (see Table 29). The pattern of results that emerged was that control 1 participants ($M = 24.22$, $SD = 8.76$) predicted working on their project longer than all of the RCF conditions ($M_{Combined} = 20.23$, $SD = 9.85$, all ps $< .035$).

The number of predicted working hours was weakly related to completion predictions, $r = -.10$, $p = .028$, suggesting that earlier completion predictions were associated with slightly fewer predicted working hours.

Part 2.

Past Completion Times, Predictions. Descriptive and test statistics for the Part 2 dependent variables can be seen in Table 34. Reports of past completion times no longer differed by condition in the Part 2 sample, $F(5, 267) = 1.62$, $p < .155$, $\eta_p^2 = .025$. The ANOVA for predictions remained significant $F(5, 267) = 2.46$, $p = .034$, $\eta_p^2 = .044$. The pattern of means was similar as for Part 1. Predictions were the earliest in control group 1 ($M = 2.15$, $SD = 2.16$), followed by control group 2 ($M = 1.67$, $SD = 1.70$). Predictions in RCF 1 ($M = 1.28$, $SD = 1.83$), RCF 3 ($M = 0.88$, $SD = 2.36$) and RCF 4 ($M = 1.17$, $SD = 1.57$) predictions were significantly later than those of the control group 1, $t(90) = 2.08$, $p = .040$, $d = 0.43$, $t(95) = 2.75$, $p = .007$, $d = 0.56$, $t(90) = 2.49$, $p = .015$, $d = 0.51$, respectively. RCF group 2 ($M = 1.63$, $SD = 1.98$) predictions didn't differ from those in control group 1, $t(90) = 1.16$, $p = .249$, $d = 0.25$. The four RCF conditions did not differ significantly from one another (all ps $> .100$), or from control group 2 (all ps $> .200$), with the exception of the RCF 3 group which had marginally later

completion predictions than the control 2 group, $t(92) = 1.83$, $p = .070$, $d = 0.38$. The pooled RCF ($M = 1.22$, $SD = 1.97$) and control ($M = 1.92$, $SD = 1.96$) groups differed significantly, $t(271) = 2.75$ $p = .006$, $d = 0.35$.

The 2 (Initial prediction) by 2 (Mean and Instruction) ANOVA revealed no main effect of Initial Predictions, $F(1, 180) = 0.05$, $p = .816$, $\eta_p^2 < .001$, Mean and Instructions, $F(1, 180) = 0.05$, $p = .816$, $\eta_p^2 < .001$, nor a significant interaction, $F(1, 180) = 0.48$, $p = .488$, $\eta_p^2 = .003$.

Comparing recalled mean past completion times to prediction bias, overall, participants indicated that they tended to finish 0.09 days ($SD = 1.92$) before their predictions (i.e., the same day as predicted) and reported finishing this project on average, 0.14 days ($SD = 2.28$) before their predictions, a non-significant difference, $t(272) = 0.28$, $p = .778$, $d = -0.02$, indicating that overall, participants finished this project approximately when they indicated finishing past projects. These differences did differ by condition, $F(5, 267) = 2.56$, $p = .028$, $\eta_p^2 = .046$. Participants in control group 1 ($M = 1.16$. $SD = 2.70$) reported finishing their current project later than they recalled finishing in the past compared to each of the RCF groups ($M_{combined} = -0.48$. SD = 3.26), all ps < .050. The RCF groups did not differ significantly from one another, all ps > .300. The RCF groups did not differ from control group 2 ($M = 0.47$. $SD = 2.97$), all ps < .100, except for RCF 3 ($M = -0.85$. $SD = 2.82$), $t(87) = 2.15$, $p = .034$, $d = 0.45$. The two control groups did not differ, $t(87) = 1.15$, $p = .254$, $d = 0.24$.

Working Time, Work Time Bias. Overall, participants reported spending on average 33.76 hours ($SD = 44.32$) working on their projects, and these did not differ by condition, $F(5, 272) = 1.02$, $p = .405$, $\eta_p^2 = .019$. Unlike in the previous studies, they <u>underestimated</u> the time they would spend working on the project, by 13.21 hours ($SD = 41.44$), a difference that was

significantly different from 0, $t(272) = -5.25$, $p < .001$, $d = 0.32$, and did not differ across conditions, $F(5, 27) = 0.67$, $p = .644$, $\eta_p^2 = .013$.

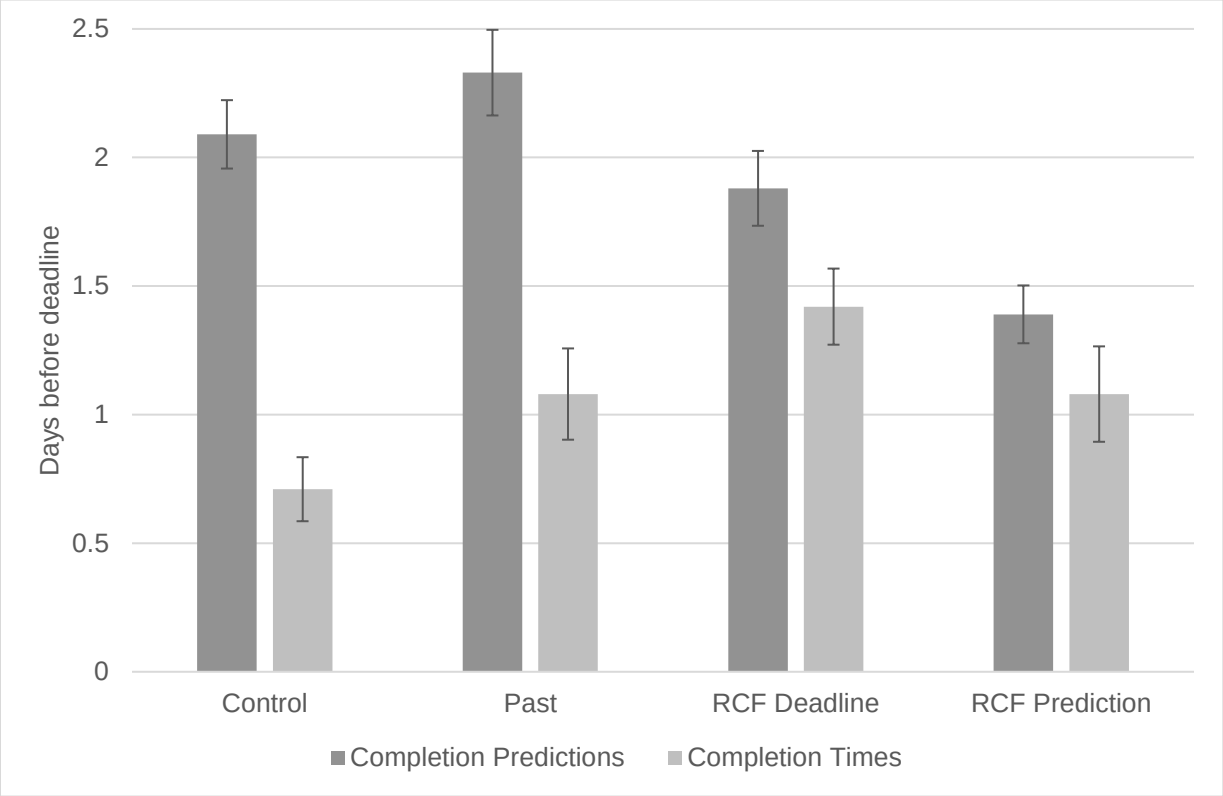


Figure 1. Bar graph showing differences between completion predictions and completion times by condition for Study 3 (Part 2).

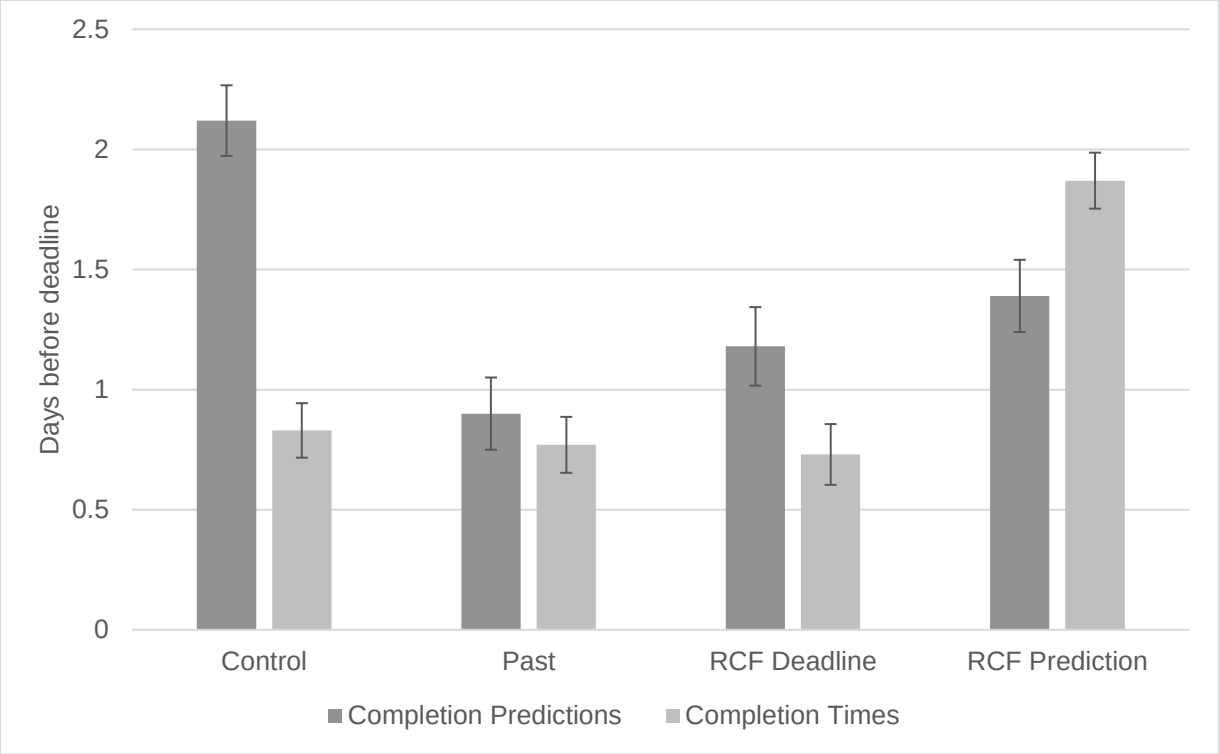


Figure 2. Bar graph showing differences between completion predictions and completion times by condition for Study 4 (Part 2).

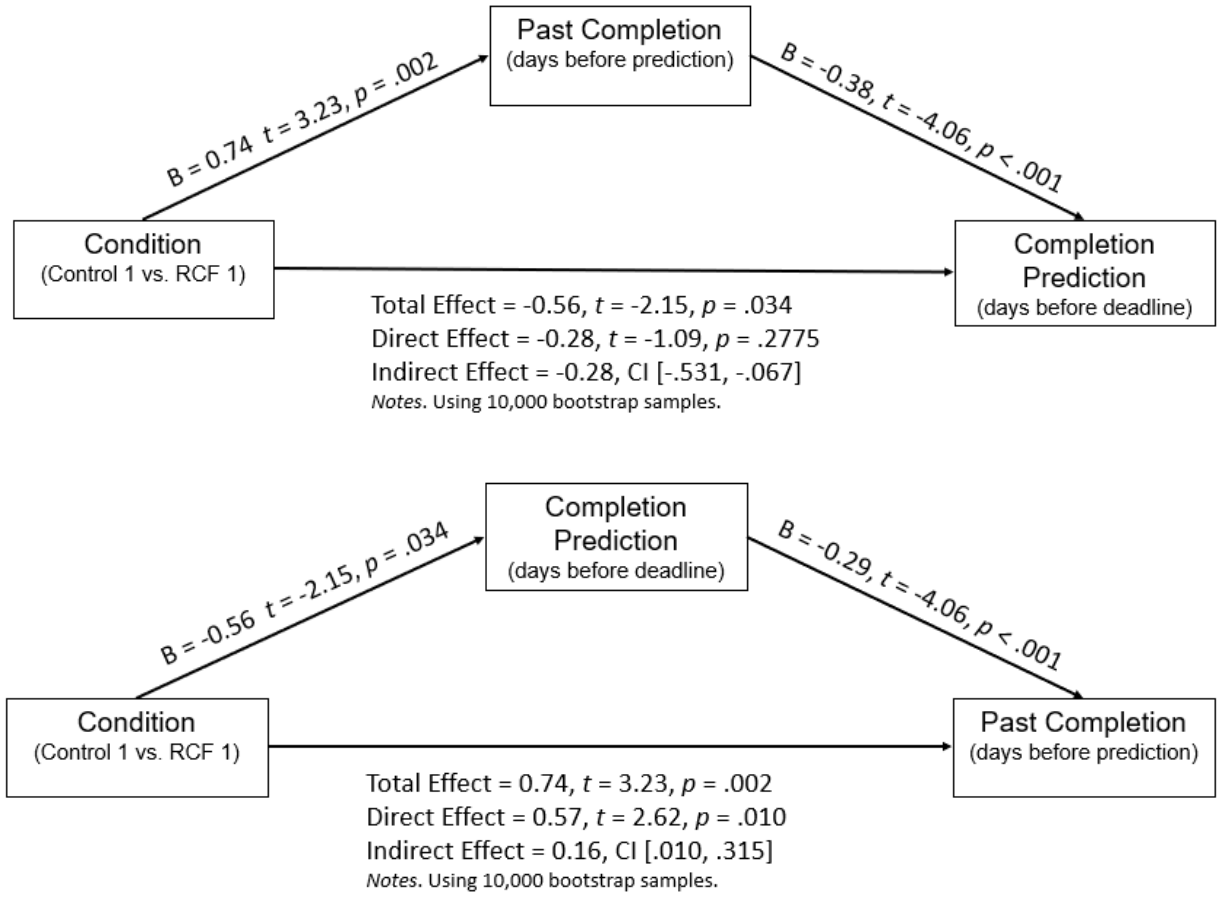


Figure 3. Mediation models testing whether the effect of the RCF 1 procedure (vs. Control 1) leads to later completion predictions through past completion times (top) or vice versa (bottom).

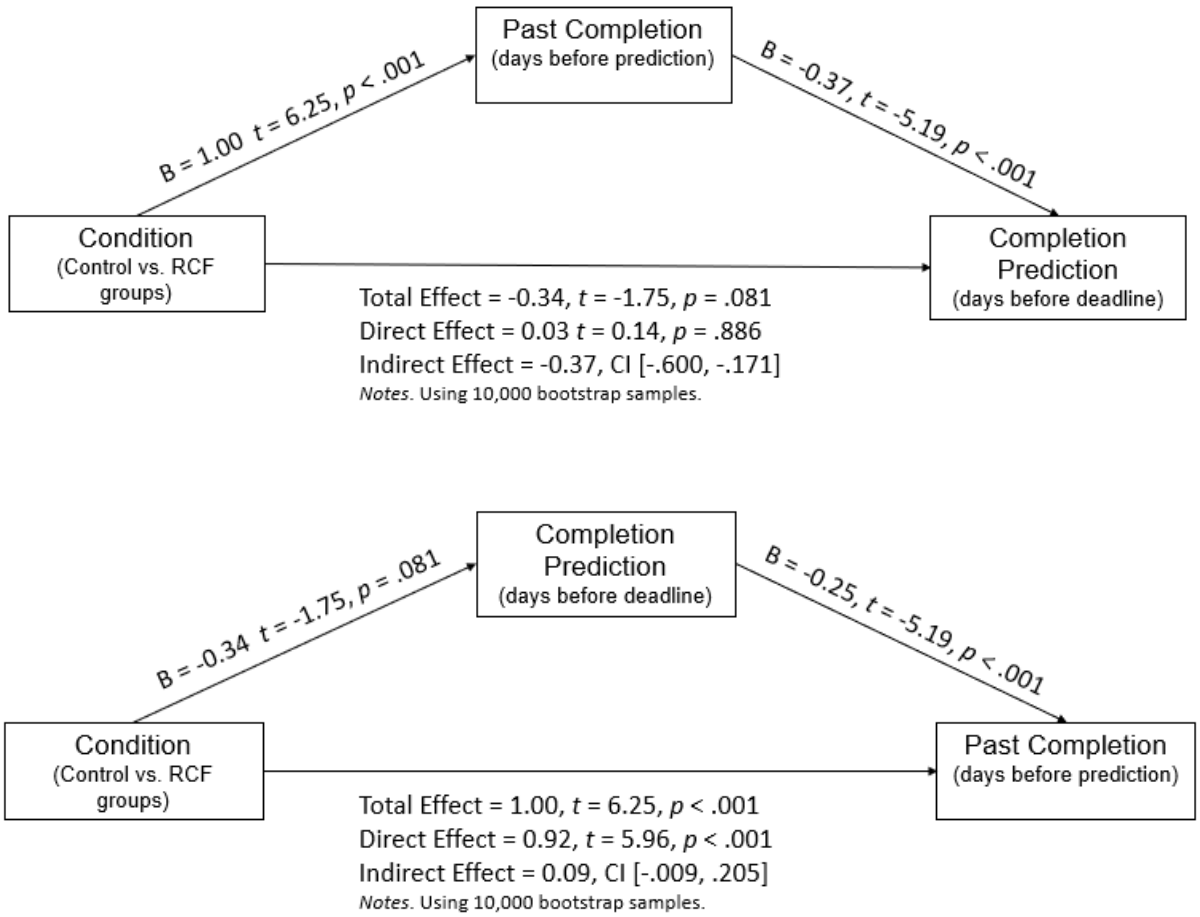


Figure 4. Mediation models testing whether the effect of the RCF groups combined (vs. control groups) leads to later completion predictions through past completion times (top) or vice versa (bottom).

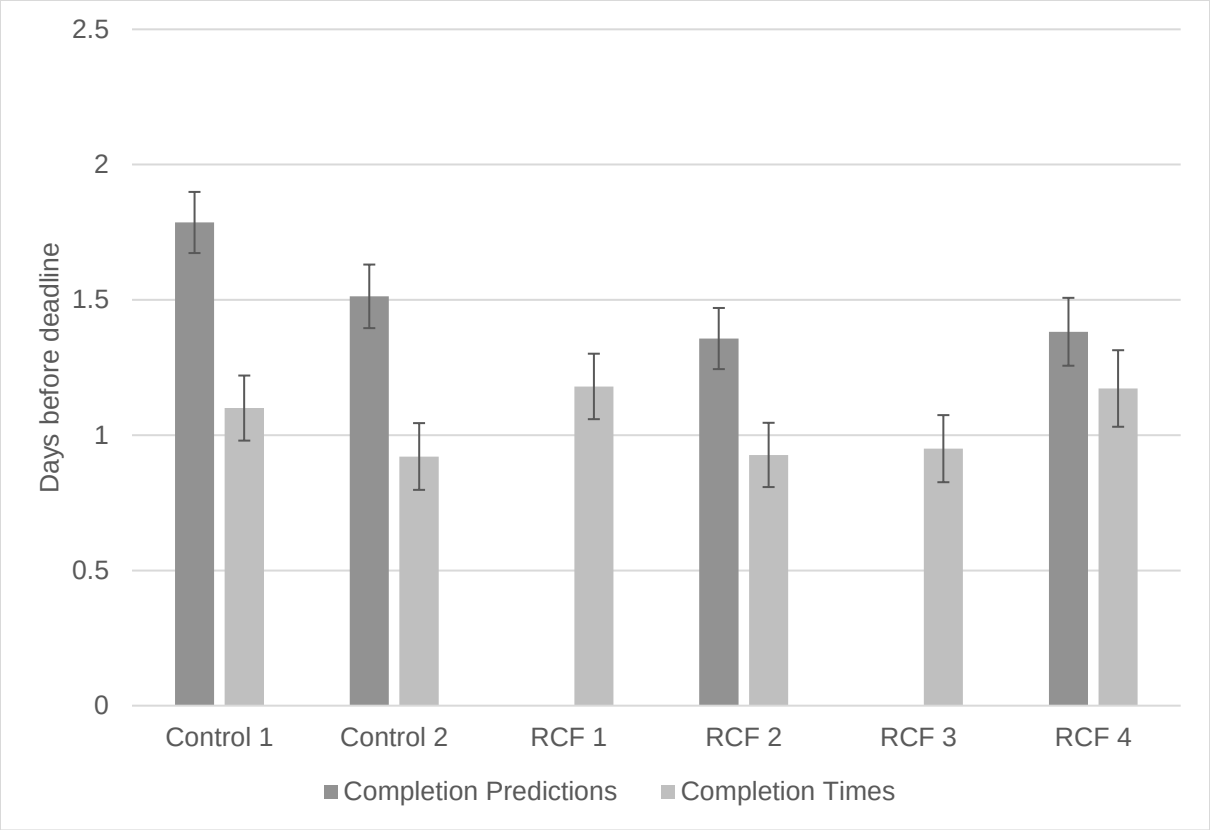


Figure 5. Bar graph showing differences between completion predictions and completion times by condition for Study 5 (Part 2).

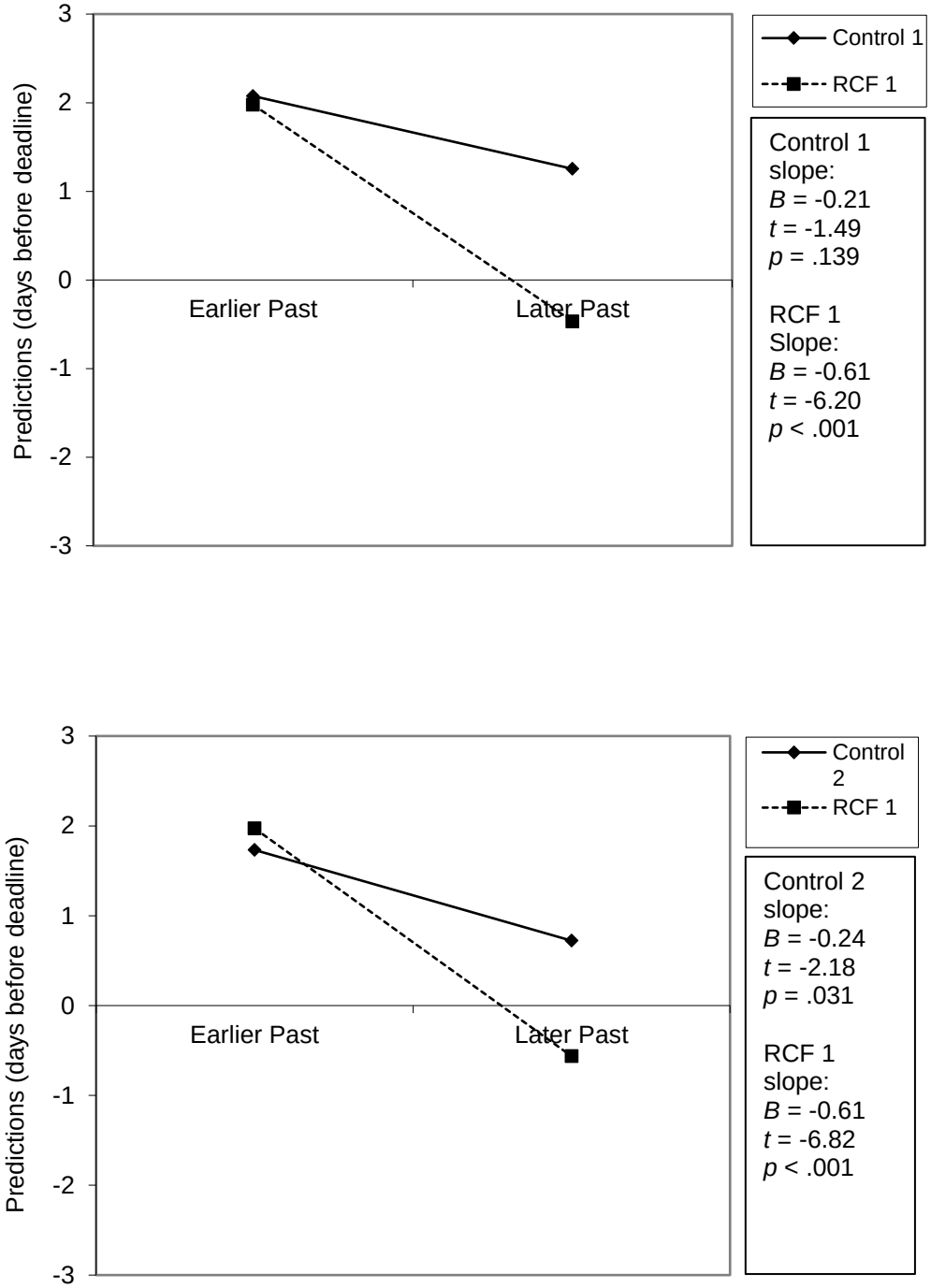


Figure 6. Moderation results for RCF 1 vs. Control 1 and Control 2 – past completion times and predictions are more related in the RCF 1 than the control groups (Study 6).

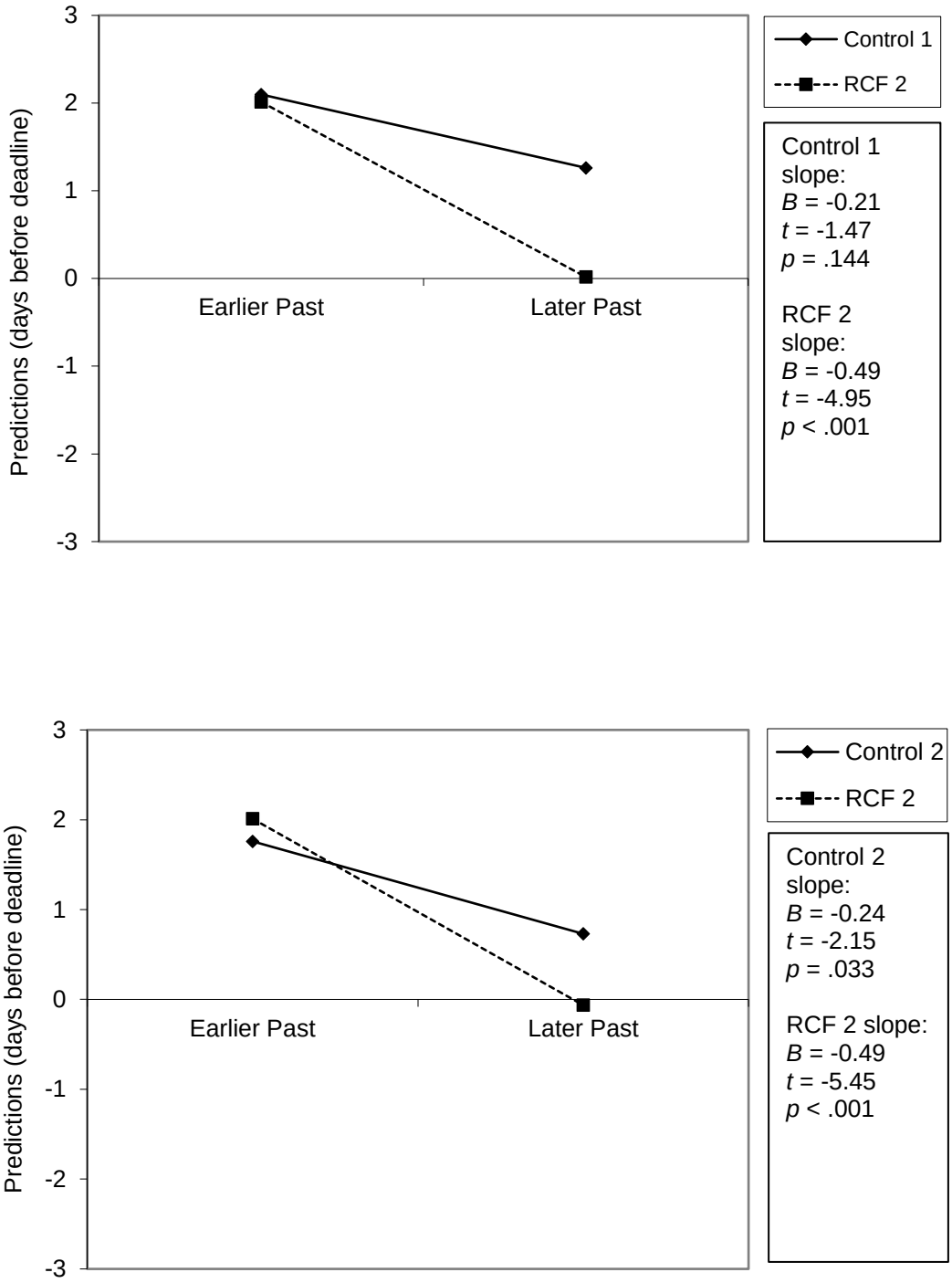


Figure 7. Moderation results for RCF 2 vs. Control 1 and Control 2 – past completion times and predictions are more related in the RCF 2 than the control groups (Study 6).

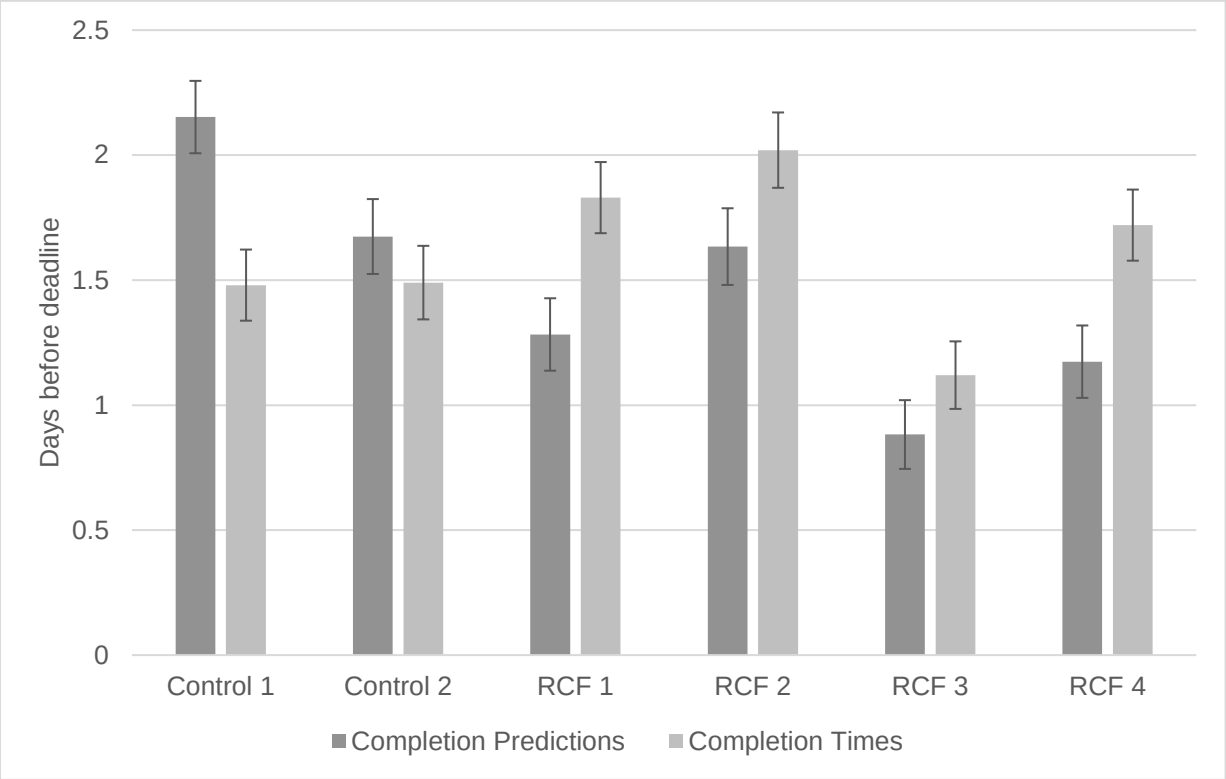


Figure 8. Bar graph showing differences between completion predictions and completion times by condition for Study 6 (Part 2).

Table 1

Planning Fallacy Beliefs Ratings by Condition (Study 1)

Variable	Must-do Projects		Want-to-do Projects		Main Effects and Interactions		
	Self *M* (*SD*)	Other *M* (*SD*)	Self *M* (*SD*)	Other *M* (*SD*)	*F*	*p*	η_p^2
Beliefs	3.62 (1.40)	3.03 (1.22)	3.55 (1.38)	3.58 (1.14)		Project Type	
	t(94) = 3.68, *p* < .001, *d* = 0.45		*t*(90) = -0.21, *p* = .837, *d* = -0.02		2.42	.122	.013
						Target	
					5.77	.017	.031
						Project Type x Target	
					7.40	.007	.039
						All other effects	
					< 1.00	> .300	< .005
% late	34.99 (23.82)	43.56 (21.28)	37.08 (22.05)	39.37 (20.65)		Project Type	
	t(94) = -2.99, *p* = .004, *d* = -0.38		*t*(90) = -0.90, *p* = .373, *d* = -0.11		0.15	.669	.001
						Target	
					7.59	.006	.040
						Project Type x Target	
					2.65	.105	.014
						All other effects	
					< .075	> .375	<.005
% early	26.09 (17.69)	22.93 (14.02)	23.51 (17.40)	25.05 (16.86)		Project Type	
	t(94) = 1.71, *p* = .091, *d* = 0.20		*t*(90) = -0.87, *p* = .388, *d* = -0.09		0.02	.880	<.001
						Target	
					0.39	.532	.002
						Project Type x Target	
					3.34	.069	.018
						All other effects	
					<.075	>.400	<.004
% on time	38.92 (18.97)	33.51 (17.39)	39.42 (21.54)	35.57 (18.59)		Project Type	
	t(94) = 2.36, *p* = .020, *d* = 0.29		*t*(90) = 1.54, *p* = .128, *d* = .19		0.34	.559	.002
						Target	
					7.10	.008	.038
						Project Type x Target	
					0.22	.642	.001
						All other effects	
					< 1.25	> .275	< .007
% bias	8.91 (37.53)	20.63 (35.17)	13.57 (33.37)	14.32 (32.78)		Project Type	
	t(94) = -2.76, *p* = .007, *d* = -.34		*t*(90) = -0.21, *p* = .838, *d* = -0.02		0.03	.866	<.001
						Target	
					4.74	.031	.025
						Project Type x Target	
					3.82	.052	.021
						All other effects	
					< .300	>.500	<.003

Table 2

Mean Completion Predictions (days after study date) by Condition (Study 1)

	Project Type		
	Must-do	Want-to-do	Total
Prediction Order	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Self-nominated Project			
Beliefs first	9.36 (6.08)	14.94 (10.04)	11.49 (8.24)
Predictions first	11.27 (7.28)	13.35 (6.74)	12.21 (7.08)
Total	10.26 (6.71)	14.07 (8.38)	11.85 (7.66)
Hypothetical Project			
Beliefs first	7.39 (4.49)	7.19 (4.37)	7.29 (4.41)
Predictions first	8.44 (6.28)	7.83 (4.71)	8.13 (5.50)
Total	7.89 (5.41)	7.53 (4.53)	7.71 (4.99)

Table 3

Zero-Order Correlations (r) Between Planning Fallacy Beliefs and Predictions (Study 1)

	Beliefs (self)	Beliefs (other)	% late (self)	% late (other)	% early (self)	% early (other)	% bias (self)	% bias (other)	% accurate (self)	% accurate (other)
Prediction (self-nominated)	-.06	.14†	-.02	-.17*	-.03	.09	.001	-.16*	.06	.12†
Prediction (hypothetical)	-.08	.01	.05	-.05	-.09	-.04	.08	-.01	.03	.08

Notes. * $p < .05$, † $p < .10$.

Table 4

Multiple Regression Analyses Testing for Beliefs by Beliefs Order Interactions on Predictions (Study 1)

	Self-nominated Projects*		Hypothetical Projects	
	Self Beliefs 95% CI for *B*	Other Beliefs 95% CI for *B*	Self Beliefs 95% CI for *B*	Other Beliefs 95% CI for *B*
Beliefs Order	-1.57, 0.77	-1.57, 0.78	-3.22, 1.31	-3.24, 1.29
Beliefs	-0.74, 0.11	-0.50, 0.49	-1.19, 0.43	-0.38, 1.55
Beliefs Order x Beliefs	-0.67, 1.10	-1.14, 0.84	-0.77, 2.50	-1.52, 2.35
Beliefs Order x Beliefs x Project type	-2.37,1.23	-1.26, 2.87	-3.42, 3.15	-6.04, 1.71
Beliefs Order	-1.56, 0.77	-1.56, 0.79	-3.24, 1.29	-3.16, 1.34
% late	-0.01, 0.04	-0.04, 0.02	-0.38, 1.55	-0.10, 0.01
Beliefs Order x % late	-0.09, 0.02	-0.07, 0.05	-1.52 2.35	-0.12, 0.01
Beliefs Order x % late x Project type	-0.08, 0.13	-0.13, 0.10	-6.04, 1.71	-0.06, 0.37
Beliefs Order	-1.57, 0.78	-1.55, 0.79	-3.20, 1.34	-3.20, 1.30
% bias	-0.01, 0.02	-0.03, 0.01	-0.04, 0.03	-0.07, 0.01
Beliefs Order x % bias	-0.07, 0.001	-0.05, 0.02	-0.07, 0.06	-0.07, 0.07
Beliefs Order x % bias x Project type	-0.08, 0.06	-0.09, 0.05	-0.05, 0.22	-0.03, 0.24

Notes. *Results controlling for deadline date, $N = 163$. Hypothetical projects, $N = 178$. To avoid capitalizing on chance when running many analyses, only Likert beliefs, % late, and % bias items (and not also % early and % accurate) were analyzed.

Table 5

Planning Fallacy Beliefs Ratings by Condition (Study 2)

Variable	No Article (Control)		Article		Main Effects and Interactions		
	Self *M* (*SD*)	Other *M* (*SD*)	Self *M* (*SD*)	Other *M* (*SD*)	*F*	*p*	η_p^2
Beliefs	3.10 (1.27)	3.11 (1.14)	2.74 (1.14)	2.17 (0.99)		Article	
	No article vs. article for self: t(89) = 1.37, *p* = .175, *d* = 0.29				13.24	<.001	.074
	No article vs. article for other: t(76) = 3.92, *p* < .001, *d* = 0.87					Target	
					2.49	.117	.015
						Article x Target	
					2.76	.099	.016
% late	40.37 (24.89)	43.96 (18.84)	42.59 (19.61)	55.36 (16.47)		Article	
	No article vs. article for self: t(89) = -0.05, *p* = .646, *d* = -0.10				4.55	.034	.027
	No article vs. article for other: t(76) = -2.85, *p* = .006, *d* = -0.64					Target	
					6.56	.011	.038
						Article x Target	
					2.06	.153	.012
% early	24.00 (18.77)	26.39 (15.67)	22.31 (16.38)	16.90 (11.31)		Article	
	No article vs. article for self: t(89) = 0.45, *p* = .654, *d* = -0.10				5.10	.025	.030
	No article vs. article for other: t(62.87) = 3.03, *p* = .004, *d* = 0.70					Target	
					0.37	.543	.002
						Article x Target	
					2.48	.117	.015
% on time	35.63 (21.83)	29.65 (15.35)	35.10 (18.35)	27.74 (11.80)		Article	
	No article vs. article for self: t(89) = 0.12, *p* = .902, *d* = 0.03				0.06	.801	<.001
	No article vs. article for other: t(76) = 0.62, *p* = .536, *d* = 0.14					Target	
					5.97	.016	.035
						Article x Target	
					0.20	.655	.001
% bias	16.37 (38.30)	17.57 (30.96)	20.28 (31.12)	38.45 (25.67)		Article	
	No article vs. article for self: t(89) = -0.52, p = .603, *d* = -0.11				6.11	.014	.036
	No article vs. article for other: t(76) = -3.26, *p* = .002, *d* = -0.72					Target	
					3.73	.055	.022
						Article x Target	
					2.86	.093	.017

Table 6

Mean Completion Predictions (days before deadline) by Condition (Study 2)

	Project Type		
	No Article (Control)	Planning Fallacy Article	Total
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Self	10.51 (2.67)	10.52 (2.77)	10.52 (2.71)
Other	9.97 (3.34)	11.93 (2.57)	11.03 (3.10)
Total	10.30 (3.02)	11.25 (2.70)	10.75 (2.90)

Table 7

Zero-Order Correlations (r) Between Planning Fallacy Beliefs and Predictions (Study 2)

	Beliefs (self)	Beliefs (other)	% late (self)	% late (other)	% early (self)	% early (other)	% bias (self)	% bias (other)	% accurate (self)	% accurate (other)
Predictions	-.17	-.17	-.02	.11	-.08	-.10	.03	.12	.09	-.06

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

Table 8

Mediation Models Testing the Effect of the Article Manipulation on Completion Predictions through Planning Fallacy Beliefs (Study 2)

	"a" Path		"b" Path		"c" (Total Effect)		"c'" (Direct Effect)		"a*b" (Indirect Effect)	Boot CI
	β	*p*	β	*p*	β	*p*	β	*p*	B (*SE*)	LL, UL
Beliefs (Likert item)	-.66	<.001	-.35	.068	0.95	.033	0.72	.116	**0.24 (.13)**	**.0184, .5552**
% late	7.37	.024	.004	.696	0.95	.033	0.92	.042	0.03 (.09)	-.1109, .2785
% bias	12.85	.012	0.01	.512	0.95	.033	0.89	.049	0.06 (.10)	-.0914, .3233

Notes. Mediation analyses conducted using 10,000 bootstrap samples. To avoid capitalizing on chance when running many analyses, only Likert beliefs, % late, and % bias items (and not also % early and % accurate) were analyzed.

Table 9

Moderated Mediation Models Testing the Effect of the Article Manipulation on Completion Predictions through Planning Fallacy Beliefs as Moderated by Target Type (Study 2)

	"a*b" (Indirect Effect) Self		"a*b" (Indirect Effect) Other		"a*b" highest order interaction	
	B (*SE*)	Boot CI	B (*SE*)	Boot CI	B (*SE*)	Boot CI
Beliefs (Likert item)	0.10 (.11)	-.0241, .4315	0.27 (.19)	-.0391, .7471	0.17 (.16)	-.0244, .6708
% late	-.006 (.06)	-.1417, .1156	-.003 (.14)	-.2742, .2919	-.003 (.12)	-.2840, .2462
% bias	0.01 (.06)	-.0807, .1898	0.04 (.15)	-.2536, .3841	0.03 (.14)	-.2031, .3888

Notes. Mediation analyses conducted using 10,000 bootstrap samples. To avoid capitalizing on chance when running many analyses, only Likert beliefs, % late, and % bias items (and not also % early and % accurate) were analyzed.

Table 10

Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample, Study 3)

	Control	Past	RCF deadline	RCF prediction			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Average past project completion	--	1.71a (1.54)	1.69a (1.81)	-0.66b (1.57)	--	--	--
Prediction (days before deadline)	**2.02a (1.67)**	**1.96a (1.67)**	**1.73a (1.62)**	**1.18b (1.06)**	**3.59**	**.014**	**.047**
Additional Variables							
Predicted work time (hrs.)	14.37 (10.95)	13.31 (9.43)	13.11 (9.44)	13.49 (9.43)	0.32	.811	.004
Extent based prediction on:							
-when I ideally want to be finished	4.84 (1.72)	5.29 (1.35)	5.02 (1.67)	5.09 (1.47)	0.59	.625	.008
-project steps and how long each will take	4.89 (1.55)	5.29 (1.35)	5.02 (1.45)	5.15 (1.56)	0.59	.622	.008
-potential obstacles	5.14 (1.55)	5.27 (1.44)	5.08 (1.56)	5.15 (1.53)	0.39	.760	.005
-other demands on my time	5.52 (1.36)	5.59 (1.22)	5.73 (1.16)	5.29 (1.45)	1.16	.327	.016
-past experiences with similar projects	5.54 (1.36)	5.59 (1.37)	5.58 (1.18)	5.87 (1.29)	0.68	.565	.009
Difficulty pre-manipulation	4.66 (1.07)	4.69 (1.19)	4.73 (0.94)	4.44 (1.50)	(between) 1.43	.234	.026
Difficulty post-manipulation	4.73 (1.21)	4.86 (1.04)	5.11 (0.99)	4.67 (1.42)	(within) 0.48	.487	.002
					(interaction) 0.94	.421	.013
Importance pre-manipulation	5.80 (1.46)	6.08 (0.91)	6.00 (1.15)	5.81 (1.08)	(between) 0.89	.453	.012
Importance post-manipulation	5.62 (1.48)	6.08 (1.04)	5.85 (1.07)	5.91 (1.20)	(within) 2.33	.128	.011

					(interaction) 0.78	.506	.011
Control pre-manipulation	5.46 (1.24)	5.78 (1.07)	6.02 (0.88)	5.89 (1.03)	(between) 1.90	.131	.025
Control post-manipulation	5.48 (1.39)	5.69 (0.94)	5.73 (1.12)	5.78 (1.03)	(within) 3.20	.075	.015
					(interaction) 1.06	.365	.014
Busyness pre-manipulation	5.29 (1.50)	5.51 (1.19)	5.45 (1.18)	5.27 (1.13)	(between) 0.88	.450	.012
Busyness post-manipulation	5.15 (1.50)	5.43 (1.16)	5.66 (1.04)	5.25 (1.31)	(within) 2.58	.110	.012
					(interaction) 1.20	.310	.016
Complete as quickly as possible pre-manipulation	4.79 (1.44)	5.02 (0.97)	4.76 (1.17)	4.80 (1.11)	(between) 0.46	.713	.006
Complete as quickly as possible post-manipulation	5.05 (1.21)	5.10 (1.14)	5.02 (1.19)	5.25 (1.28)	(within) 0.42	.516	.002
					(interaction) 0.93	.428	.013
Importance of quality pre-manipulation	6.04 (1.18)	6.31 (0.82)	6.23 (0.86)	6.02 (1.11)	(between) 1.07	.361	.015
Importance of quality post-manipulation	6.09 (1.13)	6.12 (1.03)	6.15 (0.99)	5.89 (1.13)	(within) 0.01	.913	< .001
					(interaction) 0.75	.525	.010
Confidence that prediction is accurate	5.48 (1.14)	5.65 (1.09)	5.45 (1.08)	5.42 (0.96)	0.59	.622	.008
Similarity of past projects (low - high)	--	5.49 (1.04)	5.24 (1.05)	5.51 (1.07)	1.12	.328	.014
Complexity of past projects (past less - more)	--	4.55a (1.00)	4.12b (1.06)	4.42ab (0.93)	2.79	.065	.035

Scope of past projects (past smaller - larger)	--	4.11 (1.26)	4.10 (0.97)	4.47 (0.99)	2.06	.130	.026
Work time available* (past more - current more)	--	4.40 (1.04)	4.40 (1.03)	4.43 (1.31)	0.02	.978	.000
Relevance of past projects for current prediction (low – high)	--	5.34 (1.20)	5.33 (1.16)	5.13 (1.26)	0.50	.610	.006
Exercise straightforward	--	5.37 (1.14)	5.05 (1.11)	5.24 (1.02)	1.31	.272	.016
Exercise difficult	--	3.61 (1.61)	3.81 (1.64)	3.53 (1.60)	0.62	.538	.008
Exercise useful	--	5.04 (1.33)	4.77 (1.14)	5.02 (1.13)	1.03	.360	.013
Exercise annoying	--	3.22a (1.63)	3.98b (1.66)	3.60ab (1.71)	2.80	.064	.034
Importance of accurate completion predictions	5.46ab (1.06)	5.63a (1.11)	5.18b (1.15)	5.24ab (1.24)	2.14	.097	.029
Motivated to make accurate completion predictions	5.13 (1.28)	5.29 (1.17)	5.03 (1.37)	5.15 (1.24)	0.43	.734	.006
Underestimation problematic	4.85 (1.60)	5.08 (1.43)	4.89 (1.67)	4.81 (1.66)	0.36	.779	.005

Notes. Control $n = 56$, Past $n = 49$, RCF deadline $n = 62$, RCF prediction $n = 55$.

Table 11

Zero-Order Correlations Between Completion Predictions and Additional Measures (Part 1 sample, Study 3)

	Completion prediction (days before deadline)
	R
Predicted working time (hrs.)	-.18**
Extent based prediction on when want to be finished	.25***
Extent based prediction on project steps, how long each would take	.05
Extent based prediction on anticipated obstacles	.08
Extent based prediction on other demands on my time	.01
Extent based prediction on past experiences	-.19**
Difficulty (2)	-.13*
Importance (2)	-.08
Control (2)	.02
Busyness (2)	.05
Finish as quickly as possible (2)	.09
Importance of quality (2)	-.01
Confidence in prediction accuracy	-.13†
Past projects similarity	-.12
Past projects similarly complex	.07
Past projects similar in scope	.00
Past projects similar in working time available	.13†
Past projects relevant	-.02
Exercise straightforward	-.01
Exercise difficult	.14†
Exercise useful	.19*
Exercise annoying	.01
Importance of accurate completion predictions	.15*
Motivated to make accurate completion predictions	.12†
Underestimation problematic	-.04

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, †$p < .10$.

Table 12

Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample A; Study 3)

	Control	Past	RCF deadline	RCF prediction			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Average past project completion	--	2.29a (1.64)	1.85a (1.76)	-0.80b (1.89)	--	--	--
Prediction (days before deadline)	**2.09 (1.55)**	**2.33 (1.63)**	**1.88 (1.67)**	**1.39 (1.08)**	**1.66**	**.181**	**.044**
Completion time (days before deadline)	0.71 (1.45)	1.08 (1.74)	1.42 (1.70)	1.08 (1.78)	0.47	.704	.013
Bias (prediction -completion time)	**1.38 (1.72)**	**1.25 (1.75)**	**0.45 (1.79)**	**0.30 (1.72)**	**1.65**	**.181**	**.044**
Bias binary (finish on/before prediction vs. after)	14 vs. 26	11 vs. 18	21 vs. 22	16 vs.15	X^2 = 2.87	.412	
Additional Measures							
Predicted work time (hrs.)	14.29 (11.53)	14.33 (10.45)	13.24 (9.99)	11.48 (7.49)	0.85	.467	.023
Work time (hs.)	10.56 (13.53)	9.44 (7.90)	8.95 (7.14)	9.98 (14.43)	0.07	.976	.002
Work time bias (predicted - work time) (hrs.)	3.74 (14.26)	4.90 (9.02)	4.36 (8.54)	1.50 (12.76)	0.83	.480	.023

Notes. Control n = 34, Past n = 24, RCF deadline n = 33, RCF prediction n = 23.

Table 13

Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample B; Study 3)

	Control	Past	RCF deadline	RCF prediction			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Average past project completion	--	2.10a (1.64)	1.74a (1.78)	-0.86b (1.68)	--	--	--
Prediction (days before deadline)	**2.02 (1.56)**	**2.00 (1.67)**	**1.81 (1.56)**	**1.26 (1.00)**	**1.88**	**.136**	**.039**
Bias (prediction -completion time)	**1.44 (1.70)**	**1.10 (1.63)**	**0.72 (1.72)**	**0.45 (1.52)**	**1.49**	**.220**	**.031**
Bias binary (finish on/before prediction vs. after)	12 vs. 29	9 vs. 18	11 vs. 33	14 vs. 9	X^2 = 6.10	.107	
Additional Measures							
Predicted work time (hrs.)	14.76 (11.31)	13.72 (9.86)	13.23 (9.97)	12.45 (8.66)	0.66	.577	.014
Work time (hrs.)	10.79 (12.77)	9.62 (8.00)	9.10 (7.08)	9.44 (12.52)	0.08	.972	.002
Work time bias (predicted - work time) (hrs.)	3.96 (13.35)	4.10 (8.75)	4.19 (8.39)	3.02 (12.07)	0.27	.844	.006

Notes. Control n = 41, Past n = 29, RCF deadline n = 43, RCF prediction n = 31.

Table 14

Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times by Condition (Part 2 sample A; Study 3)

	2. Completion Prediction	3. Actual Completion Time
Control		
1. Average Past Project Completion	--	--
2. Completion Prediction	--	
3. Actual Completion Time	.34*	--
Past		
1. Average Past Project Completion	.50*	.33
2. Completion Prediction	--	
3. Actual Completion Time	.46*	--
RCF deadline		
1. Average Past Project Completion	.49**	.30†
2. Completion Prediction	--	
3. Actual Completion Time	.44*	--
RCF prediction		
1. Average Past Project Completion	-.33	.19
2. Completion Prediction	--	
3. Actual Completion Time	.36†	--

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, †$p < .10$.

Table 15

Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample A; Study 3)

	Completion prediction	"Actual" completion time
Predicted working time (hrs.)	-.10	-.08
Extent based prediction on when want to be finished	.32**	.24**
Extent based prediction on project steps, how long each would take	.08	.11
Extent based prediction on anticipated obstacles	.12	.18†
Extent based prediction on other demands on my time	.12	.12
Extent based prediction on past experiences	-.24*	.10
Difficulty (2)	-.10	-.16†
Importance (2)	-.20	.08
Control (2)	-.08	.21*
Busyness (2)	.14	.14
Finish as quickly as possible (2)	.21*	.16†
Importance of quality (2)	-.05	.07
Confidence in prediction accuracy	-.23*	-.02
Importance of accurate completion predictions	.16†	.15
Motivated to make accurate completion predictions	.15	.13
Underestimation problematic	-.16†	-.02
Working time (hrs.)	-.11	-.15

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, † $p < .10$.

Table 16

Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample; Study 4)

	Control	Past	RCF deadline	RCF prediction			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Average past project completion	--	1.28 (2.06)	1.71 (2.32)	0.78 (2.12)	--	--	--
Prediction (days before deadline)	**1.83a (2.00)**	**1.11b (2.34)**	**1.48a (1.98)**	**1.06b (1.96)**	**2.16**	**.093**	**.023**
Additional Measures							
Predicted work time (hrs.)	17.58 (10.27)	18.63 (9.81)	18.50 (10.41)	20.69 (10.44)	1.16	.328	.012
Extent based prediction on:							
-when I ideally want to be finished	5.64 (1.53)	5.45 (1.38)	5.50 (1.49)	5.50 (1.64)	0.21	.890	.002
-project steps and how long each will take	5.78 (1.32)	5.75 (1.43)	5.77 (1.43)	5.97 (1.15)	0.33	.806	.004
-potential obstacles	5.49 (1.50)	5.39 (1.48)	5.25 (1.56)	5.35 (1.53)	0.39	.764	.004
-other demands on my time	5.54 (1.34)	5.76 (1.28)	5.64 (1.36)	5.38 (1.47)	1.33	.267	.014
-past experiences with similar projects	6.01ab (1.08)	6.08a (1.08)	5.81ab (1.39)	5.68b (1.55)	2.56	.055	.027
Difficulty pre-manipulation	4.87 (1.29)	4.55 (1.33)	4.70 (1.29)	4.50 (1.23)	(between) 0.61	.608	.007
Difficulty post-manipulation	4.83 (1.38)	4.82 (1.339)	4.81 (1.53)	4.68 (1.28)	(within) 0.89	.348	.003
					(interaction) 1.37	.252	.015
Importance pre-manipulation	6.13 (1.09)	5.96 (1.09)	6.42 (0.91)	5.99 (1.40)	(between) 2.48	.062	.026
Importance post-manipulation	6.26 (1.01)	5.70 (1.52)	6.19 (1.22)	6.10 (1.27)	(within) 22.47	<.001	.076

					(interaction) 3.49	.016	.037
Control pre-manipulation	5.85 (1.31)	5.70 (1.36)	5.90 (1.19)	5.81 (1.25)	(between) 0.05	.984	.001
Control post-manipulation	5.83 (1.26)	5.77 (1.45)	5.81 (1.38)	5.85 (1.25)	(within) 0.45	.503	.002
					(interaction) 0.44	.725	.005
Busyness pre-manipulation	5.36 (1.39)	5.28 (1.31)	5.25 (1.41)	4.96 (1.44)	(between) 0.88	.450	.012
Busyness post-manipulation	5.46 (1.45)	5.20 (1.23)	5.45 (1.23)	5.31 (1.44)	(within) 5.30	.022	.019
					(interaction) 2.38	.070	.025
Complete as quickly as possible pre-manipulation	5.67 (1.23)	5.61 (1.45)	5.69 (1.27)	5.68 (1.14)	(between) 0.10	.957	.001
Complete as quickly as possible post-manipulation	5.63 (1.30)	5.46 (1.38)	5.69 (1.33)	5.66 (1.24)	(within) 14.90	<.001	.051
					(interaction) 0.55	.647	.006
Importance of quality pre-manipulation	6.40 (0.84)	6.10 (1.29)	6.31 (1.21)	6.44 (0.95)	(between) 2.32	.076	.025
Importance of quality post-manipulation	6.41 (0.84)	6.04 (1.36)	6.44 (0.87)	6.37 (1.04)	(within) 1.68	.196	.006
					(interaction) 0.86	.461	.009
Confidence that prediction is accurate	5.56 (1.05)	5.63 (1.20)	5.52 (1.31)	5.97 (0.95)	2.22	.086	.024
Similarity of past projects (low - high)	--	5.32 (1.33)	5.60 (1.06)	5.72 (1.11)	1.39	.251	.014
Complexity of past projects (past less - more)	--	4.14 (1.05)	4.19 (1.19)	4.22 (1.04)	0.04	.965	.000

Scope of past projects (past smaller - larger)	--	4.11 (1.05)	4.10 (1.19)	4.47 (1.04)	0.08	.927	.001
Work time available* (past more - current more)	--	4.13 (1.04)	4.06 (1.03)	4.12 (1.31)	0.02	.979	.000
Relevance of past projects for current prediction (low – high)	--	5.24 (1.20)	5.48 (1.16)	5.58 (1.26)	0.50	.607	.005
Exercise straightforward	--	5.61 (1.27)	5.42 (1.45)	5.53 (1.19)	1.00	.369	.010
Exercise difficult	--	3.66 (1.70)	3.81 (1.80)	3.56 (1.81)	0.33	.716	.003
Exercise useful	--	5.21 (1.39)	5.36 (1.58)	5.31 (1.34)	0.10	.905	.001
Exercise annoying	--	2.76 (1.90)	3.05 (1.91)	2.74 (1.43)	0.85	.430	.009
Importance of accurate completion predictions	6.00 (0.97)	5.75 (1.37)	5.82 (1.34)	5.82 (1.11)	0.84	.473	.009
Motivated to make accurate completion predictions	5.95 (1.05)	5.86 (1.15)	5.97 (1.23)	5.82 (1.08)	0.40	.755	.004
Underestimation problematic	4.08 (2.05)	4.03 (2.01)	4.54 (2.04)	4.36 (2.01)	0.86	.462	.009

Notes. Control $n = 80$, Past $n = 72$, RCF deadline $n = 64$, RCF prediction $n = 69$.

Table 17

Zero-Order Correlations Between Completion Predictions and Additional Measures (Part 1 sample; Study 4)

	Completion prediction (days before deadline)
	r
Predicted working time (hrs.)	-.03
Extent based prediction on when want to be finished	.23***
Extent based prediction on project steps, how long each would take	.14*
Extent based prediction on anticipated obstacles	-.01
Extent based prediction on other demands on my time	.07
Extent based prediction on past experiences	-.03
Difficulty (2)	-.05
Importance (2)	.09
Control (2)	.07
Busyness (2)	-.04
Finish as quickly as possible (2)	.14*
Importance of quality (2)	.12*
Confidence in prediction accuracy	.09
Past projects similarity	-.02
Past projects similarly complex	-.09
Past projects similar in scope	-.09
Past projects similar in working time available	-.04
Past projects relevant	-.10
Exercise straightforward	.01
Exercise difficult	-.11
Exercise useful	-.03
Exercise annoying	-.09
Importance of accurate completion predictions	.00
Motivated to make accurate completion predictions	.04
Underestimation problematic	-.07

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, $^{\dagger} p < .10$.

Table 18

Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample; Study 4)

	Control	Past	RCF deadline	RCF prediction			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Average past project completion	--	1.25 (1.86)	1.39 (2.38)	1.03 (2.16)	--	--	--
Prediction (days before deadline)	**2.12a (1.78)**	**0.90b (2.33)**	**1.18b (1.86)**	**1.39b (1.42)**	**3.24**	**.024**	**.063**
Completion time (days before deadline)	0.83a (1.58)	0.77a (1.20)	0.73a (1.64)	1.87b (1.34)	5.55	.001	.103
Bias (prediction - completion time)	**1.29a (2.22)**	**0.13b (2.19)**	**0.45ab (2.15)**	**-0.47b (1.75)**	**4.84**	**.003**	**.091**
Bias binary (finish on/before prediction vs. after)	21 vs. 27	25 vs. 21	20 vs.19	32 vs. 11	X^2 = 9.23	.026	
Additional Measures							
Predicted work time (hrs.)	19.20 (10.41)	19.59 (9.38)	21.03 (10.38)	20.53 (9.94)	0.20	.898	.004
Work time (hrs.)	13.07a (11.27)	21.15b (26.47)	24.24b (23.59)	18.32ab (12.31)	2.15	.096	.043
Work time bias (predicted - work time) (hrs.)	6.12 (13.32)	-1.56 (23.77)	-3.21 (20.91)	2.21 (12.44)	1.90	.132	.038

Notes. Control n = 41, Past n = 39, RCF deadline n = 33, RCF prediction n = 39.

Table 19

Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times (Part 2 sample, Study 4)

	2. Completion Prediction	3. Actual Completion Time
Control		
1. Average Past Project Completion		
2. Completion Prediction	--	
3. Actual Completion Time	.13	--
Past		
1. Average Past Project Completion	.48**	.26
2. Completion Prediction	--	
3. Actual Completion Time	.37*	--
RCF deadline		
1. Average Past Project Completion	.80***	.15
2. Completion Prediction	--	
3. Actual Completion Time	.25	--
RCF prediction		
1. Average Past Project Completion	.23†	.07
2. Completion Prediction	--	
3. Actual Completion Time	.36†	--

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, †$p < .10$.

Table 20

Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample, Study 4)

	Completion prediction	"Actual" completion time
Predicted working time (hrs.)	-.14†	-.16*
Extent based prediction on when want to be finished	.25*	.08
Extent based prediction on project steps, how long each would take	.05	.11
Extent based prediction on anticipated obstacles	-.06	-.05
Extent based prediction on other demands on my time	-.02	.06
Extent based prediction on past experiences	-.02	-.15†
Difficulty (2)	-.02	.003
Importance (2)	-.03	.04
Control (2)	-.02	-.002
Busyness (2)	-.13	-.05
Finish as quickly as possible (2)	.08	.09
Importance of quality (2)	-.02	.09
Confidence in prediction accuracy	-.23*	-.02
Importance of accurate completion predictions	-.03	-.002
Motivated to make accurate completion predictions	.06	-.05
Underestimation problematic	.003	.004
Working time (hrs.)	-.06	.02

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, † $p < .10$.

Table 21

Components of the Prediction Intervention Conditions in Study 5 and 6

	Condition 1 (Control 1)	Condition 2 (Control 2)	Condition 3 (RCF 1)	Condition 4 (RCF 2)	Condition 5 (RCF 3)	Condition 6 (RCF 4)
Initial Prediction		☑	☑	☑		
Past Times			☑	☑	☑	☑
Mean and Instructions			☑		☑	
Final Prediction	☑	☑	☑	☑	☑	☑

Table 22

Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample; Study 5)

	Control 1 - single prediction only	Control 2 - initial prediction	RCF 1 - initial prediction, past mean	RCF 2 - initial prediction, no past mean	RCF 3 - no initial prediction, past mean	RCF 4 - no initial prediction, no past mean			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Deadline (days from study date)	11.64 (5.31)	12.59 (5.84)	12.53 (5.63)	13.48 (5.92)	12.43 (6.84)	12.81 (5.29)	0.71	.616	.009
Initial Prediction (days before deadline)	1.83 (1.67) [initial, final same]	1.82 (1.87)	1.70 (1.73)	1.86 (1.84)	--	--	0.10	.958	.001
Past Mean (days after prediction)	-0.89a (1.43)	-1.05a (1.36)	-0.13b (1.30)	-0.15b (1.23)	0.40b (1.75)	0.23b (1.21)	11.74	<.001	.129
Prediction (days before deadline)	**1.83a (1.67)**	**1.51ab (1.70)**	**(missing due to error)**	**1.27b (1.38)**	**(missing due to error)**	**1.38ab (1.66)**	**1.53**	**.207**	**.017**
Additional Measures									
Predicted working hours	12.54 (9.14)	13.88 (10.34)	12.58 (9.00)	12.18 (9.12)	10.69 (8.290	13.31 (10.42)	0.86	.511	.011
Extent based prediction on:									
-when I ideally want to be finished	5.37 (1.53)	5.35 (1.52)	5.51 (1.48)	5.39 (1.69)	5.47 (1.42)	5.41 (1.37)	.115	.989	.001
-project steps and how long each will take	5.14 (1.54)	5.08 (1.46)	5.10 (1.57)	5.10 (1.39)	5.10 (1.52)	5.22 (1.57)	1.515	.184	.019
-potential obstacles	5.20 (1.53)	5.08 (1.67)	5.12 (1.67)	4.68 (1.68)	4.58 (1.94)	5.09 (1.75)	1.295	.265	.016
-other demands on my time	5.85 (1.24)	5.53 (1.49)	5.49 (1.54)	5.25 (1.63)	5.73 (1.50)	5.63 (1.50)	.328	.896	.004
-past experiences w. projects	5.66 (1.28)	5.67 (1.69)	5.49 (1.55)	5.77 (1.23)	5.63 (1.31)	5.70 (1.23)	.076	.996	.001

Notes. Control 1 n = 65, Control 2 n = 67, RCF 1 n =71, RCF 2 n = 71, RCF 3 n = 62, RCF 4 n = 65.

Table 23

Effects of the Reference Class Forecasting Manipulations on Past Project Completion Times (Part 1 sample; Study 5)

	M (SD)		*t*	*p*	*d*
Control 1 vs. Control 2	-0.88 (1.43)	-1.05 (1.36)	0.65	.517	0.12
Control 1 vs. RCF 1	-0.88 (1.43)	-0.13 (1.30)	-3.25	.001	-0.55
Control 1 vs. RCF 2	-0.88 (1.43)	-0.15 (1.23)	-3.23	.002	-0.55
Control 1 vs. RCF 3	-0.88 (1.43)	0.40 (1.75)	-4.57	<.001	-0.80
Control 1 vs. RCF 4	-0.88 (1.43)	0.23 (1.21)	-4.83	<.001	-0.83
Control 2 vs. RCF 1	-1.05 (1.36)	-0.13 (1.30)	-4.06	<.001	-0.69
Control 2 vs. RCF 2	-1.05 (1.36)	-0.15 (1.23)	-4.07	<.001	-0.69
Control 2 vs. RCF 3	-1.05 (1.36)	0.40 (1.75)	-5.29	<.001	-0.92
Control 2 vs. RCF 4	-1.05 (1.36)	0.23 (1.21)	-5.72	<.001	-0.99
RCF 1 vs. RCF 2	-0.13 (1.30)	-0.15 (1.23)	0.10	.923	0.02
RCF 1 vs. RCF 3	-0.13 (1.30)	0.40 (1.75)	-2.01	.046	-0.35
RCF 1 vs. RCF 4	-0.13 (1.30)	0.23 (1.21)	-1.69	.094	-0.28
RCF 2 vs. RCF 3	-0.15 (1.23)	0.40 (1.75)	-2.13	.035	-0.37
RCF 2 vs. RCF 4	-0.15 (1.23)	0.23 (1.21)	-1.83	.069	-0.31
RCF 3 vs. RCF 4	0.40 (1.75)	0.23 (1.21)	0.64	.522	-0.31

Notes. Control 1 n = 65, Control 2 n = 67, RCF 1 n =71, RCF 2 n = 71, RCF 3 n = 62, RCF 4 n = 65.

Table 24

Multiple Regression Analyses Testing for Moderation Effects (Study 5)

Main Effect or Interaction Term	*t*	*p*	95% CI for *B*		R^2_{Change}
Step 1					
Control 1 vs. RCF 2	-2.15	.034	-1.08	-0.04	.033
Step 2 (add)					
Control 1 vs. RCF 2	-1.09	.278	-0.79	0.23	
Mean Past Completion Times	-4.06	.000	-0.57	-0.20	.107
Step 3 (add)					
Condition x Past Completion Times	-0.30	.763	-0.43	0.32	.001
Step 1					
Control 1 vs. RCF 4	-1.53	.129	-1.02	0.13	.018
Step 2 (add)					
Control 1 vs. RCF 4	-0.43	.667	-0.75	0.48	
Mean Past Completion Times	-2.56	.012	-0.49	-0.06	.048
Step 3 (add)					
Condition x Past Completion Times	0.85	.397	-0.25	0.62	.005
Step 1					
Control 1 vs. Control 2	-1.10	.273	-0.90	0.26	.009
Step 2 (add)					
Control 1 vs. Control 2	-1.44	.153	-0.94	0.15	
Mean Past Completion Times	-4.55	.000	-0.65	-0.25	.137
Step 3 (add)					
Condition x Past Completion Times	-0.98	.327	-0.59	0.20	.006
Step 1					
Control 2 vs. RCF 2	-0.91	.364	-0.76	0.28	.006
Step 2 (add)					
Control 2 vs. RCF 2	0.77	.443	-0.31	0.70	
Mean Past Completion Times	-5.18	.000	-0.67	-0.30	.165
Step 3 (add)					
Condition x Past Completion Times	0.73	.468	-0.24	0.51	.003
Step 1					
Control 2 vs. RCF 4	-0.42	.675	-0.70	0.46	.001
Step 2 (add)					
Control 2 vs. RCF 4	1.18	.241	-0.25	0.99	
Mean Past Completion Times	-3.50	.001	-0.60	-0.17	.087
Step 3 (add)					
Condition x Past Completion Times	1.74	.085	-0.05	0.82	.021
Step 1					
Control 1 & 2 vs. RCF 2 & 4	-1.75	.081	-0.73	0.04	.011
Step 2 (add)					
Control 12 vs. RCF 24	0.14	.886	-0.37	0.42	
Mean Past Completion Times	-5.19	.000	-0.51	-0.23	.091
Step 3 (add)					
Condition x Past Completion Times	1.10	.274	-0.13	0.44	.004

Table 25

Zero-Order Correlations Between Completion Predictions and Additional Measures (Part 1 sample; Study 5)

	Completion prediction (days before deadline)
	r
Predicted working time (hrs.)	-.03
Extent based prediction on when want to be finished	.24***
Extent based prediction on anticipated obstacles	.09
Extent based prediction on other demands on my time	.07
Extent based prediction on past experiences	-.07
Extent based prediction on project steps, how long each would take	.15*

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, $^{\dagger} p < .10$.

Table 26

Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample; Study 5)

	Control 1 - single prediction only	Control 2 - initial prediction	RCF 1 - initial prediction, past mean	RCF 2 - initial prediction, no past mean	RCF 3 - no initial prediction, past mean	RCF 4 - no initial prediction, no past mean			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Average past project completion	-0.94 (1.42)	-0.96 (1.27)	-0.24 (1.05)	-0.26 (1.09)	0.37 (1.38)	0.09 (1.27)	7.27	<.001	.134
Prediction (days before deadline)	**1.83 (1.52)**	**1.51 (1.65)**	(missing due to error)	**1.39 (1.05)**	(missing due to error)	**1.30 (1.47)**	**0.89**	**.443**	**.018**
Completion time (days before deadline)	1.10 (1.46)	0.92 (1.30)	1.18 (1.45)	0.93 (1.33)	0.95 (1.53)	1.17 (2.04)	0.25	.942	.006
Bias (prediction - completion time)	**0.73 (1.66)**	**0.50 (1.20)**	(missing due to error)	**0.46 (1.45)**	(missing due to error)	**-0.07 (1.75)**	**1.58**	**.197**	**.032**
Bias binary* (finish on/before prediction vs. after)	22 vs. 18	22 vs. 16	(missing due to error)	23 vs. 18	(missing due to error)	19 vs. 10	X^2 = 0.88	.830	V = .077
Additional Measures									
Predicted work time (hrs.)	12.60 (9.38)	11.44 (8.88)	13.09 (8.63)	11.31 (8.55)	10.00 (8.30)	14.12 (10.73)	0.99	.424	.021
Work time (hrs.)	7.57 (5.98)	8.08 (8.51)	9.94 (6.76)	7.94 (8.37)	6.87 (5.38)	11.09 (10.82)	1.60	.161	.033
Work time bias (predicted - work time) (hrs.)	5.02 (7.30)	3.36 (7.76)	3.15 (8.79)	3.37 (6.56)	3.13 (7.94)	3.03 (8.69)	0.38	.862	.008

Notes. Control 1 n = 40, Control 2 n = 38, RCF 1 n = 39, RCF 2 n = 41, RCF 3 n = 37, RCF 4 n = 29.

Table 27

Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times (Part 2 sample, Study 5)

	2. Completion Prediction	3. Actual Completion Time
Control 1		
1. Average Past Project Completion	-.46**	-.57***
2. Completion Prediction	--	
3. Actual Completion Time	.38*	--
Control 2		
1. Average Past Project Completion	-.54***	-.48**
2. Completion Prediction	--	
3. Actual Completion Time	.66***	--
RCF 1		
1. Average Past Project Completion	(missing)	-.49***
2. Completion Prediction	--	
3. Actual Completion Time	(missing)	--
RCF 2		
1. Average Past Project Completion	-.01	-.59***
2. Completion Prediction	--	
3. Actual Completion Time	.36†	--
RCF 3		
1. Average Past Project Completion	(missing)	-.03
2. Completion Prediction	--	
3. Actual Completion Time	(missing)	--
RCF 4		
1. Average Past Project Completion	-.29	-.11
2. Completion Prediction	--	
3. Actual Completion Time	.52**	--

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, † $p < .10$.

Table 28

Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample, Study 5)

	Completion prediction	"Actual" completion time
Predicted working time (hrs.)	-.002	-.09
Extent based prediction on when want to be finished	.25**	.07
Extent based prediction on anticipated obstacles	-.02	-.11†
Extent based prediction on other demands on my time	.04	-.02
Extent based prediction on past experiences	-.02	-.05
Extent based prediction on project steps, how long each would take	.11	.03
Working time (hrs.)	.03	-.20**

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, †$p < .10$.

Table 29

Effects of the Reference Class Forecasting Manipulations on Completion Predictions and Other Variables (Part 1 sample; Study 6)

	Control 1 - single prediction only	Control 2 - initial prediction	RCF 1 - initial prediction, past mean	RCF 2 - initial prediction, no past mean	RCF 3 - no initial prediction, past mean	RCF 4 - no initial prediction, no past mean			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η_p^2
Deadline (days from study date)	13.33 (4.04)	14.22 (4.68)	14.07 (4.48)	13.28 (3.96)	12.59 (4.60)	13.29 (4.33)	1.51	.185	.016
Initial Prediction (days before deadline)	--	1.34 (1.83)	1.67 (2.08)	1.38 (2.53)	--	--	0.76	.517	.007
Past Mean (days after prediction)	-0.31a (1.63)	-0.15a (1.84)	0.54b (2.24)	0.41b (2.34)	0.46b (1.94)	0.54b (2.28)	2.70	.020	.028
Prediction (days before deadline)	**1.76a (2.03)**	**1.31ab (1.69)**	**0.50c (2.40)**	**0.83bc (2.33)**	**0.64c (2.17)**	**0.40c (2.24)**	**4.85**	**<.001**	**.048**
Additional Measures									
Predicted working hours	24.22a (8.76)	20.63b (9.70)	21.52b (10.17)	19.81b (9.88)	20.99b (9.88)	18.59b (9.47)	3.13	.009	.032

Notes. Control 1 n = 82, Control 2 n = 80, RCF 1 n = 84, RCF 2 n = 78, RCF 3 n = 80, RCF 4 n = 80.

Table 30

Effects of the Reference Class Forecasting Manipulations on Past Project Completion Times (Part 1 sample; Study 6)

	M (SD)		*t*	*p*	*d*
Control 1 vs. Control 2	-0.31 (1.63)	-0.15 (1.84)	-0.59	.554	-0.09
Control 1 vs. RCF 1	**-0.31 (1.63)**	**0.54 (2.24)**	**-2.80**	**.006**	**-0.43**
Control 1 vs. RCF 2	**-0.31 (1.63)**	**0.41 (2.34)**	**-2.23**	**.025**	**-0.36**
Control 1 vs. RCF 3	**-0.31 (1.63)**	**0.46 (1.94)**	**-2.72**	**.007**	**-0.43**
Control 1 vs. RCF 4	**-0.31 (1.63)**	**0.54 (2.28)**	**-2.71**	**.007**	**-0.43**
Control 2 vs. RCF 1	**-0.15 (1.84)**	**0.54 (2.24)**	**-2.14**	**.034**	**-0.33**
Control 2 vs. RCF 2	**-0.15 (1.84)**	**0.41 (2.34)**	**-1.69**	**.095**	**-0.27**
Control 2 vs. RCF 3	**-0.15 (1.84)**	**0.46 (1.94)**	**-2.02**	**.045**	**-0.32**
Control 2 vs. RCF 4	**-0.15 (1.84)**	**0.54 (2.28)**	**-2.09**	**.038**	**-0.33**
RCF 1 vs. RCF 2	0.54 (2.24)	0.41 (2.34)	0.34	.735	0.06
RCF 1 vs. RCF 3	0.54 (2.24)	0.46 (1.94)	0.25	.806	-0.27
RCF 1 vs. RCF 4	0.54 (2.24)	0.54 (2.28)	0.001	.999	0.00
RCF 2 vs. RCF 3	0.41 (2.34)	0.46 (1.94)	-0.12	.904	-0.02
RCF 2 vs. RCF 4	0.41 (2.34)	0.54 (2.28)	-0.33	.741	-0.06
RCF 3 vs. RCF 4	0.46 (1.94)	0.54 (2.28)	-0.24	.811	-0.04

Notes. Control 1 n = 81, Control 2 n = 80, RCF 1 n =84, RCF 2 n = 78, RCF 3 n = 80, RCF 4 n = 80.

Table 31

Effects of the Reference Class Forecasting Manipulations on Predicted Project Completion Times (Part 1 sample; Study 6)

	M (SD)		*t*	*p*	*d*
Control 1 vs. Control 2	1.76 (2.03)	1.31 (1.69)	-0.59	.554	-0.24
Control 1 vs. RCF 1	**1.76 (2.03)**	**0.50c (2.40)**	**-2.80**	**.006**	**-0.56**
Control 1 vs. RCF 2	**1.76 (2.03)**	**0.83 (2.33)**	**-2.23**	**.025**	**-0.42**
Control 1 vs. RCF 3	**1.76 (2.03)**	**0.64 (2.17)**	**-2.72**	**.007**	**-0.53**
Control 1 vs. RCF 4	**1.76 (2.03)**	**0.40 (2.24)**	**-2.71**	**.007**	**-0.63**
Control 2 vs. RCF 1	**1.31 (1.69)**	**0.50 (2.40)**	**-2.14**	**.034**	**-0.39**
Control 2 vs. RCF 2	**1.31 (1.69)**	**0.83 (2.33)**	**-1.69**	**.095**	**-0.24**
Control 2 vs. RCF 3	**1.31 (1.69)**	**0.64 (2.17)**	**-2.02**	**.045**	**-0.34**
Control 2 vs. RCF 4	**1.31 (1.69)**	**0.40 (2.24)**	**-2.09**	**.038**	**-0.46**
RCF 1 vs. RCF 2	0.50 (2.40)	0.83 (2.33)	0.34	.735	0.14
RCF 1 vs. RCF 3	0.50 (2.40)	0.64 (2.17)	0.25	.806	0.06
RCF 1 vs. RCF 4	0.50 (2.40)	0.40 (2.24)	0.001	.999	-0.04
RCF 2 vs. RCF 3	0.83 (2.33)	0.64 (2.17)	-0.12	.904	-0.08
RCF 2 vs. RCF 4	0.83 (2.33)	0.40 (2.24)	-0.33	.741	-0.19
RCF 3 vs. RCF 4	0.64 (2.17)	0.40 (2.24)	-0.24	.811	-0.11

Notes. Control 1 n = 81, Control 2 n = 80, RCF 1 n =84, RCF 2 n = 78, RCF 3 n = 80, RCF 4 n = 80.

Table 32

Multiple Regression Analyses Testing for Moderation Effects (Study 6)

Main Effect or Interaction Term	*t*	*p*	95% CI for *B*		R^2_{Change}
Step 1					
Control 1 vs. RCF 1	-3.61	.000	-1.94	-.567	.074
Step 2 (add)					
Control 1 vs. RCF 1	-2.62	.010	-1.49	-.210	
Mean Past Completion Times	-5.83	.000	-0.63	-.313	.161
Step 3 (add)					
Condition x Past Completion Times	**-2.39**	**.018**	**-0.74**	**-.071**	**.026**
Step 1					
Control 2 vs. RCF 1	-2.50	.014	-1.46	-.169	.037
Step 2 (add)					
Control 2 vs. RCF 1	-1.67	.096	-1.07	.089	
Mean Past Completion Times	-6.57	.000	-0.61	-.326	.203
Step 3 (add)					
Condition x Past Completion Times	**-2.57**	**.011**	**-0.65**	**-.085**	**.030**
Step 1					
Control 1 vs. RCF 2	-2.65	.009	-1.61	-.234	.043
Step 2 (add)					
Control 1 vs. RCF 2	-1.92	.057	-1.28	.019	
Mean Past Completion Times	-4.87	.000	-0.56	-.234	.126
Step 3 (add)					
Condition x Past Completion Times	**-1.66**	**.098**	**-0.62**	**.053**	**.015**
Step 1					
Control 2 vs. RCF 2	-1.48	.140	-1.12	.159	.014
Step 2 (add)					
Control 2 vs. RCF 2	-0.86	.392	-0.85	.334	
Mean Past Completion Times	-5.57	.000	-0.53	-.254	.164
Step 3 (add)					
Condition x Past Completion Times	**-1.71**	**.089**	**-0.53**	**.038**	**.015**
Step 1					
Control 1 vs. RCF 3	-3.36	.001	-1.77	-.460	.066
Step 2 (add)					
Control 1 vs. RCF 3	-2.84	.005	-1.62	-.291	
Mean Past Completion Times	-2.29	.023	-0.39	-.029	.030
Step 3 (add)					
Condition x Past Completion Times	-0.05	.958	-0.38	.360	.000
Step 1					
Control 2 vs. RCF 3	-2.20	.030	-1.282	-.068	.030
Step 2 (add)					
Control 2 vs. RCF 3	-1.76	.080	-1.14	.065	
Mean Past Completion Times	-2.86	.005	-0.39	-.070	.048
Step 3 (add)					

Condition x Past Completion Times	0.17	.862	-0.289	.345	.000
Step 1					
Control 1 vs. RCF 4	-4.01	.000	-2.020	-.686	.092
Step 2 (add)					
Control 1 vs. RCF 4	-3.63	.000	-1.931	-.570	
Mean Past Completion Times	-1.42	.159	-.290	.048	.011
Step 3 (add)					
Condition x Past Completion Times	0.70	.483	-.230	.485	.003
Step 1					
Control 2 vs. RCF 4	-2.91	.004	-1.532	-.293	.051
Step 2 (add)					
Control 2 vs. RCF 4	-2.58	.011	-1.438	-.191	
Mean Past Completion Times	-1.90	.060	-.292	.006	.021
Step 3 (add)					
Condition x Past Completion Times	1.07	.286	-.140	.470	.007
Step 1					
Control 1 vs. Control 2	-1.49	.138	-1.024	.143	.014
Step 2 (add)					
Control 1 vs. Control 2	-1.39	.166	-.977	.169	
Mean Past Completion Times	-2.70	.008	-.392	-.061	.043
Step 3 (add)					
Condition x Past Completion Times	-0.22	.823	-.372	.297	.000
Step 1					
Control 1 & 2 vs. RCF 1, 2, 3 & 4	-4.53	.000	-1.354	-.535	.041
Step 2 (add)					
Control 12 vs. RCF 1234	-3.52	.000	-1.101	-.312	
Mean Past Completion Times	-7.26	.000	-.420	-.241	.095
Step 3 (add)					
Condition x Past Completion Times	-1.21	.227	-.340	.081	.003

Table 33

Mediation Models Testing the Effect of the RCF procedure on Completion Predictions through Past Project Completion Beliefs (Study 6)

	"a" Path		"b" Path		"c" (Total Effect)		"c'" (Direct Effect)		"a*b" (Indirect Effect)	Boot CI
	β	*p*	β	*p*	β	*p*	β	*p*	B (*SE*)	LL, UL
Outcome: Predictions										
Control 1 vs. RCF 1	0.85	.006	-0.47	<.001	-1.25	<.001	-0.85	.010	**-0.40 (0.18)**	**-.7816, -.0996**
Control 1 vs. RCF 2	0.73	.024	-0.39	<.001	-0.92	.009	-0.63	.057	**-0.29 (0.14)**	**-.5844, -.0404**
Control 1 vs. RCF 3	0.77	.007	-0.21	.023	-1.12	.001	-0.95	.005	**-0.16 (0.10)**	**-.4075, -.0062**
Control 1 vs. RCF 4	0.85	.007	-0.12	.159	-1.35	.0001	-1.25	.0004	-0.10 (0.10)	-.3132, .0839
Control 2 vs. RCF 1	0.17	.034	-0.47	<.001	-0.81	.014	-0.49	.096	**-0.32 (0.17)**	**-.6998, -.0265**
Control 2 vs. RCF 2	0.56	.094	-0.39	<.001	-0.48	.140	-0.26	.392	-0.22 (0.14)	-.5245, .0365
Control 2 vs. RCF 3	0.61	.045	-0.23	.005	-0.68	.030	-0.54	.080	-0.14 (0.10)	-.3773, .0030
Control 2 vs. RCF 4	0.69	.038	-0.14	.060	-0.91	.011	-0.81	.011	-0.10	-.2917, .0514
Control 1 & 2 vs. RCF 1, 2, 3 & 4	0.72	<.001	-0.33	<.001	-0.94	<.001	-0.71	<.001	**-0.24 (0.08)**	**-.3953, -.1066**
Outcome: Past										
Control 1 vs. RCF 1	-1.25	<.001	-0.37	<.001	0.85	.006	0.39	.179	**0.46 (0.17)**	**.1715, .8216**
Control 1 vs. RCF 2	-0.92	.009	-0.33	<.001	0.73	.024	0.42	.170	**0.31 (0.12)**	**.0882, .5695**
Control 1 vs. RCF 3	-1.12	.001	-0.15	.067	0.77	.007	0.60	.040	**0.17 (0.12)**	**.0109, .4573**
Control 1 vs. RCF 4	-1.35	<.001	-0.10	.159	0.85	.007	0.71	.031	0.14 (0.14)	-.0806, .4517
Control 2 vs. RCF 1	-0.81	.014	-0.45	<.001	0.69	.034	0.32	.275	**0.37 (0.18)**	**.0733, .7572**
Control 2 vs. RCF 2	-0.48	.140	-0.42	<.001	0.56	.094	0.36	.243	0.20 (0.14)	-.0687, .5043
Control 2 vs. RCF 3	-0.68	.030	-0.22	.005	0.61	.045	0.46	.124	0.15 (0.12)	-.0012, .4577
Control 2 vs. RCF 4	-0.91	.004	-0.16	.060	0.69	.038	0.54	.105	0.14 (0.13)	-.0432, .4665
Control 1 & 2 vs. RCF 1, 2, 3, & 4	-0.94	<.001	-0.30	<.001	0.72	<.001	0.44	.024	**0.28 (0.08)**	**.1412, .4649**

Notes. Mediation analyses conducted using 10,000 bootstrap samples.

Table 34

Effects of Reference Class Forecasting on Predictions, Completion Times, Prediction Bias, and Additional Variables (Part 2 sample; Study 6)

	Control 1 - single prediction only	Control 2 - initial prediction	RCF 1 - initial prediction, past mean	RCF 2 - initial prediction, no past mean	RCF 3 - no initial prediction, past mean	RCF 4 - no initial prediction, no past mean	*F*	*p*	η_p^2
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)			
Average past project completion	-0.48 (1.36)	-0.28 (1.86)	-0.34 (1.94)	-0.16 (2.03)	0.36 (1.74)	0.31 (2.39)	1.62	.155	.029
Prediction (days before deadline)	**2.15a (2.16)**	**1.67ab (1.70)**	**1.28bc (1.83)**	**1.63bc (1.98)**	**0.88c (2.36)**	**1.17bc (1.57)**	**2.46**	**.034**	**.044**
Completion time (days before deadline)	1.48 (2.36)	1.49 (1.96)	1.83 (2.00)	2.02 (1.93)	1.12 (1.18)	1.72 (2.02)	1.26	.282	.023
Bias (prediction - completion time)	**0.67 (2.40)**	**0.19 (2.13)**	**-0.54 (2.59)**	**-0.39 (2.25)**	**-0.23 (2.05)**	**-0.54 (2.08)**	**2.09**	**.068**	**.038**
Bias binary* (finish on/before prediction vs. after)	27 vs. 18	27 vs. 16	29 vs. 15	28 vs. 13	40 vs. 11	34 vs. 12	X^2 = 5.25	.386	
Additional Measures									
Predicted work time (hrs.)	23.72 (9.01)	20.35 (9.65)	20.87 (10.15)	17.92 (10.54)	20.94 (10.61)	19.07 (9.51)	1.718	.131	.031
Work time (hrs.)	44.85 (53.01)	33.95 (46.55)	30.33 (34.02)	23.87 (22.89)	33.69 (47.22)	34.39 (51.48)	1.022	.405	.019
Work time bias (predicted - work time) (hrs.)	-21.13 (49.87)	-13.60 (45.25)	-9.46 (31.06)	-5.95 (19.66)	-12.75 (43.45)	-15.33 (48.85)	.673	.644	.013

Notes. Control 1 n = 46, Control 2 n = 43, RCF 1 n = 46, RCF 2 n = 41, RCF 3 n = 51, RCF 4 n = 46.

Table 35

Zero-Order Correlations (r) Between Past Project Completion Times, Predictions, and Actual Completion Times (Part 2 sample, Study 6)

	2. Completion Prediction	3. Actual Completion Time
Control 1		
1. Average Past Project Completion	-.09	-.14
2. Completion Prediction	--	
3. Actual Completion Time	.44**	--
Control 2		
1. Average Past Project Completion	-.32*	-.16
2. Completion Prediction	--	
3. Actual Completion Time	.33*	--
RCF 1		
1. Average Past Project Completion	-.54***	-.01
2. Completion Prediction	--	
3. Actual Completion Time	.09	--
RCF 2		
1. Average Past Project Completion	-.64***	-.12
2. Completion Prediction	--	
3. Actual Completion Time	.34*	--
RCF 3		
1. Average Past Project Completion	-.13	-.10
2. Completion Prediction	--	
3. Actual Completion Time	.50***	--
RCF 4		
1. Average Past Project Completion	.25†	-.02
2. Completion Prediction	--	
3. Actual Completion Time	.35*	--

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, †$p < .10$.

Table 36

Zero-Order Correlations (r) Between Completion Predictions and Completions Times, and Other Variables (Part 2 sample, Study 6)

	Completion prediction	"Actual" completion time
Predicted working time (hrs.)	-.06	-.03
Working time (hrs.)	-.13*	-.07

Notes. ***$p < .001$, **$p < .01$, * $p < .05$, $^{\dagger} p < .10$.

Table 37

Overview of Main Study Results by Hypothesis Tested

	H_1 Belief in planning fallacy for self	**H_2** Reminder of past completion times relative to predictions leads to later predictions	**H_3** Effect of past reminders is moderated by planning fallacy beliefs	**H_4** RCF leads to later, less optimistically biased completion predictions	**H_5** RCF will be more effective if past is recalled relative to predictions	**H_6** Effect of RCF is moderated by planning fallacy beliefs
S1	☑	☒	☒	--	--	--
S2	☑	☒	☒	--	--	--
S3	☑	--	--	☑	☑	--
S4	☒	--	--	☑ -compared to control 1 ☒ -compared to "past" group	☒	--
S5	☒	--	--	☒ (-some marginal effects)	--	☒
S6	☑ (-only slightly)	--	--	☑	--	☑ -RCF groups 2, 3 ☒ -RCF groups 4, 5

Table 38

Overview and Meta-analysis of the Effect Sizes for Planning Fallacy Beliefs and for the RCF Interventions on Predictions and Bias

Planning Fallacy Beliefs	***M* (*SD*)**	***d***
Study 1 (n = 186)	3.59 (1.39)	-0.30
Study 2 (n = 169)	3.10 (1.27)	-0.71
Study 3 (n = 55)	0.66 (1.57)	-0.42
Study 4 (n = 69)	-0.78 (2.12)	0.37
Study 5 (n = 268)	-0.46 (1.40)	0.33
Study 6 (n = 484)	0.25 (2.08)	-0.12
Meta-analytic		**-0.12**

	Predictions		**Prediction Bias**	
	Mean Difference	***Effect Size***	***Mean Difference***	***Effect Size***
S3 (Control vs. RCF prediction condition)	2.05 vs. 1.17 days before deadline	*d* = -0.56 (medium)	1.21 vs. 0.44 days later than predicted	*d* = -0.32 (small)
S4 (Control vs. RCF prediction condition)	1.82 vs. 1.04 days before deadline	*d* = -0.42 (medium)	1.28 days later vs. 0.47 days earlier than predicted	*d* = -0.82 (large)
S5 (Control 1 vs. RCF groups combined)	1.68 vs. 1.32 days before deadline	*d* = -0.22 (small)	0.73 vs. 0.24 days later than predicted	*d* = -0.30 (small)
S6 (Control 1 vs. RCF groups combined)	1.76 vs. 0.59 days before the deadline	*d* = -0.52 (medium)	0.67 vs. 0.42 days earlier than predicted	*d* = -0.48 (medium)
Meta-analytic		***d* = -0.40**		***d* = -0.44**
		CL = 0.61		*CL = 0.62*

Notes. Effect sizes were categorized via Cohen's (1998) suggested values. CL represents the common language effect size which represents the probability that a randomly sampled person from the control group will have a higher observed measurement (earlier prediction or greater prediction bias) than a randomly sampled person from the RCF group(s) (Grissom and Kim, 2005; McGraw and Wong, 1992).