Online Supplemental Materials

More Time, More Work: How Time Limits Bias Estimates of Project Duration and Task Scope

Contents

Appendix A: Study Materials (including Stimuli and Instructions)

Study 1

Jigzone.com interface which was integrated in a Qualtrics Survey in Study 1. Both workers and judges saw this picture. The same picture was used in Study 4.

Additional picture of the completed 20-piece jigsaw puzzle shown to workers and judges in Study 1. The same picture was used in Study 4.

Please read the following information SLOWLY AND CAREFULLY. This will be important to the choice you will be asked to make as well as for a bonus you can earn (details are provided later), and we will confirm your understanding at the end of the survey.

We conducted an actual study earlier this year, in which the workers were primarily students at commuter colleges participating in a research lab in downtown Chicago.

Workers were given the puzzle shown above, and were asked to solve the puzzle **either with no restriction on the maximum time** they could take, or a **maximum time of 5 minutes**, or a **maximum time of 15 minutes** to solve the puzzle. The time limit each person got was assigned completely at random, and workers did not get to choose the time limit.

Regardless of the time limit they had or how long it took, as soon as a worker finished the puzzle, that worker was

paid $3 and, after answering a few follow-up questions, could either leave or participate in other unrelated paid studies in that lab.

Next, we will ask you a few questions in which you will estimate the time it took people to finish solving the puzzle.

You can earn up to an additional $1 bonus payment, based on your accuracy. We will pick one of your estimates below at random and compare your response to the average actual time in the study we ran, described on the previous pages.

If your estimate matches the average time exactly, you will get an extra $1. For every minute of difference between your estimated time and the actual average time, we will deduct ten cents from the bonus. Any bonus you earn will be paid as a bonus payment via mechanical turk.

So, you will earn a larger bonus if you are as accurate as you can be in making your estimate on the next page.

Common instructions in Study 1.

Consider a typical worker for whom there was **no restriction on the maximum time** that could be taken to solve the 20 piece puzzle shown earlier.

Do you think that the typical worker spent **more or less than 5 minutes** working on the puzzle?

- ○ The typical worker spent LESS than five minutes
- ○ The typical worker spent five minutes
- ○ The typical worker spent MORE than five minutes

How much time do you think the average worker with no time limit took to solve the puzzle?

Remember you can earn an additional reward of up to $1.00 for making an accurate prediction.

Please enter the time in MINUTES in the box below:

Questions asked to investigate potential anchoring and adjustment account of time limits in Study 1. The figure shows the 5 minutes (short time limit) condition as an illustrative example.

Consider a typical worker for whom there was a **maximum time of 5 minutes** to solve the 20 piece puzzle shown earlier.

How much time do you think the average worker with a maximum time limit of 5 minutes took to solve the puzzle?

Remember you can earn an additional reward of up to $1.00 for making an accurate prediction.

Please enter the time in MINUTES in the box below:

The five-minutes condition in Study 1

Consider a typical worker for whom there was a **maximum time of 15 minutes** to solve the 20 piece puzzle shown earlier.

How much time do you think the average worker with a maximum time limit of 15 minutes took to solve the puzzle?

Remember you can earn an additional reward of up to $1.00 for making an accurate prediction.

Please enter the time in MINUTES in the box below:

The fifteen-minutes condition in Study 1

How do you think workers were assigned to get the unrestricted time condition, or the 5 minute maximum time condition, or the 15 minute maximum time condition?

○ They were assigned completely at random

○ They were assigned based on how fast and proficient they were in solving jigsaw puzzles, or some other factor

End-of-survey comprehension check questions in Study 1

Study 2

Assume, that we have a 100 piece jigsaw puzzle. It need not be of the same picture which was earlier shown to you as an example.

We administered this puzzle to a group of individuals in a mid-western university and each person took less than 28 minutes to solve it.

As part of this study, we would like to administer the same puzzle to another group of people who are similar to those in our previous group.

To confirm the number of pieces in the puzzle which we are going to use in this experiment, please enter the number below to proceed further.

Common instructions to all judges in Study 2

Assume that there is no limit on the amount of time individuals in the group have to solve the puzzle.

This means that an individual can take as much time as they need to solve the jigsaw puzzle.

Instructions in the control (unlimited time) condition in Study 2

Assume, that our study has two conditions based on the total amount of time each individual is given to solve the puzzle.

Condition 1: Total time available to solve the puzzle is 30 minutes.
Condition 2: Total time available to solve the puzzle is 45 minutes.

A fair coin is tossed and if the result is HEADS, the person is assigned to Condition 1. Otherwise, if the coin comes up TAILS, the individual is assigned to Condition 2.

Therefore, every individual in the new group will have an equal chance of getting assigned to either of the two conditions.

Once assigned to a condition, individuals will work until they solve the problem or the time limit is over, whichever occurs earlier.

Instructions in the treatment (i.e., deadline) conditions in Study 2

However, here is an IMPORTANT point (please read carefully).

A person assigned to either of the above mentioned conditions is not told that there is a time limit.
So, participants in the 30 minutes condition and in the 45 minutes condition do not know about these time limits.

The experimenter who has assigned them to either of the two conditions is just going to stop a participant from proceeding any further in case he or she is not done with solving the puzzle in the allotted experimental time.

The participant is only told to solve the puzzle at his or her own pace.

Additional instructions in the treatment (i.e., deadline) conditions in Study 2

What were the maximum time available to a person to solve the jigsaw puzzle in the two conditions described in the study?

Please enter the values in MINUTES below.

Condition 1 (Max. time in MINUTES) [ ]

Condition 2 (Max. time in MINUTES) [ ]

Did people working under 30 minutes time limit and those working under 45 minutes time limit solve the same puzzle?

- ○ Yes, both solved the same jigsaw puzzle
- ○ No, they solved different jigsaw puzzles
- ○ Don't Know/Can't Say

How do you think people were assigned to one of the two time limit conditions?

- ○ They were assigned at random
- ○ They were assigned based on how fast and proficient they were in solving jigsaw puzzles

How many pieces were there in the jigsaw puzzle? Please enter the value in the box below.

[ ]

When we had tested the 100 piece jigsaw puzzle with the first group of individuals in a mid-western university, what was the maximum time people took to solve the puzzle?

Please enter the value in MINUTES below.

Time in MINUTES

Regardless of how much time they actually had, how much time do you think participants in the study **thought** they had to solve the puzzle?

Participants thought they had _______ Minutes:

- ○ Unlimited time
- ○ No more than _____ minutes:

Condition: No more than _____ minutes: Is Selected. Skip To: How many minutes did they think they ....

How many minutes did they think they had?

End-of-survey comprehension check questions in Study 2

Study 3

As you might know, **a jigsaw puzzle requires the assembly of numerous small, often oddly shaped, interlocking pieces to produce a complete picture**.

The pictures below show an example jigsaw puzzle (with the pieces shuffled, Fig. 1), and how it looks when it is solved (Fig. 2)

Fig. 1: Jigsaw puzzle

Fig. 2: Solved puzzle

Online jigsaw puzzles (like the ones used in this survey) are played using software that initially shuffles the puzzle pieces, and allows the player to drag them together. The software also gives feedback (via an audible snap sound) when two valid pieces are brought together. This ensures that a particular puzzle has only one solution, i.e., the correct solution.

The online software also automatically (and accurately) records the time taken to solve a jigsaw puzzle.

Imagine that we ask actual participants to solve an online jigsaw puzzle and the software accurately records the time taken. In this survey, your task is to predict these completion times for a few jigsaw puzzles. We will then ask you a few followup questions about your thinking process. That's all.

In the next few pages, we will provide you with information about the participants who work on the puzzles, the situation they face including details of the jigsaw puzzles they attempt to solve.

Common instructions in Study 3

We plan to recruit several **students from commuter colleges** in a large Midwestern city in the United States.

So, **our typical participant is not a jigsaw-puzzle expert**, but an ordinary person like you and me. Also, **participants solve only one puzzle** so they do not get an opportunity to practice and hone their skills.

We plan to pay every participant a fixed flat fee for the work. After they are done working, they collect their payment and leave the venue.

IMPORTANT NOTE: In the past, we have done some preliminary tests asking such participants to solve similar jigsaw puzzles. In those tests, **every individual participant was able to solve the puzzle that he/she was assigned. Also, NONE of the participants took more than 30 minutes to solve any puzzle.**

About the situation facing the participants

There is an assigned time limit within which every participant is expected to solve the puzzle. **The time limit is random**: chosen between one of two time limits BASED ON A COIN TOSS.

If the coin turns up HEADS, the participant will have a maximum time limit of **35 minutes** to solve the assigned puzzle.

If the coin turns up TAILS, the participant will have a maximum time limit of **50 minutes** to solve the assigned puzzle.

As you would realize, **every participant has an equal chance of getting one of the two time limits**.

NOTE CAREFULLY:

Participants are told about their time limit before they start working on the puzzle. In particular, they are informed that they would work until they are done or the assigned time limit expires, whichever occurs earlier.

However, the way the time limit is set (i.e., using a coin toss) is for administrative convenience and **participants are NOT told that the assigned time limit was randomly determined**.

Additional instructions in the condition where workers knew about the time limits in Study 3. In the condition where workers *did not know* the time limit, the last block was changed.

Before you proceed to the next page, please answer a few yes/no questions to indicate your comprehension. All the information required for answering these questions is there in the instructions you have read so far.

	No	Yes
The jigsaw puzzles are administered to participants using an online software	○	○
Only experts are recruited as participants for solving the jigsaw puzzles	○	○
Some participants are asked to solve more than one jigsaw puzzles	○	○
Participants are paid a fixed flat fee to work on a jigsaw puzzle	○	○
There is an assigned time limit for every participant within which he/she is expected to finish solving the puzzle	○	○
The assigned time limit for any participant is randomly determined	○	○
Participants know about their assigned time limit before starting to work	○	○
In the past, some participants took more than 30 minutes to solve the assigned puzzle	○	○
In the past, every participant was able to successfully solve the assigned puzzle	○	○

Comprehension questions about the experimental setup in Study 3

The following is one puzzle that participants solve, and **we want you to predict the completion time for a typical participant.**

Fig.1: Jigsaw puzzle

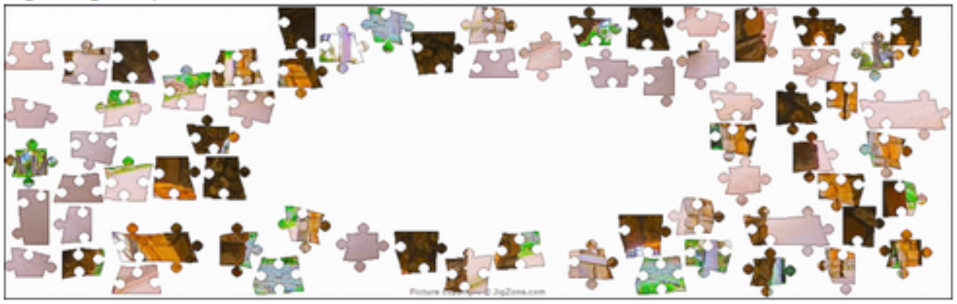


Fig. 2: Solved puzzle

IMPORTANT NOTE: The particular individual for which you are predicting is assigned a time limit of 35 minutes, based on a coin toss, to solve the puzzle.

Also, **the participant knows that he/she has a time limit of 35 minutes** before starting to work on the puzzle.

Please spend some time on this page carefully thinking about the jigsaw puzzle. How might the participant approach this task? How much work do you think solving this puzzle would entail given the situation the participant is facing?

You will be asked to enter your prediction (time in minutes, the participant took to solve the puzzle) on the next page.

Click Next only when you are ready to enter your prediction.

The first decision scenario in Study 3 (here the first decision was for a task that had a 35 minutes time limit; the task and the time limit was counterbalanced in the experiment)

Now it is time to enter your prediction.

How much time do you think the person will take to solve the puzzle? Remember that **the participant has a maximum time limit of 35 minutes** and he/she cannot work after the time limit expires.

Please enter the time in MINUTES in the box below.

Please enter only a valid number (i.e., no texts).

Eliciting judges' estimation in Study 3. A similar question followed after judges were introduced to the second decision scenario shown below.

The following is **another puzzle that a separate group of participants solves.** We now want you to predict the completion time for this puzzle for a typical participant.

Please read carefully as the situation might have changed.

Fig.1: Jigsaw puzzle

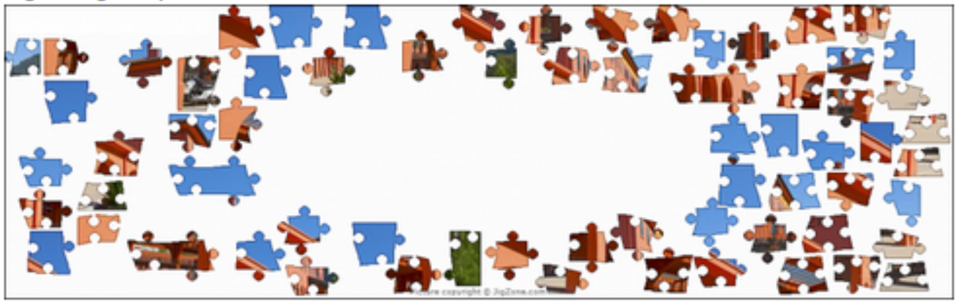


Fig. 2: Solved puzzle

IMPORTANT NOTE: The particular individual for which you are predicting is assigned a time limit of 50 minutes, based on a coin toss, to solve the puzzle.

Also, **the participant knows that he/she has a time limit of 50 minutes** before starting to work on the puzzle.

Please spend some time on this page carefully thinking about the jigsaw puzzle. How might the participant approach this task? How much work do you think solving this puzzle would entail given the situation the participant is facing?

You will be asked to enter your prediction (time in minutes, the participant took to solve the puzzle) on the next page.

The second decision scenario in Study 3

Recall that you were asked to make prediction decisions about these two jigsaw puzzles.

	
Puzzle 1 (maximum time limit = 35 mins) The survey is administered using an online interface that accurately records the time taken. Participants are commuter college students who are randomly assigned to one of two time-limits. The time limit is known to the participants.	**Puzzle 2 (maximum time limit = 50 mins)** The survey is administered using an online interface that accurately records the time taken. Participants are commuter college students who are randomly assigned to one of two time-limits. The time limit is known to the participants.

Which puzzle, according to you, did a typical participant find more interesting and enjoyable?

1. Puzzle 1 was more interesting and enjoyable	2	3	4. Equally interesting and enjoyable	5	6	7. Puzzle 2 was more interesting and enjoyable

Post-decision follow-up question in Study 3. Each of the follow-up questions was asked on a separate page with back navigation disabled.

Recall that you were asked to make prediction decisions about these two jigsaw puzzles.

	
Puzzle 1 (maximum time limit = 35 mins) The survey is administered using an online interface that accurately records the time taken. Participants are commuter college students who are randomly assigned to one of two time-limits. The time limit is known to the participants.	**Puzzle 2 (maximum time limit = 50 mins)** The survey is administered using an online interface that accurately records the time taken. Participants are commuter college students who are randomly assigned to one of two time-limits. The time limit is known to the participants.

According to you, "how much work" is there in solving **Puzzle 1**? Please drag the bar below to indicate your answer (1=A Little Work; 100= A Lot of Work).

1 11 21 31 41 51 60 70 80 90 100

Puzzle 1

According to you, "how much work" is there in solving **Puzzle 2**? Please drag the bar below to indicate your answer (1=A Little Work; 100= A Lot of Work).

1 11 21 31 41 51 60 70 80 90 100

Puzzle 2

Post-decision follow-up question in Study 3. Each of the follow-up questions was asked on a separate page with back navigation disabled.

Recall that you were asked to make prediction decisions about these two jigsaw puzzles.

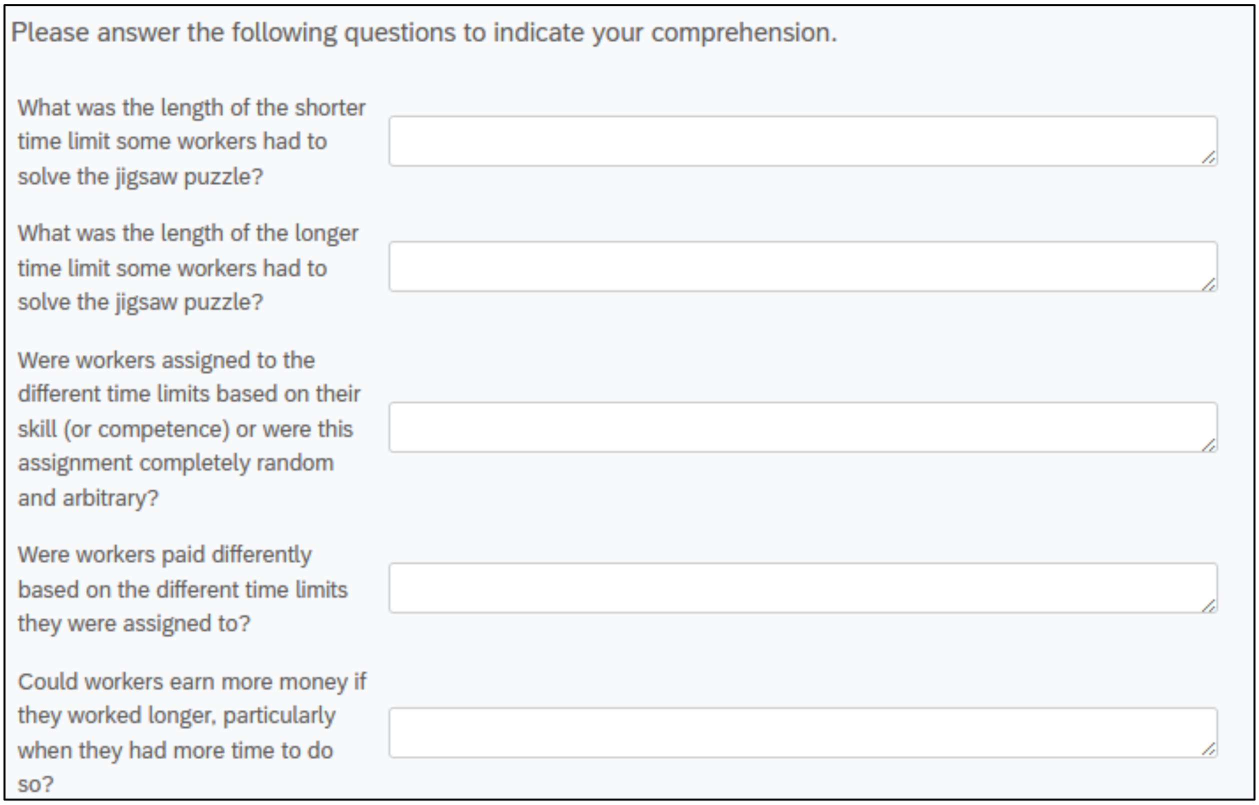
	
Puzzle 1 (maximum time limit = 35 mins) The survey is administered using an online interface that accurately records the time taken. Participants are commuter college students who are randomly assigned to one of two time-limits. The time limit is known to the participants.	**Puzzle 2 (maximum time limit = 50 mins)** The survey is administered using an online interface that accurately records the time taken. Participants are commuter college students who are randomly assigned to one of two time-limits. The time limit is known to the participants.

Which jigsaw puzzle do you think had more number of pieces?

1. More pieces in Puzzle 1	2	3	4. Same number of pieces	5	6	7. More pieces in Puzzle 2

Post-decision follow-up question in Study 3. Each of the follow-up questions was asked on a separate page with back navigation disabled.

Study 4



Please answer the following questions to indicate your comprehension.

What was the length of the shorter time limit some workers had to solve the jigsaw puzzle?

What was the length of the longer time limit some workers had to solve the jigsaw puzzle?

Were workers assigned to the different time limits based on their skill (or competence) or were this assignment completely random and arbitrary?

Were workers paid differently based on the different time limits they were assigned to?

Could workers earn more money if they worked longer, particularly when they had more time to do so?

End-of-survey comprehension check questions in Study 4

Study 5

Imagine that you have hired a third party vendor which sends out customized mailers to promote your business to prospective customers. The vendor will use their own list and they specialize in customizing the mailers based on other information they have about the individuals.

Once the deal is finalized, the vendor calculates how many mailers can be sent depending on how many prospects they find in their database. The vendor then spends 4 weeks customizing the mailers and then sends them out.

Common instructions in Study 5

However, you just came across an industry report which says that direct mail is least effective in late summer. Based on this information, you tell the vendor to delay the mailing by two weeks so that it goes out in the early Fall.

This was a last minute decision after the list of prospective customers was already finalized. But, because of this change, the vendors will instead get 6 weeks for customizing the mailers before sending them out.

Additional instructions in the long time limit condition (6 weeks) in Study 5

As mentioned earlier, the vendor has **4 weeks** to prepare the mailers.

Also as mentioned earlier, the vendor prepares the master list of prospective customers first and then does the customization for each. The vendor sends a customized mailer to each person in this list.

How many mailers do you think the vendor would **actually** send out? Just give us your best guess.

- ○ Less than 5000
- ○ 5000 to 10000
- ○ 10000 to 30000
- ○ 30000 to 50000
- ○ 50000 to 80000
- ○ 80000 to 100000
- ○ More than 100000

Eliciting scope of work in the short time limit condition in Study 5

How many customized mailers do you think an average employee of the vendor who is working only on this project is preparing per day? Just give us your best guess.

- Less than 100
- 100 to 200
- 200 to 500
- 500 to 800
- 800 to 1000
- Greater than 1000

Eliciting rate of work in both the short and the long time limit condition in Study 5

As mentioned earlier, the vendor has **4 weeks** to prepare the mailers.

How much time do you think the vendor would **actually** take to prepare the customized mailers? Please indicate which of these comes closest to your best guess.

- 25% of the available time (1 week)
- 37.5% of the available time (1.5 weeks)
- 50% of the available time (2 weeks)
- 62.5% of the available time (2.5 weeks)
- 75% of the available time (3 weeks)
- 87.5% of the available time (3.5 weeks)
- 100% of the available time (4 weeks)

Elicitation procedure for time estimation in the short time limit condition in Study 5. In the long time limit condition (6 weeks) the percentages of available time were the same (25%, 37.5%, etc.), and the absolute numbers were calculated accordingly (1.5 weeks, 2.25 weeks, etc.)

Appendix B: Data Handling and Comprehension Dropouts

Data Exclusion: All online participants with duplicate IP addresses were excluded from analysis (in all studies) before looking at the data. In addition, judges in Studies 1-4 answered the following attention check question towards the end of the survey, and those who failed this check were excluded from analysis (in all studies) before looking at the data.

People vary in the amount they pay attention to these kinds of surveys. Some take them seriously and read each question, whereas others go very quickly and barely read the questions at all. If you have read this question carefully, please write the word yes in the blank Other box below.

1 Not at all
2
3
4
5 A great deal
6 Other ________________________________

Following are the drop rates in the online studies. We also examine if these rates are different by the experimental conditions. Please note the following:

1. In order to be consistent, we compute the drop rates on the base of all Ps from whom some responses were captured when the survey ended. Some of these responses were incomplete. As a result, dropout rates might look high. For example, in the pre-registered study 3, we requested for 350 respondents and had 347 usable responses at the end (< 1% drop rate). But, based on the number of respondents in our data set (i.e., 392), the drop rate looks larger.
2. Phase 1 of Study 1 (i.e., the workers' phase) used lab participants, and there were no dropouts. Study 5 used a paid online panel, and there were no dropouts.
3. In Study 3, the time limit was a within-subjects factor. The only between-subjects factor was whether the workers knew their time limits (yes, no), which we use below.

The results confirm that the initial dropouts did not vary by experimental conditions.

Study #	Overall drop rate	Experimental Conditions (between/subjects)	Chi-Square results
1	8.2%	5, 15, 5-Anchor, 15-Anchor	$\chi 2(3)= 1.44, p= .694$
2	15.7%	Untimed, 30, 45	$\chi 2(2)= 3.04, p= .219$
3	11.5%	Deadline known, not known	$\chi 2(1)<1$
4	10%	5, 15	$\chi 2(1)<1$
4	10%	2(Time Limit: 5,15) x 2(Beliefs: C, Q)	$\chi 2(3)<1$

Dropouts because of comprehension failure: In all studies reported in the manuscript, we included everyone who passed the initial exclusion criteria reported above. However, we examined the robustness of our reported findings, in the online appendix, by looking at the subset of participants who correctly answered the comprehension check questions. Here we examine whether the dropout on account of comprehension check failures varies by experimental conditions.

Two points to note:

1. Study 5 used real managers and did not have any comprehension checks
2. In Study 3, time limit was a within-subjects factor. The only between-subjects factor was whether the workers knew their time limits (yes, no).

The results confirm that these dropouts did not vary by experimental conditions. The results of Study 4 show a marginal difference when we consider all the four cells in the 2x2 design. We do not think this small difference is systematic.

Study #	Number of comprehension check questions	Experimental Conditions (between/subjects)	Chi-Square results
1	1	5, 15, 5-Anchor, 15-Anchor	$\chi2(3)<1$
2	6	Untimed, 30, 45	$\chi2(2)<1$
3	2 (pre-registered analysis)	Deadline known, not known	$\chi2(1)<1$
3	9	Deadline known, not known	$\chi2(1)<1$
4	5	5,15	$\chi2(1)<1$
4	4	2(Time Limit: 5,15) x 2(Beliefs: C, Q)	$\chi2(3)=7.23$, $p=.065$

Appendix C: Additional Analysis

Study 1

Handling Extreme Values: While the time estimates (under deadlines) were bounded to the maximum available time limit, Study 1 also employed an anchoring manipulation where judges could enter a numerical estimate that was not bounded. We examined the robustness of our conclusions after replacing extreme estimate values in the anchoring conditions with minimum or maximum non-outlier values.

Outlier Handling Strategy	1-way ANOVA	Anchoring vs. Time Limit Estimation
Log transformation	$F(3,598)=40.64$, $p<.001$	5 mins: $t(300)=+5.78$, $p<.001$ 15 mins: $t(298)=+1.32$, $p=.186$
Truncation to maximum available time limit	$F(3,598)=80.51$, $p<.001$	5 mins: $t(300)=+4.79$, $p<.001$ 15 mins: $t(298)=+1.09$, $p=.275$
Iterative Grubbs test	$F(3,598)=50.09$, $p<.001$	5 mins: $t(300)=+7.35$, $p<.001$ 15 mins: $t(298)=+1.59$, $p=.111$
Winsorizing 90%	$F(3,598)=53.17$, $p<.001$	5 mins: $t(300)=+7.49$, $p<.001$ 15 mins: $t(298)=+1.64$, $p=.101$
Winsorizing 95%	$F(3,598)=48.39$, $p<.001$	5 mins: $t(300)=+7.23$, $p<.001$ 15 mins: $t(298)=+1.58$, $p=.113$

Within-Estimates: After estimating task completion time for a particular time limit, judges were asked to do a similar (but unanticipated) estimation for workers working under the other time limit (within-subject estimates) for the exact same task. Judges revised their estimates when the time limit changed. After estimating task completion time for the 5 minute time limit ($M_{Shorter}$ = 3.49), judges estimated that workers would take an average of M_{Longer} = 6.36 when time limits are longer ($t(148)=13.54$, $p<.001$). Likewise, when judges first estimated the completion time for longer time limit (M_{Longer} = 6.89), they then estimated a shorter time for the shorter time limit ($M_{Shorter}$ = 3.63; $t(154)=13.10$, $p<.001$). The order of the time limits did not affect the estimates significantly in this study ($\beta=0.38$, $t=1.16$, $p=.244$). However, the absolute differences in the within-subject estimates were directionally smaller (Δs = 2.87 and 3.25) than the between-subject estimates ($\Delta=3.40$). These results suggest that judges updated their subjective perception of the scope of the task when time limits changed, and accordingly revised their estimates of others' completion times.

Potential Moderating Variables: The tables below examine potential moderation of the effect of time limits (short vs. long) on judges' time estimates by various variables. Some of these variables were measured (e.g., time to read instructions), whereas others were asked as follow-up questions.

Table 1: Log of time to read instructions

	β	SE	*t*	*p*
(Intercept)	4.105	0.966	4.249	<.001
Long Time	4.214	1.194	3.527	<.001
Log of Time To Read Instructions	-0.151	0.232	-0.654	.513
Long Time x Log of Time To Read Instructions	-0.216	0.289	-0.747	.455

Table 2: Judges' self-rated knowledge about jigsaw puzzles

	β	SE	*t*	*p*
(Intercept)	3.88	0.69	5.621	<.001
Long Time	5.106	0.992	5.148	<.001
Judges' Self-Rated Knowledge About Jigsaws	-0.130	0.218	-0.595	.552
Long Time x Judges' Self-Rated Knowledge	-0.584	0.319	-1.833	.068

Interpretation: As self-rated knowledge increased, the difference in judges' estimates in the short vs. long time limit condition reduced marginally (see figure below). We do not replicate this effect systematically in other studies.

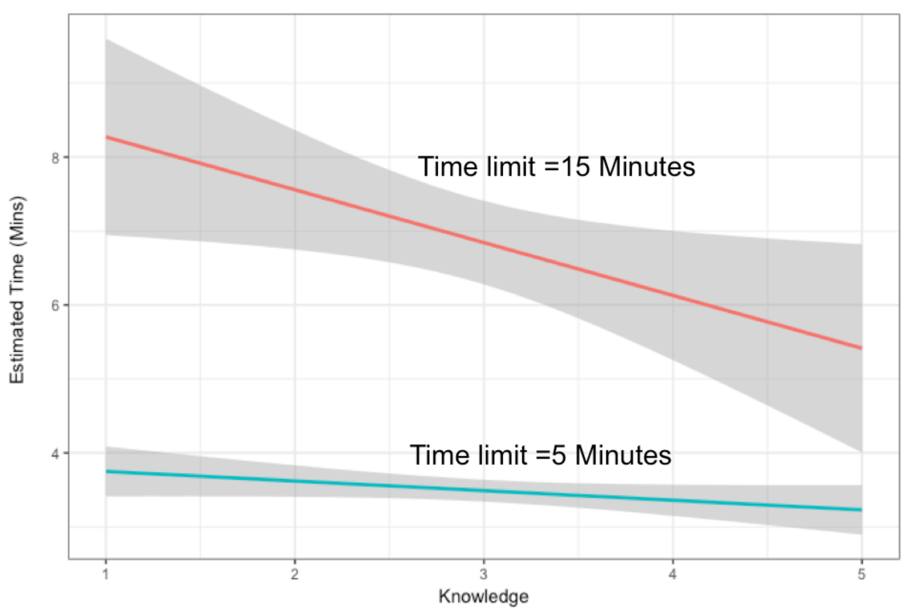


Table 3: Judges' self-rated frequency of solving jigsaw puzzles

	β	SE	t	p
(Intercept)	3.913	0.728	5.377	<.001
Long Time	3.727	1.099	3.392	<.001
Judges' Self-Rated Frequency of Solving Jigsaws	-0.181	0.297	-0.610	.542
Long Time x Judges' Self-Rated Frequency	-0.135	0.449	-0.300	.764

Table 4: Judges' belief that workers felt accountable to finish ASAP

	β	SE	t	p
(Intercept)	4.142	1.105	3.747	<.001
Long Time	3.234	1.502	2.153	.032
Workers Felt Accountable to Finish ASAP	-0.17	0.282	-0.602	.548
Long Time x Accountable to Finish ASAP	0.045	0.382	0.118	.906

Table 5: Judges' belief that workers wanted to take longer to enjoy more

	β	SE	t	p
(Intercept)	3.105	0.503	6.195	<.001
Long Time	2.575	0.714	3.604	<.001
Workers wanted to take longer to enjoy more	0.244	0.288	0.848	.397
Long Time x Workers wanted to take longer to enjoy more	0.558	0.421	1.325	.186

Table 6: Judges' belief that workers found the jigsaw puzzle interesting

	β	SE	t	p
(Intercept)	3.573	0.89	4.014	<.001
Long Time	3.842	1.206	3.185	.002
Workers Found Puzzle Interesting	-0.024	0.243	-0.097	.923
Long Time x Workers Found Puzzle Interesting	-0.128	0.334	-0.384	.701

Table 7: Judges' belief that people will work slower when more time is available

	β	SE	*t*	*p*
(Intercept)	3.524	0.375	9.397	<.001
Long Time	3.245	0.525	6.177	<.001
Beliefs Consistent with Parkinson's Law (PL)	-0.053	0.460	-0.114	.909
Long Time x Beliefs Consistent with PL	0.240	0.644	0.372	.710

Study 2

Potential Moderating Variables: The tables below examine potential moderation of the effect of time limits (short vs. long) on judges' time estimates by various variables.

Table 8: Judges' Self-Rated Knowledge About Jigsaw Puzzles

	β	SE	*t*	*p*
(Intercept)	27.691	1.658	16.696	<.001
Long Time	0.004	2.225	0.002	.998
Judges' Self-Rated Knowledge About Jigsaws	-0.427	0.512	-0.834	.405
Long Time x Judges' Self-Rated Knowledge	1.132	0.691	1.639	.103

Table 9: Judges' Self-Rated Frequency of Solving Jigsaw Puzzles

	β	SE	*t*	*p*
(Intercept)	26.814	1.499	17.877	<.001
Long Time	0.081	2.182	0.037	.970
Judges' Self-Rated Frequency of Solving Jigsaws	-0.196	0.622	-0.315	.753
Long Time x Judges' Self-Rated Frequency	1.485	0.910	1.632	.104

Study 3

Potential Moderating Variable: The table below examines the potential moderation of the relationship between workers' knowledge of time limits (known vs. not-known) and deadlines on estimates.

Table 10: Judges' Self-Rated Knowledge About Jigsaws

	β	SE	t	p
(Intercept)	20.69	1.91	10.85	<.001
Long Time	1.25	1.40	0.90	.371
Time Limit Known = True	-2.31	2.85	-0.81	.418
Judges' Self-Rated Knowledge About Jigsaws	0.02	0.70	0.03	.975
Long Time x Time Limit Known	1.25	2.09	0.60	.552
Long Time x Judges' Self-Rated Knowledge	0.36	0.51	0.71	.482
Time Limit Known x Judges' Self-Rated Knowledge	1.36	1.06	1.29	.199
Long Time x Time Limit Known x Knowledge	-0.26	0.78	-0.34	.737

Table 11: Judges' Self-Rated Frequency of Solving Jigsaw Puzzles

	β	SE	t	p
(Intercept)	19.39	2.19	8.87	<.001
Long Time	3.37	1.61	2.10	.037
Time Limit Known = True	-0.85	3.14	-0.27	.786
Judges' Self-Rated Frequency of Solving Jigsaw Puzzles	0.60	0.91	0.65	.513
Long Time x Time Limit Known	-2.18	2.31	-0.95	.345
Long Time x Judges' Self-Rated Frequency	-0.53	0.67	-0.79	.430
Time Limit Known x Judges' Self-Rated Frequency	0.87	1.32	0.66	.508
Long Time x Time Limit Known x Frequency	1.22	0.97	1.26	.207

Table 12: Judges' Age

	β	SE	*t*	*p*
(Intercept)	18.02	2.21	8.17	<.001
Long Time	4.62	1.61	2.87	.004
Time Limit Known = True	2.49	3.30	0.76	.450
Judges' Age	0.07	0.05	1.30	.193
Long Time x Time Limit Known	-1.72	2.41	-0.71	.476
Long Time x Judges' Age	-0.06	0.04	-1.61	.108
Time Limit Known x Judges' Age	-0.03	0.09	-0.40	.691
Long Time x Time Limit Known x Age	0.06	0.06	0.96	.338

Table 13: Judges' Gender (Female vs. Male)

	β	SE	*t*	*p*
(Intercept)	21.28	0.97	22.04	< .001
Long Time	2.48	0.71	3.49	.001
Time Limit Known = True	0.10	1.33	0.08	.940
Judges' Age	-1.17	1.42	-0.82	.411
Long Time x Time Limit Known	-0.59	0.98	-0.60	.550
Long Time x Judges' Age	-0.67	1.04	-0.64	.523
Time Limit Known x Judges' Age	2.61	2.01	1.29	.196
Long Time x Time Limit Known x Age	2.69	1.48	1.82	.069

Examining mediation using bootstrapping:

Approach: The bootstrapping code runs the following models for 500 times and computes $(\mathrm{b} - \mathrm{b}')$ in each iteration. This difference in coefficients is saved and sorted to calculate the 95% CI. Hierarchical regressions are used in all cases, and X denotes the covariate whose mediating effect is being studied.

Base model: $compensation\ scheme\ choice = a_0 + b * Time\ Limit$
Mediation model: $compensation\ scheme\ choice = a_1 + b' * Time\ Limit + c * X$

The table shows the 95% CI for each of the mediating covariates. All the above mediations were partial.

Mediating Covariate	**Indirect Effect (bootstrapped 95% CI)**
beliefs about whether jigsaw puzzle with 50-mins had more number of pieces	[0.0598, 0.3769]
beliefs about task scope	[0.0636, 0.5759]
Composite index of the two above variables (after standardizing using z-scores)	[0.1852, 0.69271

Proportion of judges whose estimates were higher than the time limit as a function of whether time limits were known or not

35 Minutes		**50 Minutes**	
Time Limit known	Time Limit not known	Time Limit known	Time Limit not known
0.57%	1.1%	0.56%	0%
$\chi^2(1)$= 0.0006, p= .98		NS	

The above results further suggest (as in Studies 1, S2, S4) that it is unlikely that judges censored or truncated their distributional beliefs about workers' completion times in the shorter time limit (vs. the longer time limit).

Effect of order of time limits on completion time estimates

In Study 3, judges saw two different (but very similar) puzzles that were counterbalanced and randomly assigned to one of two time limits. As expected, there was no effect of puzzle order (i.e., which jigsaw puzzle was seen first) on how time limits affected completion time estimates (interaction β= 0.59, t=0.81, p=.418). However, the order of time limits (i.e., whether judges estimated completion times for the shorter or the longer time limit first) moderated the effect of time limits on completion time estimates (β= 3.29, t=4.64, p<.001). In particular, longer deadlines exerted a stronger effect on judges' estimation. After estimating completion time for a shorter time limit, judges revised their estimations significantly when they faced a decision involving a longer time limit ($M_{Shorter}$ = 20.45 minutes vs. M_{Longer} = 24.60; β= 4.15, t=8.94, p<.001). However, after having estimated time completion for a task with a longer deadline, judges did not revise their estimates significantly when they subsequently encountered a task with a shorter deadline (M_{Longer} = 22.97 vs. $M_{Shorter}$ = 22.11; β = 0.86, t=1.61, p=.109).

Robustness Check using Judges who passed all the nine comprehension check questions before answering any questions in the survey:

In this analysis, we used only those judges who passed all the nine comprehension check questions before answering any questions in the survey (N=203).

Judges estimated longer completion time when the externally assigned time limit was longer ($M_{Shorter}$ = 20.42 minutes vs. M_{Longer} = 22.15; β= 1.72, t=2.73, p=.007). Furthermore, the effect persisted even when judges were told that workers did not know about the time limits ($M_{Shorter}$ = 18.74 minutes vs. M_{Longer} = 19.91; β= 1.17, t=2.35, p=.021). In fact, the increase in estimation as a result of longer time limits was identical irrespective of whether the workers knew about the time limits or not (interaction β= 0.56, t=0.69, p=.488), further confirming that judges' beliefs about how deadlines affect workers' behavior were not responsible for our findings. Given that there was no difference in estimates because of this experimental manipulation, the rest of the analysis uses the combined data.

Consistent with our hypothesis that longer deadlines increase the perceived scope of work, a composite z-score index computed from these two measures of task scope, partially mediated the effect of time limit on completion time estimates (bootstrapped 95% CI = [0.18, 0.81]). Therefore, using only a subset of judges who passed a rigorous set of nine comprehensive checks, we successfully replicated the important findings and found substantive evidence in favor of our hypothesized process mechanism.

The results also replicate if we look at the judges who correctly answered the two questions included in our pre-registered analysis: time limits were random and workers either were aware or were not aware of the time limits, depending on the conditions (N=268).

Study 4

Potential Moderating Variables: The table below examines the potential moderation of the relationship between belief-manipulation and deadlines on estimates.

Table 14: Judges' Self-Rated Knowledge About Jigsaws

	β	SE	*t*	*p*
(Intercept)	3.63	0.82	4.43	<.001
Judges' Self-Rated Knowledge About Jigsaws	0.00	0.24	-0.01	.996
Long Time	3.52	1.36	2.59	.010
Belief Manipulation = Question	0.32	1.26	0.26	.799
Judges' Self-Rated Knowledge x Long Time	0.12	0.42	0.30	.767
Judges' Self-Rated Knowledge x Question	-0.08	0.40	-0.21	.833
Time Limit Known x Question	-0.07	1.88	-0.04	.969
Long Time x Question x Knowledge	-0.35	0.59	-0.60	.551

Table 15: Judges' Self-Rated Frequency of Solving Jigsaw Puzzles

	β	SE	*t*	*p*
(Intercept)	3.31	0.81	4.09	<.001
Judges' Self-Rated Frequency of Solving Jigsaw Puzzles	0.13	0.32	0.40	.687
Long Time	5.88	1.29	4.57	<.001
Belief Manipulation = Question	0.16	1.27	0.13	.900
Judges' Self-Rated Frequency x Long Time	-0.83	0.52	-1.61	.109
Judges' Self-Rated Frequency x Question	-0.03	0.51	-0.07	.947
Time Limit Known x Question	-2.11	1.84	-1.15	.251
Long Time x Question x Frequency	0.41	0.74	0.56	.578

Table 16: Judges' Age

	β	SE	*t*	*p*
(Intercept)	3.51	0.76	4.64	<.001
Judges' Age	0.03	0.19	0.16	.870
Long Time	6.17	1.21	5.08	<.001
Belief Manipulation = Question	-0.01	1.21	-0.01	.993
Judges' Age x Long Time	-0.59	0.30	-1.95	.052
Judges' Age x Question	0.02	0.30	0.07	.943
Long Time x Question	-4.08	1.71	-2.39	.017
Long Time x Question x Age	0.77	0.42	1.81	.071

Table 17: Judges' Education

	β	SE	*t*	*p*
(Intercept)	3.46	0.83	4.19	<.001
Judges' Education	0.04	0.18	0.20	.839
Long Time	4.71	1.31	3.60	<.001
Belief Manipulation = Question	0.16	1.29	0.13	.899
Judges' Education x Long Time	-0.19	0.29	-0.64	.524
Judges' Education x Question	-0.02	0.29	-0.07	.943
Long Time x Question	-1.02	1.83	-0.56	.576
Long Time x Question x Education	-0.04	0.41	-0.10	.917

Robustness Check using Judges correctly answered all the five comprehension check questions:

After judges indicated their completion time estimates, they answered five questions to recall details of the instructions provided to them earlier. Specifically, they recalled the two different time limits, how the workers were assigned to one of the time limits, whether the workers were paid differently based on the assigned time limit, and whether workers could potentially earn more money by working longer. Judges who passed *all* these checks (N=257) were affected by the experimental conditions ($F(3, 253)=45.16$, $p<.001$). The manipulation of lay beliefs had no discernible effect on judges' estimates in the shorter time limit conditions ($M_{Shorter, Question}= 3.61$ vs. $M_{Shorter, Confirm} = 3.51$, $t(125)<1$, $p=.549$). However, judges' estimates in the longer time limit condition were marginally lower in the questioning lay belief condition compared to the confirming lay belief condition ($M_{Longer, Question}= 6.43$ vs. $M_{Longer, Confirm} = 7.41$, $t(128)=1.77$, $p=.078$). The 2-way interaction was also marginally significant ($F(1, 253)=3.48$, $p=.063$), demonstrating that prompting judges to question the over-generalized belief reduced the time limit bias, as we saw with the entire sample of judges used in the study.

Study 5

Potential Moderating Variable: The tables below examine the potential moderating effect of prior experience in running Direct Marketing campaigns on the effect of time limits on judges' estimates of work scope and completion time. All the results are from ordinal regressions.

Table 18: Judges' Self-Rated Experience (Yes, No) in running Direct Marketing Campaigns on Time Estimate

	β	SE	*z*	*p*
Long Time	2.187	0.487	4.487	<.001
Have Experience with running Direct Marketing campaigns	0.313	0.571	0.548	0.584
Long Time x Had Experience Running such campaigns	-0.323	0.758	-0.426	0.670

Table 19: Judges' Self-Rated Experience (Yes, No) with running Direct Marketing Campaigns on Scope of Work

	β	SE	*z*	*p*
Long Time	0.722	0.473	1.526	0.127
Have Experience with running Direct Marketing campaigns	0.478	0.661	0.723	0.470
Long Time x Had Experience Running such campaigns	-0.326	0.836	-0.390	0.696

Table 20: Judges' Self-Rated Experience (Yes, No) with running Direct Marketing Campaigns on Rate of Work. This regression uses the entire data since all judges in the study estimated the rate of work.

	β	SE	*z*	*p*
Long Time	-0.583	0.313	-1.861	0.062
Have Experience with running Direct Marketing campaigns	0.169	0.407	0.415	0.678
Long Time x Had Experience Running such campaigns	0.467	0.537	0.869	0.384

Examining rate of work by conditions: Did judges' estimates of the rate of work vary based on whether they estimated completion time or scope of work?

Table 21: Effect of time limits on the rate of work estimates as a function on the task assigned

	β	SE	*z*	*p*
Long Time	-0.364	0.355	-1.026	0.305
Estimated Scope of Work	0.309	0.358	0.863	0.388
Long Time x Estimated Scope of Work	0.025	0.499	0.051	0.959

Robustness check: Effect of time limits on scope/completion time estimates after controlling for covariates

Table 22: Effect of time limits on completion time estimates controlling for estimates rate of work

	β	SE	*z*	*p*
Long Time	2.110	0.400	5.265	<.001
Estimated Rate of Work	0.0004	0.0005	0.859	0.390

Table 23: Effect of time limits on completion time estimates controlling for estimates rate of work and population density (based on zip code)

	β	SE	*z*	*p*
Long Time	2.129	0.402	5.290	<.001
Estimated Rate of Work	0.0004	0.0005	0.881	0.378
Population Density	-0.00001	0.00002	-0.597	0.550

Table 24: Effect of time limits on scope of work estimates controlling for estimates rate of work

	β	SE	*z*	*p*
Long Time	1.062	0.413	2.570	0.010
Estimated Rate of Work	0.002	0.0005	4.536	<.001

Table 25: Effect of time limits on scope of work estimates controlling for estimates rate of work and population density (based on zip code)

	β	SE	z	p
Long Time	1.109	0.417	2.657	0.007
Estimated Rate of Work	0.002	0.0005	4.343	<.001
Population Density	0.00004	0.00001	2.660	0.007

Appendix D: Additional Studies

Study S1: Effect of deadlines on completion time estimates for everyday activities

Method

We tested for the prevalence of the proposed association between time limits and estimated task completion times for common everyday activities (N=29 adult online participants). We asked participants to list five household chores and the average time they take to complete each one. The five most frequently mentioned chores were house cleaning (30%), washing dishes (9%), laundry (8%), vacuuming (8%), and cooking (7%). Participants were then asked, for each chore they had listed, to estimate how long another typical person would take to do the same work in two different conditions: first for a short and then longer available time limit, within-subjects. The short time limit presented was 1.5 times the participants' own self-reported completion time, while the long time limit was three times the participants' own self-reported completion time.

Results

We divided the estimated completion time for others by the time that participants reported for themselves for each of the five tasks. We then compared the average of these five standardized time estimates for each participant across the two (short vs. long) time limit conditions. Averaging across five tasks, participants estimated an average standardized time of 1.07 (SD=0.19) in the shorter time limit condition, and 1.64 (SD=0.57) in the longer time limit condition, a statistically significant 53 percentage point increase in estimated time ($t(28)=5.86$, $p<.001$). We get the same results when we use a regression model with standard errors clustered at the person level. The majority of participants (79%) gave a longer estimated completion time when more time was available for at least one of the five chores they listed.

The results provide evidence that, across a range of participant-chosen tasks, people tend to estimate that others will take longer to complete a task when there is a longer time limit. Therefore, the effect described in the paper is very robust and happens even with everyday activities.

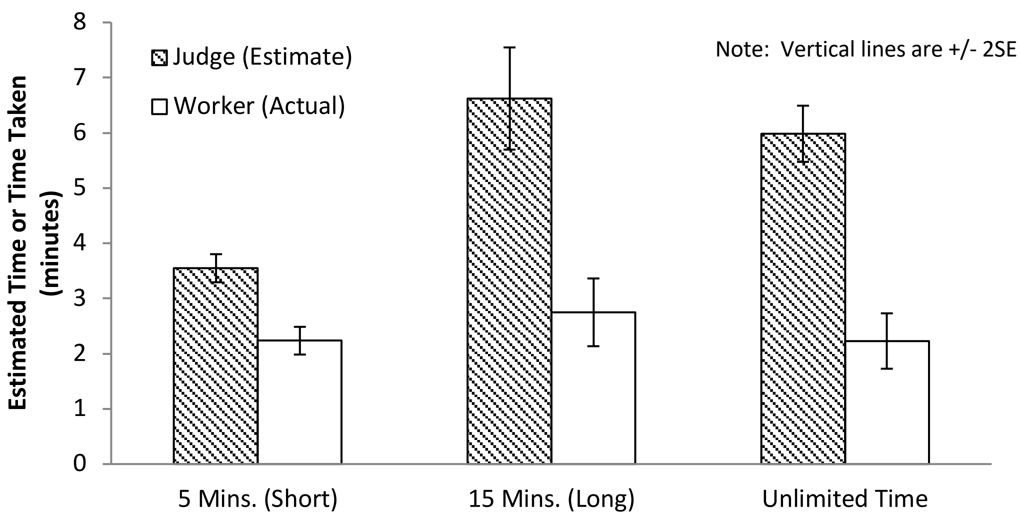
Study S2: Replication of Judges' results in Study 1 (using the workers of Phase 1)

Method

In Phase 2, a sample of online participants (N=103) was assigned the role of *judges* and was provided with detailed information about the Phase 1 Study, including two pictures of the puzzle which the workers had solved (like Study 1). The other instructions, likewise, were exactly similar to Study 1. Judges were asked to predict the task completion time for an average worker under one of the three different time limits (short, long, untimed; order counter-balanced). Similar to Study 1, judges could earn a bonus of up to $1 based on how accurately they predicted the task completion time: for every 1 minute of deviation from the actual average time in a particular time limit condition, 10 cents were deducted from the maximum bonus amount (i.e., a linear incentive for accuracy). The only difference with Study 1 was that this study did not have additional conditions to examine the standard anchoring account.

Results

We strongly replicated the basic results of Study 1. The judges' first estimates did not accurately predict the workers' times. Judges' estimates were significantly higher than the actual workers' times in each of the three time limit conditions (Short Time: $M_{Judges} = 3.55$ vs. $M_{Workers} = 2.24$, $t(68) = 7.21$, $p<.001$; Long Time: $M_{Judges} = 6.62$ vs. $M_{Workers} = 2.75$, $t(77) = 6.86$, $p<.001$; Unlimited Time: $M_{Judges} = 5.98$ vs. $M_{Workers} = 2.23$, $t(68) = 6.33$, $p<.001$). More importantly, in the between-subjects comparison of judges' estimates, the over-prediction increased with longer time limits (interaction $F(2,213) = 8.36$, $p<.001$). As shown in the figure below, this differential over-estimation can be seen when comparing the short time limit condition to either the long time limit ($F(1,145) = 16.67$, $p<.001$) or the unlimited condition ($F(1,136) = 15.38$, $p<.001$). There was no significant difference in over-prediction between the long and unlimited time limit conditions.



Like in Study 1, the bias in completion time estimation does not seem to be attributable to judges' lack of attention, lack of relevant experience or beliefs about workers' state of mind. The amount of time judges took to read the instructions, and their self-reported knowledge or experience with puzzles did not affect estimates or moderate the effect of time limits. Judges' beliefs about differences in either how accountable workers felt to finish the puzzle as soon as possible or about workers' task goals (to finish quickly or to take longer and enjoy it) in the different time limit conditions could not explain the findings. Furthermore, judges were well-calibrated for a diagnostic cue, correctly predicting that workers with low self-rated knowledge of puzzles would take longer to complete under each of time limits.

A majority (70%) of judges stated that people take more time when more time is available. However, this belief in "Parkinson's law" (Parkinson, 1955) did not moderate the effect of time limits on completion time estimates, suggesting that beliefs about differences in rates of work cannot fully explain the results, like in Study 1. Furthermore, if judges simply recoded all completion times above five minutes to five, the estimates in the short time limit condition should be similar to the bounded estimates in the unlimited time condition. To test this, we truncated those estimates from the unlimited time condition which were greater than 5 minutes to 5 minutes. However, judges' estimates were significantly higher in the unlimited time limit condition, even after truncating to five minutes, than in the short time limit condition (M_{5Mins} = 3.55 vs. $M_{Unlimited}$ = 4.37, $t(60)=4.35$, $p<.01$). We find the same result comparing data in the 15 minute time condition after bounding at 5 minutes (M_{15Mins} = 4.50) to the short time limit condition ($t(69)=4.95$, $p<.01$).

As an additional test, we analyzed the second type of estimate the judges made. After their time estimates, judges also estimated what proportion of workers they thought would have completed the work in less than 4 minutes in each of the different deadline conditions (within-subjects). If the observed effect of time limits on time estimates was due to truncation, the findings should be driven by differences in the predictions for those taking more than four minutes, and the estimated proportion of workers who finished in less than 4 minutes should be the same in all conditions. However, we find instead that judges estimated a significantly higher proportion of workers completing the work in less than 4 minutes under the short time limit condition than in the long time limit condition (58% versus 46%; $t(102)=6.34$, $p<.001$), incompatible with a truncation account.

Study S3: Further examination of the role of incidental anchors

Method

Online participants acting as judges (N=120) participated in an estimation game where they were required to estimate the time taken by an imaginary worker to 'solve' a jigsaw puzzle of a known number of pieces. Judges were told that in an untimed pre-test, a 67-piece jigsaw puzzle was solved by a group of workers and no one took more than 31 minutes to solve it. Judges were required to re-enter the maximum time taken by the workers in a text box to proceed. This was done to make sure that the judges registered the information. Subsequently, judges were told that the *same* puzzle was being administered to a new worker. Half the judges were told that based on a coin-flip the worker was assigned to either a short time limit (35 minutes) or a long time limit (60 minutes) condition. These judges were first asked to estimate the task completion time for the worker in a particular time limit condition (between-subjects). The other half were either told that the worker had no time limit (control), or that the worker had no time limit *and* the worker spends 60 minutes every day commuting to work (control with an *incidental* anchor). Judges estimated the task completion time for one of the two control conditions. Before making their estimate, judges were told that they could earn an additional incentive of up to $1 (50% of their base pay) if their estimate was accurate. After the study, the imaginary worker's time was drawn from a uniform distribution of [1, 31] minutes and judges were paid their base pay along with additional bonus based on a linear payoff rule (5 cent deduction for every minute of deviation from the time drawn).

Results

In the between-subject evaluation, judges predicted significantly higher task completion time for workers who had long time limit than those who had short time limit (M_{Short} = 25.83 vs. M_{Long} = 33.96, $t(61)=4.92$, $p<.001$) replicating prior results. However, the estimates of predicted task completion times in the two control conditions did not differ ($M_{Control}$ = 28.96 minutes vs. $M_{Control\ with\ anchor}$ = 27.42, $t<1$). Therefore, the introduction of an incidental anchor, which was in the same dimension as the quantity to be estimated, did not influence the predicted task completion time.

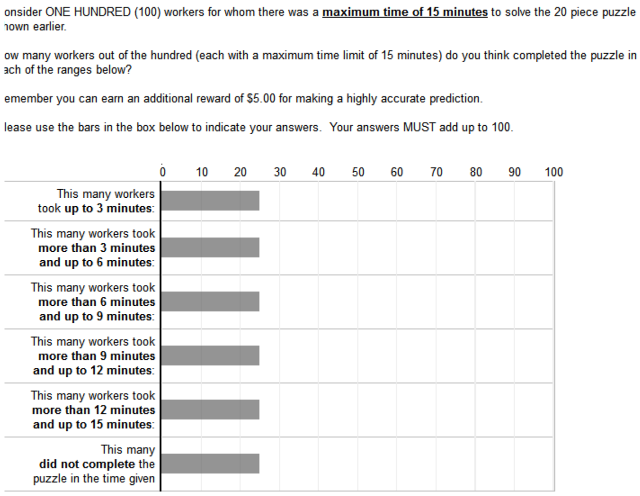
Study S4: Further examination of judges' beliefs about the distribution of worker times

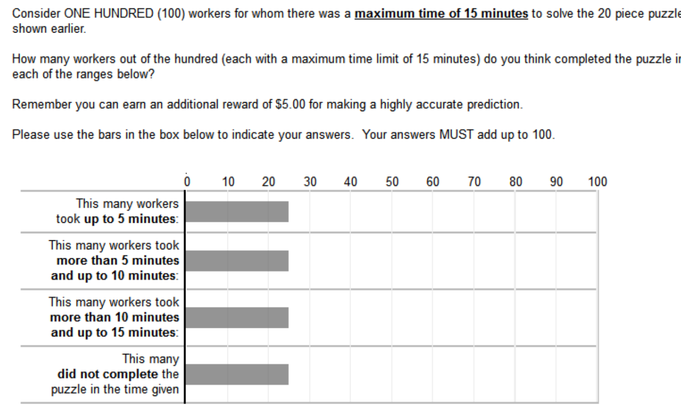
Method

Online participants (N=88) acted as judges and were assigned to two between-subject (time limit: shorter vs. longer) conditions. As in Study 1, judges were told that each worker was randomly assigned to either a maximum time of 5 minutes (the shorter time limit) or a maximum time of 15 minutes (the longer time limit) to solve the puzzle and that workers could not choose or influence their time limit. The judges' task was to estimate the time it took different workers to finish solving the puzzle by indicating the proportion of workers in each of a set of time ranges using an adjustable histogram.

Judges were told that the three participants with the most accurate estimates would receive a bonus payment of $5 (5 times the base payment).

For the shorter time limit condition, judges estimated how many of 100 workers' times would fall into each of five one-minute long intervals (i.e., "This many workers took up to 1 minute," "This many workers took more than 1 minute and up to 2 minutes," etc.), or into the "did not complete" category. For the longer time limit condition, judges estimated either how many of 100 workers' times would fall into each of three five-minute long intervals or into each of five three-minute-long intervals (between subjects; see the figure below). The online interface included a "did not complete" (DNC) category and required the sum of all the allocations to equal 100. In effect, this study used three between-subjects conditions. Therefore, judges allocated workers into six bins in the short time limit condition, and into either four or six bins in the long time limit condition. Lastly, the judges answered a few follow-up questions.





Results

There were no significant differences in the two 15-minute sub-conditions, so we combined them when we computed the average imputed completion time. Based on the elicited distributions, we calculated each judges' estimated average task completion time in the two time-limit conditions. We did this by taking the mid-point of each of the time-bins and computing the weighted average using the proportion of workers allocated to each bin as weights. The imputed means did not differ in the two longer-time sub-conditions ($M_{Longer,\ 3\ bins} = 6.69$ vs. $M_{Longer,\ 5\ bins} = 7.03$, $t(40)<1$) and therefore, we combined these two sub-conditions when calculating the completion time estimate for the longer time limit condition. Replicating Study 1, judges estimated a longer mean completion time for workers in the longer time limit condition ($M_{Shorter} = 2.85$ vs. $M_{Longer} = 6.86$, $t(86)=11.78$, $p<.001$). As in Study 1, this effect was not moderated by judges' experience or their beliefs about workers' state of mind. This replication of Study 1 using a different elicitation approach also demonstrates that our findings are not an artifact of the measurement method.

To test the censoring account, we compared the number of workers estimated to take 5 minutes or less in the longer vs. the shorter time limit condition (i.e., we compared the 5-minutes condition and the second 15-minutes sub-condition). This meant using the estimated 'completes' in the shorter time limit condition and the proportion of workers assigned to the first time-bin in the four-bin condition in the longer time limit. If the censoring account explains the effect of time limits, these two estimates should not differ. However, judges estimated that, on average, 92% of the workers would complete the puzzle in under 5 minutes in the shorter time limit condition, but only 37% of the workers would complete the puzzle in 5 minutes or less time in the longer time limit condition ($t(65)=11.14$, $p<.001$). This result is not consistent with a censoring account.

We had tested and did not find support for a truncation account in Study 1. To test the truncation account yet again in this study, we compared the proportion of workers estimated to finish in up to 3 minutes in the longer vs. shorter time limit conditions. This meant using the proportion of workers assigned to the first three time-bins in the shorter time limit condition and the proportion of workers assigned to the first time-bin in the 6-bin condition in the longer time limit (i.e., the first 15-minutes sub-condition). Under a truncation process, these two estimates should not differ. However, consistent with our proposed scope perception account, judges estimated that fewer workers would complete the puzzle in up to 3 minutes in the shorter than longer time limit condition ($M_{Shorter}$ = 47% vs. M_{Longer} = 18%, $t(65)=4.50$, $p<.001$).

Study S5: Further examination of the role of irrelevant time limits using an incentive-compatible budgeting game

Method

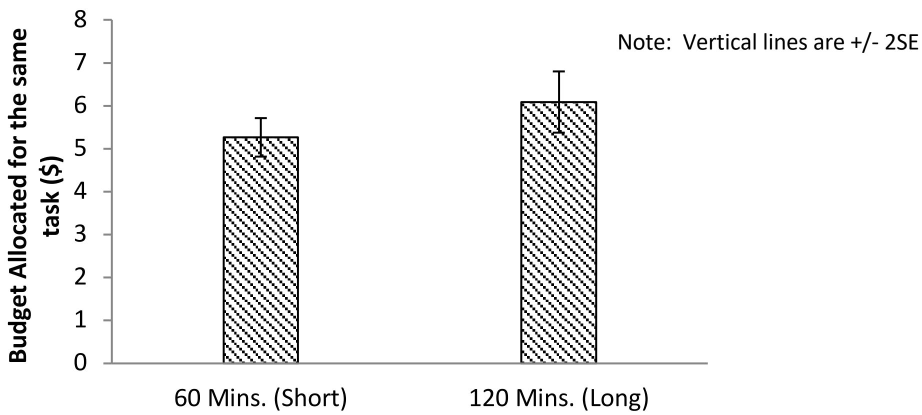
This study was conducted in a classroom setting in two sessions (one for each condition) using both verbal and written instructions. The participants (N=33) were undergraduate students at a large mid-western university who each participated in one session of the Study as part of an Economics course requirement and could earn additional bonus credits based on their performance in the study. Participants played the role of judges (e.g., project managers) in a budget-setting exercise. They needed to budget for a hypothetical worker, who was paid a constant wage rate of 10 cents per minute for the time taken to finish the job, to paint a 20 feet by 10 feet wall. In the scenario, the organization had set a time limit to complete the project – either a short time limit (60 minutes) or a long time limit (120 minutes), varied between subjects – that the hypothetical worker did not know about. The judges were informed that in this game, the worker's time to complete the task would be determined by drawing a number randomly from a uniform distribution between 30 minutes and 90 minutes.

Judges were then asked to budget for the task, by choosing how much money to allocate for the worker's compensation (based on the time to complete the project and the constant wage rate) from the $12.00 available. Judges were incentivized to not over-budget or under-budget. They would earn more if they had budgeted less and the project was still completed. However, if they budgeted less money than turned out to be necessary (based on the randomly drawn task completion time), the participant, having "failed" the budgeting exercise, would not receive any bonus.

Judges, therefore, had an incentive to provide as low a time estimate as possible (i.e., by budgeting as low an amount as possible) without under-guessing (in which case they would not be eligible for any bonus at all). Most importantly, the optimal strategy depended only on the randomly drawn worker time and was independent of the time limit. Thus, there was no incentive for rational judges to incorporate the time limit into their estimate since the optimal bid was determined only by the payoffs and the known distribution of randomly generated worker times. The optimal bid for risk-neutral judges in either time limit condition (60 minutes or 120 minutes) was \$4.63, which corresponds to predicting that the worker would take approximately 46 minutes (see calculations below the Results section). This guess would have earned the judges a bonus of \$0.88 in the game, on average.

Results

Comprehension checks suggested that judges understood that the completion times were drawn from a uniform distribution between 30 minutes to 90 minutes.[1] Despite the fact the optimal bid based on the information known to the judges was the same in both conditions, judges bid significantly more in the longer time limit condition than in the shorter time limit condition, implying a longer time estimate (M_{Short} = \$5.26 vs. M_{Long} = \$6.09, $t(31) = 2.05$, p =.049; see the figure below). Therefore, the longer time limit influenced judges to budget more money for the task, even though they knew that the workers were not aware of the time limit, and the time limit did not even affect the randomly drawn time used to determine the bonus.



This study suggests that the influence of time limits on managers' time estimates can affect their decisions (e.g. budgeting), even when the time limits are irrelevant to the decision. These findings cannot be explained by a motivational lay theory, since judges were told that workers in the scenario did not know about the time limit.

Optimal Strategy for the Budgeting Game

The pay-off structure was as follows:

Offer/Bid = *\$y* Wages Payable = \$w	Wages payable is over-budget If w > y	Wages payable is under-budget If w <= y
Actual Wage Paid in \$	w	y
Bonus Paid in \$	0	0.50 + max(12.00 – 2*y, 0)

[1] In the 60-minute time limit condition there were five judges whose bids represented a completion time of more than 60 minutes. We truncated their bids to the maximum time available for the reported analysis, and we get the same results even if we discard these participants from the analysis.

We know, $t \sim U(30, 90)$, where t is time taken by contractor in minutes to complete the work. This means, $w \sim U(3, 9)$, where w is the wages payable to contractor in Dollars … (1)

Expected Bonus:

$$\text{E(b)} = \Pr(going\ overbudget) * 0 + \Pr(going\ underbudget) * [0.50 + max(12 - 2y, 0)]$$

$$= \Pr(going\ underbudget) * [0.50 + max(12 - 2y, 0)]$$

$$= \Pr(w \le y) * [0.50 + max(12 - 2y, 0)]$$

… (2)

Case I: What if $y < 3$. As per (1) this would tantamount to the manager going under-budget in which case he will not earn any bonus. Hence, $y < 3$ is not possible.

Case II: What if $y > 9$. In this case as per (1) $\Pr(going\ underbudget) = 1$. However, knowing (1) the manager would strictly prefer a bid $y \le 9$ to $y > 9$.

Case III: From the above discussion it is clear that $y \in [3, 9]$ is the feasible range of the manager's bid.

Therefore given an Uniform distribution, $\Pr(w \le y) = \frac{y-3}{9-3} = \frac{y-3}{6}$. … (3)

Using (3) in (2), we get

$$E(b) = \frac{y-3}{6} * [0.50 + max(12 - 2y, 0)]$$

$$= \frac{1}{6} * [\frac{y-3}{2} + (\text{y} - 3) * \max(12 - 2y, 0)] \quad \dots (4)$$

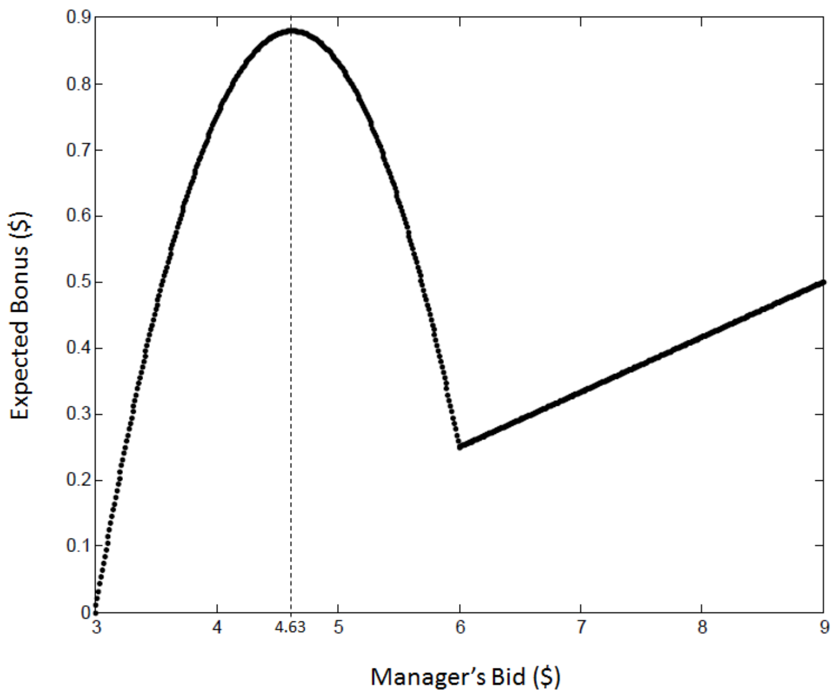
To solve the maximization problem in (4), we divide y into two ranges $(i)\ 3 \le y \le 6$ and $(ii)\ 6 < y \le 9$

(i) For $3 \le y \le 6$: $E(b) = \frac{1}{6} * (y - 3) * [\frac{1}{2} + 12 - 2y]$
$=> E(b) = \frac{1}{6} * [(y - 3)(12.5 - 2y)]$
This is a concave function in y. Solving the FOC gives us $y = 4.63$. This means $E(b) = 0.88$.

(ii) For $6 < y \le 9$: $E(b) = \frac{1}{6} * \frac{y-3}{2}$

This is an increasing function of y, hence is maximized at $y = 9$. This means $E(b) = 0.50$.

Therefore, the optimal bid is $4.63 in which case the expected value of manager's bonus is $0.88. The figure below depicts expected bonus as a function of the bid.



Complete instructions for the Budgeting Game (conducted in a classroom)

I. Basic Instructions

1. You are a Project Manager in an Organization and you are responsible for budgeting for a task that will be done by a Contractor.

2. The task is very simple – it is painting a 20 ft by 10 ft wall with red color for a promotional activity.

3. The Contractor has already been selected and your job is to set the budget in US Dollars based on your estimate of time which the task will take to complete.

4. If the project goes over-budget (i.e. the project takes longer than you had budgeted for), the contractor will be paid for their time (as per a fixed wage rate), but you will not receive a bonus.

5. If the project comes in under budget, the contractor will receive the budgeted payment, and you will receive a bonus which depends on the difference between the Cap and the budget you set *plus* some fixed amount.

6. The Cap is the maximum money your Organization has set apart for this task.

7. The Organization also has a maximum time which it has set for completing this task.

8. The Contractor does not know anything about either the Cap or the maximum time set by your Organization to complete the task. That information is available only to you as a Project Manager.

9. The Contractor has strong incentive to do it as quickly as they can else they are worried about future consequences (i.e. getting hired for future jobs). So, you can be assured that the contractor will not take any longer than it takes.

II. Contract Details

1. The Cap set by your Organization to complete the task is <u>$12.00</u>.

2. The maximum time set by your Organization to complete the task is <u>60 minutes (120 minutes)</u>.

3. The Contractor's wage rate is <u>10 cents per minutes</u> and this is fixed.

4. If the Contractor takes ***t*** minutes to complete the work – the total wages payable is **$w=** ***t*0.10***. If the work remains unfinished till the total time available for its completion – the money payable to the contractor is the time for the entire duration i.e. 60*0.10 = $6.00 (120*0.10 = $12.00).

5. As the Project Manager, you make an offer to the Contractor which is either greater than or equal to the total wages payable. If your offer is ***$y*** and it is <u>not less than</u> the actual wages payable ***$w***, then the Contractor gets ***$y***, as per your offer, and you get a bonus which is detailed in the payoff matrix below.

6. If you make an offer, which is <u>less than</u> the actual wages payable, then the Contractor gets the wages payable ***$w*** and you do not get anything since you did not do a good job in estimating the time.

7. Here is the payoff matrix for you:

	Goes Over Budget	Completed Under Budget
Your Offer = **$y**; Actual wages payable to the Contractor = **$w**	IF w > y	IF w <= y
Contractor gets USD (**as wage**)…	w	y
You get USD (**as bonus**)…	0	$0.50 + max(C-2y,0)

C=Cap i.e. $12.00.

III. Logistics

For the purpose of this exercise, we are going to make a <u>random draw</u> and estimate the Contractor's wages payable at the end of the Study. This will be done in front of you all after you are done answering the questions.

IV. Additional Private Information

You have leant from your reliable sources that in the past other Contractors have taken anywhere between 30 minutes to 90 minutes to complete similar tasks and any duration within this interval were equally likely.

Please ensure that you have understood the above directions completely and let the experimenter know in case you have any doubts before proceeding to answer the questions below.

Study S6: Further examination of the effect of time limits on estimated task scope

Method

Online participants (N=118) acting as judges were told about the range of puzzle sizes (between 6 and 247 pieces) available on *www.jigzone.com* and were also given information about the best and average solving times among visitors to the site. Judges read a hypothetical scenario in which one such puzzle had been selected and administered to a group of students who took an average of 28 minutes to solve it. However, unlike the prior studies, judges did not know which puzzle it was and, more specifically, how many pieces it had. Judges were assigned to one of three between-subject conditions: either an unlimited time condition, a shorter (30 minutes) time limit, or a longer (45 minutes) time limit.

In the scenario used in the study, judges were told that the *same* puzzle was administered to another person (the worker), who was either described as working under no time limit or under a time limit. In the time limit conditions, judges were told that the worker had been randomly assigned (using a coin flip) to either a shorter time limit (30 minutes) or a longer time limit (45 minutes). Each judge then estimated the number of pieces in the puzzle, as a proxy for the scope of the work, and then, on a separate screen, the worker's task completion time. With this information, we could also calculate judges' believed rate of work (the number of puzzle pieces solved per minute) implied by their estimates.

Results

Replicating our results, judges' estimates of task completion time were affected by time limits ($F(2,115)=3.58$, $p=.030$). In particular, judges predicted significantly more time in the longer time limit condition than in the shorter time limit condition ($M_{Shorter} = 24.28$ vs. $M_{Longer} = 36.58$, $t(70)=8.56$, $p<.001$). The estimated task completion time in the unlimited condition ($M_{Unlimited} = 31.97$) was somewhere between the time limit conditions.

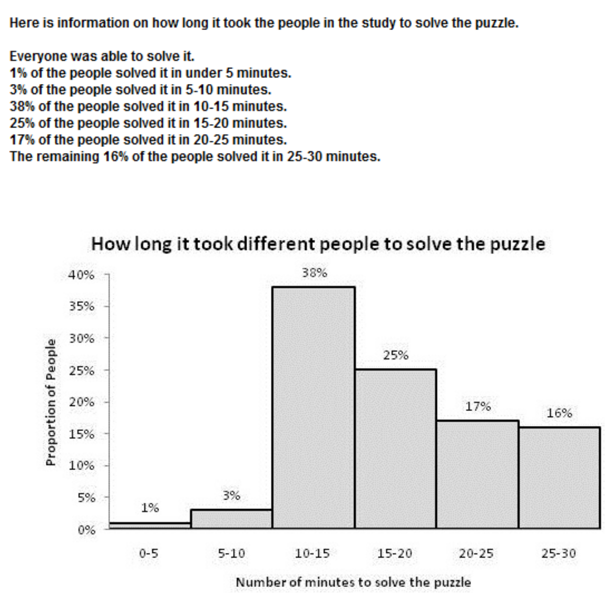
According to our proposed scope perception account, time limits affect completion time estimates because of how the task is perceived. Consistent with this prediction, judges' estimates of the number of pieces in the puzzle varied as a function of the external time limit ($F(2,115)=7.16$, $p=.001$). More specifically, we argue that observing a larger time limit makes the task subjectively seem larger. Indeed, judges estimated that the puzzle had significantly more pieces when the deadline was 45 minutes than when the deadline was 30 minutes ($M_{Shorter} = 130.83$ vs. $M_{Longer} = 177.25$, $t(70)= 2.46$, $p=.016$). The effect of longer time limit on larger estimated puzzle size was confirmed using a non-parametric test (Kruskal-Wallis $\chi 2(1)= 4.56$, $p= .032$). The findings provide direct evidence that people's perception of the scope of work is affected by time limits, such that longer external deadlines cause people to think of the task as larger in scope.

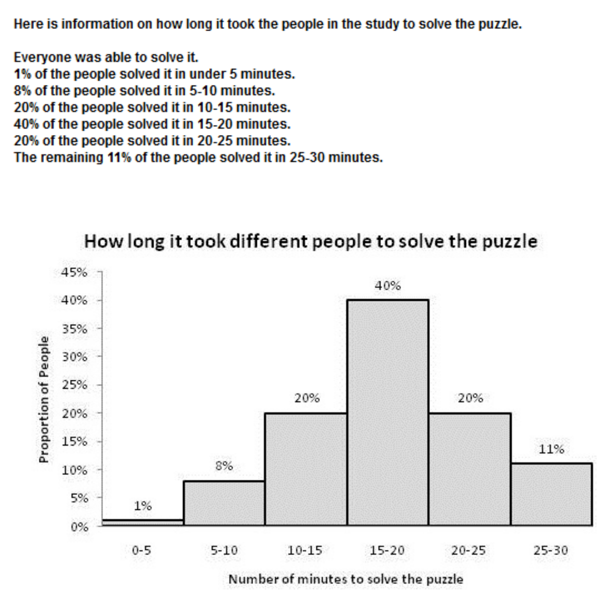
Given that we elicited judges' beliefs about time taken as well as the number of pieces in the jigsaw puzzle, we could compute their implicit beliefs about the rate of work. However, the implied rate of work (estimated puzzle pieces divided by estimated minutes to complete the task) was not statistically different across the different time limit conditions ($F(2,115)=1.37$, $p=.257$), further ruling out beliefs about the effects of time limits on workers' behavior as an alternative explanation behind the observed findings. Ninety-three percent of the judges (N=110) correctly recalled the two time limits used in the study, and that workers were randomly assigned to one of these two time limits. All the reported results hold if we only look at this subset of judges.

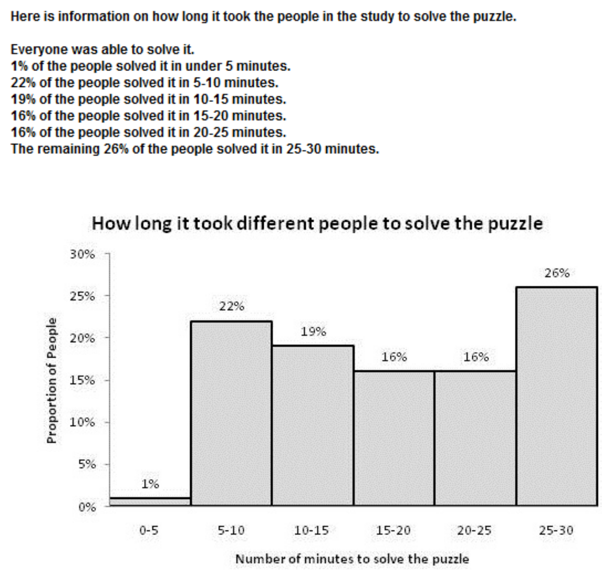
Study S7: Does providing judges with a complete distribution of task completion time debias the effect

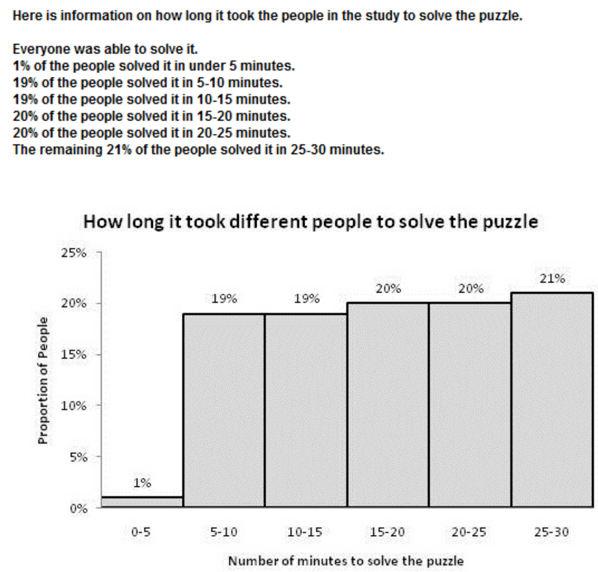
Method

Online participants (N=122) acting as judges were informed about the range of jigsaw puzzle pieces in the website *jigzone.com* along with the best and average time taken to complete puzzles of various sizes.









In the scenario described to the judges, *one* puzzle (unknown to the judge) was taken from the website and was administered to a group of workers. Half the judges were told that, based on a coin flip, workers were either given 30 minutes (short time limit) or 45 minutes (long time limit) to solve the puzzle, and the other half were told that the workers had unlimited time to solve the puzzle.

Judges were first asked to determine the number of pieces there were in the chosen puzzle in one of the time limit conditions (between-subjects). *Before* the judges made their prediction, they were given the distribution of time it took hypothetical workers to complete the task, in the form of a histogram. Four different histograms were used judges saw only one histogram (randomly chosen) with an implied mean of 17.6 minutes (the histograms used are shown above). The same set of histograms was used in both the time limits. After the judges estimated the number of pieces in the puzzle, they were asked to estimate the average time it took the workers to complete the work in a particular condition. Using the reported scope and the task completion time, we could impute judges' beliefs about workers' rate of work.

Results

When judges were given the distribution of task completion times in the various conditions, judges estimated a directionally higher number of puzzle pieces when the time limit was longer, but the difference was no longer statistically significant (M_{Short} = 104.47 vs. M_{Long} = 122.12, $t(78)$=1.10, p=0.274). The corresponding estimate in the unlimited condition was 121.81. Likewise, when asked to estimate the average time a worker took to solve the puzzles, judges' estimates were very similar (M_{Short}= 20.19 vs. M_{Long} = 18.81 vs. $T_{Unlimited}$ = 17.89). Therefore, providing exhaustive information, in the form of distribution of past completion times, can help arrest the bias reported in the paper.

Imputing the rate of work from these estimates suggests that the judges' implied rate of work (pieces per minute) was marginally faster for the workers in the longer time limit condition than the shorter time limit condition (M_{Short} = 5.46 vs. M_{Long} = 7.19, $t(78)$=1.63, p=.106). These results are further inconsistent with a lay theory account in which judges predict that people work at a slower pace when they have a longer time limit.

Appendix E: Effect Sizes Meta-analysis

The overall effect size for the completion-time estimation effect and the perceived scope of work effect is below. The analysis using only between-subjects data, and therefore excludes Study 3 that uses a within-subjects design.

Time Estimation

	Shorter time limit			Longer time limit						
Study	N	Time	SD Time	N	Time	SD Time	s within	d	Var d	Weight
1	149	3.489	0.918	155	6.894	3.592	2.644	1.287	0.016	62.936
2	112	26.366	3.948	122	29.844	5.856	5.034	0.691	0.018	55.111
4 confirm	89	3.624	0.972	70	7.529	3.318	2.317	1.686	0.034	29.022
4 question	68	3.701	0.983	90	6.433	2.926	2.302	1.187	0.030	33.030
5	50	1.620	0.805	51	2.735	1.271	1.066	1.046	0.045	22.211
S2 (apx)	30	3.548	0.697	41	6.622	2.964	2.301	1.336	0.070	14.228
S3 (apx)	31	25.839	5.466	32	33.969	7.460	6.555	1.240	0.076	13.207
S4 (apx)	46	2.851	0.726	42	6.861	2.181	1.595	2.514	0.081	12.274
S5 (apx)	19	52.632	9.771	14	60.857	13.341	11.405	0.721	0.132	7.579
S6 (apx)	36	24.278	5.230	36	36.583	6.859	6.099	2.018	0.084	11.930

Scope Estimation

	Shorter time limit			Longer time limit						
Study	N	Scope	SD Scope	N	Scope	SD Scope	s within	d	Var d	Weight
5	51	12254.900	24100.080	51	17450.980	25497.500	24808.631	0.209	0.039	25.361
S6 (apx)	36	130.833	58.821	36	177.250	96.860	80.131	0.579	0.058	17.275
S7 (apx)	38	104.474	63.786	42	122.119	77.976	71.596	0.246	0.051	19.800

Source: Borenstein, Hedges, Higgins and Rothstein (2009) p. 26-27; 65-66

The meta-analytic effect size for the time estimation is: $d = 1.247, z = 20.173, p < .001$
The meta-analytic effect size for the scope estimation is: $d = 0.324, z = 2.556, p = .011$