A. Appendix - For Online Publication

A.1. Choice-Based Sampling

Innovation scholars often seek to predict rare events—outcomes that remain unrealized most of the time (*zeroes*) and are realized only rarely (*ones*). For example, Singh and Marx (2013) predict whether a focal patent is cited by another patent from the universe of all patents, while Sorenson and Stuart (2001) examine whether venture capitalist firms invest in a focal company. The *rare* event we are concerned with is whether or not evaluators generate ideas on a given day, with many more days on which they do *not* generate ideas than days on which they do.

Applying logistic and Poisson regression to such rare events data risks underestimating an event's likelihood (Imbens, 1992; King and Zeng, 2001a; King and Zeng, 2001b). Choice-based sampling helps to overcome this problem: We sample a fraction α of ones and a fraction β of zeroes, with α being much larger than β, and include weights to account for the oversampling of ones. These weights denote the number of nonsampled zeroes represented by the sampled zeroes.

Let N_i^{one} be the number of days on which evaluator i generates ideas, and N_i^{zero} the number of days on which evaluator i does not generate any ideas, with $N_i^{one} * 2 \leq N_i^{zero}$. For every evaluator i, we sample all days N_i^{one} on which they generate ideas, and $N_i^{one} * 2$ days on which they do not generate any ideas. Thus, $\alpha_i = \dfrac{N_i^{one}}{N_i^{one}} = 1$, and $\beta_i = \dfrac{N_i^{one} * 2}{N_i^{zero}}$. Correspondingly, evaluator-specific weights are the inverse of α_i and β_i: $w_i^{\alpha} = 1$ and $w_i^{\beta} = \dfrac{N_i^{zero}}{N_i^{one} * 2}$.

A.2. Treatment Effect Heterogeneity

A growing literature is showing that TWFE estimates may be biased even in the absence of pre-trends (Borusyak et al., 2024; Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021). TWFE estimates are the weighted sum of the treatment effects of each unit and period. If treatment effects vary across units and periods, some weights can be negative, potentially leading to TWFE estimates that have the opposite sign of the true treatment effect. These problematic negative weights emerge from "forbidden comparisons" (Goodman-Bacon, 2021), where units that have already received treatment are incorrectly used as controls for units currently undergoing treatment.

We address concerns about treatment effect heterogeneity in two ways. First, following de Chaisemartin and D'Haultfœuille (2024), we assess the extent to which our TWFE estimates are influenced by negative weights[2]. We find that only a very small fraction of the weights—0.001%—are negative, which reassures us that negative weights have minimal impact on our estimates.

Second, we apply the heterogeneity-robust estimator proposed by de Chaisemartin and D'Haultfœuille (2024). Unlike earlier estimators, such as those by Callaway and Sant'Anna (2021) and Sun and Abraham (2021), this estimator accommodates non-absorbing treatments (i.e., treatments that are turned on and turned off again) as well as dynamic treatment effects. Thus, it is well suited for addressing potential treatment effect heterogeneity in our empirical design. The intuition underlying the estimator is to compare the outcome trajectory of units (in this case, evaluators) that start off with the same treatment in their same "period one," with some later changing their treatment (*switchers*) and others not (yet) changing their treatment (*nonswitchers*).

The estimator requires a sufficiently large cohort of evaluators who share the same "period one" and are observed over the same subsequent periods. Our original sample does not meet

[2] We compute negative weights using the R package *twowayfeweights* developed by de Chaisemartin and D'Haultfœuille (2024) and provided on their GitHub repository here.

this criterion, as the evaluators enter and exit the sample on different days and remain in the sample for varying durations. Thus, we construct a sample of evaluators who are present in the sample on January 1, 2017, designating this day as a common "period one." Evaluators enter the sample with as many days as they are in the panel following January 1, 2017, yielding a sample of 676,011 evaluator-day observations.

Based on this sample, we re-estimate specifications 1 and 2 using both the non-robust TWFE estimator from above as well as the robust TWFE estimator proposed by de Chaisemartin and D'Haultfœuille (2024). Figure A.5 and Table A.10 show that the robust TWFE estimator produces even larger coefficient estimates, which reduces concerns around treatment effect heterogeneity and suggests that our original estimates even underestimate the full magnitude of idea crowding-in.

A.3. Robustness Checks

In the following, we perform several empirical tests to verify the robustness of our findings.

A.3.1. Instrumental Variable Strategy

A potential concern with our empirical strategy is that evaluators have some discretion over when they complete their assigned evaluation tasks. As a result, unobserved variables at the evaluator-day level (e.g., workload, illness) may simultaneously influence both the likelihood of evaluating and generating ideas, confounding the effect of idea evaluation on idea generation. For example, during periods of low workload, evaluators may have more time for both evaluating and generating ideas. In this case, the observed post-evaluation increase in idea generation may reflect lower workload rather than the impact of the evaluation process itself.

We address endogeneity concerns by employing an instrumental variable strategy. We instrument our treatment, *Post Evaluation*, with the indicator variable *Reminder* for whether evaluators received e-mails about open evaluation tasks on any given day. They receive a first e-mail on the day they are assigned a new evaluation task, with follow-up e-mails sent at first every three weeks and then every four weeks if the task remains incomplete.

For *Reminder* to be a valid instrument, it must satisfy both the relevance condition and the exclusion restriction. The relevance condition requires that *Reminder* be correlated with *Post Evaluation*. Extensive research has demonstrated that reminders and deadlines effectively motivate individuals to complete tasks (Altmann et al., 2021; Balasubramanian et al., 2018; Cohen et al., 2021). Consistent with this, we find that *Reminder* has a strong, positive, and statistically significant effect on the evaluators' likelihood of evaluating ideas. They are 317.7% more likely to evaluate ideas on days when they receive reminders compared to days without reminders (column 1 of Table A.25). The strength of *Reminder* as an instrument is further supported by a Kleibergen-Paap F-statistic well above the threshold of 10.

The exclusion restriction requires that *Reminder* influences *Idea Generation* solely by increasing evaluators' likelihood of evaluation, without having a direct effect on *Idea Generation*. A potential concern is that *Reminder* may independently stimulate evaluators' creativity, increasing idea generation even when no evaluation occurs. To address this, we conduct placebo tests in the spirit of, for example, Choudhury et al. (2024). The logic behind our placebo test is as follows: If *Reminder* boosted *Idea Generation* by stimulating creativity, we would expect a positive and statistically significant correlation between *Reminder* and *Idea Generation* even when evaluators do not act on the reminder and fail to evaluate ideas. However, if *Reminder* influenced *Idea Generation* only by prompting evaluators to evaluate ideas, the positive and statistically significant correlation should hold only when evaluators respond to the reminder by evaluating ideas. We report the results from this placebo test in Table A.26. We find that our instrument, *Reminder*, has a positive and statistically significant effect on *Idea Generation* if and only if evaluators act on the reminder and evaluate ideas, suggesting that *Reminder* satisfies the exclusion restriction.

We present the results from the instrumented specification in column 4 of Table A.25. The 2SLS estimates predict a 182.7% increase in evaluators' likelihood of generating ideas within the week following an idea evaluation, compared to outside this post-evaluation period. This effect is even larger than that predicted by the original OLS estimates (column 3 of Table A.25), suggesting that our initial analysis may have underestimated the full extent of idea crowding-in.

A.3.2. Different Samples

We established our main findings using a choice-based sample of evaluators for the reasons outlined in section 4. Here, we demonstrate that our findings can be reproduced with alternative samples. First, we re-estimate our event study specification using the full sample of evaluator-day observations from 2015 onward. (For computational reasons, we could not use the full sample of 26,425,003 evaluator-day observations between 2004 and 2018.) Figure A.12 plots the coefficient estimates, which are very close to those obtained in the previous analyses. Second, we newly include non-evaluators in addition to evaluators. Again, Figure A.13 shows coefficient estimates very close to those from the main analysis.

A.3.3. Placebo Test

We conduct a placebo test by randomly shifting the de facto evaluation dates into the future or the past. If these fake evaluation dates produced results similar to our main findings, it would suggest that the observed patterns reflect a spurious correlation between idea evaluation and generation, rather than a causal effect. However, if the fake evaluation dates failed to replicate the original results, it would reinforce the robustness of our findings and would support a causal interpretation of the effect of idea evaluation on idea generation. Table A.27 shows that *Post Evaluation* has no effect on *Idea Generation* when based on fake evaluation dates, further supporting the validity of our results.

A.3.4. Business Days vs. Weekdays

Our main analysis predicts evaluators' idea generation based on the number of weekdays—rather than business days—since idea evaluation. We focused on weekdays because the firm operates its plant 24/7, with employees working in shifts, meaning that ideas are also generated and evaluated on weekends, though on a smaller scale (Figure A.14). In an alternative estimation, we predict evaluators' idea generation based on the number of *business days* since idea evaluation. The results, plotted in Figure A.15, are nearly identical to those in Figure A.3, indicating that it makes little difference whether we consider weekdays or business days.

A.3.5. Matching

The evaluators may be assigned evaluation tasks based on their expected ideation productivity. Idea crowding-in would then reflect evaluators living up to this expectation rather than a causal effect of idea evaluation on idea generation. To address this concern, we apply CEM (Iacus et al., 2012; Iacus et al., 2019). We match treated units (i.e., the evaluators who evaluate ideas on day t) with control units (i.e., the evaluators who do not evaluate ideas on day t) based on observable covariates. Under the assumption that these covariates are predictive of the evaluators' expected ideation productivity, any differences in their ideation productivity after t can then be attributed to the evaluation event.

We seek to match an evaluator i who evaluates an idea on day t with another evaluator j who is almost identical to evaluator i regarding pre-evaluation characteristics except that they do not evaluate an idea on day t. The treated and the control evaluators are matched based on the following covariates: *# Prev. Generated Ideas*, *Cumulative Idea Value*, *# Prev. Evaluated Ideas*, *Hierarchy*, *Prev. Mobility*, *Permanent*, *Gender*. We coarsen these covariates and construct strata based on the coarsened covariates. We assign evaluators to the strata and discard strata for which we do not find at least one treated and one control evaluator. Every evaluator-day observation is matched to just one other evaluator-day observation (one-to-one matching). An evaluator-day observation must not be a control for more than one evaluator-day observation from the treatment group. If there are multiple control candidates, we randomly draw one.

We find matches for 56,180 evaluator-day observations, yielding a matched sample of 112,360 observations (56,180 treated and 56,180 control). Table A.28 shows that matched and unmatched evaluator characteristics are very well balanced between the treatment and the control groups, suggesting that our matching is effective.

We re-estimate specifications 1 and 2 using the matched sample. Table A.29 shows that evaluators have a 98.4% greater likelihood of generating ideas on the day of idea evaluation, and 63.2%, 55.4%, and 82.1% higher likelihoods of generating ideas one, two, and three days after evaluation. Similarly, Table A.30 indicates a 90.6% greater likelihood of idea generation within one week after evaluation. These estimates fall within the range of the original estimates, supporting the robustness of our findings.

A.4. Boundary Condition: Is There Too Much Idea Evaluation?

In the main analysis, we explored how idea crowding-in depends on *daily* evaluation workloads. Here, we replicate this analysis for *weekly* evaluation workloads. Based on Figure A.6, we distinguish between *Post Evaluation Low*$_{it}$—which is 1 if evaluator i evaluated at least one but fewer than six ideas within the week preceding day t, and 0 otherwise—and *Post Evaluation High*$_{it}$—which is 1 if evaluator i evaluated six or more ideas within the week preceding day t, and 0 otherwise. Table A.11 shows that, while a 1-SD increase in *Post Evaluation Low*$_{it}$ raises evaluators' likelihood of idea generation by 14.7%, the same increase in *Post Evaluation High*$_{it}$ raises it by only 3.6%. This supports our earlier finding (section 5.3) that low evaluation workloads are more effective at eliciting evaluator ideas than high workloads.

These findings cannot be explained with high evaluation workloads delaying instead of reducing idea crowding-in. High workloads may stimulate idea generation just as much as low workloads, but because evaluators prioritize completing their evaluation tasks, they delay idea submission. If this were true, we would expect crowding-in from high workloads to surpass that from low workloads over time, as evaluators eventually catch up on idea generation. However, this does not align with the coefficient estimates shown in Figure A.7, which remain consistently smaller for high workloads, even with greater temporal distance from evaluation events.

A.5. Assignment of Topics to Ideas Using Natural Language Processing

To investigate the nature of knowledge recombination (sectin 6.1) and measure the novelty of ideas (section 7.1), it is crucial that we identify the topic set to which ideas relate. Although we have access to a set of pre-specified idea topics from the firm, these only cover the period after 2015 and are fairly broad. To achieve the nuance required for our analysis, we instead infer idea topics by leveraging recent advances in NLP. The steps are as follows:

1. **Total Set of M Idea Topics**. We identify the entire set of M topics covered by the universe of ideas. First, we extract all the unique words that appear in the title of any idea. We stem these words and remove stopwords (e.g., and, or), which yields a dictionary of words. Next, we extract the word embedding of each word in this dictionary using the Python package *SentenceTransformer* (Reimers and Gurevych, 2019), one of the most advanced and widely used tools for extracting word embeddings. A word embedding

represents a word in a continuous vector space, where words with similar meanings are mapped to nearby points. This technique captures semantic relationships between words, mapping "hammer" and "wrench" to nearby points in the vector space.

We rely on these embeddings to identify a total of M topic clusters using a KMeans algorithm. This algorithm assigns words with nearby points in the vector space to the same topic. For example, the words "hammer" and "wrench" are assigned to the topic "tool." Once each word is assigned to a topic, we calculate the centroid of each topic cluster, which is the average over the vector embeddings of the words in that topic cluster. For example, if "hammer" has the word embedding [1, 2, 3] and "wrench" has the word embedding [2, 2, 2], and "hammer" and "wrench" are the only words assigned to the topic "tool," then the topic centroid is [1.5, 2, 2.5].

The KMeans algorithm allows us to flexibly set M. Our choice of M is informed by Criscuolo et al. (2017), who use 574 keywords to span the topic space, and our values fall within a similar range. For the main analysis, we set $M = 1{,}000$. To ensure that our findings are not sensitive to this specific choice, we replicate the analyses related to knowledge recombination and idea novelty using $M = 500$. The results are very similar, as shown in Tables A.13, A.14, A.15, and A.21.

2. **Assigning N Topics to Each Idea**. Next, we assign to each idea the ($N < M$) topics that best describe it. To do this, we extract the embedding of each idea title, again using *SentenceTransformer* (Reimers and Gurevych, 2019). We then calculate the cosine similarity between the embedding of each idea title and the topic centroids, resulting in a vector of M cosine similarities. We assign the N topics with the highest cosine similarities to each idea.

   Notably, since we characterize ideas by the embedding of their titles rather than the embeddings of the individual words within the titles, we can more precisely locate ideas in the vector space. For example, the embedding of the full title "gadget with battery technology" provides a more accurate location in the vector space than the separate embeddings for the individual words "gadget," "battery," and "technology."

   Again, we have flexibility in our choice of N. In the main analysis, we set $N = 5$, assuming that ideas can be meaningfully represented by five topics. To test the robustness of our results, we replicate the analyses related to knowledge recombination and idea novelty using $N \in \{3, 7\}$. This range is informed by Boudreau et al. (2016), who assign an average of 12.42 keywords per idea. While their number is higher than ours, it is still within a comparable range. The results are very similar, as shown in Tables A.13, A.14, A.15, and A.21.

A.6. Similarity Between Evaluated and Generated Ideas

We measure *Idea Similarity* in three ways. While the first two measures rely on NLP models, the third uses idea topics pre-assigned by the firm. The first measure is the *Cosine Similarity* between ideas k and l, which we calculate based on the word embeddings of their titles (Aceves and Evans, 2024; Guzman and Li, 2023; Hoberg and Phillips, 2018). This measure is 1 for identical titles, and 0 for entirely disparate ones. For the second measure, *Share Same Topic—NLP-Assigned*, we assign topics to ideas using NLP and clustering models, and calculate the share of topics shared by the evaluated and the generated ideas. This measure ranges from 1 (when two ideas share all the topics) to 0 (when there is no topic overlap). Finally, the third measure, *Share Same Topic—Pre-Assigned*, again indicates the number of topics shared by the evaluated and the generated ideas, this time using the topics pre-specified by the firm. Since 2016, the firm requires its employees to indicate which one of seven topics their idea pertains to. A disadvantage of this measure is that, because the number of available topics is small, the similarity measure is necessarily coarse.

A.7. Problem-Based vs. Solution-Based Recombination

To better characterize knowledge recombination, we investigate whether the similarity between the evaluated and the generated ideas arises from evaluators building on the problems raised or the solutions proposed by the ideas they evaluate. We calculate the problem- and solution-based similarity between dyads of evaluated and generated ideas as follows:

1. **Problem and Solution Components of Ideas**: We seek to identify the parts of idea titles that relate specifically to problems vs. solutions by splitting titles along linking words commonly employed to distinguish between the two (e.g., for, addressing, etc.). For example, we split the title "Automated system for monitoring air quality" into the solution component "Automated system" and the problem component "monitoring air quality." Ideas for which problem and solution components cannot be clearly differentiated are excluded from this analysis.

2. **Word Embeddings for Problem and Solution Components**: We extract word embeddings for every solution and problem component using the Python package *SentenceTransformer* (Reimers and Gurevych, 2019).

3. **Calculating Similarities**: For every evaluator, we create all possible dyads of evaluated and generated ideas. For these dyads, we calculate their *solution-based similarity* and *problem-based similarity*. *Solution-based similarity* is the cosine similarity between the word embeddings of the solution components of evaluated and generated ideas, while *problem-based similarity* is the cosine similarity between the word embeddings of the problem components of the evaluated and the generated ideas.

A.8. Idea Evaluation and Idea Quality: Empirical Approach

In section 7, we examine how idea evaluation affects the quality (i.e., the novelty and the value) of subsequently generated ideas. Here, we provide more details on our empirical approach. The idea is to compare the quality of ideas from the same evaluator depending on whether they are generated shortly after evaluation or at other times. This comparison requires that evaluators generate ideas in the first place. Thus, we construct a sample at the evaluator-idea-day level that conditions on idea generation: For every evaluator i, we only consider the days t on which they generate at least one idea j. We closely follow specification 2 from section 5.1, and estimate the following:

$$y_{ijt} = \alpha * Post\ Evaluation_{it} + \beta \boldsymbol{X}_{it} + \gamma \boldsymbol{Z}_j + \delta_i + \zeta_t + \epsilon_{ijt} \tag{4}$$

y denotes the quality (novelty and value) of idea j. As above, $Post\ Evaluation_{it}$ is a dummy variable that takes the value 1 if day t is within one week since evaluator i evaluates ideas, and 0 otherwise. $\boldsymbol{X}_{it}$ is a vector of time-variant evaluator characteristics. $\boldsymbol{Z}_j$ is a vector of idea characteristics. δ_i are evaluator fixed effects that adjust for some employees generating generally more novel or valuable ideas than others. ζ_t are day fixed effects to control for time trends in idea quality. Our objective is to estimate α, which captures whether and how the ideas generated by evaluators within one week after an idea evaluation differ from those they generate in the absence of idea evaluation, i.e., outside the one-week post-evaluation period.

A.9. Tables and Figures

Table A.1: Evaluators and Ideators

# Employees (%)		Did Employee Generate Ideas?		
		Yes ('Ideator')	No ('Non-Ideator')	Total
Did Employee Evaluate Ideas?	Yes ('Evaluator')	N=5,267 (22.5%)	N=6,802 (29.1%)	12,069 (51.6%)
	No ('Non-Evaluator')	N=11,339 (48.4%)	-	11,339 (48.4%)
Total		16,606 (70.9%)	6,802 (29.1%)	23,408 (100%)

Notes: This table shows the distribution of employees across evaluator and ideator roles. Employees are included in the IMS only if they evaluated or generated at least one idea, which is why we cannot determine how many employees were part of the firm but did not evaluate or generate ideas during the observation period.

Table A.2: Descriptive Statistics - Non-Evaluators vs. Evaluators

	All Employees		Non-Evaluators		Evaluators	
Sample coverage						
# Employees	23,408		11,339		12,096	
# Ideas	185,681		130,072		94,637	
Variable	Mean	Std.	Mean	Std.	Mean	Std.
# Ideas Generated	20.42	48.78	13.46	29.70	35.40	72.64
Years Active as Ideator	6.68	4.74	5.73	4.42	8.73	4.76
# Ideas Generated per Year	2.63	4.03	2.27	2.94	3.41	5.64
# Co-Ideators	4.68	5.87	5.68	6.77	2.54	1.85
Permanent	0.84	0.34	0.78	0.39	0.96	0.14
Hierarchy	1.83	0.38	1.76	0.40	1.98	0.25
Prev. Mobility	0.47	0.60	0.38	0.54	0.66	0.69
Gender	0.12	0.32	0.12	0.33	0.10	0.30

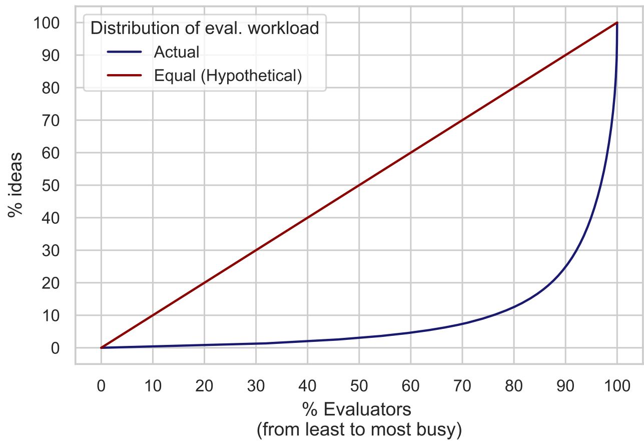
Notes: This table compares non-evaluators with evaluators based on their ideation productivity and background characteristics (*Permanent*, *Hierarchy*, *Prev. Mobility*, *Gender*). Variables related to ideation productivity are defined as follows. *# Ideas Generated*: the total number of ideas generated by the employee; *Years Active as Ideator*: the number of years during which the employee generated ideas; and *# Ideas Generated per Year*: *# Ideas Generated* divided by *Years Active as Ideator*. Background characteristics are measured at the point of idea submission and defined in Table A.6.

Table A.3: Descriptive Statistics - Evaluators' Evaluation Activities

	Mean	Std.	Min.	Median	Max.
# Ideas Evaluated	24.61	119.97	1.00	3.00	6,145
# Evaluation Reports Filed	32.73	176.23	1.00	4.00	9,487
Years Active as Evaluator	5.23	4.58	1.00	4.00	15.00
# Ideas Evaluated per Year	3.01	10.14	0.13	1.00	454.00
# Evaluation Reports Filed per Year	3.90	15.13	0.13	1.00	713.92

Notes: This table provides descriptive statistics for evaluators' evaluation activities, $N = 12,096$. Variables are defined as follows. *# Ideas Evaluated*: the total number of ideas generated by the evaluator; *# Evaluation Reports Filed*: the total number of evaluation reports filed by the evaluator; *Years Active as Evaluator*: the number of years during which the evaluator evaluated ideas; *# Ideas Evaluated per Year*: *# Ideas Evaluated* divided by *Years Active as Evaluator*; and *# Evaluation Reports Filed per Year*: *# Evaluation Reports Filed* divided by *Years Active as Evaluator*.

Figure A.1: Distribution of Evaluation Workload

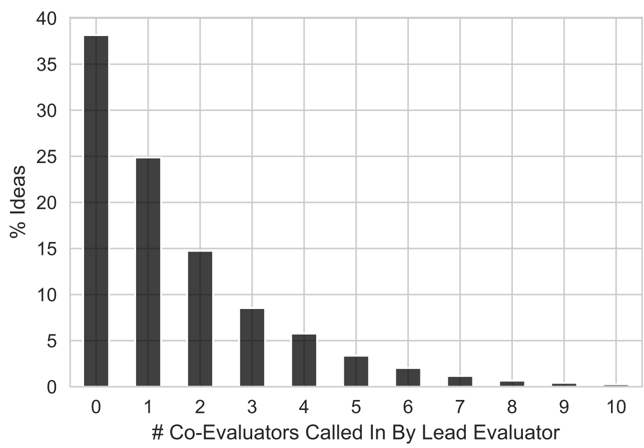


Notes: This figure illustrates the distribution of the evaluation workload across evaluators. The blue line plots the share of evaluators against the share of ideas they evaluate, and the red line a hypothetical, perfectly equal distribution of the evaluation workload across evaluators.

Table A.4: Summary Statistics, Choice-Based Panel, $N = 419,134$

Variable	Mean	Std.	Min.	50%	Max.
Idea Generation	0.33	0.47	0.00	0.00	1.00
# Ideas Generated	0.44	0.92	0.00	0.00	61.00
1 Day Before Eval.	0.01	0.10	0.00	0.00	1.00
Day of Eval.	0.02	0.12	0.00	0.00	1.00
1 Day After Eval.	0.01	0.11	0.00	0.00	1.00
Post Evaluation	0.06	0.23	0.00	0.00	1.00
# Prev. Generated Ideas	73.56	130.91	0.00	31.00	1,963
Previous Mobility	1.07	1.07	0.00	1.00	8.00
Hierarchy	1.97	0.25	1.00	2.00	5.00
Permanent	0.97	0.18	0.00	1.00	1.00
# Prev. Evaluated Ideas	25.95	178.50	0.00	1.00	9,263

Figure A.2: Distribution of Co-Evaluation



Notes: This figure illustrates the share of ideas involving co-evaluators.

Table A.5: Correlations, Choice-Based Panel, $N = 419,134$

	Variable	1	2	3	4	5	6	7	8	10	11	12
1	Idea Generation	1.00										
2	# Ideas Generated	0.69	1.00									
3	1 Day Before Eval.	0.01	0.01	1.00								
4	Day of Eval.	0.06	0.04	0.11	1.00							
5	1 Day After Eval.	0.02	0.01	0.09	0.11	1.00						
6	Post Evaluation	0.04	0.02	0.13	0.50	0.44	1.00					
7	# Prev. Generated Ideas	-0.01	0.07	0.01	0.01	0.01	0.02	1.00				
8	Cumulative Idea Value	-0.01	0.03	0.02	0.02	0.02	0.04	0.56	1.00			
9	Previous Mobility	-0.03	-0.02	0.02	0.03	0.03	0.05	0.21	0.17			
10	Hierarchy	0.01	0.01	0.02	0.03	0.02	0.06	0.06	0.07	1.00		
11	Permanent	0.04	0.03	0.02	0.02	0.02	0.04	0.09	0.08	0.73	1.00	
12	# Prev. Evaluated Ideas	-0.01	-0.01	0.17	0.18	0.18	0.21	0.02	0.07	0.02	0.03	1.00

Table A.6: Description of Variables

Variable	Description
Dependent	
Idea Generation	1 if evaluator i generates idea on day t, 0 otherwise
# Ideas	Number of ideas that evaluator i generates on day t
Idea Novelty	Percentage share of topic combinations made by idea j that have not been made by any previous idea
Idea Value	First-year savings or additional revenue generated by idea j minus implementation costs
Idea Similarity	Similarity between evaluated idea and generated idea
Independent	
E^{τ}	Event study lead and lag indicators. 1 if evaluator i evaluated an idea τ days ago, 0 otherwise
Post Evaluation	1 if day t is within one week since evaluator i evaluated an idea, 0 otherwise
Control	
# Previous Ideas	Number of ideas that evaluator i generated until day t
Cumulative Idea Value	Idea value that evaluator i generated until day t
Permanent	1 if evaluator i is on a permanent contract by day t, 0 otherwise
Hierarchy	Evaluator i's position in the organizational hierarchy; 5 levels: 1-lowest, 5-highest
Prev. Mobility	Number of times evaluator i has previously moved between organizational units
Gender	1 if male, 0 if female
Mechanism	
Years Since IMS Interaction	Number of years since evaluator i last logged onto the Idea Management System
Post Evaluation ...	1 if day t is within one week since evaluator i evaluated an idea ...
... Same Unit	... by an employee from own unit, 0 otherwise
... Different Unit	... by an employee from different unit, 0 otherwise
... Same Gender	... by a same-gender employee, 0 otherwise
... Different Gender	... by a different-gender employee, 0 otherwise
... Same Mobility	... by an employee who previously moved between units for the same number of times, 0 otherwise
... Different Mobility	... by an employee who previously moved between units for a different number of times, 0 otherwise
... Unfamiliar	... related to an unfamiliar topic, 0 otherwise
... Familiar	... related to a familiar topic, 0 otherwise

Table A.7: Effect of Idea Evaluation on Idea Generation, Event Study - Logistic and Poisson Regression

	Logistic Regression		Poisson Regression	
DV	Idea Generation		# Ideas	
	(1)	(2)	(3)	(4)
Days Since Evaluation				
-3	−0.006	−0.013	−0.031	−0.034
	[0.994]	[0.987]	[0.969]	[0.967]
	(0.047)	(0.047)	(0.054)	(0.051)
-2	0.026	0.008	0.049	0.031
	[1.026]	[1.008]	[1.050]	[1.032]
	(0.049)	(0.049)	(0.056)	(0.054)
0	0.962***	0.954***	0.929***	0.910***
	[2.616]***	[2.596]***	[2.533]***	[2.485]***
	(0.040)	(0.040)	(0.040)	(0.044)
1	0.261***	0.247***	0.287***	0.270***
	[1.298]***	[1.280]***	[1.333]***	[1.309]***
	(0.041)	(0.041)	(0.046)	(0.043)
2	0.215***	0.202***	0.185***	0.170***
	[1.240]***	[1.224]***	[1.203]***	[1.185]***
	(0.041)	(0.042)	(0.050)	(0.050)
3	0.172***	0.158***	0.154**	0.138**
	[1.187]***	[1.171]***	[1.166]**	[1.148]**
	(0.047)	(0.047)	(0.052)	(0.051)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls		X		X
DV Weighted Mean	0.005	0.005	0.007	0.007
# Evaluators	5,267	5,267	5,267	5,267
# Days	4,446	4,446	4,446	4,446
# Obs.	412,332	412,332	412,332	412,332
Pseudo R^2	0.153	0.156	0.183	0.186
Wald	46.857***	55.502***	44.093***	50.739***

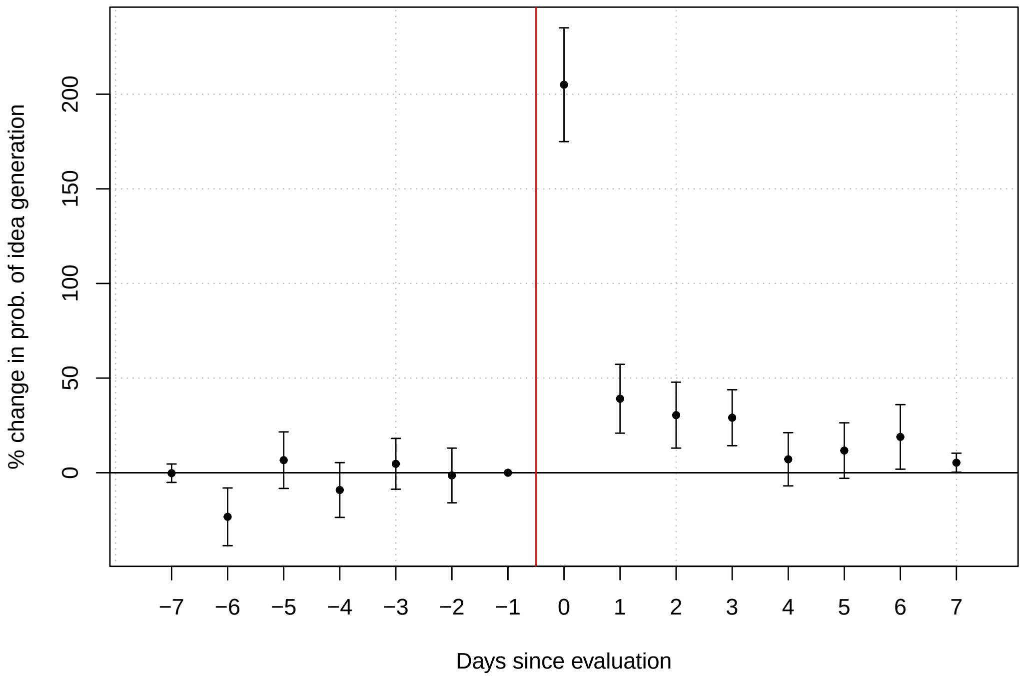
Notes: This table reports event study results from specification 1 using logistic and Poisson regression. The analysis is at the evaluator-day level. Coefficient estimates indicate changes in the log odds of idea generation (col. 1 and 2) and the log number of ideas generated (col. 3 and 4). Exponentiated coefficient estimates, shown in square brackets, indicate changes in the odds of idea generation (col. 1 and 2) and the number of ideas generated (col. 3 and 4). Robust standard errors, clustered at the evaluator level, are in parentheses. All regressions include weights to adjust for choice-based sampling. * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.8: Effect of Evaluation on Idea Generation, Event Study - OLS Regression

	OLS Regression			
DV	Idea Generation		# Ideas	
	(1)	(2)	(3)	(4)
Days Since Evaluation				
-3	0.000	0.000	0.000	0.000
	[4.704]	[3.087]	[2.591]	[0.674]
	(0.000)	(0.000)	(0.001)	(0.001)
-2	0.000	0.000	0.000	0.000
	[−1.445]	[−3.108]	[5.285]	[4.318]
	(0.000)	(0.000)	(0.001)	(0.001)
0	0.011***	0.011***	0.014***	0.014***
	[205.015]***	[204.791]***	[266.556]***	[267.811]***
	(0.001)	(0.001)	(0.001)	(0.001)
1	0.002***	0.002***	0.003***	0.003***
	[39.092]***	[37.919]***	[57.739]***	[57.072]***
	(0.000)	(0.000)	(0.001)	(0.001)
2	0.002***	0.001**	0.002*	0.002*
	[30.428]***	[28.098]**	[35.906]*	[33.569]*
	(0.000)	(0.000)	(0.001)	(0.001)
3	0.002***	0.001***	0.002**	0.002**
	[29.070]***	[28.545]***	[35.507]**	[35.032]**
	(0.000)	(0.000)	(0.001)	(0.001)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls		X		X
DV Weighted Mean	0.005	0.005	0.007	0.007
# Evaluators	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134
Adj. R^2	0.022	0.022	0.019	0.019

Notes: This table reports event study results from specification 1 using OLS regression. The analysis is at the evaluator-day level. Coefficient estimates indicate percentage point changes in the likelihood of idea generation (col. 1 and 2) and in the number of ideas generated (col. 3 and 4). Percentage changes, shown in square brackets, are calculated by dividing percentage point changes by the weighted mean of the respective dependent variable. Robust standard errors, clustered at the evaluator level, are in parentheses. All regressions include weights to adjust for choice-based sampling. This tables includes more observations than Table A.7 because, unlike logistic and Poisson regression, OLS regression does not drop employees who never generate any ideas. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A.3: Effect of Idea Evaluation on Idea Generation, Event Study - OLS Regression



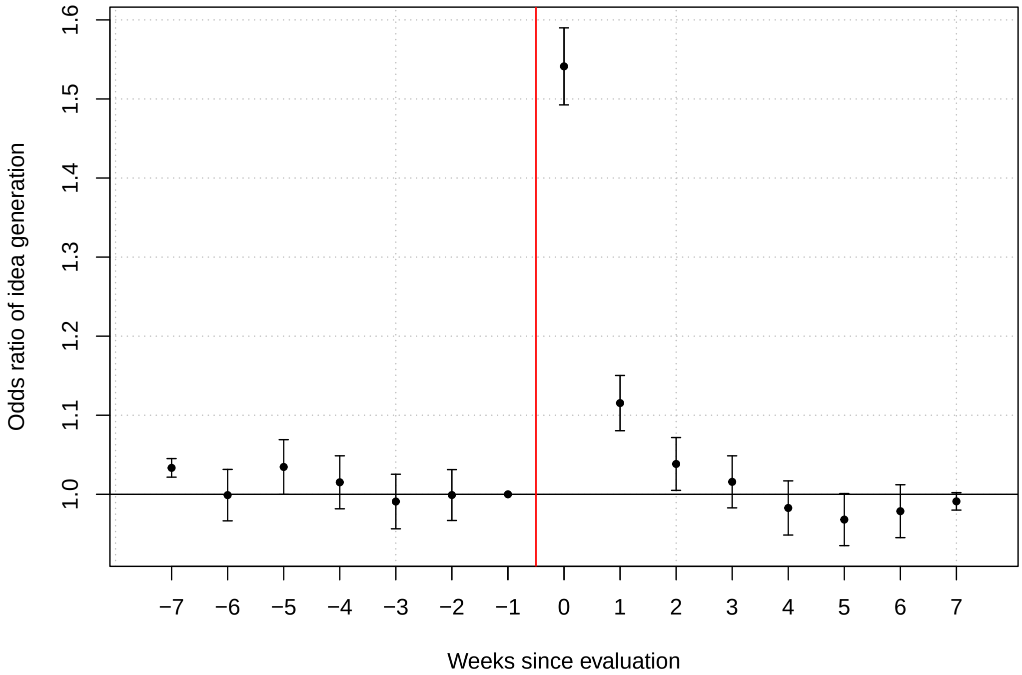
Notes: This figure displays percentage changes in the probability of idea generation as reported in col. 1 of Table A.8. Error bars indicate 95% confidence intervals.

Table A.9: Effect of Evaluation on Idea Generation, Single Diff.-in-Diff. - OLS Regression

	OLS Regression			
DV	Idea Generation		# Ideas	
	(1)	(2)	(3)	(4)
Post Evaluation	0.003***	0.003***	0.005***	0.005***
	[66.255]***	[64.065]***	[89.832]***	[89.302]***
	(0.000)	(0.000)	(0.001)	(0.001)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls		X		X
DV Weighted Mean	0.005	0.005	0.007	0.007
# Evaluators	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134
Adj. R^2	0.022	0.022	0.019	0.019

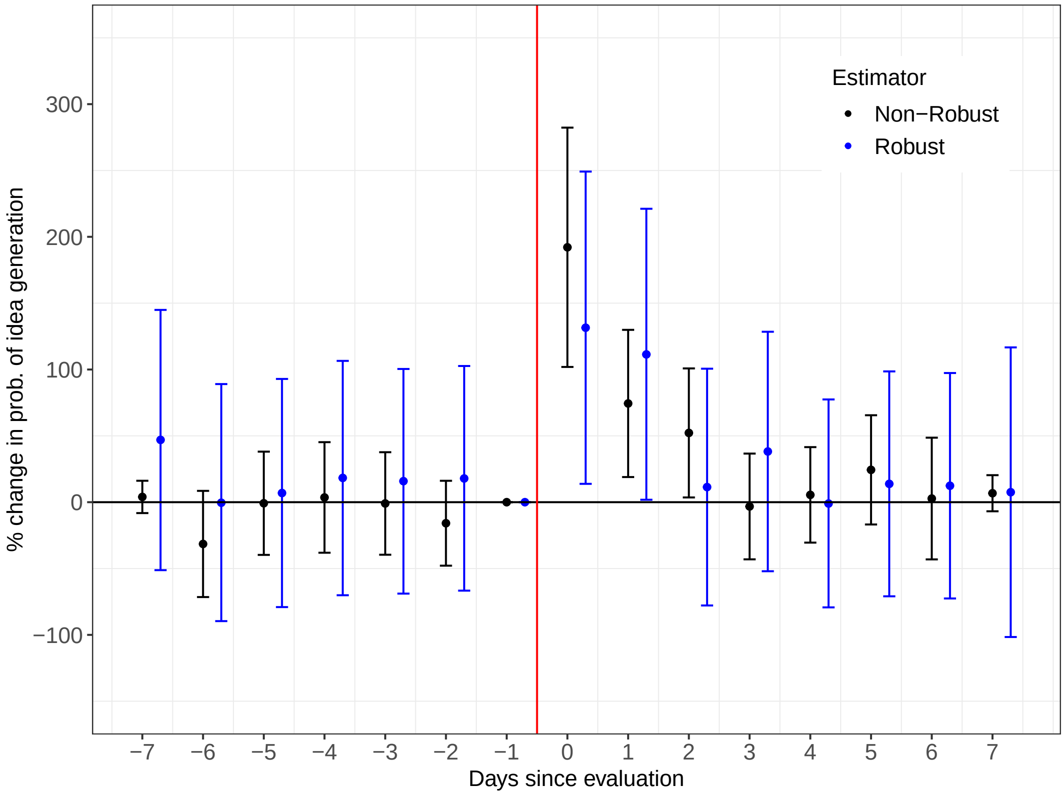
Notes: This table reports results from specification 2 using OLS regression. The analysis is at the evaluator-day level. Coefficient estimates indicate percentage point changes in the likelihood of idea generation (col. 1 and 2) and in the number of ideas generated (col. 3 and 4). Percentage changes, shown in square brackets, are calculated by dividing percentage point changes by the weighted mean of the respective dependent variable. Robust standard errors, clustered at the evaluator level, are reported in parentheses. All regressions include weights to adjust for choice-based sampling. This tables includes more observations than Table 1 because, unlike logistic and Poisson regression, OLS regression does not drop employees who never generate any ideas. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A.4: Effect of Idea Evaluation on Idea Generation, Evaluator-Week Level



Notes: This figure reports event study results from specification 1 at the evaluator-week level. Estimates are obtained from logistic regression, with robust standard errors clustered at the evaluator level and weights to adjust for choice-based sampling. Controls are as usual. Error bars represent 95% confidence intervals.

Figure A.5: Accounting for Treatment Effect Heterogeneity



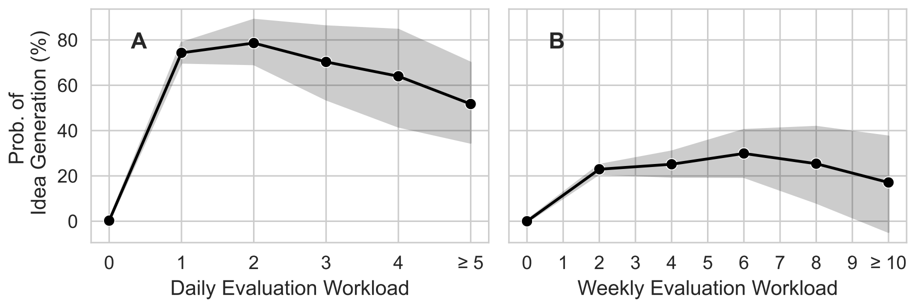
Notes: This figure compares non-heterogeneity-robust event study OLS estimates (in black) with those obtained using the heterogeneity-robust estimator proposed by de Chaisemartin and D'Haultfœuille (2024) (in blue). The estimates are derived from a sample of 676,011 evaluator-day observations, using January 1, 2017, as an artificial common 'period-one'. Evaluators are included in the sample if they were part of the panel on that day. The predicted percentage changes in idea generation closely resemble those shown in Figure A.3. Error bars represent 95% confidence intervals, which are generally wider than those in the main analysis due to the shorter period and smaller sample of employees covered in this analysis.

Table A.10: Accounting for Treatment Effect Heterogeneity

	OLS Regression			
	Non-Robust		Robust	
DV	Idea Gen.	# Ideas	Idea Gen.	# Ideas
	(1)	(2)	(3)	(4)
Post Evaluation	0.003*** [55.573]*** (0.001)	0.004*** [86.689]*** (0.001)	0.006*** [131.503]*** (0.003)	0.012*** [189.383]*** (0.006)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls	X	X	X	X
Dep. Var. Mean	0.005	0.006	0.005	0.006
# Evaluators	4,014	4,014	4,014	4,014
# Days	181	181	181	181
# Obs.	676,011	676,011	676,011	676,011
Adj. R^2	0.049	0.087	0.029	0.060

Notes: This table compares non-heterogeneity-robust OLS estimates (col. 1 and 2) with those obtained using the heterogeneity-robust estimator proposed by de Chaisemartin and D'Haultfœuille (2024) (col. 3 and 4). Coefficient estimates indicate percentage point changes in the likelihood of idea generation (col. 1 and 3) and in the number of ideas generated (col. 2 and 4). Percentage changes, shown in square brackets, are calculated by dividing percentage point changes by the weighted mean of the respective dependent variable. The estimates are derived from a sample of 676,011 evaluator-day observations, using January 1, 2017, as an artificial common "period-one". Evaluators are included in the sample if they were part of the panel on that day. Non-robust coefficient estimates closely resemble those reported in Table A.9. Robust coefficient estimates are larger than the non-robust ones, suggesting that the non-robust estimates even underestimate the full magnitude of idea crowding-in. Robust standard errors, clustered at the evaluator level, are shown in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A.6: Boundary Condition, Relationship between Idea Generation and Evaluation Workload



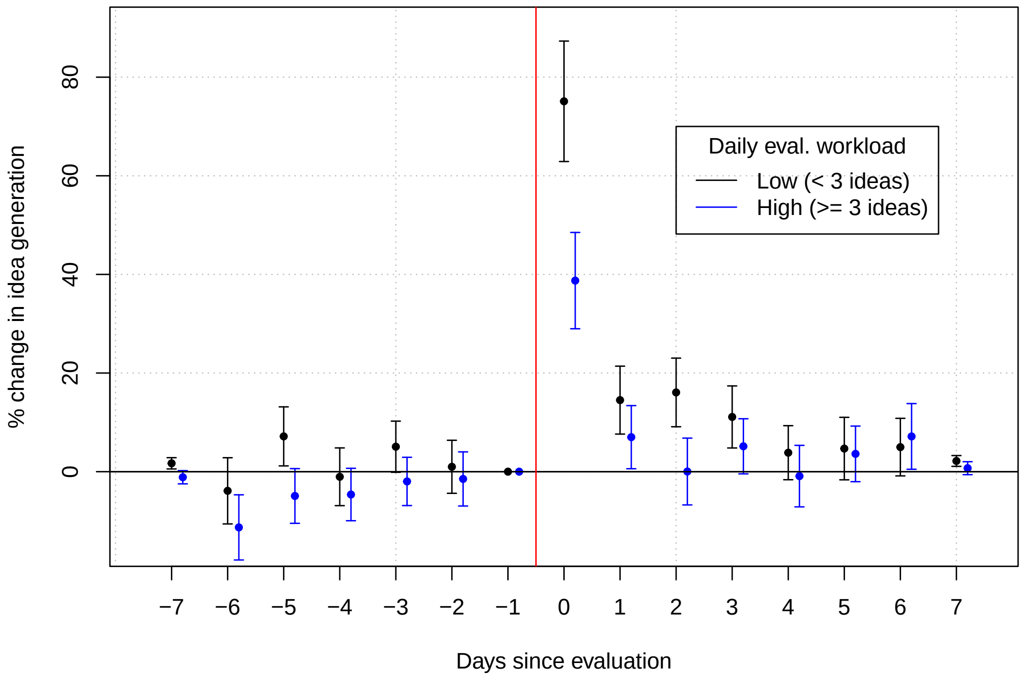
Notes: This figure plots evaluators' probability of generating ideas on a given day as a function of the number of ideas evaluated on the same day (Panel A) and within the preceding week (Panel B). Probabilities are relative to days with no ideas evaluated. Shaded areas represent 1-SD bands. The figure is based on the sample of evaluator-day observations used in Table A.11.

Table A.11: Boundary Condition, Weekly Evaluation Workload

	OLS Regression					
DV	Idea Generation			# Ideas		
	(1)	(2)	(3)	(4)	(5)	(6)
Post Eval. Low	0.0009***		0.0009***	0.0012***		0.0012***
	[16.8104]***		[17.0615]***	[23.2680]***		[23.6701]***
	(0.0001)		(0.0001)	(0.0001)		(0.0001)
Post Eval. High		0.0002*	0.0002**		0.0003**	0.0003***
		[3.0427]*	[4.1746]**		[5.1177]**	[6.6880]***
		(0.0001)	(0.0001)		(0.0001)	(0.0001)
Day FE	X	X	X	X	X	X
Evaluator FE	X	X	X	X	X	X
Controls	X	X	X	X	X	X
DV Weighted Mean	0.0052	0.0052	0.0052	0.0071	0.0071	0.0071
# Evaluators	12,069	12,069	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134	419,134	419,134
Adj. R^2	0.0217	0.0217	0.0217	0.0190	0.0189	0.0190

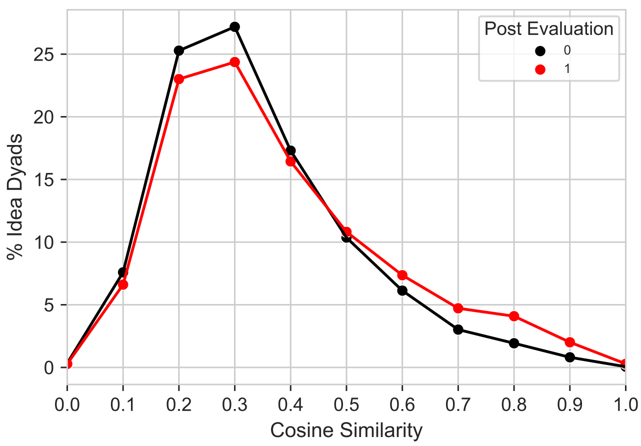
Notes: This table reports results from specification 2 using OLS regression. The analysis is at the evaluator-day level. *Post Eval. Low* takes the value 1 if an evaluator evaluated between one and six ideas within the preceding week, and 0 otherwise. *Post Eval. High* takes the value 1 if an evaluator evaluated more than six ideas within the preceding week, and 0 otherwise. Coefficient estimates indicate percentage point changes in the likelihood of idea generation (col. 1-3) and in the number of ideas generated (col. 4-6) in response to a 1-SD increase in *Post Eval. Low* and *Post Eval. High*, respectively. Percentage changes, shown in square brackets, are calculated by dividing percentage point changes by the weighted mean of the dependent variable. Standard errors, clustered at the evaluator level, are in parentheses. All regressions include weights to adjust for choice-based sampling. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A.7: Boundary Condition, Daily Evaluation Workload



Notes: This figure reports event study results from specification 1 at the evaluator-day level. Estimates are obtained from OLS regression, with robust standard errors clustered at the evaluator level and weights to adjust for choice-based sampling. Controls are as usual. Error bars represent 95% confidence intervals.

Figure A.8: Mechanism, Recombination - Cosine Similarity between Evaluated and Generated Ideas



Notes: This figure compares the distribution of cosine similarities between evaluated and generated ideas for ideas that were generated within one week after idea evaluation (Post Evaluation = 1) and ideas generated outside this time window (Post Evaluation = 0).

Table A.12: Mechanism, Recombination - Similarity between Evaluated and Generated Idea

	OLS Regression					
DV	Cosine Similarity		Share Same Topic (NLP-Assigned)		Share Same Topic (Pre-Assigned)	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Evaluation	0.035***	0.031***	0.034***	0.029***	0.050***	0.048***
	[11.859]***	[10.635]***	[52.393]***	[44.689]***	[7.368]***	[7.003]***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.010)	(0.009)
Controls		X		X		X
Dep. Var. Mean	0.292	0.292	0.065	0.065	0.679	0.679
# Idea Dyads	6,267,633	6,267,633	6,267,633	6,267,633	114,708	114,708
# Obs.	8,827,612	8,827,612	8,827,612	8,827,612	155,307	155,307
Adj. R^2	0.000	0.008	0.000	0.013	0.000	0.011

Notes: This table reports results from OLS regressions of the similarity between evaluated and generated ideas on *Post Evaluation*. Coefficient estimates indicate changes in idea similarity following idea evaluation. Percentage changes, shown in square brackets, are calculated by dividing the coefficient estimates by the mean of the dependent variable. We control for the number of years between idea generation and evaluation, as well as whether the evaluated and the generated idea was selected. The dependent variables are detailed in section 6.1. Robust standard errors, clustered at the evaluator level, are provided in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.13: Mechanism, Recombination - Similarity between Evaluated and Generated Idea, Sensitivity Analysis

	DV = Share Same Topic (NLP-Assigned), OLS Regression					
Total # topics	1000			500		
# Topics per idea	3	5	7	3	5	7
	(1)	(2)	(3)	(4)	(5)	(6)
Post Evaluation	0.024***	0.029***	0.030***	0.028***	0.031***	0.034***
	[49.471]***	[44.689]***	[37.725]***	[37.438]***	[30.313]***	[27.831]***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	X	X	X	X	X	X
Dep. Var. Mean	0.049	0.065	0.080	0.074	0.103	0.123
# Idea Dyads	6,267,633	6,267,633	6,267,633	6,267,633	6,267,633	6,267,633
# Obs.	8,827,612	8,827,612	8,827,612	8,827,612	8,827,612	8,827,612
Adj. R^2	0.009	0.013	0.018	0.013	0.017	0.019

Notes: This table replicates column 4 of Table A.12, varying both the total number of topics used to span the topic space, M, and the number of topics assigned to each idea, N. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.14: Mechanism, Recombination - Exposure to Unfamiliar Knowledge, Sensitivity Analysis: Total # Topics = 1,000

	OLS Regression, DV = Idea Generation								
# Topics assigned to idea	3			5			7		
Topic exposure	any time past	< 3 yrs. ago	< 1 yr. ago	any time past	< 3 yrs. ago	< 1 yr. ago	any time past	< 3 yrs. ago	< 1 yr. ago
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post Evaluation Familiar	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***
	[4.0068]***	[4.4148]***	[4.0543]***	[4.0030]***	[4.3216]***	[4.1288]***	[3.9858]***	[4.2228]***	[4.0319]***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Post Evaluation Unfamiliar	0.0006***	0.0005***	0.0005***	0.0006***	0.0005***	0.0005***	0.0006***	0.0006***	0.0005***
	[11.1623]***	[10.2545]***	[10.2719]***	[11.3432]***	[10.4042]***	[10.1686]***	[11.5885]***	[10.6377]***	[10.2846]***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Evaluator FE	X	X	X	X	X	X	X	X	X
Day FE	X	X	X	X	X	X	X	X	X
Controls	X	X	X	X	X	X	X	X	X
DV Weighted Mean	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052
# Evaluators	12,069	12,069	12,069	12,069	12,069	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134	419,134	419,134	419,134	419,134	419,134
Adj. R^2	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217

Notes: This table replicates Table 2, setting $M = 1{,}000$ while varying the number of topics assigned to each idea, N. * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.15: Mechanism, Recombination - Exposure to Unfamiliar Knowledge, Sensitivity Analysis: Total # Topics = 500

	OLS Regression, DV = Idea Generation								
# Topics assigned to idea	3			5			7		
Topic exposure	any time past	< 3 yrs. ago	< 1 yr. ago	any time past	< 3 yrs. ago	< 1 yr. ago	any time past	< 3 yrs. ago	< 1 yr. ago
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Post Evaluation Familiar	0.0002***	0.0002***	0.0002***	0.0002***	0.0002***	0.0002**	0.0002***	0.0002***	0.0002***
	[4.4861]***	[4.5219]***	[4.0995]***	[4.1050]***	[4.2612]***	[3.9702]**	[4.1691]***	[4.3678]***	[4.0869]***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Post Evaluation Unfamiliar	0.0006***	0.0005***	0.0005***	0.0006***	0.0006***	0.0005***	0.0006***	0.0006***	0.0005***
	[11.0107]***	[10.4082]***	[10.3153]***	[11.8834]***	[10.8821]***	[10.4881]***	[12.1906]***	[10.9513]***	[10.4433]***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Controls	X	X	X	X	X	X	X	X	X
Evaluator FE	X	X	X	X	X	X	X	X	X
Day FE	X	X	X	X	X	X	X	X	X
DV Weighted Mean	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052
# Evaluators	12,069	12,069	12,069	12,069	12,069	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134	419,134	419,134	419,134	419,134	419,134
Adj. R^2	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217	0.0217

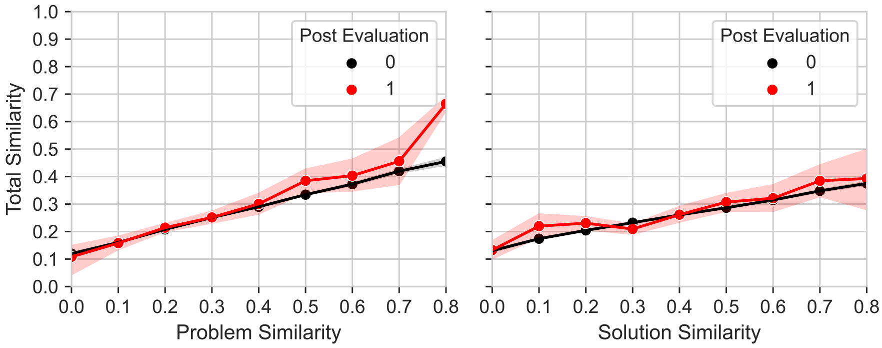
Notes: This table replicates Table 2, setting $M = 500$ while varying the number of topics assigned to each idea, N. * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.16: Mechanism, Recombination - Co-Evaluation

	OLS Regression, DV = Idea Generation					
	(1)	(2)	(3)	(4)	(5)	(6)
Post Co-Evaluation	0.0007***		0.0006***			
	[12.5330]***		[11.3524]***			
	(0.0001)		(0.0001)			
Post Solo-Evaluation		0.0005***	0.0004***			
		[8.7382]***	[6.8813]***			
		(0.0001)	(0.0001)			
Post Co-Evaluation, w/ Access to Reports				0.0006***		0.0005***
				[10.9089]***		[10.2861]***
				(0.0001)		(0.0001)
Post Co-Evaluation, w/o Access to Reports					0.0003***	0.0002**
					[5.1826]***	[3.4649]**
					(0.0001)	(0.0001)
Evaluator FE	X	X	X	X	X	X
Day FE	X	X	X	X	X	X
Controls	X	X	X	X	X	X
DV Weighted Mean	0.0052	0.0052	0.0052	0.0052	0.0052	0.0052
# Evaluators	12,069	12,069	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134	419,134	419,134
Adj. R^2	0.0216	0.0215	0.0216	0.0215	0.0215	0.0215

Notes: This table reports results from specification 2 using OLS regression. The analysis is at the evaluator-day level. *Post Co-Evaluation*$_{it}$ equals 1 if evaluator i evaluated an idea involving co-evaluators during the week preceding day t, and 0 otherwise. *Post Solo-Evaluation*$_{it}$ equals 1 if evaluator i evaluated an idea without co-evaluators during the week preceding day t, and 0 otherwise. *Post Co-Evaluation, w/ Access to Reports*$_{it}$ and *Post Co-Evaluation, w/o Access to Reports*$_{it}$ indicate whether evaluator i co-evaluated an idea while having access to previous evaluation reports on the same idea by other evaluators, or not. We normalize these variables to have a mean of zero and a standard deviation of one such that coefficient estimates indicate percentage point changes in the likelihood of idea generation in response to a 1-SD increase in the independent variables. Values in square brackets indicate percentage changes. Robust standard errors, clustered at the evaluator level, are reported in parentheses. All regressions include weights to adjust for choice-based sampling. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A.9: Mechanism, Problem- vs. Solution-Based Recombination



Notes: This figure plots the total (cosine) similarity between evaluated and generated ideas against their problem (cosine) similarity (left) and solution (cosine) similarity (right), before evaluation (black) and after evaluation (red). Dots indicate averages, shaded areas indicate 1-SD bands.

Table A.17: Mechanism, Problem- vs. Solution-Based Recombination

	OLS Regression, DV = Total Similarity				
	(1)	(2)	(3)	(4)	(5)
Problem Sim.	0.447***		0.446***		0.424***
	[181.819]***		[181.699]***		[172.360]***
	(0.002)		(0.002)		(0.002)
Solution Sim.		0.300***		0.300***	0.276***
		[122.059]***		[122.033]***	[112.228]***
		(0.002)		(0.002)	(0.001)
Problem Sim. x Post Eval.			0.104**		0.076*
			[42.339]**		[31.085]*
			(0.034)		(0.030)
Solution Sim. x Post Eval.				0.015	−0.016
				[6.213]	[−6.460]
				(0.027)	(0.023)
DV Mean	0.246	0.246	0.246	0.246	0.246
# Evaluators	2,407	2,407	2,407	2,407	2,407
# Idea Dyads	113,168	113,168	113,168	113,168	113,168
# Obs.	161,362	161,362	161,362	161,362	161,362
Adj. R^2	0.284	0.187	0.284	0.187	0.442

Notes: This table reports OLS regressions at the evaluated–generated idea dyad level. *Total Similarity* is the cosine similarity between evaluated and generated ideas; *Problem Similarity* is the cosine similarity between the problem components of evaluated and generated ideas; and *Solution Similarity* is the cosine similarity between the solution components of evaluated and generated ideas. Percentage changes, shown in square brackets, are calculated by dividing coefficient estimates by the mean of *Total Similarity*. Estimations are based on the sample of idea dyads for which we could unambiguously identify problem and solution components. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.18: Mechanism, Recombination - Problem Exploitation and Problem Distance

	OLS Regression, DV = Total Similarity			
	Problem Exploitation		Problem Distance	
	Low	High	Low	High
Problem Sim.	0.527***	0.373***	0.387***	0.461***
	[237.435]***	[137.749]***	[137.426]***	[217.658]***
	(0.003)	(0.002)	(0.002)	(0.003)
Solution Sim.	0.201***	0.346***	0.341***	0.199***
	[90.662]***	[127.921]***	[121.238]***	[94.093]***
	(0.002)	(0.002)	(0.002)	(0.002)
Problem Sim. x Post Eval.	0.145**	−0.005	0.028	0.133**
	[65.326]**	[−1.670]	[9.898]	[62.647]**
	(0.045)	(0.043)	(0.046)	(0.051)
Solution Sim. x Post Eval.	−0.025	−0.026	−0.028	−0.040
	[−11.108]	[−9.768]	[−9.867]	[−19.052]
	(0.033)	(0.032)	(0.032)	(0.034)
DV Mean	0.222	0.271	0.282	0.212
# Evaluators	1,939	1,817	1,549	1,643
# Idea Dyads	56,382	56,619	52,684	50,293
# Obs.	79,201	79,204	73,065	73,066
Adj. R^2	0.410	0.458	0.457	0.323

Notes: This table replicates the analysis from Table A.17, but splits the data into subsamples based on *Problem Exploitation*, which is the extent to which the problem in the evaluated idea has been previously exploited (col. 1–2), and *Problem Distance*, which is the distance between the evaluated problem and the problems addressed in the evaluator's prior idea generation (col. 3–4). We define *Problem Exploitation* as the average cosine similarity between the problem of the evaluated idea and the problems of all ideas submitted to the firm in the previous month. We define *Problem Distance* as one minus the average cosine similarity between the evaluated problem and the problems from the evaluator's own previously submitted ideas. In each case, we split the data at the median value of *Problem Exploitation* and *Problem Distance*, respectively. The number of observations is slightly smaller than in Table A.17, because we exclude cases where *Problem Exploitation* or *Problem Distance* could not be calculated due to the absence of prior ideas for comparison. Percentage changes, shown in square brackets, are calculated by dividing coefficient estimates by the mean of *Total Similarity*. Estimations are based on the sample of idea dyads for which we could unambiguously identify problem and solution components. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.19: Mechanism, Attention-Shifting

	OLS Regression, DV = Idea Generation					
	Full Sample		Recent		Long Ago	
	(1)	(2)	(3)	(4)	(5)	(6)
Post Evaluation	0.0007***	0.0007***	0.0003*	0.0003*	0.0008***	0.0008***
	[13.1094]***	[12.9623]***	[2.5461]*	[2.5537]*	[32.4887]***	[32.7008]***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Post Evaluation	0.0014**	0.0014**				
x Years Since IMS Interaction	[27.4134]**	[27.1402]**				
	(0.0005)	(0.0004)				
Day FE	X	X	X	X	X	X
Employee FE	X	X	X	X	X	X
Controls		X		X		X
DV Weighted Mean	0.0051	0.0051	0.0118	0.0118	0.0025	0.0025
# Employees	10,451	10,451	6,752	6,752	8,423	8,423
# Days	4,443	4,443	4,443	4,443	4,411	4,411
# Obs.	413,218	413,218	207,295	207,295	208,209	208,209
Adj. R^2	0.0208	0.0208	0.0397	0.0397	0.0098	0.0098

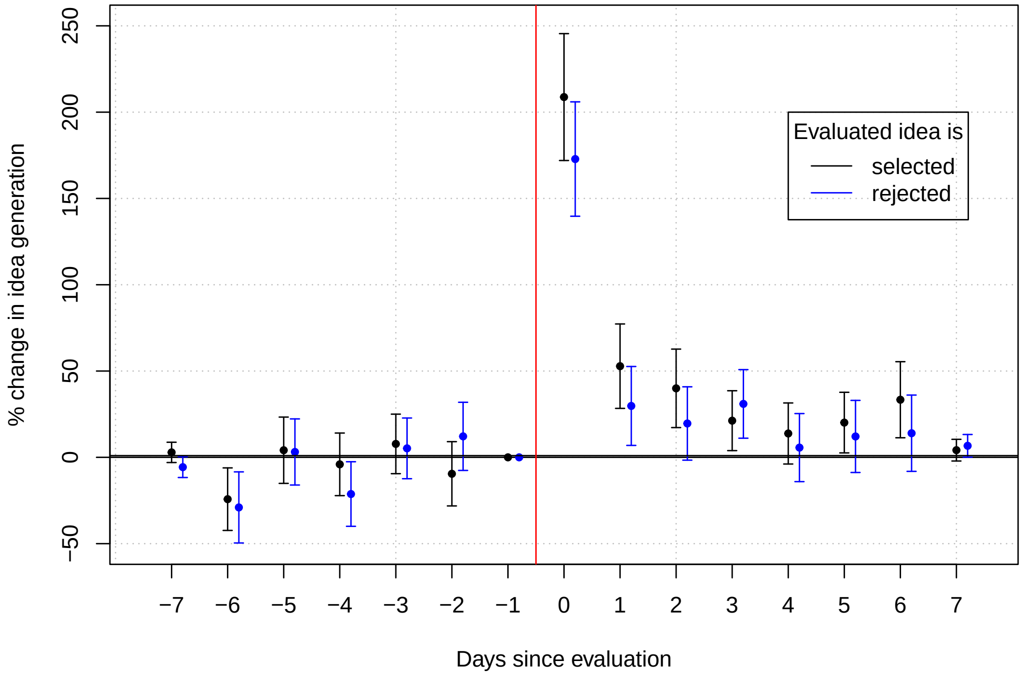
Notes: This table reports results from specification 2 using OLS regression. The analysis is at the evaluator-day level. *Years Since IMS Interaction* indicates the number of years since employees' most recent IMS interaction. *Post Evaluation* is standardized to have a mean of zero and a standard deviation of one, so coefficient estimates indicate percentage point changes in the likelihood of idea generation in response to a one-standard-deviation increase in *Post Evaluation*. The values in square brackets show the corresponding percentage changes in *Idea Generation*. While col. 1–2 report results based on the full sample, col. 3–6 present split-sample analyses using the median value of *Years Since IMS Interaction* as the cutoff, which corresponds to 0.10 years or 37 days. Standard errors, clustered at the evaluator level, are in parentheses. All regressions include weights to adjust for choice-based sampling. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.20: Mechanism, Social Comparison, OLS Regression

DV	OLS Regression		
	Idea Generation		
	(1)	(2)	(3)
Post Eval. Same Unit	0.0005**		
	[8.8905]**		
	(0.0002)		
Post Eval. Diff. Unit	0.0003***		
	[6.2921]***		
	(0.0001)		
Post Eval. Same Gender		0.0003***	
		[5.3892]***	
		(0.0001)	
Post Eval. Diff. Gender		0.0002*	
		[3.8777]*	
		(0.0001)	
Post Eval. Same Mobility			0.0002**
			[4.1230]**
			(0.0001)
Post Eval. Diff. Mobility			0.0003**
			[4.9313]**
			(0.0001)
Controls	X	X	X
Evaluator FE	X	X	X
Day FE		X	X
Org. Unit*Year FE	X		
DV Weighted Mean	0.0052	0.0052	0.0052
# Evaluators	5,267	12,069	12,069
# Days	5,400	4,446	4,446
# Obs.	408,953	419,134	419,134
Adj. R^2	0.0240	0.0217	0.0217

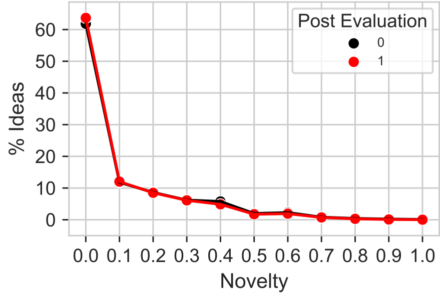
Notes: This table reports results from specification 2 using OLS regression. The analysis is at the evaluator-day level. Coefficient estimates indicate percentage point changes in the likelihood of idea generation in response to a 1-SD increase in the independent variables. Percentage changes, shown in square brackets, are calculated by dividing percentage point changes by the weighted mean of the dependent variable. The independent variables indicate whether a given day is within one week since an evaluator evaluated ideas by socially proximate (or distant) evaluatees. Robust standard errors, clustered at the evaluator level, are reported in parentheses. All regressions include weights to adjust for choice-based sampling. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure A.10: Alternative Mechanism, Plagiarism



Notes: This figure compares event study results from specification 1 at the evaluator-day level. Coefficient estimates in black represent percentage changes in the likelihood of idea generation following idea selection, while estimates in blue represent percentage changes following idea rejection. Estimates are obtained from OLS regression, with robust standard errors clustered at the evaluator level and weights to adjust for choice-based sampling. Controls are as usual. Error bars represent 95% confidence intervals.

Figure A.11: Relationship between Evaluation and Novelty of Generated Ideas



Notes: This figure compares the novelty distribution of ideas generated within one week after idea evaluation against the value distribution of ideas generated outside this time window.

Table A.21: Effect of Idea Evaluation on Idea Novelty, OLS Regressions

	DV = Idea Novelty, OLS Regression											
Total # topics	1,000						500					
# Topics per idea	3		5		7		3		5		7	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Post Evaluation	0.002	0.003	−0.001	−0.001	0.001	0.001	0.001	0.001	−0.001	−0.001	0.000	0.000
	[1.621]	[1.769]	[−0.586]	[−1.302]	[2.608]	[2.214]	[1.014]	[0.950]	[−1.320]	[−1.898]	[−0.285]	[−1.235]
	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Day FE	X	X	X	X	X	X	X	X	X	X	X	X
Evaluator FE	X	X	X	X	X	X	X	X	X	X	X	X
Controls		X		X		X		X		X		X
DV Mean	0.154	0.154	0.105	0.105	0.075	0.075	0.104	0.104	0.054	0.054	0.032	0.032
# Evaluators	5,267	5,267	5,267	5,267	5,267	5,267	5,267	5,267	5,267	5,267	5,267	5,267
# Days	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446	4,446
# Obs.	186,475	186,475	186,475	186,475	186,475	186,475	186,475	186,475	186,475	186,475	186,475	186,475
Adj. R^2	0.177	0.178	0.257	0.257	0.325	0.326	0.206	0.206	0.320	0.320	0.411	0.412

Notes: This table reports results from specification 4 using OLS regression. The analysis is at the idea-evaluator-day level. Total # topics is the number of topics used to span the topic space; # topics per idea is the number of topics assigned to each idea. Coefficient estimates indicate nominal changes in *Idea Novelty*, values in square brackets indicate percentage changes in *Idea Novelty* following idea evaluation. Standard errors, clustered at the evaluator level, are in parentheses. Controls are as usual (see Table A.6). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.22: Effect of Idea Evaluation on Idea Value, Long-Term Effect

	OLS Regression				
DV	log(Idea Value + 1)				
Evaluation is ... ago	1 Week	2 Weeks	1 Month	6 Months	1 Year
	(1)	(2)	(3)	(4)	(5)
Post Evaluation	0.201**	0.139*	0.077	0.106	0.091
	[21.189]**	[14.964]*	[8.001]	[11.158]	[9.474]
	(0.062)	(0.065)	(0.073)	(0.073)	(0.079)
Evaluator FE	X	X	X	X	X
Day FE	X	X	X	X	X
Controls	X	X	X	X	X
# Evaluators	4,633	4,633	4,633	4,633	4,633
# Days	4,284	4,284	4,284	4,284	4,284
# Obs.	108,938	108,938	108,938	108,938	108,938
Adj. R^2	0.385	0.382	0.382	0.382	0.382

Notes: This table replicates column 2 of Table 3, using alternative post-evaluation time windows. * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.23: Effect of Idea Evaluation on Likelihood of Hit Ideas, Long-Term Effect

	OLS Regression				
DV	Top-1				
Evaluation is ... ago	1 Week	2 Weeks	1 Month	6 Months	1 Year
	(1)	(2)	(3)	(4)	(5)
Post Evaluation	0.006**	0.005*	0.003	−0.002	−0.001
	[59.128]**	[52.944]*	[33.237]	[−19.664]	[−6.346]
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Evaluator FE	X	X	X	X	X
Day FE	X	X	X	X	X
Controls	X	X	X	X	X
# Evaluators	4,633	4,633	4,633	4,633	4,633
# Days	4,284	4,284	4,284	4,284	4,284
# Obs.	108,938	108,938	108,938	108,938	108,938
Adj. R^2	0.238	0.233	0.233	0.233	0.233

Notes: This table replicates column 3 of Table 3, using alternative post-evaluation time windows. * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.24: Effect of Idea Evaluation on Idea Value, Quantile Regressions

	Quantile Regression, DV = Log(Idea Value + 1)				
	Q75	Q80	Q85	Q90	Q95
	(1)	(2)	(3)	(4)	(5)
Post Evaluation	0.357***	0.585***	0.371**	0.315***	0.296**
	(0.064)	(0.103)	(0.116)	(0.095)	(0.105)
Day FE	X	X	X	X	X
Evaluator FE	X	X	X	X	X
Controls	X	X	X	X	X
# Evaluators	4,633	4,633	4,633	4,633	4,633
# Days	4,284	4,284	4,284	4,284	4,284
# Obs.	108,938	108,938	108,938	108,938	108,938
Adj. R^2	0.341	0.285	0.221	0.163	0.111

Notes: This table reports results from specification 4 using quantile regression. The analysis is at the idea-evaluator-day level. Estimations are based on the sample of selected ideas, as the value is only known for these ideas. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.25: Instrumental Variable Strategy - First and Second Stage

	First Stage		OLS	2SLS
DV	Post Evaluation		Idea Generation	
	(1)	(2)	(3)	(4)
Reminder	0.172***	0.170***		
	[320.461]***	[317.740]***		
	(0.010)	(0.010)		
Post Evaluation			0.003***	0.009***
			[63.462]***	[182.692]***
			(0.000)	(0.002)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls		X	X	X
Weighted DV Mean	0.054	0.054	0.005	0.005
# Evaluators	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134
Adj. R^2	0.453	0.454	0.022	0.022
F-stat.	267.373	264.728		

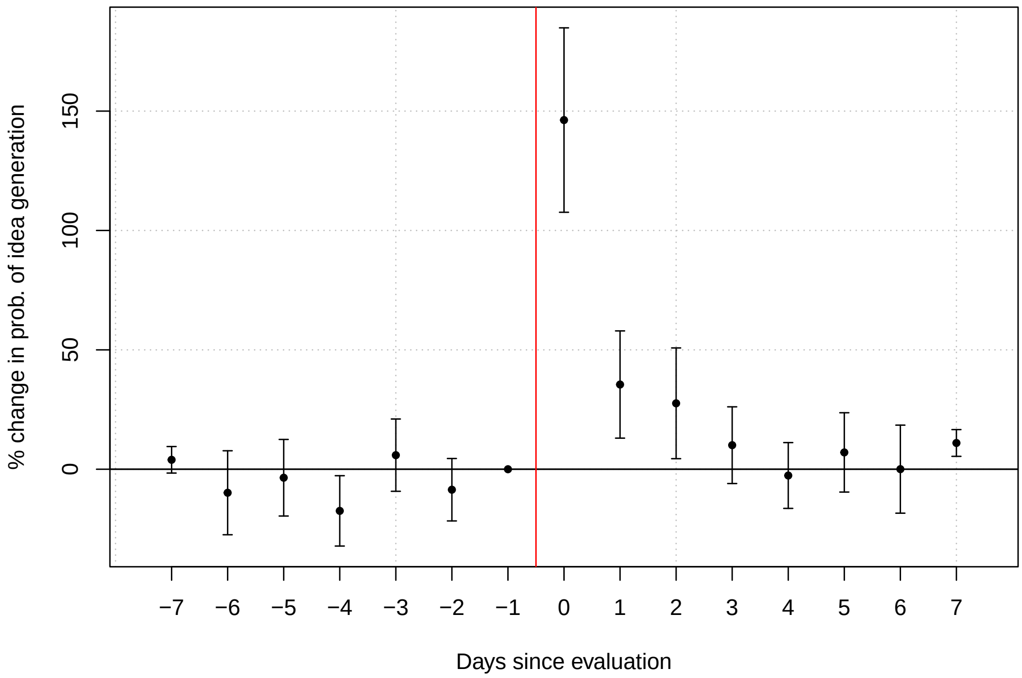
Notes: This table presents an instrumental variable estimation of specification 2. Col. 1 and 2 present coefficient estimates from the first-stage regression at the evaluator-day level: $Post\ \widehat{Evaluation}_{it} = \alpha * Reminder_{it} + \beta \boldsymbol{X}_{it} + \gamma_i + \delta_t + \epsilon_{it}$. Col. 3 presents original OLS estimates, col. 4 2SLS estimates. The F-stat. indicates the first-stage Kleibergen-Paap F-statistic. Coefficient estimates indicate percentage point changes in the dependent variable. Percentage changes, shown in square brackets, are calculated by dividing percentage point changes by the weighted mean of the dependent variable. Robust standard errors, clustered at the evaluator level, are shown in parentheses. Controls are as usual. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.26: Instrumental Variable Strategy - Placebo Test

	OLS Regression		
DV	Idea Generation		
	(1)	(2)	(3)
Reminder Complier	0.005***		0.005***
	[102.580]***		[102.193]***
	(0.001)		(0.001)
Reminder Non-Complier		0.000	0.000
		[−9.095]	[−4.668]
		(0.000)	(0.000)
Day FE	X	X	X
Evaluator FE	X	X	X
Controls	X	X	X
Weighted DV Mean	0.005	0.005	0.005
# Evaluators	12,069	12,069	12,069
# Days	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134
Adj. R^2	0.022	0.022	0.022

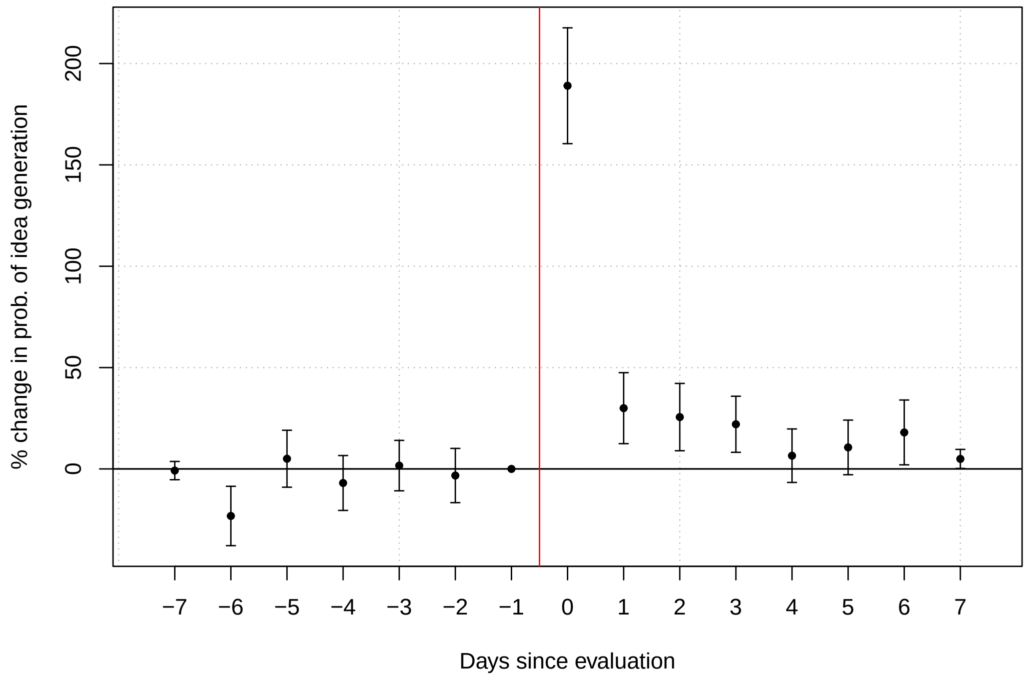
Notes: This table reports results from a placebo test to check whether the instrument, *Reminder*, satisfies the exclusion restriction. *Reminder Complier*$_{it}$ is 1 if evaluator i receives a reminder on day t and evaluates an idea on that same day (i.e., complies with the reminder), and 0 otherwise. *Reminder Non Complier*$_{it}$ is 1 if evaluator i receives a reminder on day t but does not evaluate an idea on that same day (i.e., does not comply with the reminder), and 0 otherwise. Robust standard errors, clustered at the employee level, are shown in parentheses. Controls are as usual. * p < 0.05, ** p < 0.01, *** p < 0.001.

Figure A.12: Robustness, All Evaluator-Day Observations from 2015 onwards



Notes: This figure reports event study results from specification 1 at the evaluator-day level using all evaluator-day observations from 2015 onwards. Estimates are obtained from OLS regression, with robust standard errors clustered at the evaluator level and weights to adjust for choice-based sampling. Controls are as usual. Error bars represent 95% confidence intervals.

Figure A.13: Robustness, Sample Including Non-Evaluators



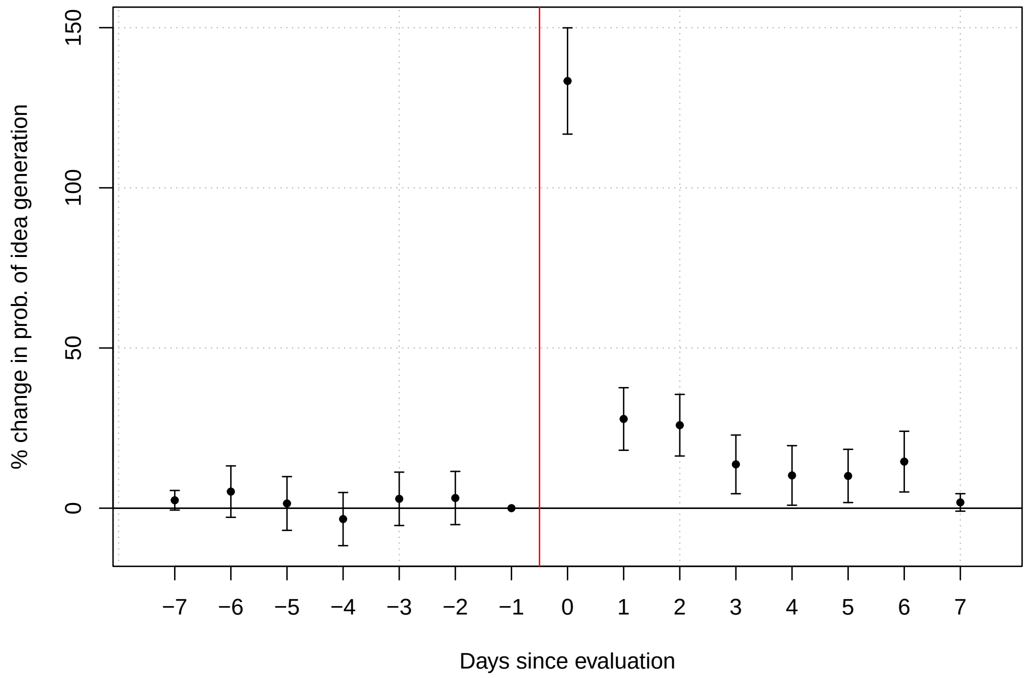
Notes: This figure reports event study results from specification 1 at the evaluator-day level using a choice-based sample that includes non-evaluators in addition to evaluators. Estimates are obtained from OLS regression, with robust standard errors clustered at the evaluator level and weights to adjust for choice-based sampling. Controls are as usual. Error bars represent 95% confidence intervals.

Table A.27: Robustness, Placebo Test Based on Fake Evaluation Dates

	OLS Regression			
DV	Idea Generation		# Ideas	
	(1)	(2)	(3)	(4)
Post Evaluation	0.0001	0.0001	0.0003	0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls		X		X
DV Weighted Mean	0.0052	0.0052	0.0071	0.0071
# Evaluators	12,069	12,069	12,069	12,069
# Days	4,446	4,446	4,446	4,446
# Obs.	419,134	419,134	419,134	419,134
Adj. R^2	0.02151	0.02169	0.01872	0.01892

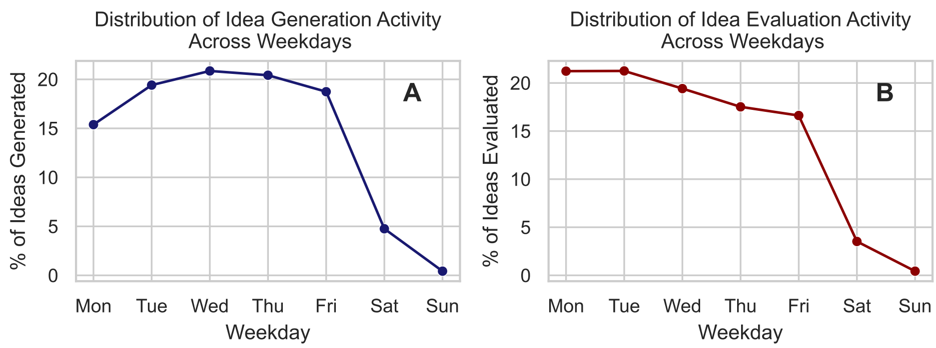
Notes: This table replicates Table A.9 based on fake evaluation dates. * p < 0.05, ** p < 0.01, *** p < 0.001.

Figure A.15: Robustness, Business Days Instead of Weekdays, Event Study



Notes: This figure reports event study results from specification 1 at the evaluator-day level using the number of business days (rather than weekdays) between idea generaion and idea evaluation. Estimates are obtained from OLS regression, with robust standard errors clustered at the evaluator level and weights to adjust for choice-based sampling. Controls are as usual. Error bars represent 95% confidence intervals.

Figure A.14: Robustness, Business Days Instead of Weekdays, Activity Distribution



Notes: This figure plots the distribution of idea generation (Panel A) and idea evaluation (Panel B) activity across weekdays.

Table A.28: Robustness, Coarsened Exact Matching - Covariate Balance

	Matched sample				Unmatched sample			
Variable	Treated	Control	Diff	p-value	Treated	Control	Diff	p-value
# Prev. Gen. Ideas	5.485	5.497	-0.012	0.88	81.752	73.434	8.318	0.01
Cum. Idea Value	8,527	8,003	524.198	0.088	83,265	59,614	23,651	0.01
# Prev. Eval. Ideas	46.976	46.832	0.143	0.713	280.508	21.995	258.513	0.01
Hierarchy	2.133	2.133	0.000	1.0	2.023	1.966	0.057	0.01
Permanent	0.996	0.996	0.000	1.0	0.999	0.966	0.034	0.01
Prev. Mobility	0.645	0.645	0.000	1.0	1.295	1.064	0.231	0.01
Gender	0.084	0.084	0.000	1.0	0.055	0.053	0.002	0.567
N	112,360				419,134			

Notes: This table compares covariate balances in the matched sample and the unmatched sample. The differences in means between the treated and control evaluators are much smaller in the matched than the unmatched sample and statistically insignificant, which provides support for the quality of the matching. We match based on the following covariates: *# Prev. Generated Ideas*, *Cumulative Idea Value*, *# Prev. Evaluated Ideas*, *Hierarchy*, *Prev. Mobility*, *Permanent*, *Gender*. We coarsen *# Prev. Generated Ideas* and *# Prev. Evaluated Ideas* by splitting them into intervals of 5: [0, 5), [5, 10), ... [max.-5, max.]. We coarsen *Cumulative Idea Value* and *Hierarchy* by splitting them into quartiles. We do not coarsen the remaining variables.

Table A.29: Robustness, Coarsened Exact Matching - Event Study

	OLS Regression			
DV	Idea Generation		# Ideas	
	(1)	(2)	(3)	(4)
Days Since Evaluation				
-3	0.001	0.001	0.002	0.002
	[15.471]	[17.694]	[25.894]	[30.315]
	(0.002)	(0.002)	(0.002)	(0.002)
-2	−0.001	−0.001	−0.002	−0.002
	[−9.555]	[−8.150]	[−34.010]	[−31.339]
	(0.001)	(0.001)	(0.002)	(0.002)
0	0.007***	0.007***	0.009***	0.008***
	[98.390]***	[98.033]***	[128.188]***	[127.521]***
	(0.001)	(0.001)	(0.001)	(0.001)
1	0.004*	0.004*	0.005*	0.005*
	[63.156]*	[64.697]*	[77.488]*	[80.579]*
	(0.002)	(0.002)	(0.002)	(0.002)
2	0.004*	0.004*	0.005	0.005
	[55.377]*	[57.414]*	[72.055]	[76.055]
	(0.002)	(0.002)	(0.003)	(0.003)
3	0.005**	0.006**	0.005*	0.005*
	[82.048]**	[83.333]**	[74.214]*	[76.711]*
	(0.002)	(0.002)	(0.002)	(0.002)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls		X		X
DV Mean	0.007	0.007	0.008	0.008
# Evaluators	8,690	8,690	8,690	8,690
# Days	5,129	5,129	5,129	5,129
# Obs.	112,360	112,360	112,360	112,360
Adj. R^2	0.188	0.189	0.205	0.207

Notes: This table replicates Table A.8 based on the matched sample. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.30: Robustness, Coarsened Exact Matching - Single Difference-in-Differences

	OLS Regression			
DV	Idea Generation		# Ideas	
	(1)	(2)	(3)	(4)
Post Evaluation	0.006*** [90.574]*** (0.001)	0.006*** [89.567]*** (0.001)	0.008*** [120.800]*** (0.001)	0.008*** [120.011]*** (0.001)
Day FE	X	X	X	X
Evaluator FE	X	X	X	X
Controls		X		X
DV Mean	0.007	0.007	0.008	0.008
# Evaluators	8,690	8,690	8,690	8,690
# Days	5,129	5,129	5,129	5,129
# Obs.	112,360	112,360	112,360	112,360
Adj. R^2	0.188	0.189	0.205	0.207

Notes: This table replicates Table A.9 based on the matched sample. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.