Online appendix to “Motivated procrastination”

Charlotte Cordes (LMU Munich), Jana Friedrichsen (Kiel University),
Simeon Schudy (Ulm University)
published in *Management Science*

A Additions to the theoretical framework

A.1 Nuanced signal structure

The conclusions drawn from the model presented in Section 3 naturally extend to a setting with a more nuanced signal structure like the one implemented in our experiment. Consider the case with four possible signal realizations, where a signal can be very negative ($s = 0$), somewhat negative ($s = 1$), somewhat positive ($s = 2$) or very positive ($s = 3$).

We modify Assumptions 1 and 2 to accommodate this belief structure by continuing to assume that agents update symmetrically around the prior and more strongly so after the more extreme signals and that the probability of forgetting a signal is smaller the more positive the signal is (see also Footnote 4).

Assumption 1′. *Agents update symmetrically in the direction of the signal:* $p_2^3 - p_1 = p_1 - p_2^0 = \Delta > \Delta' = p_2^2 - p_1 = p_1 - p_2^1 > 0$.

Assumption 2′. *Memory is imperfect, and news is more likely to be recalled the more positive it is:* $0 < q^0 < q^1 < q^2 < q^3 < 1$.

Recall that an agent's posterior p_2 equals her prior p_1 when she forgets her signal. Then, the average posterior belief in LowScope is:

$$\frac{1}{4}\sum_{s=0}^{3} p_2^s \tag{1}$$

and the average posterior belief in HighScope is:

$$\frac{1}{4}\left(\sum_{s=0}^{3} q^s p_2^s + \sum_{i=0}^{3}(1-q^s)p_1\right) \tag{2}$$

Thus, posterior beliefs are more optimistic in HighScope if and only if

$$\sum_{s=0}^{3} q^s(p_2^s - p_1) > \sum_{s=0}^{3}(p_2^s - p_1). \tag{3}$$

Note that the right-hand side equals 0 by Assumption 1′. Rearranging and using the notation from Assumption 1′, average beliefs are more positive in HighScope if and only if

$$-q^0\Delta - q^1\Delta' + q^2\Delta' + q^3\Delta > 0 \Leftrightarrow (q^3 - q^0)\Delta + (q^2 - q^1)\Delta' > 0. \tag{4}$$

By Assumption $2'$, this inequality always holds. Thus, under the modified Assumptions $1'$ and $2'$, we obtain an analogue to Prediction 1 for the more nuanced signal strucutre: On average, posterior beliefs are more optimistic in HighScope than in LowScope. Again, one could replace Assumption $1'$ with a more direct assumption on the relative asymmetry of the recall probabilities and the updating process.

A.2 Non-uniform signal distribution

Suppose a participant receives a positive signal with likelihood α and a negative signal with likelihood $(1-\alpha)$. Then, on average the belief of participants in LowScope is:

$$\alpha p_2^{pos} + (1-\alpha) p_2^{neg}. \tag{5}$$

and on average the belief of a participant in HighScope is:

$$\alpha \left(q^{pos} p_2^{pos} + (1-q^{pos}) p_1 \right) + (1-\alpha) \left(q^{neg} p_2^{neg} + (1-q^{neg}) p_1 \right). \tag{6}$$

Condition (3), which states the condition under which posterior beliefs in HighScope are more optimistic on average than posterior beliefs in LowScope, becomes:

$$\frac{1-q^{neg}}{1-q^{pos}} > \frac{p_2^{pos}-p_1}{p_1-p_2^{neg}} \frac{\alpha}{(1-\alpha)}. \tag{7}$$

In comparison to the condition under uniform probabilities (3), the new condition (7) takes the relative likelihood of receiving positive versus negative signals $\alpha/(1-\alpha)$ into account. Assuming that agents update symmetrically around the prior (Assumption 1), Prediction 1 continues to hold as long as the joint probability of receiving and forgetting negative news exceeds the joint probability of receiving and forgetting positive news.

Intuitively, Prediction 1 would fail if relatively many people receive positive news because this reduces the scope for memory-induced optimism. A relatively higher asymmetry in forgetting can counteract this effect. Our experiment rules out such an imbalance in signals by distributing positive and negative news with the same theoretical probability across participants.

A.3 Microfoundation for asymmetric memory

To derive the assumption that positive news is recalled with a higher likelihood compared to negative news (Assumption 2), we follow the dual-self-approach by Bénabou and Tirole (2002) as further developed by Hagenbach and Koessler (2022). They model agents with psychological utility (of which anticipatory utility is one example) who may want to distort their memory by manipulating the probability of recalling negative signals.[1] In the dual-self

[1]The idea of this model has also been used by Balietti *et al.* (forthcoming) to derive predictions about the recall of information about an unpleasant event. An alternative microfoundation of the asymmetric memory effect, based on endogenous rehearsal decisions as in Conlon (2024), would yield equivalent predictions. Specifically, optimistic beliefs in our main experiment could result from a participant's intrinsic preference for rehearsing positive rather than negative signals about their workload that leads them to recall positive signals more often than negative signals. The absence of an asymmetric memory effect in our observer experiment would then mean that participants are indifferent with respect to rehearsing negative or positive signals about a workload that does not directly affect them so that recall is symmetric.

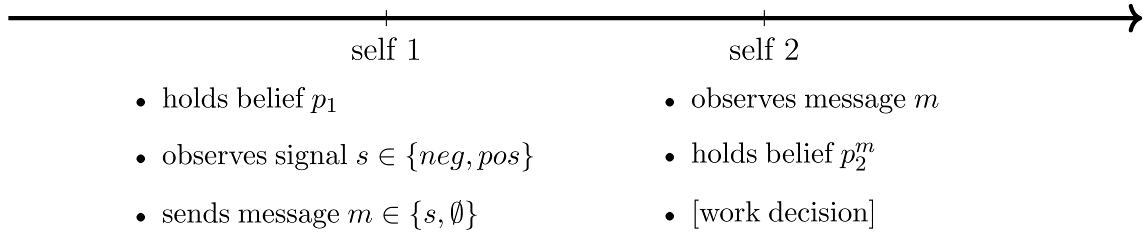


Figure A.1: Timeline of the model

model, self 1 receives information and decides whether or not to disclose this information to self 2. Self 2 forms a posterior belief and is later asked to choose an action, based on her belief. As in the experiment, the existence of this action choice is unknown to self 2 when forming her posterior belief, and can thus be modeled separately (see Subsection 3.3).

The timeline of the model is displayed in Figure A.1. Self 1 holds a prior belief p_1 and receives the signal s. Then she decides whether or not to disclose the signal to self 2 (i.e., whether or not to remember the signal). In particular, self 1 can send the message $m = s$ or the message $m = \emptyset$, but self 1 cannot confabulate and send a signal different from s. When self 1 decides whether or not to remember the signal, she faces a trade-off between holding more favorable beliefs about future workload and the costs associated with memory distortion. These costs k_i may vary between individuals.[2] For illustrative purposes, we assume that these individual costs are drawn independently from a normal distribution: $k_i \sim \mathcal{N}(\mu_k, \sigma_k)$ with $\mu_k, \sigma_k > 0$.

After observing message $m = s$, self 2 forms the posterior p_2^m, which may differ from the Bayesian posterior but systematically increases in the signal s (as alluded to above). To keep the model concise, we assume that the agent is naive in a way that self 2 does not infer any information from the message $m = \emptyset$. Her posterior in this case equals her prior, $p_2^\emptyset = p_1$.[3]

The two selves' incentives in this intrapersonal game are aligned. In particular, self 1 and self 2 share the same utility function which depends on the posterior belief p_2 through the expected work load ($E[w] = b + E[x]$): $U(p_2) = E_{x\sim p_2}[-c(b + x)]$, where $c(w)$ is the effort cost, increasing and convex in the work load w. Using the simplification to binary workloads introduced above, we write this expectation as $E_{x\sim p_2}[-c(b+x)] = -p_2 c(b+x_L) - (1 - p_2)c(b + x_H)$ and note that utility of self 1 and self 2 is increasing in p_2.

Intuitively, this as-if model captures the idea that individuals may distort their memory by suppressing negative information to enjoy the prospect of a lower workload. The agent

[2]In our experimental setup, such costs can result not only from the cognitive effort of maintaining incorrect beliefs but also from incentives to report accurate beliefs. Further, these costs may become negative if recall is cognitively more costly than suppression of information. Finally, as discussed earlier, we refrain from modeling feedback effects between action and beliefs (see Brunnermeier *et al.*, 2017) and thus from including an additional cost component that stems from anticipated effects of beliefs on the allocation decision.

[3]When signals are perfectly reliable, a fully sophisticated self 2 will not hold optimistic beliefs because it correctly identifies no news as negative news. However, extending the model by allowing for exogenous failures of message delivery (= unplanned forgetting) would allow also sophisticated agents to hold optimistic beliefs (see Hagenbach and Koessler, 2022).

will trade off the utility realized from recalling the signal against the utility realized from not recalling the signal but paying the cost of memory distortion. In terms of the dual-self model, self 1 will send message $m = s$ (i.e., the agent will recall) if and only if:

$$E_{x\sim p_2^s}[-c(b+x)] \geq E_{x\sim p_2^\emptyset}[-c(b+x)] - k_i \tag{8}$$

$$\Leftrightarrow k_i \geq (p_2^s - p_1)\left[c(b+x_L) - c(b+x_H)\right] \tag{9}$$

Inequality (8) ensures that self 2 (weakly) benefits from receiving the message. Inserting the expectations about effort costs with and without signals and rearranging yields inequality (9). Clearly, the second term on the right-hand side of (9) is always negative, as the effort costs from high workload outweigh the effort costs from low workload. The belief differential $(p_2^s - p_1)$ will be positive for positive signals (and negative for negative signals). For each signal, there exists a cost threshold $\tilde{k}^s$ such that self 1 optimally discloses the signal s if and only if $k_i \geq \tilde{k}^s$. By our assumption on the effect of signals on beliefs, it holds that $\tilde{k}^{neg} < 0 < \tilde{k}^{pos}$. Thus, a positive signal will be recalled by all agents who experience any positive costs from suppressing a signal. For a negative signal, it depends on the cost of suppression whether a signal is still disclosed. The normal distribution with a strictly positive mean further implies that the majority of agents may forget negative signals but will recall positive signals.[4]

The described model, where self 1 can successfully manipulate self 2's beliefs by sending empty messages, mirrors the HighScope condition in our experimental design, where participants have the possibility to suppress news or forget. In contrast, the LowScope condition, where participants are reminded of the signal about their future workload, is best understood as a situation in which self 1 cannot manipulate self 2's beliefs because self 2 always learns the true signal s from an external information source, irrespective of the message sent by self 1.

B Additional analyses

B.1 Prior beliefs

The prior beliefs have a mean of 48.95 (not significantly different from the rational benchmark of 50, $p = 0.354$). We display the cumulative distribution function in Figure B.1. Around half (48.8 percent of participants) state the modal and median belief $p_1 = 50$. The standard deviation of the prior p_1 is 21.6 and it's distribution is symmetric around the rational benchmark of $p_1 = 50$. While 26.7 percent of participants state a belief below 50, 24.5 percent state a belief above 50. This variance reflects homegrown optimism or pessimism.

[4]With more nuanced signal, the costs of suppression would also have more possible outcomes. In particular, it would hold that: $\tilde{k}^0 < \tilde{k}^1 < 0 < \tilde{k}^2 < \tilde{k}^3$. Also here, positive signals will be recalled by all agents who experience any positive costs from suppressing a signal. The more negative the signal is, the larger the cost of suppression must be for inequality (8) to hold and thus for self 1 to decide to still disclose the signal. As we assume k_i to be continuously distributed, a signal is less likely to be recalled the more negative it is. A minority of agents with (very) negative cost draws will recall neither negative nor even (very) positive signals.

The prior belief does not differ by Scope (LowScope: 49.76, HighScope: 48.14, $p = 0.473$) nor news valence (pos. news: 49.28, neg. news 48.65, $p = 0.782$).

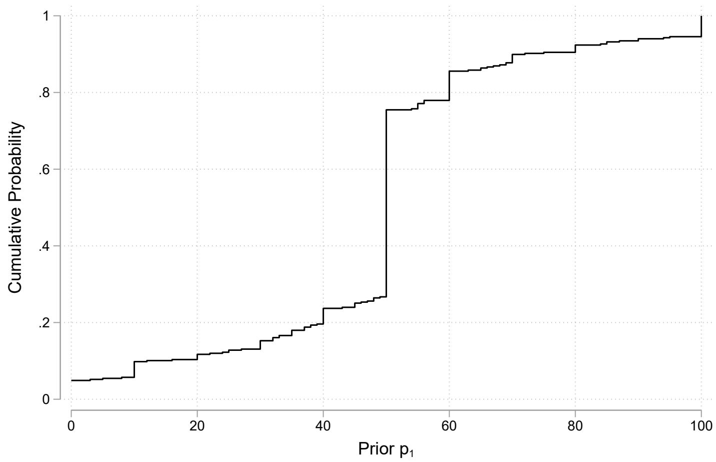


Figure B.1: Distribution of prior beliefs

Notes: The figure shows the cumulative distribution function of participants' prior beliefs in Session 1 (p_1).

B.2 Posterior beliefs

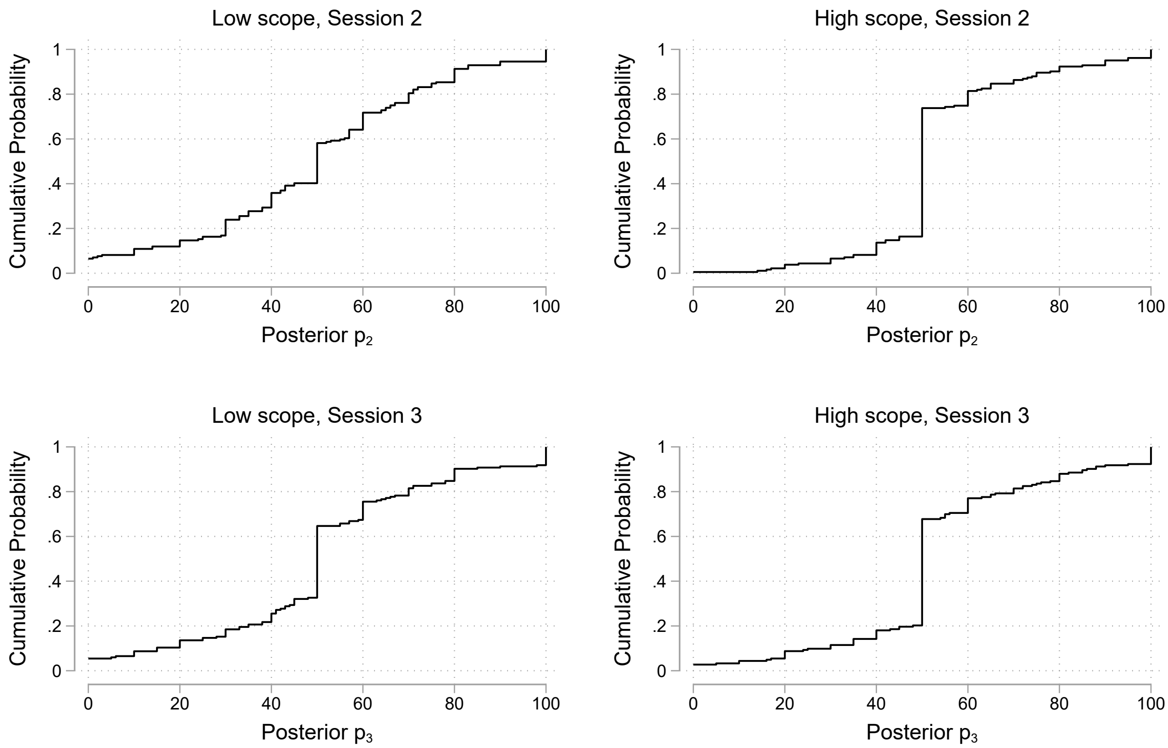


Figure B.2: Distribution of posterior beliefs

Notes: The figure shows the cumulative distribution function of participants' posterior beliefs by treatment and Session.

B.3 Main results based on exact feedback

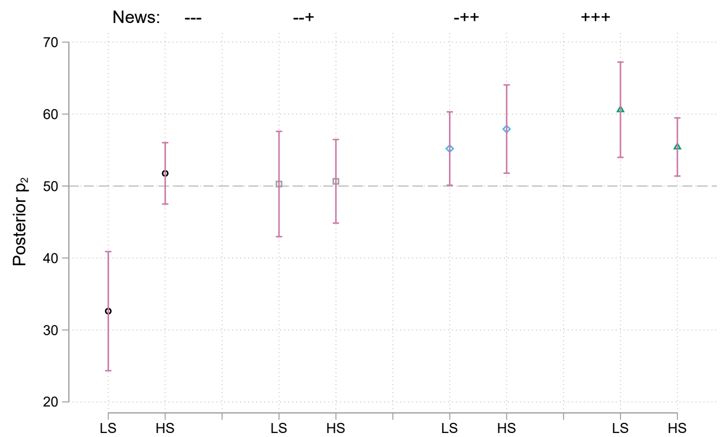


Figure B.3: Posterior beliefs

Notes: The figure shows participants' posterior beliefs in Session 2 (p_2) across treatment conditions and news (very negative: $---$, negative: $--+$, positive: $-++$, very positive: $+++$) received. The pink bars indicate 95%-confidence intervals.

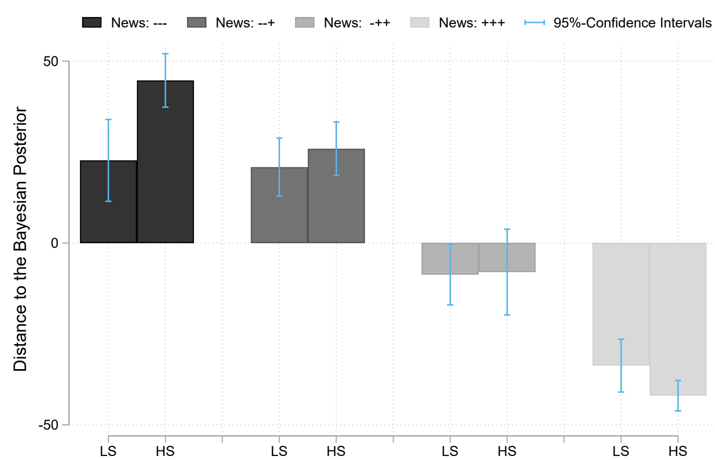


Figure B.4: Distance to the Bayesian posterior

Notes: The figure shows participants' optimism in Session 2 ($p_2 - p_{\mathcal{B}}$) across treatment conditions and news (very negative: $---$, negative: $--+$, positive: $-++$, very positive: $+++$) received. The blue bars indicate 95%-confidence intervals.

Table B.1: Regression results: Effects on posterior beliefs and optimism

Panel A: Posterior beliefs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HighScope	-5.173	-5.523	-5.219	-5.258	-5.155	-5.093	-5.542
	(3.84)	(3.91)	(3.87)	(3.85)	(3.86)	(3.87)	(3.98)
1 neg. signal	-5.396	-5.729	-5.530	-5.342	-5.377	-5.567	-5.921
	(4.14)	(4.19)	(4.21)	(4.14)	(4.19)	(4.17)	(4.30)
2 neg. signals	-10.322**	-10.478**	-10.330**	-10.468**	-10.279**	-10.446**	-10.723**
	(4.89)	(4.93)	(4.91)	(4.90)	(5.02)	(4.91)	(5.09)
3 neg. signals	-27.988***	-28.021***	-28.040***	-27.881***	-27.981***	-28.228***	-28.198***
	(5.26)	(5.24)	(5.28)	(5.27)	(5.27)	(5.20)	(5.22)
HS X 1 neg. signal	7.886	8.333	8.081	7.992	7.869	7.867	8.529
	(5.50)	(5.60)	(5.56)	(5.50)	(5.52)	(5.52)	(5.68)
HS X 2 neg. signals	5.548	5.660	5.557	5.660	5.510	5.485	5.660
	(6.01)	(6.03)	(6.03)	(6.00)	(6.08)	(6.03)	(6.11)
HS X 3 neg. signals	24.318***	24.844***	24.305***	24.288***	24.322***	24.272***	24.792***
	(6.02)	(5.99)	(6.02)	(6.03)	(6.03)	(6.05)	(6.04)
Patience		0.884					0.940
		(1.05)					(1.06)
Procrastination scale			-0.471				-0.240
			(1.16)				(1.20)
Suppression factor				-0.630			-0.693
				(1.01)			(1.02)
Reappraisal factor					-0.067		-0.111
					(1.13)		(1.13)
Pref. for Information						-0.599	-0.686
						(1.12)	(1.14)
Constant	60.605***	60.755***	60.658***	60.619***	60.584***	60.721***	60.907***
	(3.28)	(3.30)	(3.32)	(3.29)	(3.33)	(3.31)	(3.42)
N	367	367	367	367	367	367	367

** table continues on next page **

Panel B: Optimism

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HighScope	-8.233**	-7.710*	-8.417**	-8.329**	-8.657**	-8.238**	-8.319*
	(4.14)	(4.20)	(4.20)	(4.15)	(4.16)	(4.14)	(4.28)
1 neg. signal	25.055***	25.552***	24.524***	25.117***	24.595***	25.065***	24.819***
	(5.49)	(5.53)	(5.60)	(5.47)	(5.54)	(5.55)	(5.71)
2 neg. signals	54.577***	54.810***	54.546***	54.414***	53.568***	54.583***	53.777***
	(5.34)	(5.35)	(5.37)	(5.36)	(5.57)	(5.38)	(5.63)
3 neg. signals	56.426***	56.477***	56.217***	56.545***	56.280***	56.439***	56.352***
	(6.65)	(6.68)	(6.65)	(6.61)	(6.68)	(6.55)	(6.60)
HS X 1 neg. signal	8.911	8.243	9.687	9.030	9.301	8.912	9.328
	(8.25)	(8.23)	(8.35)	(8.22)	(8.28)	(8.25)	(8.32)
HS X 2 neg. signals	13.282*	13.115*	13.317*	13.407**	14.174**	13.285*	14.115**
	(6.77)	(6.78)	(6.80)	(6.79)	(6.84)	(6.77)	(6.89)
HS X 3 neg. signals	30.201***	29.415***	30.149***	30.167***	30.107***	30.203***	29.152***
	(7.87)	(7.93)	(7.86)	(7.87)	(7.87)	(7.91)	(7.95)
Patience		-1.319					-1.526
		(1.50)					(1.53)
Procrastination scale			-1.867				-1.715
			(1.44)				(1.53)
Suppression factor				-0.705			-0.500
				(1.38)			(1.46)
Reappraisal factor					1.587		1.553
					(2.04)		(2.07)
Pref. for Information						0.033	0.298
						(1.67)	(1.70)
Constant	-33.717***	-33.941***	-33.506***	-33.701***	-33.239***	-33.723***	-33.363***
	(3.59)	(3.60)	(3.67)	(3.60)	(3.67)	(3.64)	(3.80)
N	367	367	367	367	367	367	367

Notes: The table shows results from OLS regressions. The dependent variable in Panel A is participants' posterior belief about the probability to face low workload (p_2). The dependent variable in Panel B is participants' optimism ($p_2 - p_{\mathcal{B}}$). The main explanatory variables are the treatment dummy HighScope, dummies for the number of negative signals received and the interaction between the number of negative signals and HighScope. The omitted category are 0 neg. signals (= 3 pos. signals). The control variables are standardized continuous measures resulting from the respective questionnaires. Robust standard errors clustered at the day level reported in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table B.2: Regression results: beliefs on the work decision

Panel A: The number of tasks participants complete in Session 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Posterior (instrumented)	-0.156*	-0.153*	-0.156*	-0.156*	-0.153*	-0.155*	-0.150*
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Patience		-0.159					-0.226
		(0.63)					(0.63)
Procrastination scale			-0.221				-0.139
			(0.67)				(0.71)
Suppression factor				0.059			0.041
				(0.61)			(0.62)
Reappraisal factor					0.644		0.646
					(0.68)		(0.69)
Pref. for Information						-0.122	-0.099
						(0.67)	(0.67)
Constant	37.724***	37.590***	37.717***	37.707***	37.590***	37.667***	37.388***
	(4.49)	(4.52)	(4.49)	(4.49)	(4.47)	(4.44)	(4.47)
N	367	367	367	367	367	367	367
Mean dependent variable	29.67	29.67	29.67	29.67	29.67	29.67	29.67

Panel B: The probability to solve the maximum number of tasks (40) in Session 2

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Posterior (instrumented)	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Patience		-0.005					-0.005
		(0.03)					(0.03)
Procrastination scale			-0.006				-0.001
			(0.03)				(0.03)
Suppression factor				-0.033			-0.032
				(0.03)			(0.03)
Reappraisal factor					0.008		0.009
					(0.03)		(0.03)
Pref. for Information						0.008	0.004
						(0.03)	(0.03)
Constant	0.703***	0.702***	0.702***	0.710***	0.701***	0.710***	0.709***
	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
N	367	367	367	367	367	367	367
Mean dependent variable	0.44	0.44	0.44	0.44	0.44	0.44	0.44

Notes: The table shows results from IV regressions using the general method of moments estimator. The posterior belief measured on a scale from 0 - 100 (p_2) about facing low workload is instrumented with the treatment dummy for HighScope, a dummy for each possible signal (number of positive comparisons) and the interaction of both. The dependent variable in Panel A is the number of tasks participants complete in Session 2. The dependent variable in Panel B is the probability to solve the maximum number of tasks (40) in Session 2. The control variables are standardized continuous measures resulting from the respective questionnaires. Robust standard errors clustered at the day level reported in parentheses, and *** p<0.01, ** p<0.05, * p<0.1.

Table B.3: ATE of HighScope on the work decision

Panel A: The number of tasks participants complete in Session 2

	Logit		Probit	
	1 Neighbor	2 Neighbors	1 Neighbor	2 Neighbors
	(1)	(2)	(3)	(4)
ATE				
HighScope	-1.674	-1.297	-1.674	-1.297
	(1.24)	(1.21)	(1.24)	(1.21)
N	367	367	367	367

Panel B: The probability to solve the maximum number of tasks (40) in Session 2

	Logit		Probit	
	1 Neighbor	2 Neighbors	1 Neighbor	2 Neighbors
	(1)	(2)	(3)	(4)
ATE				
HighScope	-0.122**	-0.107**	-0.122**	-0.107**
	(0.05)	(0.05)	(0.05)	(0.05)
N	367	367	367	367

Notes: The table shows results from propensity score matching. The analysis matches individuals on priors (p_1) and signals (number of positive comparisons) using nearest neighbor matching, with replacement when > 1 neighbor. The dependent variable in Panel A is the number of tasks participants complete in Session 2. The dependent variable in Panel B is the probability to solve the maximum number of tasks (40) in Session 2. Columns (1) and (2) use 1 neighbor, while Column (2) uses 2 neighbors. Columns (1) and (2) use logit while Columns (3) and (4) use probit to estimate propensity scores. Abadie-Imbens robust standard errors clustered at the day level reported in parentheses, and *** p<0.01, ** p<0.05, * p<0.1..

B.4 Exploratory analyses of sticky beliefs and belief dynamics

As discussed in the main text, 42.5% of our participants do not adjust their expectations from Session 2 to Session 3 ($p_2 = p_3$). This stickiness may have different explanations of which we discuss two in more detail here.

First, our design potentially facilitates consistency in the sense of stating identical beliefs across sessions because the decision screens for the belief elicitation looked very similar in Sessions 2 and 3. However, general concerns for consistency *per se* cannot explain why we observe sticky beliefs significantly more often in the HighScope than in the LowScope treatment.[5] About two thirds of those participants with "sticky beliefs" hold a posterior belief of 50% and thereby seem to completely ignore the signal in both sessions. The remaining third of participants with "sticky beliefs" hold beliefs different from 50%. These participants have invested cognitive resources to form and express their belief in Session 2, which may have increased their adjustment costs in Session 3 akin to the findings by Falk and Zimmermann (2018). Pushing this line of thought further, this result may hint to a belief formation process in which beliefs are internalized once they are formed. In a subsequent belief elicitation, the internalized belief will simply be retrieved instead of the participant forming a new one. Indeed, rational individuals have no reason to adjust their beliefs if there is no new information.

Second, participants may use their actions in Session 2 to *ex post* (in Session 3) impute the beliefs they must have held when allocating work as in Heidhues *et al.* (2023). Such behavior will yield optimistic beliefs exactly for those participants who chose a low workload (which might have been a suboptimal result based on an optimistic belief in Session 2) because optimistic beliefs are needed to rationalize this choice. In the theory by Heidhues *et al.* (2023), individuals may accurately recall their past actions (here: work decision) but not what led to their decision (here: signals received). As a result, actions and beliefs are consistent and procrastination can occur in equilibrium without individuals learning over time. That is, optimistic beliefs in Session 3 may simply appear consistent to participants who chose to work rather little in Session 2. Importantly, the latter intuition cannot explain the observed asymmetry in posteriors due to negative news in HighScope in Session 2.

B.5 Additional descriptive results on the work choice

In this section, we briefly discuss the raw data of the number of tasks solved in Session 2. First, we show a histogram of the chosen number of tasks in Figure B.5. The histogram shows that, although the modal choice was 40 tasks, there is substantial variation in work allocation decisions. Second, we find that participants choose to work on more tasks after receiving negative news than after receiving positive news. Positive as compared to negative news lead participants to complete on average 2.72 tasks less in Session 2 (30.95 average tasks after negative news, 28.23 after positive news, t–test $p = 0.023$). Third, participants tend to work less in HighScope (28.92 vs. 30.41, t–test $p = 0.215$), which is in line with our finding that the HighScope treatment induced some individuals to hold optimistic beliefs.

[5] We find that 54% of participants in HighScope report the same belief in Session 2 and 3, while only 32% do so in the LowScope condition (t–test: $p < 0.001$).

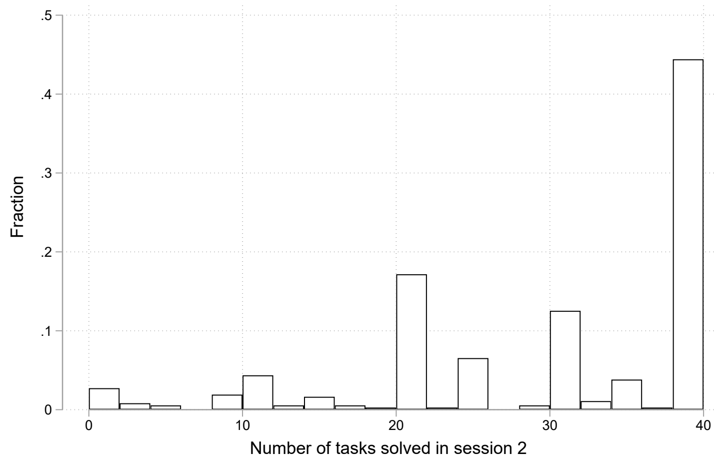


Figure B.5: Histogram of the number of tasks chosen to be solved in Session 2

These findings are in line with the idea that more optimistic beliefs (through positive news and the scope to manipulate beliefs) provide lower incentives to exert effort in Session 2 because these beliefs suggest a lower total required work effort to complete the task.

B.6 Robustness analyses

B.6.1 Randomization check

Following the advise by Kerwin *et al.* (2024), we conduct a test of joint orthogonality with randomization inference p-values. Reassuringly, both sources of experimental variation (Scope and signal valence) are successfully randomized ($p = 0.175$ for scope and $p = 0.881$ for signal valence). The balance test uses all demographics[6] and the Big Five personality traits (extraversion, agreeableness, conscientiousness, neuroticism, openness). A joint orthogonality test with inference based p-values leads to the same conclusion ($p = 0.125$ for scope and $p = 0.863$ for signal valence).

B.6.2 Alternative priors

Our treatment effect is robust to using rational priors that assign 10% probability to each possible workload. In Table B.4, we show the specification with subjective priors as a benchmark in Column (1) and the specification using objective priors in Column (2). The

[6]Demographics include age, a male dummy, and the number of previously completed studies (grouped into 0 [omitted category], 1–5, 6–10, and more than 10). We also control for educational attainment (high school, bachelor's, and at least a master's degree) and field of study. For the latter, we summarized the categories and included two dummy variables: one for economics or business, and one for sciences; all other fields constitute the omitted category. Finally, the demographics include the self-reported level of English proficiency.

Table B.4: Regression results: Comparison of the treatment effect on optimism with subjective and objective priors

	(1) Benchmark based on subjective priors	(2) Benchmark based on objective priors
HighScope	-5.709	-2.698
	(4.6)	(3.4)
Neg. News	42.295***	55.206***
	(4.45)	(3.29)
HighScope*Neg. News	21.530***	15.339***
	(6.33)	(4.68)
Constant	-20.501***	-27.302***
	(3.17)	(2.34)
N	367	367

Notes: The table shows results from seemingly unrelated regressions. The dependent variable is participants' optimism. In Column (1), optimism is calculated using the subjective priors that we elicited in Session 1 ($p_2 - p_{\mathcal{B}}$). Column (2) uses the objective prior where every number of tasks is equally likely to calculate optimism ($p_2 - p_{\mathcal{B}}^{O}$). The explanatory variables are the treatment dummy HighScope, a negative news dummy and the interaction. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

effects are qualitatively similar and only the coefficient for *Neg. News* (and the constant) is significantly different across the two specifications.[7]

B.6.3 Experience with the task

In the first wave, due to a technical problem, participants from the Berlin-based participant pool did not have to complete the ten trial sequences of the transcription task in Session 1. Nevertheless, these participants read the instructions for the transcription task (see Figure F.1). To test whether the lack of task experience alters our results, we employ two complementary approaches. First, we include a dummy variable (*No trial*) indicating whether participants did not experience the trial phase. Second, we run the same regression specifications as in the main text but exclude those participants who had no trial experience.

Table B.5 presents the results of these additional analyses. Column (1) shows that not having participated in the trial (*No trial*) did not significantly affect prior beliefs. The same holds true for posterior beliefs and optimism (see Column (3) and (6)). Further, including the *No trial* dummy does not change the point estimates of our relevant explanatory variables (compare Column (2) and (3) for posterior beliefs and Column (5) and (6) for optimism). The point estimates also remain largely unchanged when we run the regression on the subsample of individuals who participated in the trial (compare Column (2) and (4) for posterior beliefs and Columns (5) and (7) for optimism) even though this exclusion reduces the sample size substantially (from N=367 to N=277, see Column (6)).

Table B.6 repeats the above analyses with respect to the work decision: We find that neither controlling for the technical error (compare Column (2) to (1) and (5) to (4), respectively) nor excluding participants who could not participate in the trial task (compare Columns (3) to (1) and (6) to (4), respectively) has an impact on the relevant point esti-

[7]Constant: $p = 0.006$, *Neg. News*: $p < 0.001$, *HighScope*: $p = 0.339$, Interaction: $p = 0.160$; Wald-tests.

mates. However, due to the loss in power (sample size reduction from N=367 to N=277), the effect of beliefs on the probability to complete the maximally possible number of sequences is no longer significant.

Table B.5: Regression Results: Controlling for experience with the task

	Prior		Posterior			Optimism	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
No trial	1.110		-1.075			-1.667	
	(2.48)		(2.34)			(2.90)	
HighScope		-1.173	-1.295	-0.495	-5.709	-5.899	-3.629
		(2.71)	(2.75)	(3.21)	(4.61)	(4.65)	(5.76)
Neg. News		-16.403***	-16.503***	-15.769***	42.295***	42.141***	40.696***
		(3.54)	(3.56)	(4.19)	(4.58)	(4.62)	(5.83)
HighScope*Neg. News		11.165***	11.376***	10.985**	21.530***	21.857***	21.203***
		(4.31)	(4.38)	(5.11)	(6.39)	(6.46)	(7.98)
Constant	48.679***	57.758***	58.077***	57.391***	-20.501***	-20.006***	-20.214***
	(1.34)	(2.05)	(2.24)	(2.59)	(3.05)	(3.41)	(4.14)
N	367	367	367	277	367	367	277

Notes: The table shows results from OLS regressions. The dependent variable is participants' prior (p_1) in Column (1), posteriors (p_2) in Columns (2)-(4) and optimism ($p_2 - p_{\mathcal{B}}$) in Columns (5) - (7). The explanatory variables are a dummy that is 1 if participants did not experience the task prior to belief elicitation (*No trial*), the treatment dummy *HighScope*, a negative news dummy and the interaction. Columns (4) and (7) are based on the restricted sample excluding those participants that did not experience the task prior to belief elicitation. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.6: Regression results: Controlling for experience with the task

	Number Tasks			P(40 tasks)		
	(1)	(2)	(3)	(4)	(5)	(6)
No trial		0.263			-0.075	
		(1.34)			(0.06)	
Posterior $\hat{p}_2$	-0.232**	-0.232**	-0.220*	-0.008*	-0.007*	-0.007
	(0.10)	(0.10)	(0.12)	(0.00)	(0.00)	(0.00)
Constant	41.650***	41.568***	41.001***	0.836***	0.844***	0.835***
	(5.36)	(5.38)	(6.45)	(0.21)	(0.21)	(0.25)
N	367	367	277	367	367	277

Notes: The table shows results from IV regressions using the general method of moments estimator. The posterior belief (p_2) about facing low workload is instrumented with the treatment dummy for HighScope, a dummy for negative news and the interaction of both (for the first stage, see Table 1). The dependent variable in Columns (1) to (3) is the number of tasks participants complete in Session 2. The dependent variable in Columns (4) to (6) is the probability to solve the maximum number of tasks (40) in Session 2. Columns (4) and (7) are based on the restricted sample excluding those participants that did not experience the task prior to belief elicitation. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B.6.4 Controlling for wave and subject pools effects

Table B.7: Regression Results: Controlling for wave and subject pools

	Posterior		Optimism	
	(1)	(2)	(3)	(4)
HighScope	-1.173	-1.132	-5.709	-5.872
	(2.71)	(2.75)	(4.61)	(4.67)
Neg.News	-16.403***	-16.305***	42.295***	42.096***
	(3.54)	(3.55)	(4.58)	(4.63)
HighScope*Neg.News	11.165***	11.101**	21.530***	21.677***
	(4.31)	(4.39)	(6.39)	(6.45)
Wave1		2.543		-5.519
		(3.16)		(4.99)
Berlin		2.990		1.643
		(3.11)		(4.63)
Wave1*Berlin		-4.836		1.087
		(4.35)		(6.41)
Constant	57.758***	56.209***	-20.501***	-18.815***
	(2.05)	(2.71)	(3.05)	(4.30)
N	367	367	367	367

Notes: The table shows results from OLS regressions. The dependent variable is participants' prior (p_1) in Column (1), posteriors (p_2) in Columns (1) and (2) and optimism ($p_2 - p_{\mathcal{B}}$) in Columns (3) and (4). The explanatory variables are the treatment dummy *HighScope*, a negative news dummy and the interaction. They also include a dummy *Wave1*, that is equal to 1 if the data was collected in the first wave, a dummy *Berlin* that is equal to one if the data was collected in the Berlin-based laboratory, as well as the interaction of the two variables. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We collected data in two laboratories, the Munich Experimental Laboratory of Economic and Social Sciences (MELESSA) and the TU-WZB-laboratory in Berlin, and during two waves (Wave 1: June-July 2022, Wave 2: October-November 2022). Therefore, we test whether our main results stay robust to controlling for wave and lab. Tables B.7 and B.8 show that the coefficients of the main specifications hardly change when including as controls a dummy for the wave, a dummy for the lab as well as their interaction.

Table B.8: Regression results: Controlling for wave and subject pools

	Number Tasks		P(40 tasks)	
	(1)	(2)	(3)	(4)
Posterior (instrumented)	-0.232**	-0.237**	-0.008*	-0.008*
	(0.10)	(0.10)	(0.00)	(0.00)
Wave1		0.669		0.031
		(1.98)		(0.08)
Berlin		2.007		0.132*
		(1.87)		(0.08)
Wave1*Berlin		-1.547		-0.185*
		(2.65)		(0.11)
Constant	41.650***	40.980***	0.836***	0.810***
	(5.36)	(5.32)	(0.21)	(0.21)
N	367	367	367	367

Notes: The table shows results from IV regressions using the general method of moments estimator. The posterior belief (p_2) about facing low workload is instrumented with the treatment dummy for HighScope, a dummy for negative news and the interaction of both (for the first stage, see Table 1). The dependent variable in Columns (1) and (2) is the number of tasks participants complete in Session 2. The dependent variable in Columns (3) to (4) is the probability to solve the maximum number of tasks (40) in Session 2. In Columns (2) and (4) we control for a dummy *Wave1*, that is equal to 1 if the data was collected in the first wave, a dummy *Berlin* that is equal to one if the data was collected in the Berlin-based laboratory, as well as the interaction of the two variables. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B.6.5 Controlling for demographic characteristics

As our experiment is successfully randomized (see B.6.1), we do not control for general demographic characteristics in our main specification. However, we can show that our main results are highly robust to controlling for demographic characteristics. Table B.9 shows the results for our main specifications in columns (1) and (4), along with additional specifications where sets of control variables are introduced step-wise. First, we add controls for demographic characteristics in Columns (2) and (5). These include gender, age, number of studies previously completed, degree and field of studies. Second, we add the Big Five personality traits as controls in columns (3) and (6). Our results remain robust to the inclusion of these additional controls.

Table B.9: Regression Results: Controlling for demographics

	Posterior			Optimism		
	(1)	(2)	(3)	(4)	(5)	(6)
HighScope	-1.173	-0.941	-1.090	-5.709	-5.963	-5.691
	(2.71)	(2.75)	(2.75)	(4.61)	(4.73)	(4.77)
Neg.News	-16.403***	-16.195***	-16.306***	42.295***	42.219***	43.066***
	(3.54)	(3.52)	(3.54)	(4.58)	(4.69)	(4.67)
HighScope*Neg.News	11.165***	10.279**	10.300**	21.530***	21.878***	20.675***
	(4.31)	(4.33)	(4.31)	(6.39)	(6.64)	(6.64)
Constant	57.758***	36.210***	17.413	-20.501***	-41.152**	-76.335***
	(2.05)	(11.70)	(15.05)	(3.05)	(16.54)	(22.60)
Demographic controls		✓	✓		✓	✓
Controlling for Big 5			✓			✓
N	367	366	366	367	366	366

Notes: The table shows results from OLS regressions. The dependent variable is participants' prior (p_1) in Column (1), posteriors (p_2) in Columns (1) to (3) and optimism ($p_2 - p_{\mathcal{B}}$) in Columns (4) to (6). The explanatory variables are the treatment dummy *HighScope*, a negative news dummy and the interaction. Demographic controls include age, a male dummy, and the number of previously completed studies (grouped into 0 [omitted category], 1–5, 6–10, and more than 10). We also control for educational attainment (high school, bachelor's, and at least a master's degree) and field of study. For the latter, we summarized the categories and included two dummy variables: one for economics or business, and one for sciences; all other fields constitute the omitted category. Finally, the demographics include the self-reported level of English proficiency. The big five include seperate controls for all five factors (extraversion, agreeableness, conscientiousness, neuroticism and openness). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B.7 Participants' characteristics and prior beliefs

Heterogeneity in characteristics may systematically shape participants priors and thus bias the findings reported in Section 5.2. In Table B.10, we regress priors on participants' characteristics. We do not find statistically significant relationships between participants characteristics and their priors and thus no indication of such bias.

Table B.10: Regression results: Balance in priors

	(1)	(2)	(3)	(4)	(5)	(6)
Patience		0.858				
		(1.08)				
Procrastination scale			0.624			
			(1.21)			
Suppression factor				-0.913		
				(1.10)		
Reappraisal factor					-1.437	
					(1.37)	
Pref. for Information						-0.911
						(1.17)
Constant	48.951***	48.951***	48.951***	48.951***	48.951***	48.951***
	(1.13)	(1.13)	(1.13)	(1.13)	(1.13)	(1.13)
N	367	367	367	367	367	367

Notes: The table shows results from OLS regressions. The dependent variable is participants' prior belief about the probability to face low workload (p_1). The control variables are standardized continuous measures resulting from the respective questionnaires. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B.8 Analysis of starting times

Table B.11: Regression results: Session start times

	time start s1	time start s2	time start s3	time start s2	time start s3
	(1)	(2)	(3)	(4)	(5)
Patience measure	0.882	1.990*	0.230		
	(1.09)	(1.04)	(1.07)		
Procrastination scale	0.021	0.295	0.184		
	(0.22)	(0.23)	(0.24)		
Neg.News				-0.460	-0.711
				(0.45)	(0.67)
HighScope					-0.311
					(0.67)
HighScope*Neg.News					0.888
					(0.91)
Constant	12.712***	12.679***	13.604***	14.347***	14.055***
	(0.81)	(0.75)	(0.79)	(0.33)	(0.50)
N	308	367	367	367	367

Notes: All regressions are estimated using OLS. The dependent variable is the hour of the day a participant started with the respective session. The explanatory variables are the treatment dummy HighScope, a dummy for negative news (Neg. News) and their interaction as well as a standardized patience measure (an aggregate measure derived from hypothetical choices between money now or later, and the stated willingness to give up something that is beneficial today in order to benefit in the future (Falk *et al.*, 2023)) and the standardized procrastination scale (Steel, 2010). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B.9 The dotspot task

Our experiment also included a belief formation and memory task regarding a neutral topic: The number of red/blue dots in a graph shown to participants, see Figure B.6. In the first session, we showed participants this graph consisting of a total of 400 dots for 8 seconds. We randomized i) whether the majority of the dots was red or blue (35% and 65%, respectively) and ii) whether we asked participants to estimate the percentage of red or blue dots. In the second session, we again asked participants for the percentage of red (or blue) dots. Before doing so, in LowScope, the graph appeared again on participants' screen while it was not shown again in HighScope. Participants' guesses regarding the number of dots of a particular color was incentivized using the binarized scoring rule in both sessions. Figure B.6 shows an example of participants' screen with a graph including 35% red dots.

This task allows us to examine whether, when being asked about the probability of an event to occur, participants tend to forget or misremember signals indicating a low likelihood of that event more easily than signals suggesting a high likelihood (here: many of the dots were of the color of interest / in the main experiment: signal about a high likelihood of being

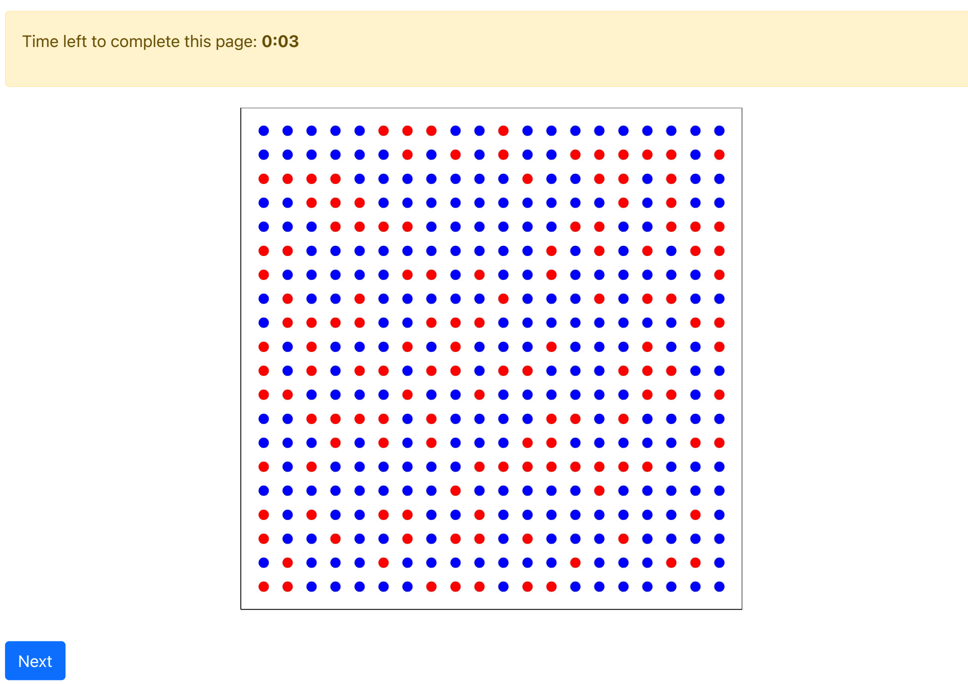


Figure B.6: Screen with the dotspot task

assigned a low number of sequences).

Table B.12 shows the results of a regression analysis, where we regress the absolute difference between the percentage entered between Session 1 and 2 on the HighScope dummy, a dummy variable that indicated whether the majority of the dots was in the color that was asked for (*ManyDots*) as well as the interaction of these two. Naturally, the absolute difference is larger in HighScope, where individuals need to rely on their memory. However, we find no differential effect by *ManyDots*. In other words, in a neutral task, we observe no difference in memory effects for news that convey a high likelihood of the event of interest occurring (as compared to a low likelihood). Consequently, the asymmetry in optimism after negative news (a signal for a low likelihood of the event considered) in HighScope in the main experiment does not stem from a general memory effect but results from participants' wishful thinking about low total workload (see also Appendix E).

For completeness, we also provide information about participants' updating in the dotspot task by showing the distributions of the difference between the percentage entered in Session 1 and 2. Figure B.7 plots these differences separately by scope and fraction of dots in the color that was asked for. While in LowScope this difference is approximately normally distributed around 0, in HighScope belief dynamics are in line with regression to the mean in a symmetric manner: Participants who saw a dotspot with 35% dots of the color asked for similarly moved towards stating 50% as did participants who saw a dotspot with 65%. The latter findings underline our previous conclusion.

Table B.12: Regression results: Absolute difference between guessed percentage of dots in Session 1 and 2 in the dotspot task

	(1)
HighScope	3.637***
	(1.40)
ManyDots	0.660
	(1.39)
HighScope× ManyDots	2.226
	(1.97)
Constant	6.096***
	(0.97)
N	367

Notes: The table shows results from OLS regressions. The dependent variable is the absolute difference between the percentage entered between Session 1 and 2. The independent variables include the HighScope dummy, a dummy that indicates whether the majority of the dots was in the color that was asked for (ManyDots) as well as the interaction of these two. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

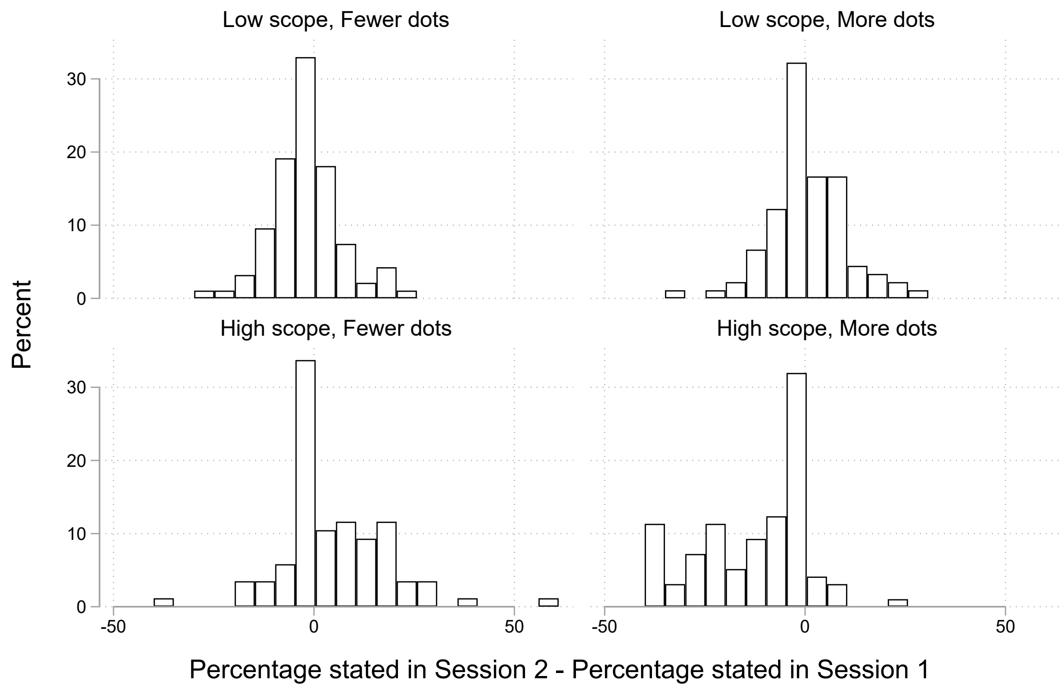


Figure B.7: Histogram of belief differences in the dotspot task

C Preregistration and deviations from it

Our preregistration comprised five hypotheses on motivated beliefs in a work context (called 'RE' for real effort in the preregistration) and four hypotheses on motivated beliefs about relative performance in a set of Raven matrices, which are used in intelligence tests (IQ). The RE and IQ experiments were independent of each other and the IQ setting is analyzed in Cordes *et al.* (2025). Here, we discuss how the preregistered hypotheses for the RE setting, relate to the analysis and results shown in this manuscript. H1, H2, and H3 were preregistered as our main hypotheses, H4 and H5 were introduced as additional hypotheses in the preregistration. For convenience, we quote the hypotheses from the preregistration, which can be found at `https://aspredicted.org/SHS_XD6`. The abbreviations LS and HS refer to our treatments LowScope and HighScope.

H1: *Participants update their prior according to the noisy signal they receive.* The hypothesized correlation between signals and updating is a prerequisite for the subsequent hypotheses to make sense and for our analyses to be meaningful. We chose not to display this statement as a numbered hypothesis or prediction in the manuscript, but it is obvious from the posteriors in LowScope shown in Figure 2. We state this observation in the first paragraph of Section 4.1.

H2: *In Session 2, the departure from Bayesian updating (based on individual priors) is more optimistic for participants in HS as compared to LS.* This hypothesis corresponds to Prediction 1 of our theoretical framework. Result 1 is in line with the hypothesis. The assumptions that a) signals move beliefs in the right direction and b) costs of suppressing a signal are continuously distributed with a positive mean together yields Prediction 2, which we test to better understand Result 1. The corresponding analysis is summarized as Result 2, lending support to this additional hypothesis.

H3: *In the third session, beliefs are less optimistic than in Session 2.* We discuss the change in beliefs from Session 2 to Session 3 in Section 5.1. The hypothesis that beliefs are less optimistic in Session 3 is not confirmed by our data.

H4: *In RE, the number of sequences solved in Session 2 correlates positively with the expected number of additional sequences to be solved until the end of Session 3.* H4 states that we expected participants to choose their effort in Session 2 (i.e., the number of sequences to solve in Session 2) in response to their beliefs about their assigned workload. As both posterior beliefs and work choices are endogenous and potentially influenced by signals and scope, we chose not to analyze the raw correlation, but use the exogenous variation in beliefs through signals and reminder in an IV approach. In line with H4, the IV analysis shows that more pessimistic individuals, who expect a higher workload, complete more sequences in Session 2.

H5: *In HS, the change in beliefs across time in the Dot-Spot task is less asymmetric as compared to the IQ/RE task.* We analyze participants' beliefs in the Dot-Spot task in Appendix B.9 and show that, as hypothesized, the HighScope treatment does not lead to similarly asymmetric updating in this task as that we observe in beliefs about the workload of sequences (i.e. in the RE task).

D Exclusion and attrition

D.1 Overview

As specified in our preanalysis plan, we excluded participants who:

1. have a low level of English (below 30% on a self-assessment scale from 0 to 100%).
2. rushed through the belief elicitation (spent less than 1 minute in total on the three pages related to the explanation of the belief elicitation and incentivization, point-belief elicitation and probabilistic belief elicitation).
3. did not pass one of our two attention checks in the first and third session (questions where we asked participants to select one specific value on a Likert-scale)

The first criterion applies to five (out of 531) participants. The second applies to two participants and the third to 37 participants. One participant met two of the exclusion criteria, resulting in a total of 43 individuals excluded based on the three preregistered criteria listed above. Our main analysis focuses on participants who completed all three sessions. Accordingly, we also discuss attrition (see Figure D.1).

While we observe some attrition at each stage of the experiment, overall attrition is relatively low for a longitudinal study spanning four weeks and appears not to be selective based on the news received. Table D.1 presents a formal analysis of differential attrition, indicating that attrition is unrelated to news, prior beliefs, or treatment assignment. This holds for participants who dropped out before Session 2 as well as for those who dropped out later. Columns (1) and (2) report estimates from a linear probability model where the dependent variable is an indicator for dropping out before Session 2, that is, either during Session 1 ($N = 12$) or prior to Session 2 ($N = 57$).[8] The estimated coefficients suggest that dropout prior to Session 2 is not systematically related to beliefs or to the exogenously assigned news. Columns (3) and (4) present results from a linear probability model where the dependent variable is an indicator for dropping out after the start of Session 2, that is, during Session 2 ($N = 5$), between Sessions 2 and 3 ($N = 36$), or during Session 3 (N = 11). Again, the estimates show no evidence that attrition is influenced by beliefs or by treatment assignment.[9]

D.2 Comparison between stay-ons and drop-outs

Although the previous subsection showed no selective attrition with respect to news received, prior beliefs, or experimental conditions, one might still ask whether participants who remained until the end differ systematically from those who dropped out. To examine this,

[8] As the random assignment of workloads and news took place at the beginning of the experiment, Neg.News is defined for all 488 participants, see Column (1). However, eight participants dropped out before their prior was elicited, leaving us with $N = 480$; see Column (2).

[9] Out of the five participants who dropped during Session 2, three did not submit their posterior belief (p_2), leaving us with $N = 416$; see Column (4).

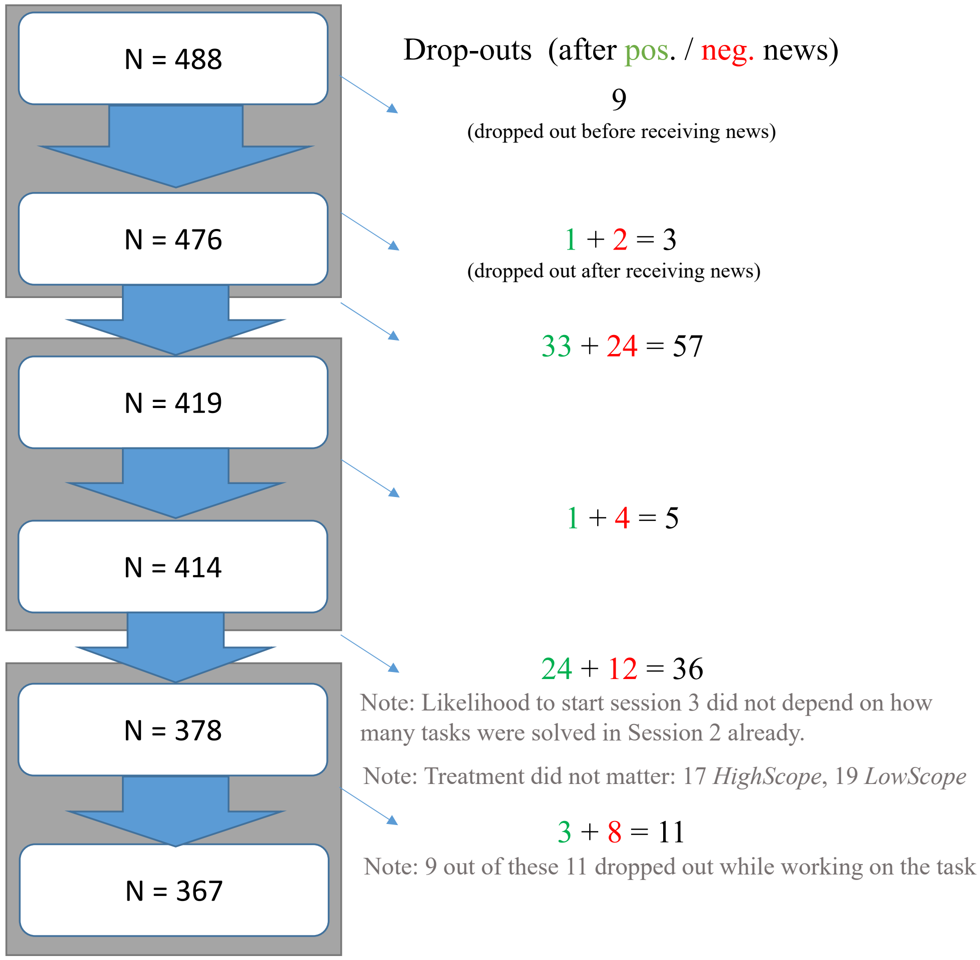


Figure D.1: Exclusion and attrition

Notes: The figure summarized the attrition conditional on passing the three exclusion criteria. Gray squares in the Figure represent the three experimental sessions. The number ($N =$) on the top of each gray square indicates how many participants started the respective session while the bottom number indicates how many participants completed the respective session. The numbers at the right of the arrows indicate how many participants drop out at the respective point in time.

we assess whether completing all sessions is correlated with any participant characteristics elicited in Session 1.[10] Figure D.2 reports coefficient estimates and 95-percent confidence intervals from regressions with a dummy that equals 1 if a participant completed all parts of the experiment on different (standardized) variables elicited in Session 1. Out of 19 coefficients only two turn out to be significantly different from zero: *Patience*, and *Giving up for future*. This points to a straightforward intuition: More patient people may be more willing to complete all parts of the experiment. As a result, participants who completed all

[10]Since most dropouts occurred before the start of Session 2, we restrict this analysis to characteristics measured in Session 1.

Table D.1: Regression Results: Attrition overview

	(1)	(2)	(3)	(4)
	Dropout before S2		Later dropout	
Prior p_1		0.0006		
		(0.001)		
Posterior p_2				-0.0002
				(0.001)
Neg.News	-0.0456		-0.0276	
	(0.029)		(0.045)	
HighScope			0.0125	
			(0.049)	
HighScope*Neg.News			-0.0042	
			(0.065)	
Constant	0.1404***	0.0897***	0.1333***	0.1269***
	(0.023)	(0.034)	(0.033)	(0.042)
N	488	480	419	416

Notes: The table shows results from OLS regressions. The dependent variable in Columns (1) and (2) is a dummy that indicates whether an individual dropped out before the start of Session 2. In Columns (3) and (4) the dependent variable is a dummy defined contingent on starting Session 2 and indicates whether an individual dropped out after the start of Session 2. The explanatory variables are the treatment dummy HighScope, a dummy for negative news (Neg. News) and their interaction as well as the posterior p_2 and the prior p_1. The sample includes participants who satisfy the three exclusion criteria and for whom the respective variables are defined. Note that the sample sizes in Columns (1) and (2) (as well as (3) and (4)) differ slightly. This is because a few people dropped out between the start of Session 1(2) and the belief elicitation. Therefore, for these individuals, the treatment variables are available but not the beliefs. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

sessions are, on average, more patient and more willing to delay gratification than those who dropped out. The next subsection presents additional robustness tests based on less strict exclusion criteria, which yield similar results. This suggests that the observed differences in patience and willingness to delay gratification do not drive our main findings.

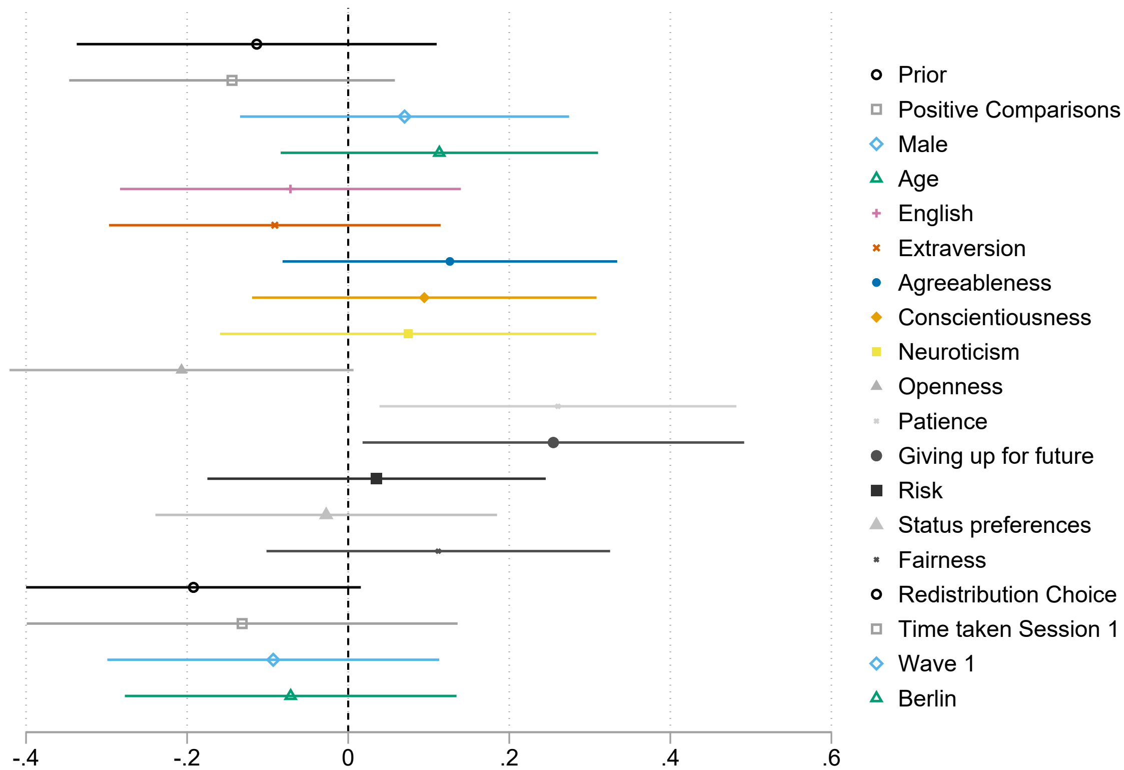


Figure D.2: Sample selection

Notes: The graph reports coefficient estimates and 95-percent confidence intervals from individual OLS regressions. The independent variable in each regression is a dummy variable that indicates whether a participant dropped out of the experiment. The dependent variables are different key variables elicited in Session 1. These include the prior (p_1), the number of positive comparisons (signal), a dummy for male participants, age, self-assessed level of English, the Big Five personality traits, our two patience measures, risk preferences, status preferences, the fairness measure and redistribution decision, minutes needed to complete Session 1, wave and subject pool fixed effects. Note that, while in all analyses above, we used a patience measure that combined results from money now vs. later decisions with the willingness to give up something today to benefit from it in the future (akin to the approach for *Time discounting* by Falk *et al.*, 2023), in this figure, we show coefficients for both measures separately. In each regression, the sample includes all participants who satisfy the three exclusion criteria and for whom the respective variables were elicited.

D.3 Robustness to more lenient exclusion criteria

Our main analysis focuses on participants who completed all three sessions ($N = 367$) to ensure that the variation in beliefs we study is not confounded by participants anticipating that they would not finish the experiment and thus reporting biased beliefs. As our main outcome variables are elicited in Session 2, we also investigate whether our results are robust to relaxing the restriction of having completed all three sessions. Specifically, we include participants who dropped out before the end of the experiment but only after the elicitation of the main variables, that is, those who at least stated their posterior belief p_2 and committed to a number of tasks to be solved in Session 2. This yields a sample of 416 observations for belief-related outcomes and 414 for work-related outcomes (see Columns (2) and (5) in Tables D.2 and D.3). In a second step, we provide an additional robustness test by further relaxing the restrictions and including individuals who failed the exclusion criteria specified in Section D.1. This results in 455 observations for belief outcomes and 453 for work decisions (see Columns (3) and (6) in Tables D.2 and D.3). Our results remain highly robust when using these larger samples, with effect sizes and significance levels largely unchanged.

Table D.2: Beliefs: Relaxing exclusion criteria

	Posterior			Optimism		
	(1)	(2)	(3)	(4)	(5)	(6)
HighScope	-1.173	-1.545	-2.116	-5.709	-6.787	-6.736
	(2.71)	(2.64)	(2.61)	(4.61)	(4.30)	(4.11)
Neg.News	-16.403***	-14.962***	-14.428***	42.295***	42.696***	45.296***
	(3.54)	(3.33)	(3.16)	(4.58)	(4.37)	(4.16)
HighScope*Neg.News	11.165***	10.902***	10.967***	21.530***	20.754***	19.318***
	(4.31)	(4.12)	(4.00)	(6.39)	(6.09)	(5.84)
Constant	57.758***	57.087***	57.211***	-20.501***	-19.910***	-20.962***
	(2.05)	(2.01)	(1.91)	(3.05)	(2.91)	(2.76)
N	367	416	455	367	416	455

Notes: The table shows results from OLS regressions. The dependent variables are participants' posterior (p_2) in Columns (1) to (3) and optimism ($p_2 - p_{\mathcal{B}}$) in Columns (4) to (6). The explanatory variables are the treatment dummy *HighScope*, a negative news dummy and the interaction. Columns (1) and (4) restrict the sample to those who completed all three sessions of the experiment and did not fulfill any of the other exclusion criteria preregistered (as in our main analysis). Columns (2) and (5) restrict the sample according to our original preregistration including participants who may not have completed the experiment. Columns (3) and (6) report results for everyone for whom the outcome variables are available. Note that the differences in the sample size between (3) and (2), as well as (6) and (5) is slightly lower than the number of participants who failed at least one of the three exclusion criteria. This is because some people who failed the exclusion criteria dropped out of the experiment before the elicitation of posterior beliefs and the work decision. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D.3: Work decision: Relaxing exclusion criteria

	Number Tasks			P(40 tasks)		
	(1)	(2)	(3)	(4)	(5)	(6)
Posterior (instrumented)	-0.232**	-0.233**	-0.233**	-0.008*	-0.007*	-0.007*
	(0.10)	(0.11)	(0.11)	(0.00)	(0.00)	(0.00)
Constant	41.650***	41.238***	41.321***	0.836***	0.815***	0.803***
	(5.36)	(5.57)	(5.49)	(0.21)	(0.22)	(0.22)
N	367	414	453	367	414	453

Notes: The table shows results from IV regressions using the general method of moments estimator. The posterior belief (p_2) about facing low workload is instrumented with the treatment dummy for HighScope, a dummy for negative news and the interaction of both (for the first stage, see Table 1). The dependent variable in Columns (1) to (3) is the number of tasks participants complete in Session 2. The dependent variable in Columns (4) to (6) is the probability to solve the maximum number of tasks (40) in Session 2. Columns (1) and (4) restrict the sample to those who completed all three sessions of the experiment and did not fulfill any of the other exclusion criteria preregistered (as in our main analysis). Columns (2) and (5) restrict the sample according to our original preregistration including participants who may not have completed the experiment. Columns (3) and (6) report results for everyone for whom the outcome variables are available. Note that the differences in the sample size between (3) and (2), as well as (6) and (5) is slightly lower than the number of participants who failed at least one of the three exclusion criteria. This is because some people who failed the exclusion criteria dropped out of the experiment before the elicitation of posterior beliefs and the work decision. *** p<0.01, ** p<0.05, * p<0.1.

E Observer experiment

To strengthen the interpretation that belief-driven procrastination in our main experiment is driven by motivated reasoning, we conducted an additional study – the Observer experiment. This experiment isolates the role of motivational incentives in shaping memory by holding the informational environment constant while removing personal stakes. Specifically, in the additional experiment, participants observed the same signals regarding the total effort required to complete the transcription task but were not required to provide this effort. Thus, we can test whether in the absence of a motive for optimism observers still exhibit asymmetric (negative) news suppression.

Design The Observer experiment mimics the structure of the main experiment (henceforth, DM experiment) and consists of three sessions two weeks apart. Each observer is matched to a DM from the DM experiment (referred to as: Participant A) and must form expectations about Participant A's assigned workload based on the signal Participant A had received. That is, in Session 1, observers see the signal their matched Participant A received. In Session 2, observers are either reminded (OLowScope) or not reminded (OHighScope) of the signal they saw two weeks earlier and have to state their subjective probability that Participant A's had to solve at most 40 additional tasks (p_2). Akin to the DM experiment, we measure observers' expectations in Session 2 and 3, with our main outcome variable being the observers' posterior belief in Session 2.[11]

[11] For exploratory reasons, we also elicit observers' second-order beliefs, normative views, expectations regarding Participant A's work allocation decisions, and (hypothetical) decisions on behalf of Participant A. Note that we elicit the main outcome variable (posterior beliefs) before the additional outcome variables and without the knowledge thereof. Therefore, the elicitation of these additional variables could not have affected the main outcome variable of interest.

Procedures We collected data with new participants from the same subject pools as for the main experiment. We collected data in January-February 2025. The preregistration can be found at `https://aspredicted.org/ktdb-n4xq.pdf`. After exclusion based on preregistered criteria, the final sample consists of 359 observers. The median participant took 57 minutes to complete all three sessions and on average, participants earned 18.07€.

Results We test our main hypothesis that the asymmetric effect of the reminder treatment, which we observed in the DM experiment, does not arise in the OBSERVER experiment.

First, we find no significant effect of the scope manipulation on posterior beliefs (two-sided t-test, $p = 0.872$) or on optimism ($p = 0.520$). This finding contrasts with Result 1 from the original experiment that scope for motivated reasoning led to optimism about (own) effort costs. Next, we examine treatment effects separately for positive and negative news. The OHIGHSCOPE condition does not significantly affect posteriors after either negative ($p = 0.341$) or positive news ($p = 0.310$), relative to OLOWSCOPE. We also find no asymmetric treatment effect on optimism, defined as the difference between subjective and Bayesian posteriors. While OHIGHSCOPE slightly increases optimism after negative news ($p = 0.108$), it simultaneously increases pessimism after positive news ($p = 0.028$). This symmetric pattern does not reflect motivated reasoning but is instead consistent with a general tendency to revert to priors when relying on memory. Again, this result contrasts with the asymmetric effect of the scope variation on beliefs found in the original experiment.

Taken together, the results from the additional observer experiment support the interpretation of Result 2 in the main text as reflecting motivated memory.

F Instructions

This section contains screenshots of the crucial parts in the experimental design. Complete instructions are available upon request.

F.1 Screenshots Session 1

Your task: Translate a sequence of numbers into a sequence of letters:

During the experiment you will have to solve a transcription task. The task works as follows:

You have to transcribe sequences of 6 numbers into sequences of 6 letters. To do so, an input field is displayed below the sequence of numbers. Here is an example:

12 16 14 16 16 1

You will transcribe the sequence of numbers with the help of a coding key, that assigns a specific letter to each number (see the example below):

Number:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Letter:	M	Y	A	L	G	Z	E	S	K	O	T	H	X	F	I	C	D	P	R	Q	V	W	U	N	J	B

Your task is to find the corresponding letter for each number and enter the resulting sequence of letters in the input field. Do not enter any spaces between the letters in the input field. Note also, the input field is not case-sensitive. That is, it does not matter whether you enter for example “k” or “K”.

In the above example you are seeing the sequence **"12 16 14 16 16 1".** Given this coding key, the solution is the letter sequence **"HCFCCM".** For this example, we have entered this solution for you in the input field.

12 16 14 16 16 1

HCFCCM

Once you submit a correct code, the computer will prompt you with another sequence.

In case you submit an incorrect code, you will be notified by the computer and have to redo the sequence. To complete the task, you have to transcribe a certain number of sequences. From sequence to sequence, both the number sequence and the coding key change.

You will now have to solve 10 such sequences for practise.

Next

Figure F.1: Explanation of the transcription task

Your current task:

Until now you have correctly transcribed **0 sequences**. This means that 10 sequences are still outstanding.

For each number, enter the appropriate letter from the code table (without spaces).

13 16 5 1 18 19

Number:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Letter:	L	R	T	M	D	S	N	X	G	A	I	K	C	H	Z	U	J	F	V	Y	E	W	O	B	P	Q

Next

Figure F.2: Example of the transcription task

Your task in 28 days:

You have now completed today's transcription task. Note that we randomly selected 9 other participants from this study who also solved a transcription task consisting of 10 sequences (just as you did). Together with these 9 participants you now form a group of 10 participants.

In 28 days, each participant in your group will have to solve another transcription task consisting of a unique number of sequences. The number of sequences varies from participant to participant.

Who has to solve how many sequences will be determined in the following way:

Each participant will have to solve 40 sequences correctly, plus a unique number of additional sequences.

There are 10 possibilities for the unique number of additional sequences a group member must solve in 28 days.

Each group member will be randomly assigned to solve either 8, 16, 24, 32, 40, 48, 56, 64, 72, or 80 additional sequences correctly.

Importantly, each possibility is only assigned once in your group, that is, no two group members will have to solve the same number of additional sequences in 28 days. For example, if one member of the group is randomly assigned to solve 72 transcription tasks in total (=40+32), no other member in your group will be assigned to solve 72 sequences tasks in total, and each possible number of sequences to be solved is equally likely to be assigned to you.

To make sure you understand the procedure, please answer two short comprehension questions.

Imagine one person out of your groups has to solve 40 additional sequences. Is it possible that you also have to solve 40 additional sequences?

What is more likely: That a participant is assigned to solve 8 additional sequences or that a participant is assigned to solve 80 additional sequences?

Show further details

Next

Figure F.3: Explanations of the work load assignment

Notes: By clicking the button "Show further details" participants could see a paragraph that explained the random draw using a pictorial description.

Guess how many additional sequences you will have to solve:

Remember: Each group member was randomly assigned a unique number of additional sequences to be solved and no group member has been assigned the same number of sequences. The possible numbers of additional sequences are 8, 16, 24, 32, 40, 48, 56, 64, 72 and 80.

Right now, we would like to know more about your own expectations about the number of additional sequences that you have to solve, and we offer a bonus payment for guessing the number correctly:

As explained before, we will ask you to make several guesses in this study. The computer will randomly choose one of these guesses you make to be payoff relevant. This is the first such guess.

For each guess you make, you maximize the chances of winning the additional 6€, if you simply state your true expectation. (The button below provides further details on the exact procedure.)

Show/Hide Details

For your guess, we will always ask you how likely you think a particular state is. For example, we may ask how likely you think it is (in percent) that you have to solve 40 or fewer additional sequences. Thus, when you make a guess, you can pick any number between 0 and 100 and the number you select indicates the chance (out of 100) you assign to the state we ask for (for example, the chance that you have to solve 40 or fewer additional sequences).

For your payoff-relevant guess, the computer will then randomly draw two numbers between 0 and 100 (including 0, 100 and all decimal numbers between) and all numbers are equally likely to be selected. The two random numbers are drawn independently. That is, irrespective of the first number drawn, the second number is again randomly chosen from the interval of 0 to 100 such that the outcome of the first draw in no way affects the outcome of the second number drawn.

You will receive the 6€ bonus if ...

- ... you actually have to solve 40 or fewer additional sequences and the number picked by you is larger than either of the two draws
- ... you actually have to solve 48 or more additional sequences and the number picked by you is smaller than either of the two draws

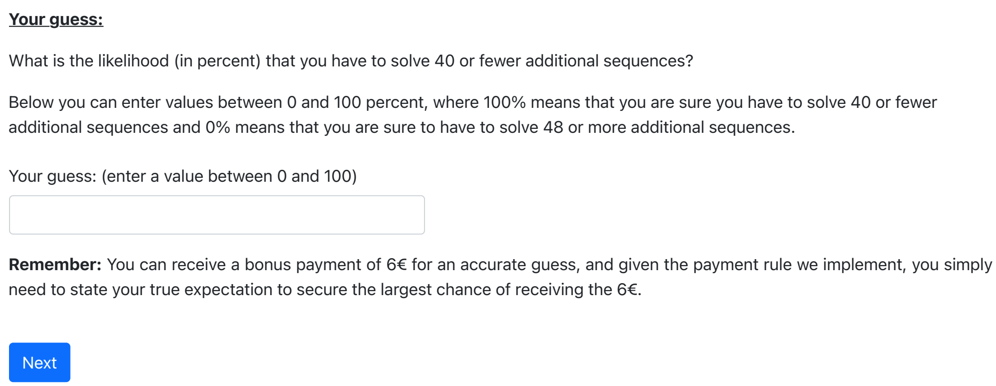
Given this rule, you can maximize the probability of winning by simply stating your true expectation.

On the next page you can enter your guess.

Next

Figure F.4: Belief instructions with details

Notes: The framed part was only shown after requesting further details.



Your guess:

What is the likelihood (in percent) that you have to solve 40 or fewer additional sequences?

Below you can enter values between 0 and 100 percent, where 100% means that you are sure you have to solve 40 or fewer additional sequences and 0% means that you are sure to have to solve 48 or more additional sequences.

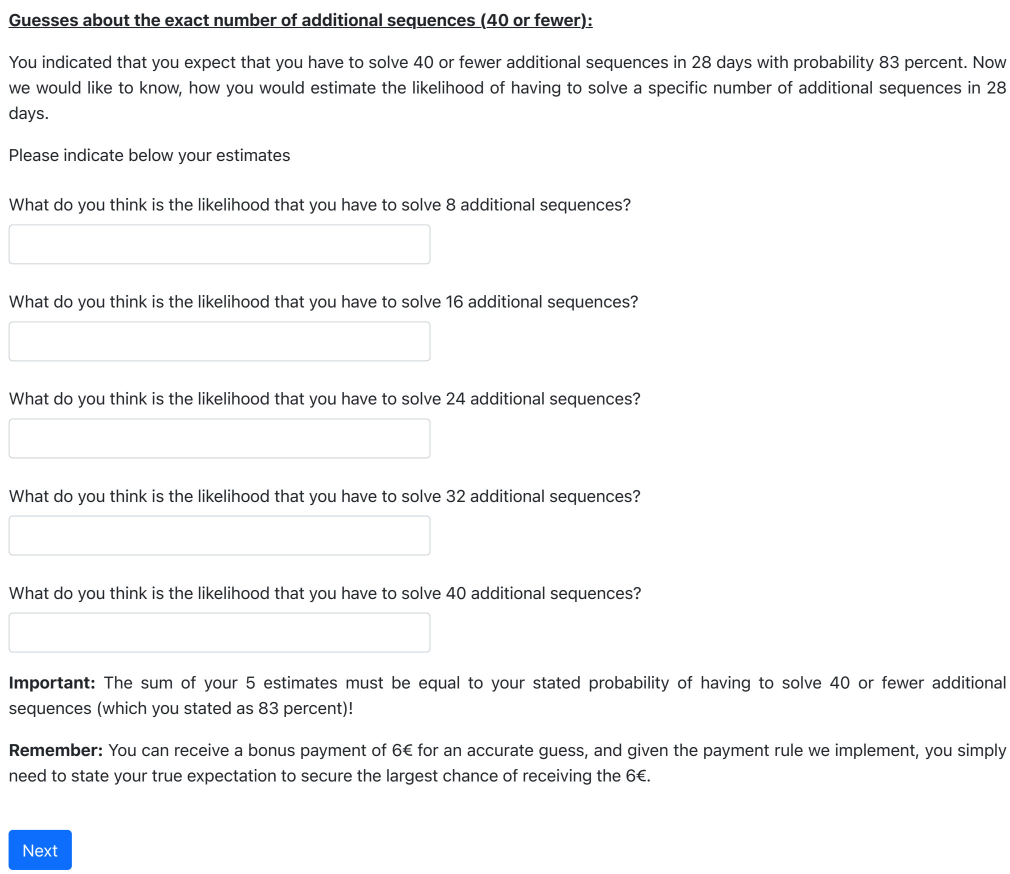
Your guess: (enter a value between 0 and 100)

Remember: You can receive a bonus payment of 6€ for an accurate guess, and given the payment rule we implement, you simply need to state your true expectation to secure the largest chance of receiving the 6€.

Next

Figure F.5: Belief elicitation of p_1

Notes: The elicitation of the posterior beliefs p_2 in Session 2 looked identical.



Guesses about the exact number of additional sequences (40 or fewer):

You indicated that you expect that you have to solve 40 or fewer additional sequences in 28 days with probability 83 percent. Now we would like to know, how you would estimate the likelihood of having to solve a specific number of additional sequences in 28 days.

Please indicate below your estimates

What do you think is the likelihood that you have to solve 8 additional sequences?

What do you think is the likelihood that you have to solve 16 additional sequences?

What do you think is the likelihood that you have to solve 24 additional sequences?

What do you think is the likelihood that you have to solve 32 additional sequences?

What do you think is the likelihood that you have to solve 40 additional sequences?

Important: The sum of your 5 estimates must be equal to your stated probability of having to solve 40 or fewer additional sequences (which you stated as 83 percent)!

Remember: You can receive a bonus payment of 6€ for an accurate guess, and given the payment rule we implement, you simply need to state your true expectation to secure the largest chance of receiving the 6€.

Next

Figure F.6: Belief elicitation of probabilistic beliefs (part 1/2)

Remember: Each group member was randomly assigned to solve either 8, 16, 24, 32, 40, 48, 56, 64, 72 or 80 additional tasks in 28 days and no group member will have to solve the same number of additional tasks.

We will now provide you with some information that could be helpful for you in order to better estimate whether you have to solve many or few additional tasks.

We randomly selected 3 out of the 9 other participants from your group. We will now inform you, whether each of these 3 participants must solve more or fewer additional tasks than you in 28 days.

Of the 3 randomly selected participants from your group...

Number of participants that need to solve **fewer** additional tasks: 0

Number of participants that need to solve **more** additional tasks: 3

Next

Figure F.7: Feedback provision

F.2 Screenshots Session 2

Reminder:

In the first session, you had to transcript 10 sequences of numbers into 10 sequences of letters. We had then randomly selected 9 other study participants with whom you formed a group of 10 participants. In 14 days, you have to correctly solve 40 sequences plus a number of additional sequences, and each group member was randomly assigned a unique number of additional sequences (either 8, 16, 24, 32, 40, 48, 56, 64, 72 or 80 additional sequences).

Next

Figure F.8: General reminder (both treatments)

Reminder:

Remember, we have also compared you with three other randomly chosen participants from your group and told you whether they have to solve more or fewer additional tasks than you.

Of the 3 randomly selected participants from your group...

Number of participants that need to solve **fewer** additional tasks: 0

Number of participants that need to solve **more** additional tasks: 3

Next

Figure F.9: Reminder (only LowScope)

Your work today

In the first part of this experiment, you have already tried out some transcription sequences. In the third part of the study you will have to complete a transcription task consisting of a certain number of these sequences: 40 plus the additional sequences that have been assigned to you.

As the number you have to solve in the end might be high, you can now choose to already complete some of the sequences today. The maximum number of sequences you can already solve today is 40.

Here, please type in the number of sequences you want to solve **today**:

Click "next" if you are ready to start.

Click here if you want to see the explanation of the task again

Next

Figure F.10: Work choice

F.3 Screenshots Session 3

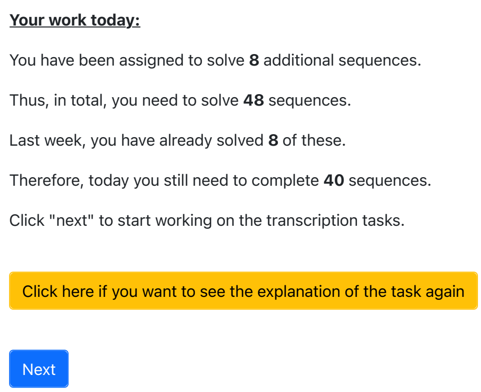


Figure F.11: Work load resolution

F.4 Screenshots from observer experiment

PARTICIPANT A's task:

You have just completed the same practice of the transcription task (10 sequences) that all participants in the past experiment completed at the beginning of their experiment.

In the following, we will ask you questions about one participant from this past experiment, let us call this participant PARTICIPANT A. PARTICIPANT A and all other participants from the past study were students recruited from the same subject pool and in the same way as you. PARTICIPANT A was grouped with 9 randomly selected other participants from the study to form a group of 10 participants. Let us call this group GROUP A.

Each participant in GROUP A was informed that they had to solve a transcription task consisting of a unique number of sequences 28 days from their session 1. The number of sequences varied from participant to participant.

Who had to solve how many sequences was described to the participants as follows (screenshot from the previous experiment):

> **Each participant will have to solve 40 sequences correctly, plus a unique number of additional sequences.**
>
> There are 10 possibilities for the unique number of additional sequences a group member must solve in 28 days.
>
> Each group member will be randomly assigned to solve either 8, 16, 24, 32, 40, 48, 56, 64, 72, or 80 additional sequences correctly.
>
> Importantly, each possibility is only assigned once in your group, that is, no two group members will have to solve the same number of additional sequences in 28 days. For example, if one member of the group is randomly assigned to solve 72 transcription tasks in total (=40+32), no other member in your group will be assigned to solve 72 sequences tasks in total, and each possible number of sequences to be solved is equally likely to be assigned to you.

To make sure you understand the procedure of the past experiment, please answer two short comprehension questions.

Note: In the past experiment, PARTICIPANT A had to answer two very similar comprehension questions to proceed.

Imagine one person in GROUP A (other than PARTICIPANT A) had to solve 40 additional sequences. Is it possible that PARTICIPANT A was also assigned 40 additional sequences to solve?

What is more likely: That a participant was assigned to solve 8 additional sequences or that a participant was assigned to solve 80 additional sequences?

Show further details

Next

Figure F.12: Session 1: Main instructions

Your guess:

What is the likelihood (in percent) that PARTICIPANT A had to solve 40 or fewer additional sequences?

Below you can enter values between 0 and 100 percent, where 100% means that you are sure PARTICIPANT A had to solve 40 or fewer additional sequences and 0% means that you are sure PARTICIPANT A had to solve 48 or more additional sequences.

Your guess: (enter a value between 0 and 100)

Remember: You can receive a bonus payment of 6€ for an accurate guess, and given the payment rule we implement, you simply need to state your true expectation to secure the largest chance of receiving the 6€.

Next

Figure F.13: Session 1: Feedback provision

Remember: Each member of group A was randomly assigned either 8, 16, 24, 32, 40, 48, 56, 64, 72 or 80 additional tasks to be solved in 28 days and no member of GROUP A had to solve the same number of additional tasks.

We will now provide you with some information that could be helpful for you in order to better estimate whether PARTICIPANT A had to solve many or few additional tasks.

We randomly selected 3 out of the 9 other participants from GROUP A. We will now inform you, whether each of these 3 participants was assigned more or fewer additional tasks to be solved in 28 days than PARTICIPANT A.

Of the 3 randomly selected participants from GROUP A...

Number of participants that had to solve **fewer** additional tasks: 1

Number of participants that had to solve **more** additional tasks: 2

Next

Figure F.14: All sessions: Belief elicitation

References

Balietti, A., Budjan, A., Eymess, T. and Soldà, A. (forthcoming). Perceived control and attendance to information: Evidence on the health risks of air pollution. *The Economic Journal*.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, **117** (3), 871–915.

Brunnermeier, M. K., Papakonstantinou, F. and Parker, J. A. (2017). Optimal time-inconsistent beliefs: Misplanning, procrastination, and commitment. *Management Science*, **63** (5), 1318–1340.

Conlon, J. J. (2024). Memory rehearsal and belief biases. *Mimeo*.

Cordes, C., Friedrichsen, J. and Schudy, S. (2025). Motivated beliefs under delayed uncertainty resolution. *CESifo Working Paper No. 12286*.

Falk, A., Becker, A., Dohmen, T., Huffman, D. and Sunde, U. (2023). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science*, **69** (4), 1935–1950.

— and Zimmermann, F. (2018). Information processing and commitment. *The Economic Journal*, **128** (613), 1983–2002.

Hagenbach, J. and Koessler, F. (2022). Selective memory of a psychological agent. *European Economic Review*, **142**, 104012.

Heidhues, P., Köszegi, B. and Strack, P. (2023). Misinterpreting yourself. *Available at SSRN 4325160*.

Kerwin, J., Rostom, N. and Sterck, O. (2024). Striking the right balance: Why standard balance tests over-reject the null, and how to fix it. *IZA Discussion Papers No. 17217*.

Steel, P. (2010). Arousal, avoidant and decisional procrastinators: Do they exist? *Personality and Individual Differences*, **48** (8), 926–934.