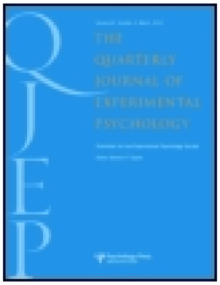

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Actors, observers, and the estimation of task duration

Michael M. Roy [a c], Nicholas J. S. Christenfeld [b] & Meghan Jones [a]
[a] Department of Psychology, Elizabethtown College, Elizabethtown, PA, USA
[b] Department of Psychology, University of California, San Diego, CA, USA
[c] School of Music, North-West University, Potchefstroom, South Africa

Accepted author version posted online: 07 Jun 2012.Version of record first published: 29 Aug 2012.

To cite this article: Michael M. Roy, Nicholas J. S. Christenfeld & Meghan Jones (): Actors, observers, and the estimation of task duration, The Quarterly Journal of Experimental Psychology, DOI:10.1080/17470218.2012.699973

To link to this article: http://dx.doi.org/10.1080/17470218.2012.699973

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, i*First*, 1–17

Actors, observers, and the estimation of task duration

Michael M. Roy[1,3], Nicholas J. S. Christenfeld[2], and Meghan Jones[1]

[1]Department of Psychology, Elizabethtown College, Elizabethtown, PA, USA
[2]Department of Psychology, University of California, San Diego, CA, USA
[3]School of Music, North-West University, Potchefstroom, South Africa

People are often wrong in estimating both how long tasks have taken in the past and how long they will take in the future. Bias could be due to factors such as task involvement, an individual's engagement or motivation in completing the task, or aspects of the task such as its relative duration or memory storage size associated with it. We examined time estimation bias in actors (likely to experience high levels of task involvement) and observers (likely to experience low levels of task involvement) for both predictions of and memory of task duration. Results suggest that bias appears to be due to memory storage size rather than to involvement with the task.

Keywords: Time estimation; Actor; Observer; Prediction; Memory.

People are often called upon to estimate duration, whether it is how long it took to drive across town or how long it will be until a major project will be completed. These estimates are often biased (see Roy, Christenfeld, & McKenzie, 2005, for review), with, frequently, systematic underestimating of task duration. When remembering how long it took to complete a task, people often think they finished more quickly than they actually did, and when predicting how long it will take to complete a task, they often think that they will finish sooner than they do. However, there are times, such as when the task is very short or unfamiliar, when, instead, overestimation is likely for both remembered and predicted duration (Boltz, Kupperman, & Dunne, 1998; Roy & Christenfeld, 2007, 2008).

Actor/observer differences in predicted and remembered duration

It has been speculated that a person's involvement with a task might be the cause of his or her bias for both predicted and remembered duration. Task involvement refers to level of engagement or effort during a task, potentially influencing motivation during the task. For prediction, people may focus too narrowly on the task at hand and ignore memories of how long similar tasks have taken in the past, resulting in an overly optimistic prediction (Kahneman & Tversky, 1982). If a person's involvement with a task contributes to the bias, then actors should exhibit more bias than observers (Buehler, Griffin, & Ross, 1994; Byram, 1997). Actors may be more susceptible to this type of bias because

Correspondence should be addressed to Michael Roy, Elizabethtown College, Department of Psychology, 1 Alpha Dr., Elizabethtown, PA, 17601, USA. E-mail: roym@etown.edu

http://www.psypress.com/qjep
http://dx.doi.org/10.1080/17470218.2012.699973

they are likely to focus on the immediate actions they are expected to take. In contrast, an observer, when predicting task duration, might be better able to objectively evaluate expected duration based on past experience because they are not required to consider taking action (Kahneman & Tversky, 1982). Additionally, a desire to see themselves in a positive light might cause actors, but not observers, to remember a task as taking less time than it actually did; when quick task performance reflects ability, participants are more likely to remember quicker completion times (Meade, 1959, 1960, 1963).

Possible causes for actor/observer differences

The tendency to underestimate duration can be seen as consistent with an overall tendency for people to be overly optimistic, as, for many tasks, people should generally be pleased with more rapid task completion (Armor & Taylor, 1998, 2002). For instance, people expecting a tax refund predict a quicker completion of their taxes than people not expecting a refund even though they did not actually submit their tax forms any earlier (Buehler, Griffin, & MacDonald, 1997). Having an optimistic view of how long it will take to finish a task may be beneficial in terms of motivating oneself to perform that task and also may serve as a form of goal setting that will speed actual performance, even if the optimistic goals are not met (Armor & Taylor, 1998, 2002). If people had a realistic idea of how long some tasks take, then they might decide not even to attempt them. Further, people might distort memory to help motivate themselves or envision a more pleasant future. For example, people seem to remember their vacations as being more pleasant than they actually were, resulting in an optimistic view of how pleasant their next vacation will be and increasing anticipation for that vacation (Mitchell, Thompson, Peterson, & Cronk, 1997; Wirtz, Kruger, Scollo, & Diener, 2003).

Optimistically biased memories might in turn cause optimistic predictions. One explanation of prediction errors, the memory bias account, holds that biased predictions are caused by biased memories (Roy et al., 2005; see also Addis, Wong, & Schacter, 2008; Schacter, Addis, & Buckner, 2008; Szpunar & McDermott, 2008, for more on the relationship between memory and prediction). In support of this view, factors that alter estimates, such as experience with a task (Roy & Christenfeld, 2007) and task duration (Roy & Christenfeld, 2008) have been shown to have parallel effects on memory and prediction, with biased memories for duration often leading to biased predictions (Roy, Mitten, & Christenfeld, 2008; Thomas, Handley, & Newstead, 2004, 2007). Actors might be more likely to remember a task as having taken them less time, because they were more involved in the task than observers, and, therefore, predict that they will complete it more quickly than they actually can. Observers might not exhibit a motivational bias because their predictions are based upon on memories not distorted by engagement or motivation.

Alternatively, differences in remembered duration between actors and observers might be due to cognitive differences in their experience with the task. Attentional models of time estimation hold that actor's memories might be distorted by the increased attention that they pay to the task during performance, which can lead to shorter estimations of duration (Thomas & Weaver, 1975; Zakay & Block, 1997). Actors would be likely to pay more attention to the task than observers and therefore be more likely to underestimate task duration. Bias may also be due to contents of memory for the task (Block & Reed, 1978; Ornstein, 1969). Memory models of time estimation hold that the larger the memory storage for a task (Ornstein, 1969) or the larger the number of remembered contextual changes (Block & Reed, 1978), the longer the estimated duration. Actors' involvement with the task may cause them to remember more or fewer aspects of the task than observers. Attentional models best explain bias in perception of time in passing while memory storage size/change models best explain bias in remembered task duration (Block & Zakay, 1997; Zakay & Block, 1997). Whether estimation is influenced by motivational or cognitive causes, a person highly involved with a task, therefore, would be more likely to be biased when estimating task duration than an uninvolved observer.

Possible reasons to not expect actor/observer differences

Of course, it is also possible that bias is not caused by task involvement, but by other nonmotivational factors, and, therefore, no differences would be expected between actors and observers. Research indicates that a number of biases that appear to be due to people's motivation to see themselves in a positive light might actually have nonmotivational causes. At times, people appear to overvalue their abilities, thinking they are better than they actually are, due to cognitive inadequacies in the way that they process information about their own ability and the ability of others (see Chambers & Windschitl, 2004, for review). Similarly, the tendency for people to seem overconfident in their responses might have more to with the type of questions being asked than with self-confidence (see Juslin, Winman, & Olsson, 2000, for review). It may be that bias in estimated duration does not have a motivational cause, but instead is due to other factors such as task novelty (Boltz et al., 1998; Hinds, 1999; Roy & Christenfeld, 2007), relative task duration (Lejeune & Wearden, 2009; Roy & Christenfeld, 2008; Yarmey, 2000), size of potential estimation anchors (König, 2005; Thomas & Handley, 2008), or duration since task completion (Ogden, Wearden, & Jones, 2008; Roy et al., 2008; Wearden & Ferrara, 1993). If bias is due to nonmotivational causes, it is possible that these factors would similarly influence actors and observers, and, therefore, both would be similarly biased in their memories and predictions. For example, actors and observers might not differ in number of remembered components for a task and would therefore be similarly biased (Block & Reed, 1978; Ornstein, 1969).

Previous research and overview of the current studies

While no previous experiments have directly examined possible difference between actors and observers for both predicted or remembered duration, five studies indirectly examined differences between actors and observers for predicted duration. Two studies found that actors were more optimistic than observers (Buehler et al., 1994; Newby-Clark, Ross, Buehler, Koehler, & Griffin, 2000). In these studies, actors estimated when a task would be completed, often in relationship to a deadline, and supplied a written description of how they made their prediction. Observers later read the instructions given to the actors and a portion of the actors' written thoughts and predicted when the actor would complete the task. Observers often predicted later completion times.

Three other studies (Byram, 1997; Hinds, 1999; Jørgensen, 2004), found no difference in bias between actors and observers predicting time needed to complete a task. In the study by Byram (1997), there was no difference between actors who estimated how long it would take them to complete a task and observers who estimated how long it would take for an average, hypothetical other to complete the same task. Similarly, Hinds (1999) had observers estimate how long it would take a novice to complete either a cell phone or a Lego task. Participants also estimated how long it took them to complete the task when they were new to the task. Participants remembered the task as taking less time than it actually had when they were new to the task and also underestimated how long it would take a novice to complete the same task. A final study by Jørgensen (2004) found no difference in bias or tendency to underestimate or overestimate, but did find that actors were more accurate overall. When predicting when a software project would be completed, participants that were going to take part in the project made more accurate predictions than participants not involved with the project.

In most of the previous studies, participants were either instructed to think about a hypothetical other or were given limited information about a previous, nonpresent participant. Making predictions about hypothetical others and tasks may involve different cognitive processes from those for time estimation regarding a task and person that are real and present. For example, Armor and Sackett (2006) found that participants were more likely to be biased when estimating the duration for hypothetical tasks than when estimating duration for real tasks. Further, in the studies where participants compare their own ability to that of others, participants are more likely to view themselves as being better than an abstract other, but not better than a concrete other (see Chambers & Windschitl, 2004). The current studies focus on bias in situations where observers do not have to try to imagine some abstract other completing a task, enabling a more direct examination of the influence of task involvement on remembered and predicted duration. Further, with the actor and observers present together, the duration that it takes the actor to complete the task can have a direct impact on the observers, where there can be no impact when estimating for a hypothetical or nonpresent other.

In two studies using three different tasks, we examine whether an actor, who is performing and therefore involved with the task, is more likely than an observer, who is not involved, to make biased estimations of future (Experiment 1) and past (Experiments 1 and 2) task duration. Actors and observers participated as yoked pairs, with both estimating task duration for the actor. Having actors and observers participate together allowed for direct comparison for possible differences in bias. Observers' estimates were not based on hypothetical situations or incomplete information. The only factor differing between the actor and observer was participation in the task and not a whole host of other potential differences. Further, we extend from previous research to examine actor and observer differences in remembered duration as well as predicted duration.

EXPERIMENT 1

Along with varying the role of the participant (actor and observer) and when the estimation was made (before or after the task), we also varied the task that participants completed. Half of the pairs were assigned a paper-counting task, which was chosen because previous experiments have found underestimation for both remembered and predicted duration for this task (Roy & Christenfeld, 2008; Roy et al., 2008). The other half of the participants was assigned a spellcheck task, a task that has been found to lead to overestimation (Burt & Kemp, 1994; although see Francis-Smythe & Robertson, 1999, for a study that found underestimation using a similar task). These tasks were chosen because motivational factors, which are likely to differ based on task involvement, may be the cause of bias for these tasks. Paper counting is a dull, manual task, and speedy completion might be desirable, leading to underestimation. On the other hand, checking for spelling errors requires a person to be detailed and thorough, and it might be desirable for participants to portray themselves as having these characteristics, leading to overestimation. It would be expected that observers would be less affected by these motivational causes and therefore make less biased estimates than actors (Buehler et al., 1994; Byram, 1997). If bias, however, is driven by factors specific to the tasks that are equally available to both the actor and observer, such as complexity or overall task duration, and not due to individual task involvement, then the patterns of bias should be similar for both actors and observers.

Method

Participants

Two-hundred and sixteen University of California, San Diego students (157 females, 59 males) participated, 108 as actors and 108 as observers. The scores for two pairs of participants performing the paper-counting task were eliminated because the actors in each pair did not understand the directions, leaving 212 total participants. Participants received course credit for their psychology classes in exchange for participation.

Design

The experiment was a three-way mixed-model design with two independent factors (when the estimation was made and type of task) and one matched factor (actor or observer). An actor performed either a paper-counting or spellchecking task while an observer looked on. The actor and observer estimated either how long it would take the actor to perform the task (prediction condition) or how long it had just taken the actor to perform the task (memory condition).

Procedure

Before participants arrived, it was randomly determined which task they would perform, paper counting or spellchecking, and whether they were in the prediction or memory condition. To ensure that they could not monitor time, participants were told that, because they would be working with their hands on the task, they should remove their watches and any rings or bracelets they were wearing. The two participants were seated directly across from each other at a 3-foot-wide table.

To determine who would be the actor and who the observer, a coin was flipped in front of both participants. To keep them naïve to the purpose of the experiment, participants were told that the coin flip was to determine the order in which they would be performing the task.

The paper-counting task consisted of counting out 250 sheets of paper by placing them into perpendicular stacks of 10. For the spellchecking task, participants were given a one-page, single-spaced short story containing spelling errors and were asked to circle all the errors they found. There were 12 typographical errors in 614 total words.

In the prediction condition, the experimenter explained the task, either paper counting or spellchecking, and gave the actor the necessary materials. After giving them sufficient time to familiarize themselves with the task (with no time limit), both participants completed questionnaires handed to them on separate clipboards that asked how long, in minutes and seconds, it would take the actor to perform the task. This question was embedded in some other background and distractor questions. Participants were asked not to discuss their answers with each other and to keep their responses private. The actor then performed the task while the experimenter surreptitiously timed him or her with a stopwatch.

In the memory condition, the actor was given the materials and, when ready, performed the task. Afterwards, both actor and observer completed questionnaires that asked how long it had taken the actor to perform the task. In both conditions, participants were then told that this was the end of the experiment and were informed of the true intent of the experiment.

Dependent variables

An index of estimation bias was created by taking the log of the ratio of estimated duration to actual duration, which we call log proportional error (Roy & Christenfeld, 2007, 2008). This index helps to simplify interpretation; a negative score indicates underestimation, a score of zero indicates perfect accuracy, and a positive score indicates overestimation. The index also normalizes the data; there was, as is generally found, a strong positive skew in estimates of duration. In addition, this index allows a comparison of the bias in estimates across tasks of different lengths.

We also examined the accuracy of actors' and observers' estimations by analysing their overall and relative error rates. Overall error, or average error disregarding sign, was measured by taking the absolute value of the difference between the estimated duration and the actual duration (absolute error). This measures how close people come to the actual duration, regardless of direction of deviation. Relative error was measured by finding the correlation between log estimated duration and log actual duration. A high correlation indicates that participants know if they are relatively slow or fast at the task compared to others. It is possible that actors might be more biased than observers, consistently under- or overestimating task duration, but are overall more accurate, making less overall and relative error. Previous research has found that measures of bias and other measures of accuracy can be largely independent (Epley & Dunning, 2006; Stone & Opel, 2000).

Results and discussion

The median duration to perform the paper-counting task was 6.8 min (interquartile range, IQR = 2.6), and for the spellchecking task the median was 3.8 min (IQR = 1.5). The median for the

Table 1. *Median estimated and actual task duration for Experiment 1*

Task	*Condition*	*Estimated duration* *Actor*	*Estimated duration* *Observer*	*Actual duration*
Paper counting	Future	5.0 (2.0)	4.0 (4.0)	7.0 (2.2)
	Past	6.0 (3.5)	6.0 (3.4)	6.6 (3.6)
Spellcheck	Future	5.5 (6.0)	6.0 (5.4)	4.0 (1.6)
	Past	4.8 (2.8)	4.9 (4.0)	3.6 (1.2)

Note: Durations in min. Interquartile range in parentheses.

estimates, taking future and past estimations and actors and observers together, was 5 min (IQR = 3.0) for the paper-counting task and 4.1 min (IQR = 2.7) for the spellchecking task (see Table 1 for full results).

Bias in estimated duration

To examine bias, a 2 × 2 × 2 mixed-model analysis of variance (ANOVA) was performed on the log proportional error (log of the ratio of estimated to actual duration). As expected, there was a significant main effect of task, $F(1, 102) = 28.47$, $p < .0001$, $\eta^2 = .18$, with both actors and observers tending to underestimate the paper-counting task and overestimate the spellchecking task (see Figure 1). However, estimates were less biased for both actors and observers in the past condition for both tasks, as indicated by a significant interaction between task and when estimation was given, $F(1, 102) = 10.17$, $p = .002$, $\eta^2 = .06$. Simple effects

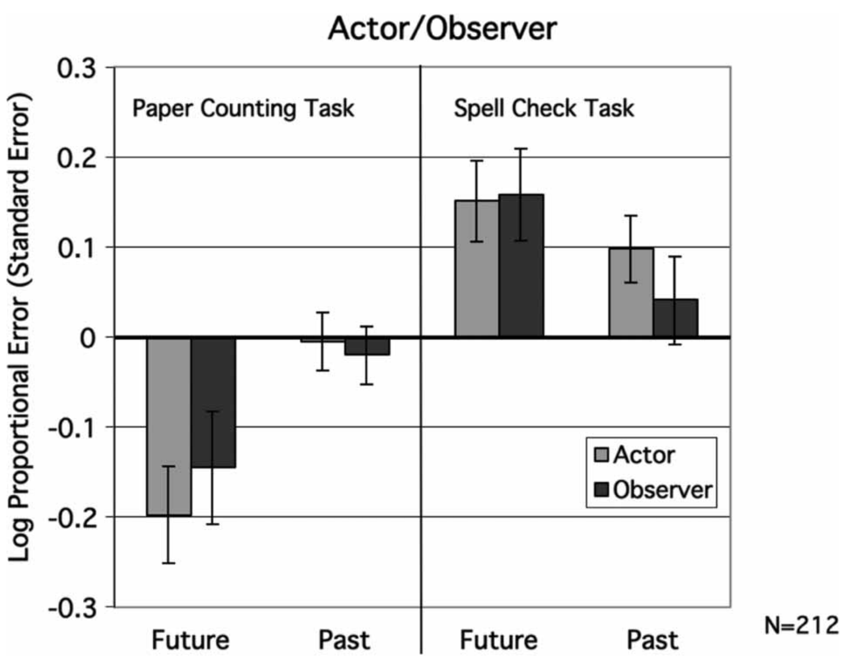

Figure 1. *Log proportional error [log(estimated duration/actual duration)] for actors and observers as a function of when estimation was given (past or future) and type of task performed (paper-counting task or spellcheck task).*

Table 2. *Percentage of participants overestimating and underestimating task duration for Experiment 1*

Task	*Condition*	*Participant*	*Underestimation*	*Overestimation*
Paper counting	Future	Actor	83	17
		Observer	70	30
	Past	Actor	59	41
		Observer	62	38
Spellcheck	Future	Actor	17	78
		Observer	30	70
	Past	Actor	23	77
		Observer	42	58

Note: One actor in the future, spellcheck condition provided a correct estimate.

test indicate significantly less bias for remembered than for predicted duration for the paper-counting task, $p = .003$, but not for the spellcheck task, $p = .13$. The majority of participants underestimated the paper-counting task and overestimated the spellchecking task; effects were not driven by a minority grossly under- or overestimating task duration (see Table 2).

Most importantly, there was no difference between actors and observers in bias. The main effect for actors and observers was not significant, $F(1, 102) = 0.01$, $p = .91$, $\eta^2 < .001$. Also, there was no significant interaction between actors and observers and whether the estimation was given before or after the task, $F(1, 102) = 1.35$, $p = .25$, $\eta^2 = .004$, or between actors and observers and the type of task performed (paper-counting task or spellcheck task), $F(1, 102) = 0.61$, $p = .44$, $\eta^2 = .002$, and no three-way interaction, $F(1, 102) = 0.001$, $p = .98$, $\eta^2 < .001$.

Absolute error in estimated duration

Average absolute error (the absolute value of estimated duration minus actual duration) was 2.7 min ($SD = 2.5$) for actors and 3.0 min ($SD = 3.8$) for observers (see Figure 2). A 2 × 2 × 2 ANOVA on absolute error indicated no differences between actors and observers and no significant interactions involving actors and observers and the other variables (all Fs < 2.9, ps > .09). The only significant difference in absolute error was a main effect of when estimation was given: Participants made less overall error when the estimate was given after the task (M = 1.9 min, $SD = 2.4$) than when it was given before the task (M = 3.8 min, $SD = 3.5$), $F(1, 102) = 17.75$, $p < .001$, $\eta^2 = .14$.

For both the log proportional error and the absolute error analyses, the lack of difference between actors and observers does not appear to be due to insufficient power. Given our large sample size (106 actors and 106 observers), power calculations indicate that there was more than sufficient power to detect an effect (Cohen, 1988; power calculation indicate that the power to detect a moderate effect was 1 for all main effects, .96 for all two-way interactions, and .73 for the three-way interaction). Further, as can be seen from the above analyses, the effect sizes for possible actor and observer differences were all very small (all $\eta^2 < .01$). Differences between actors and observers accounted for less than one percent of the variability in bias and accuracy of the estimates.

Relative error in estimated duration

To examine relative error, the correlation between log estimated and log actual duration was computed for each task separately and then averaged together (first converting to Fisher z scores, averaging together, then converting back). Results indicate that, on average, the correlation between actors' predictions of how long the task would take and the actual duration, $r = .38$, was significantly more accurate than the correlation between predicted and actual duration for observers, $r = -.06$, $Z = 2.28$, $p = .02$ (see Table 3 for full results). Actors, reasonably, were more likely than were observers to know whether they were going to be relatively slower or faster than others. While actors' predictions of task

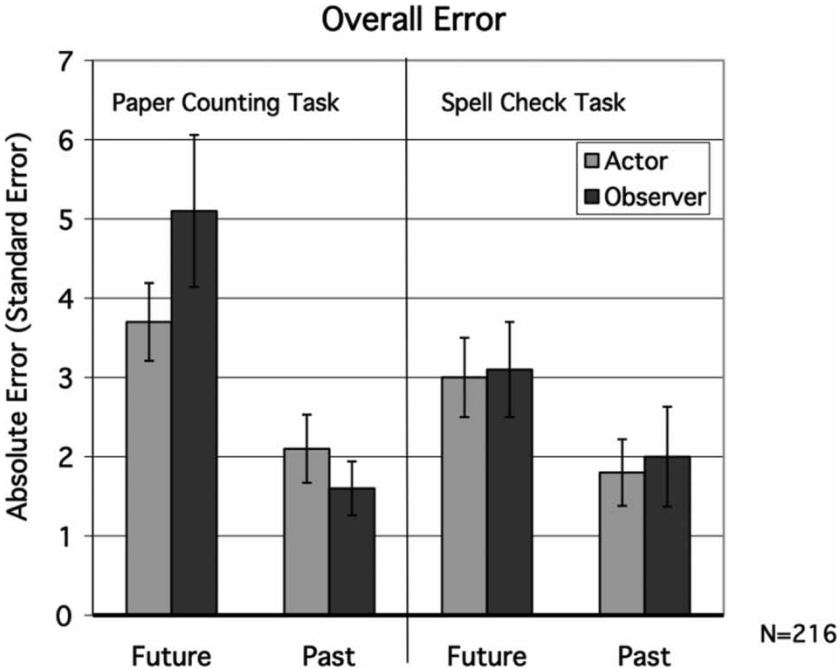

Figure 2. *Absolute error [absolute value (estimated duration – actual duration)] for actors and observers as a function of when estimation was given (past or future) and type of task performed (paper-counting task or spellcheck task).*

Table 3. *Correlation between estimated and actual task duration for Experiment 1*

Task	*Condition*	*Actor*	*Observer*
Paper counting	Future	.24	–.20
	Past	.57*	.47*
Spellcheck	Future	.53*	.12
	Past	.52*	.62*

*$p < .05$.

duration were more closely aligned to actual duration than those of observers, their estimates were not unrelated: The few minutes spent with the actor before the task provided observers with enough information when making their predictions that there was at least some level of agreement between actors and observers on how long they thought it would take that actor to complete the task, $r = .36$, $p = .007$. It is interesting to note that while the actors and observers tended to make similar predictions for how long it would take the actor to complete the task, only the actors' predictions were related to task completion time. Any cues that were visible to both the actor and observer that led them to jointly predict relatively fast or slow completion times for the actor were not related to task performance.

There were no differences between actors and observers for relative accuracy in remembered duration, with a fairly strong correspondence between estimated and actual duration for both: actors, $r = .54$ vs. observers, $r = .55$, $Z = –0.04$, $p = .97$.

Comparison to previous work

Having an audience present introduced another possible source of motivation: an attentive audience. The presence of others appears to increase motivation and performance on well-learned tasks (Zajonc, 1965) and may similarly affect prediction for how long these tasks will take. Predicting task duration in the presence of others has been found to increase bias and decrease accuracy (Pezzo, Pezzo, & Stone, 2006). Here, a desire to appear competent to an audience may increase bias in estimated duration.

However, it does not appear that the presence of an audience affected bias for either the actor or the observer, at least for the paper-counting task. Results for the paper-counting task for actors and observers were compared to the results of participants from an earlier experiment conducted in the same lab, during approximately the same time period, where participants completed the task alone (Roy & Christenfeld, 2008). A 3 (actor, observer, alone) × 2 (before, after) ANOVA on log proportional error indicated that there was no main effect of condition (actor, observer, or alone), $F(2, 156) = 0.13$, $p = .88$, $\eta^2 = .001$, and no interaction between condition and when the estimate was given (before or after the task), $F(2, 156) = 0.53$, $p = .59$, $\eta^2 = .006$ (see Figure 3). As found previously, there was a significant main effect of when the estimate was given, with estimates more likely to be biased when given before the task than when given after the task, $F(1, 156) = 11.98$, $p < .001$, $\eta^2 = .07$. Overall, actors and observers completing the paper-counting task did not differ in bias from participants that performed the task alone in a previous experiment. However, caution should be taken in interpreting the results since the actor/observer pairs and the participants that were alone were part of two different experiments. To remedy this, whether or not participants were alone or in pairs was directly manipulated in Experiment 2.

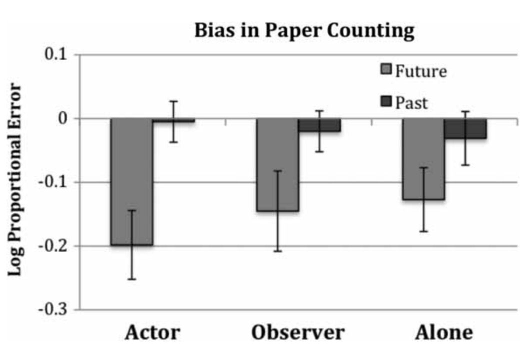

Figure 3. *Log proportional error [log(estimated duration/actual duration)] for actors, observers, and participants alone as a function of when estimation was given (past or future) for the paper-counting task.*

Possible influence of task duration on bias

The different role of actors and observers did not affect bias in estimation, and other causes of bias may have affected actors and observers equally. Previous research indicates that task duration affects bias in both memory and prediction (Roy & Christenfeld, 2008). Similarly, the current results indicate that task duration was related to bias in estimation for both actors and observers with a significant correlation between the actual duration (act) and log proportional error (lpe) for remembered duration [$r(104) = -.25$, $p = .01$, $\text{lpe} = -.02(\text{act}) + .14$] (see Figure 4) and for predicted duration [$r(104) = -.50$, $p < .001$, $\text{lpe} = -.08(\text{act}) + .43$] (see Figure 5). As actual duration increased, bias in estimation moved from overestimation to underestimation. As can be seen in Figures 4 and 5, the relationship between actual duration and bias appears to be fairly continuous between the two tasks; there was little difference in bias for participants that were slow at the spellchecking task and those that were fast at the paper-counting task. However, the results are only suggestive, and it is possible that numerous other differences in the two tasks, and not simply duration, were the cause of the shift from overestimation to underestimation. For example, differences in memory storage size associated with the completed task or previous similar tasks may have contributed to bias with the simpler task (paper counting) likely to be underestimated and the more complex task (spellchecking) likely to be overestimated. In either case, it appears that task factors other than level of involvement were the likely cause of bias for these tasks.

Summary

There was a striking level of agreement between actors and observers. Whether actors were completing a spellcheck or paper-counting task, there was no difference between actors and observers for either remembered or predicted duration. It is possible that the difference in bias for these tasks might be due to differences in the overall duration of the tasks, with the shorter

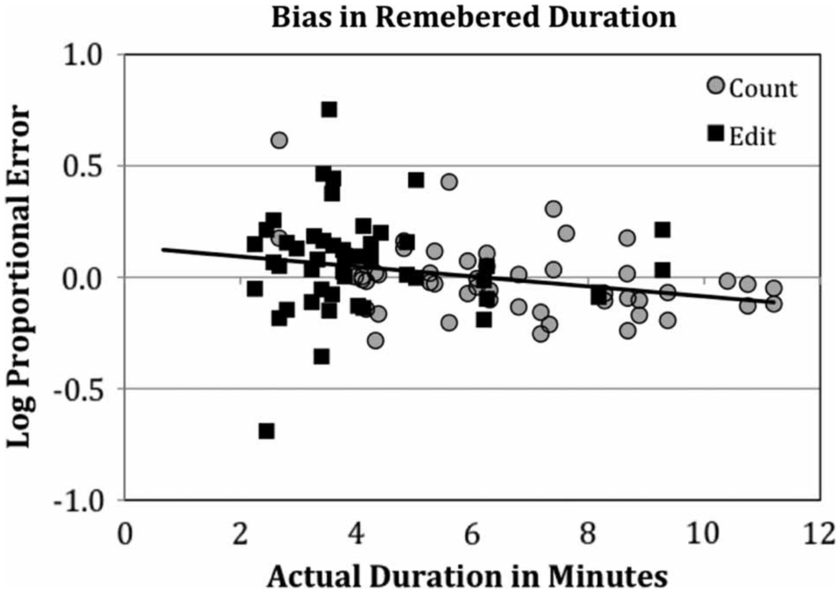

Figure 4. *Log proportional error [log(estimated duration/actual duration)] for remembered duration plotted as a function of actual duration.*

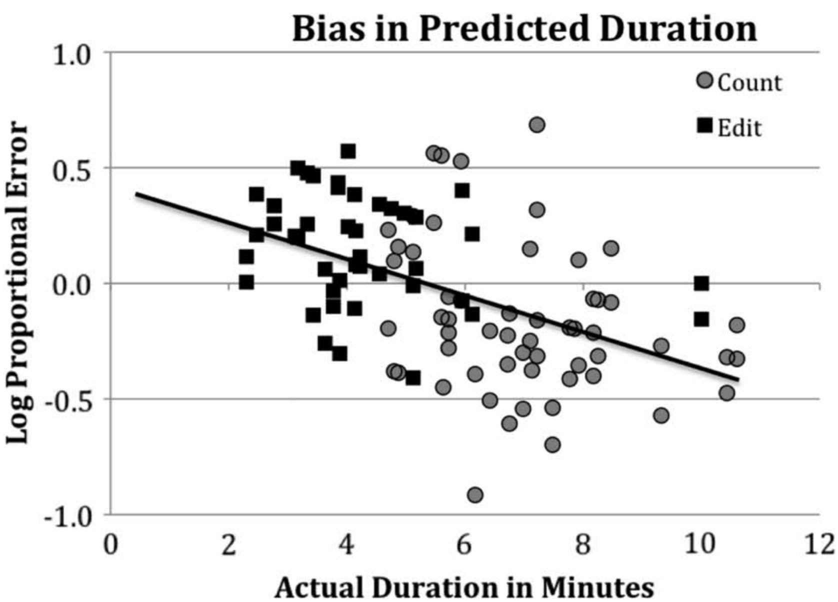

Figure 5. *Log proportional error [log(estimated duration/actual duration)] for predicted duration plotted as a function of actual duration.*

task overestimated and the longer task underestimated.

The tasks performed in these studies may not have not engaged the participants in a way that tied their self-concept to their performance, which would have led them to distort estimates to make themselves look better. Indeed, actors and observers for the paper-counting task did not differ in bias from participants that completed the task alone. Therefore, in Experiment 2 a potentially more engaging task was used.

EXPERIMENT 2

There were two goals for Experiment 2: to examine whether the findings of Experiment 1 replicate using a more engaging task, and to examine more directly the role of memory size in bias. Because memory effects are most clearly examined with retrospective judgements of duration, and because we found, in the first study and in other work, that predictions and memories were substantially similar, in the second study estimation was restricted to memory for previous task duration.

The first goal for Experiment 2 was to examine possible actor and observer differences for a more demanding task than that in the previous studies, likely inducing greater differences in task involvement between actors and observers, and to better examine possible causes of bias. Participants in Experiment 2 completed a speech task where they discussed the merits of a college education. Giving a speech is a task that many find to be highly stressful, and, for this reason, it is a commonly used task in studies examining physiological consequences of stress (Uchino, Cacioppo, & Kiecolt-Glaser, 1996). Further, college-aged students rate the ability to speak in public as being a very important skill, more important than either the ability to drive a car or to detect emotion in others (Roy, Liersch, & Broomell, 2012).

The speech task also allowed us to test potential causes of bias in estimates. Bias in estimated task duration might be due to amount of information about the task stored in memory, with tasks that have more remembered components having longer estimates of duration (Block & Reed, 1978; Ornstein, 1969). Participants were asked to record the number of distinct ideas expressed by the speaker to examine whether or not the number of remembered changes in topic related to bias in remembered duration for the speech. To examine the possible effect of engagement, role (actor or observer) and audience (whether or not another participant was present) were manipulated, and level of engagement was measured. It would be expected that having an audience would make the speech more important to the participant with the potential to make them more engaged with the speech. Estimation for Experiment 2 was limited to memory for past task duration since it would be difficult for participants to estimate number of concepts that would be discussed and level of engagment with speech before actually giving the speech, especially for observers.

Method

Participants

Eighty-one Elizabethtown College students (59 females, 22 males) participated as an actor ($n = 27$), an observer ($n = 27$), or alone ($n = 27$). One participant in the alone condition did not understand the directions and was eliminated, leaving 26 participants in the alone condition. For estimated duration and number of concepts remembered, one actor and one observer, respectively, gave non-numeric responses, but the rest of their results were retained for analysis. Participants received course credit for their psychology classes in exchange for participation.

Design

One participant, the actor, completed a speech task while in the presence of another participant, in the presence of the observer, or alone (only in the presence of the experimenter). Participants estimated how long it had just taken the actor (sometimes themselves) to complete the task, their level of engagement during the task, and the number of distinct concepts they remembered from the speech.

Procedure

Participants were seated across from one another at a table in the experimental room, and a coin was flipped in front of both to determine who would be the actor and who would be the observer. In the single condition, participants were seated at the table alone. Participants were not permitted to reference cell phones or watches during the experiment.

The experimenter explained that the task was to give a short speech about why it is important to go to college. Participants were given a few minutes to prepare the speech if they wanted and were given scrap paper to take notes. Although they were allowed to reference these notes during the speech, notes were collected immediately after the speech so that participants could not use them in estimating either duration or number of concepts covered. Participants giving the speech were told to begin when ready.

After the speech, all participants completed a questionnaire that asked how long the speech lasted and how many different distinct ideas or concepts were in the speech. The questionnaire also asked participants to rate their level of engagement during the speech on a scale of 1 (not engaged) to 9 (very engaged). Participants were asked not to discuss their answers with each other and to keep their responses private.

Results and discussion

The median duration for the speech task was 35 s (IQR = 35) while the estimated duration was 60 s (IQR = 60). Overall, participants tended to overestimate the length of the speech, with log proportional error ($M = .12$, $SD = .23$) significantly greater than 0, $t(78) = 4.52$, $p < .001$, $d = 0.51$.

Actor and observers

As can be seen in Table 4, observers tended to be slightly more biased and make less overall error, but neither of these differences were significant: log proportional error, $t(25) = -1.09$, $p = .29$, $d = -0.21$; absolute error, $t(25) = 0.44$, $p = .66$, $d = 0.09$. Also there was no significant difference in relative accuracy between actors, $r(26) = .55$, and observers, $r(27) = .68$; $Z = -0.7$, $p = .48$. Surprisingly, there were no significant differences between actors and observers in terms of reported engagement with the task, $t(26) = -1.19$, $p = .24$, $d = -0.23$. Observers did tend to remember the actor making more difference points during the speech than did the actors, $t(25) = -2.67$, $p = .01$, $d = 0.52$. For potential differences between actors and observers, power to detect a moderate size differences was .75.

Participating together or alone

Participants who were alone ($M = 61.27$, $SD = 43.40$) tended to give longer speeches than participants who had an audience ($M = 43.81$, $SD = 32.73$), $t(51) = 1.66$, $p = .10$, $d = 0.45$. While participants acting alone tended to be more biased and less accurate with more overall error, there was no significant difference due to the presence of an audience: log proportional error, $t(50) = 1.02$, $p = .31$, $d = 0.28$; absolute error, $t(50) = 1.21$, $p = .23$, $d = 0.33$ (see Table 4). Participants acting alone, $r(26) = .87$, were more likely than those with an audience, $r(27) = .55$, to know whether they were relatively fast or slow ($Z = 2.4$, $p = .02$). This could indicate a benefit to not having an audience, but the increased correlation could also be due to the fact that participants who were alone gave longer speeches with greater variability, thereby increasing the chance for a strong relationship between estimated and actual duration. Finally, there was no difference in the number of remembered concepts in the speech, $t(51) = 1.24$, $p = .22$, $d = 0.34$, or reported engagement, $t(51) = 0.24$, $p = .81$, $d = 0.07$. It should be noted

Table 4. *Means and standard deviations for Experiment 2*

	Log proportional error	*Absolute error*	*Concepts*	*Engagement*
Actor	.08 (.28)	25.69 (31.39)	3.61 (1.11)	5.70 (1.38)
Observer	.13 (.21)	23.08 (16.74)	4.38 (1.32)	6.18 (1.92)
Alone	.14 (.20)	40.50 (53.83)	4.10 (1.70)	5.81 (1.79)

Note: Standard deviations in parentheses.

that power for the between-subjects analysis was fairly low, with a .43 chance of detecting a moderate size effect.

Possible causes of bias

Unlike the previous study, there was no relationship between bias (log proportional error) and task duration, $r(79) = -.09$, $p = .43$, but there was a relationship between actual duration and overall error (absolute error), with longer speeches leading to greater error in estimation, $r(79) = .40$, $p < .001$. Neither bias nor overall error was related to reported level of engagement with the task: $r(79) = .08$, $p = .47$; $r(79) = .17$, $p = .13$, respectively. An increase in the number of remembered concepts was associated with greater overestimation of duration [log proportional error, $r(78) = .22$, $p = .05$] and a decrease in accuracy overall [absolute error, $r(78) = .34$, $p = .002$]. Further, differences in number of concepts remembered can almost completely explain the small, nonsignificant, differences between actors and observers in bias. When difference between actors and observers in the number of remembered concepts was added as a covariate to the analysis, effect size for the difference between actors and observers in log proportional error dropped from $\eta^2 = .054$ to $\eta^2 = .001$. Differences in number of memories associated with the task could also potentially explain why participants that were alone had less relative error: The absence of an audience may have removed distractions that would interfere with memory of the task. Results indicate that bias was influenced by memory for the number of different points made during the speech, but not to duration of the speech or how engaged the participant was with the speech.

Summary

Results provide support for memory storage size explanations for bias (Block & Reed, 1978; Ornstein, 1969). There was greater overestimation (and more error overall) when speeches were remembered as containing a greater number of distinct ideas. Aspects of the task itself, but not role as actor or observer, provide better accounts for bias and error in estimation of task duration.

As with Experiment 1, there was little difference between actors and observers in terms of either bias or overall error. While participants were biased in their estimates, estimating that the task took longer than it actually did, their particular role did not influence bias or error. However, it should be noted that there were no reported differences in level of engagement between roles. Overall engagement level was fairly high, approximately 6 on a scale from 1 to 9, with observers actually reporting slightly higher levels of engagement. While it could be that the speech task was equally engaging for participants to either give or listen to, this seems somewhat unlikely since studies employ giving a speech, but not listening to a speech, as a stressor (Uchino et al., 1996). It may be that actors and observers used different frames of reference when making their ratings; comparing their level of engagement to previous times they have given or listened to a speech, respectively. Ability to use different frames of reference is a weakness of Likert-type scales like the one used here (Bartoshuk et al., 2003). While it would seem that giving a speech is more engaging and stressful than listening to a speech, the lack of difference in reported engagement does leave open the possibility that our manipulation did not have the desired effect and indicates that our results should be interpreted with some caution.

GENERAL DISCUSSION

The findings from the two studies indicate that the particular role of the estimator, either that of an actor or of an observer, is not critical for consistent bias to emerge. Results from these studies indicate that different task factors might be a more likely cause of bias. In Experiment 1, bias in estimation was related to the duration of the task with overestimation for short tasks and underestimation for longer tasks (see also Lejeune & Wearden, 2009; Roy & Christenfeld, 2008; Yarmey, 2000). In Experiment 2, bias was related to number of remembered concepts in the speech, with more concepts associated with greater overestimation. These results are consistent with memory storage size explanations for bias (Block & Reed, 1978; Ornstein, 1969). Participants remembering more content of the speech (Ornstein, 1969) or more changes in topic (Block & Reed, 1978) were more likely to overestimate task duration. It is also possible that the contents of memory can explain the results for Experiment 1, with the simple, lower memory content task (paper counting) leading to underestimation, and the more complex, higher memory content task (spellcheck) leading to overestimation. The relationship between task duration and bias and Experiment 1 and between number of concepts remembered and bias in Experiment 2 are correlational and therefore should be interpreted with caution. However, the results are consistent with bias being caused by factors other than the participants' role with the task.

Over three very different tasks we found no reliable difference between actors and observers in their estimates of both remembered and predicted task duration. In terms of prediction, the results are consistent with the Byram (1997) and Hinds (1999) studies that found no differences between actors and observers in estimation of future task duration. It does not seem to matter whether or not the actor is a live other, as in the current studies, or a hypothetical other, as in the Byram and Hinds studies; there was no influence of role in bias. It should be noted that the current study and the Bryam and Hinds studies used fairly short lab tasks all lasting less than approximately one hour. In contrast, the study by Jørgensen (2004), which found that observers made less accurate predictions than actors, and those by Buehler et al. (1994) and Newby-Clark et al. (2000), which found that actors were more likely to underestimate when a task would be completed than observers, all used much longer tasks, usually lasting longer than one day. It may be that there is a lack of actor/observer differences for shorter tasks, but inconsistent actor/observer differences for long tasks. Alternatively, the different experiments can be delineated by the type of estimation that was made: The current study and the studies by Byram, Hinds, and Jørgensen all had participants predicting how long a task would take, while the Buehler et al. and Newby-Clark et al. studies had participants predicting when a task would be completed, often in relation to a deadline. It is possible that very different cognitive processes are used when estimating how long it will take to complete a task and when a task will be finished, with task completion time having to incorporate a number of other variables such as when the task will be started and when other projects will be completed (Halkjelsvik & Jørgensen, 2012). Any potential benefit that would come from being an observer seems only to be present when predicting when a task will be finished. It should be noted, however, that while the Buehler et al. and Newby-Clark studies found a decrease in the tendency to underestimate for observers, observers were no more accurate overall, often exhibiting more overall/absolute error. Taken together, the results of these studies point to little benefit, and at times a detriment, to having people predict task duration from the vantage point of an observer. Taking an alternative viewpoint is not an intervention that appears to lead to more accurate predictions. Further, the current study extends beyond the previous ones to include memory for past task duration. Here again there does not seem to be a benefit to being an observer.

It may well be possible to construct a scenario where actors and observers would make similarly biased estimates, but for very different reasons. For example, the motivational involvement of the actor could be balanced by the inattention of the observer. However, it is less likely that separate processes could lead to the pattern of results found here, a tendency toward overestimation for two of the tasks and a tendency toward underestimation for the other. Given that actors and observers showed similar patterns of bias for the different tasks, it would seem likely that both used a fundamentally similar process in forming their estimates. Further, the process that they used does not seem to be affected by task involvement, but by task factors such as number of remembered components and relative duration. It does not appear that there was a difference in the way that actors and observers experienced the task that caused them to differ in either remembered or predicted duration.

Although we found no significant difference between actors and observers in bias and overall error, results indicate that the predictions of actors do have at least some value beyond those made by observers. There was greater correspondence for actors between predicted and actual duration for the tasks. They knew whether they were likely to be relatively fast or slow at the task, while observers did not. Not surprisingly, as they had both just observed the same performance, the advantage held by actors in prediction was no longer present when the correspondence between memory and actual duration was examined.

Summary

The current results suggest that being the one performing the task, and therefore being more actively involved, is not necessary for consistent and predictable bias to emerge. Bias is possible, likely, and systematic without task involvement. It appears that task factors, such as whether the task is relatively short or long, or number of remembered components, caused both actors and observers to be biased on these tasks. These results suggest that attempts at improving estimates of task duration, whether for how long a task took or how long a task will take, should examine specific nonmotivational factors about that task that might lead to bias. Interventions aimed at changing perspective to that of an observer do not seem likely to lead to improvement in either remembered or predicted task duration. Interventions would probably be more successful if they took into account memories associated with the task.

Original manuscript received 17 June 2011
Accepted revision received 18 May 2012
First published online 30 August 2012

REFERENCES

Addis, D. R., Wong, A. T., & Schacter, D. L. (2008). Age-related changes in the episodic simulation of future events. *Psychological Science*, *19,* 33–41.

Armor, D. A., & Sackett, A. M. (2006). Accuracy, error, and bias in predictions for real versus hypothetical events. *Journal of Personality and Social Psychology*, *91,* 583–600.

Armor, D. A., & Taylor, S. E. (1998). Situated optimism: Specific outcome expectancies and self-regulation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 309–379). San Diego, CA: Academic Press.

Armor, D. A., & Taylor, S. E. (2002). When predictions fail: The dilemma of unrealistic optimism. In T. D. Gilovich, D. W. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 334–347). New York, NY: Cambridge University Press.

Bartoshuk, L. M., Duffy, V. B., Fast, K., Green, B. G., Prutkin, J., & Snyder, D. J. (2003). Labeled scales (e.g., category, Likert, VAS) and invalid across-group comparisons: What we have learned from genetic variation in taste. *Food Quality Preference*, *14,* 125–138.

Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, *4,* 656–665.

Block, R. A., & Zakay, D. (1997). Prospective and retrospective durations judgments: A meta-analytic review. *Psychonomic Bulletin & Review*, *4,* 184–197.

Boltz, M. G., Kupperman, C., & Dunne, J. (1998). The role of learning in remembered duration. *Memory & Cognition*, *26,* 903–921.

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*, *67,* 366–381.

Burt, C. D. B., & Kemp, S. (1994). Construction of activity duration and time management potential. *Applied Cognitive Psychology*, *8,* 155–168.

Byram, S. J. (1997). Cognitive and motivational factors influencing time prediction. *Journal of Experimental Psychology: Applied*, *3,* 216–239.

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: The role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin*, *130,* 813–838.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Epley, N., & Dunning, D. (2006). The mixed blessing of self-knowledge in behavioral prediction: Enhanced discrimination but exacerbated bias. *Personality and Social Psychology Bulletin*, *32,* 641–655.

Francis-Smythe, J. A., & Robertson, I. T. (1999). On the relationship between time management and time estimation. *British Journal of Psychology*, *90,* 33–347.

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.

Hinds, P. J. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied*, *5,* 205–221.

Jørgensen, M. (2004). Regression models of software development effort estimation accuracy and bias. *Empirical Software Engineering*, *9,* 297–314.

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard–easy effect. *Psychological Review*, *107,* 384–396.

Kahneman, D., & Tversky, A. (1982). Intuitive prediction: Biases and corrective procedures. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgments under uncertainty: Heuristics and biases* (pp. 414–421). Cambridge, UK: Cambridge University Press.

König, C. J. (2005). Anchors distort estimates of expected duration. *Psychological Reports*, *96,* 253–256.

Lejeune, H., & Wearden, J. H. (2009). Vierordt's *The Experimental Study of the Time Sense* (1868) and its legacy. *European Journal of Cognitive Psychology*, *21,* 941–960.

Meade, R. D. (1959). Time estimates as affected by motivational level, goal distance, and rate of progress. *Journal of Experimental Psychology*, *58,* 275–279.

Meade, R. D. (1960). Time estimates as affected by need tension and rate of progress. *Journal of Psychology*, *50,* 173–177.

Meade, R. D. (1963). Effect of motivation and progress on the estimation of longer time intervals. *Journal of Experimental Psychology*, *65,* 564–567.

Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. (1997). Temporal adjustment in the evaluation of events: The "rosy view". *Journal of Experimental Social Psychology*, *33,* 421–448.

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

Ogden, R. S., Wearden, J. H., & Jones, L. A. (2008). The remembrance of times past: Interference in temporal reference memory. *Journal of Experimental Psychology: Human Perception and Performance*, *34,* 1524–1544.

Ornstein, R. E. (1969). *On the experience of time*. Harmondsworth, UK: Penguin Books.

Pezzo, S. P., Pezzo, M. V., & Stone, E. R. (2006). The social implications of planning: How public predictions bias future plans. *Journal of Experimental Social Psychology*, *42,* 221–227.

Roy, M. M., & Christenfeld, N. J. S. (2007). Bias in memory predicts bias in estimation of future task duration. *Memory & Cognition*, *35,* 557–564.

Roy, M. M., & Christenfeld, N. J. S. (2008). Effect of task length on remembered and predicted duration. *Psychonomic Bulletin & Review*, *15,* 202–208.

Roy, M. M., Christenfeld, N. J. S., & McKenzie, C. R. M. (2005). Underestimation of future duration: Memory incorrectly used or memory bias. *Psychological Bulletin*, *131,* 738–756.

Roy, M. M., Liersch, M. J., & Broomell, S. (2012). *People believe that they are prototypically good and bad*, Manuscript submitted for publication.

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

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: Concepts, data, and application. *Annals of the New York Academy of Sciences,* Special Issue: The Year in Cognitive Neuroscience 2008*, 1124*, 39–60.

Stone, E. R., & Opel, R. B. (2000). Training to improve calibration and discrimination: The effect of performance and environmental feedback. *Organizational Behavior and Human Decision Processes*, *83,* 282–309.

Szpunar, K. K., & McDermott, K. B. (2008). Episodic future thought and its relation to remembering: Evidence from ratings of subjective experience. *Consciousness and Cognition*, *17,* 330–334.

Thomas, K. E., & Handley, S. J. (2008). Anchoring in time estimation. *Acta Psychologica*, *127,* 24–29.

Thomas, K. E., Handley, S. J., & Newstead, S. E. (2004). The effects of prior experience on estimating the duration of simple tasks. *Current Psychology of Cognition*, *22,* 83–100.

Thomas, K. E., Handley, S. J., & Newstead, S. E. (2007). The role of prior task experience in temporal misestimation. *Quarterly Journal of Experimental Psychology*, *60,* 230–240.

Thomas, E. A. C., & Weaver, W. B. (1975). Cognitive processing and time perception. *Perception and Psychophysics*, *17,* 363–367.

Uchino, B. N., Cacioppo, J. T., & Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin*, *119,* 488–531.

Wearden, J. H., & Ferrara, A. (1993). Subjective shortening in humans' memory for stimulus duration. *The Quarterly Journal of Experimental Psychology*, *46B*(2), 163–186.

Wirtz, D., Kruger, J., Scollo, C. N., & Diener, E. (2003). What to do on Spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science*, *14,* 520–524.

Yarmey, A. D. (2000). Retrospective duration estimations for variant and invariant events in field situations. *Applied Cognitive Psychology*, *14,* 45–57.

Zajonc, R. B. (1965). Social facilitation. *Science*, *149,* 269–274.

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science*, *6,* 12–16.